Forschungsbericht 2025-21

Urban Science with a View from Space - Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics

Hannes Taubenböck

Deutsches Zentrum für Luft- und Raumfahrt
Deutsches Fernerkundungsdatenzentrum
Oberpfaffenhofen

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Forschungsbericht 2025-21

Urban Science with a View from Space – Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics

Hannes Taubenböck

Deutsches Zentrum für Luft- und Raumfahrt
Deutsches Fernerkundungsdatenzentrum
Oberpfaffenhofen

697 Seiten
174 Bilder
37 Tabellen
1678 Literaturstellen

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Herausgeber:

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Wissenschaftliche Information
Linder Höhe
D-51147 Köln

ISSN 1434-8454
ISRN DLR-FB-2025-21
Erscheinungsjahr 2025
DOI: 10.57676/75k5-re54

Erklärung des Herausgebers
Dieses Werk – ausgenommen anderweitig gekennzeichneter Teile – ist lizensiert unter den Bedingungen der Creative Commons Lizenz vom Typ Namensnennung-Nicht kommerziell-Keine Bearbeitungen 4.0 International
(CC BY-NC-ND 4.0), abrufbar über https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode

Lizenz
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International

Urbanisierung, Siedlungsstrukturen, Raummuster, Erdbeobachtung, Fernerkundung, Gesellschaft

Hannes TAUBENBÖCK
DLR, Deutsches Fernerkundungsdatenzentrum, Oberpfaffenhofen

Stadtforschung mit Blick aus dem Weltraum – Dynamik, Dimensionen, Muster, Auswirkungen, Anpassung, Ethik

Im Jahr 2024 lebten 58 % der Weltbevölkerung in Städten. Städte, ob groß oder klein, sind der primäre Lebensraum für mehr als die Hälfte der Menschen und sie fungieren als politische, soziale, wirtschaftliche, kulturelle und ökologische Zentren. Mit der Art und Weise, wie wir die gebauten Strukturen und die Nutzungsmuster gestalten, haben wir einen direkten Einfluss auf verschiedene Variablen. Auf individueller Ebene bestimmen die Strukturen den Alltagsrhythmus, das Verhalten und die Wahrnehmung. Auf gesellschaftlicher Ebene zeigt sich die Relevanz der räumlichen Strukturen in ihren Auswirkungen auf den Flächenverbrauch, die Treibhausgasemissionen, den städtischen Wärmeinseleffekt, die Art der Mobilität und viele andere Themen. In den letzten Jahren haben die rasanten Entwicklungen in der Datentechnologie, der Rechenleistung und den Methoden in der Fernerkundung und anderen geografischen Bereichen es ermöglicht, urbane Strukturen mit einer noch nie dagewesenen Menge an Daten empirisch und global zu analysieren. In diesem Forschungsbericht wird die gebaute urbane Landschaft auf unserem Planeten in ihrer zeitlichen Entwicklung dokumentiert, die entstehende Dynamik systematisiert, urbane Formen verglichen, Auswirkungen der entstehenden räumlichen Muster bewertet und Möglichkeiten der städtischen Anpassung vorgestellt. Ein grundlegendes Thema ist der Begriff 'urban' an sich, wie er theoretisch und konzeptionell definiert wird und wie er raumbezogen gemessen und kartiert werden kann. Zudem werden die methodischen und thematischen Arbeiten auch unter ethischen Gesichtspunkten bewertet. Der Forschungsbericht basiert auf 24 veröffentlichten Arbeiten und einer Synthese, die aus den Einzelstudien eine Roadmap für zukünftige Forschungsrichtungen der urbanen Fernerkundung schafft.

Urbanization, settlement structures, spatial pattern, earth observation, remote sensing, society
(Published in English)

Hannes TAUBENBÖCK
German Aerospace Center (DLR), German Remote Sensing Data Center, Oberpfaffenhofen

Urban Science with a View from Space – Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics

In 2024, 58 % of the global population is urban. Cities, whether large or small, are the primary living space for the majority of humanity and are political, social, economic, cultural and ecological centers. The way we design cities has an impact on all these issues with the built city as the most direct feature. The structures of the built city and the patterns of use have a huge impact on an individual and a societal level. On individual level, the structures define everyday rhythms, behaviors and perceptions. On societal level, the relevance of spatial structures and patterns of urban land is given due to their impact on land consumption, greenhouse gas emissions, the urban heat island effect, mobility types, among many other issues. In recent years, rapid developments in data technology, computing power and methodologies in remote sensing and other geographic domains have made it possible to analyze urban structures empirically across the globe with an unprecedented amount of data. In this research report, the built urban landscape around the globe is documented in its temporal development, the emerging dynamics are systemized, urban forms are compared, effects of the emerging spatial patterns are evaluated and capabilities for urban adaptation are presented. An underlying theme is the term ‘urban’ in itself, how it can be defined theoretically and conceptionally, and how it can be measured and mapped in spatial domains. The methodological and thematic works are also evaluated from an ethical perspective. The research report is based on 24 published papers and a synthesis that creates from the individual studies a road map for future research paths in urban remote sensing.

Preface

When you walk through a lively city center on a warm summer evening, and you hear the murmur of voices that enliven the cafés on the streets and music in the background, it creates an atmosphere that I love. People come together, people exchange ideas, people look each other in the eye and the city becomes a social meeting place, a positive feeling of togetherness, a celebration of life. I am fully aware that not everyone feels this way or has the opportunity to simply participate in it in everyday life. I am also aware that this density, this social interaction, cannot always take place without conflict or with the naively positive attitude to life described above. Two of my best friends constantly try to convince me that Munich is too crowded. But for me, who grew up in the countryside, the density, diversity, simultaneity and energy of large cities, the pulse of its very different people and the associated endless possibilities were always mystical.

By now, I have had the privilege of dealing professionally with the topic of '*the urban*' for more than 20 years. Probably, I have just been trying to make this mystical space more tangible for myself. Or, if you grant me higher ideals, to contribute to better cities through better knowledge. I have to admit this endeavor was not always crowned with success; there remains a gap between what can be rationalized scientifically and the mystic atmospheres and feelings that are created by places and its people.

There is always hope – and still I do have this hope today – that our studies published in scientific journals can make a relevant contribution to knowledge. Knowledge that allows to better understand processes and developments in urban societies. Knowledge that helps make cities a better place. As interesting as some scientific questions may be, the various studies, however, often lead to loose ends. Sometimes I have the feeling that a larger, coherent scientific picture – even only within my own, very limited research work – is lost if these loose ends remain untied. And, a larger, coherent scientific picture is difficult to derive because it is usually a long time ago that thinking took place and results were left without consistently taking the next steps. This makes it all the more challenging to derive a coherent whole from the past individual parts.

With this in mind, and with the ambition to find out how far the understanding may go, I was keen to compile these papers on the topic of '*the urban*' in general, and with analyses based primarily on '*remote sensing*' data in particular. As all the studies have already been published in international journals, one cannot claim that there are any new scientific findings here. So please forgive me if I undertake this compilation primarily for myself to form out of the individual pieces a larger puzzle. This has two aims: when the individual pieces of the puzzle are put together, hopefully a more coherent picture is created that was

previously invisible. And this picture – so at least my hope – allows to think more clearly about the past research findings as a basis to identify the right future research directions.

Over the years, research work is produced that sometimes takes a more conceptual, sometimes a more methodological, sometimes a more thematic direction. Maintaining a common thread is not always easy under the current system of research funding. For me it is a great privilege to be able – despite some detours – to work on a common thread over a long period of time. The here compiled studies all focus in one way or another on the built urban space, its development over time, the structures that have emerged and their corresponding effects, and also pose the fundamental question of what '*the urban*' actually is. They represent a selection of studies published between 2019 and 2025. This compilation is neither complete, nor is the selection based on my own's assessment of their quality. They have been put together, because they fit together thematically.

Of course, I am aware that this research report could follow the path from 'unconnected and unread work' to 'connected and unread work'. And actually, I should already know by now, because when I compiled works from 2014 to 2018 in a similar form for my habilitation, the synthesis – i.e. the envisaged more coherent picture of many individual works – probably remained largely unread. And so, I come back to what was said above: I undertake this compilation primarily for myself.

And still, in the deep conviction that we need to take more targeted care of '*the urban*', which is already the living space for the majority of people on our planet today and will be the center of life for even more people in the future, the understanding of the built form is crucial. The built form has an impact in political, social, economic or ecological terms and quite simply defines all the everyday routines of city dwellers. Ultimately, the question that stands above everything else is – how do we adapt and design the cities of the future so that they are worth living in, socially balanced, as well as economically prosperous, ecologically sustainable and resilient to changes? The hope remains that the compiled studies broaden our understanding of the city, both individually, but above all in their synopsis.

Oberpfaffenhofen, 10.09.2025

Acknowledgements

Science has to question its own latest findings over and over again. This constant questioning of your results, of their embedding to other studies, of their significance for society, and ultimately of yourself can be exhausting. It is, however, an incredible privilege to question and discuss the diverse latest findings and challenges over and over again with so many highly motivated and committed scientists who are bursting with new and fresh ideas. I consider it a great honor to develop these ideas with all of you for better knowledge.

I would like to dedicate this book first and foremost to my mentor and friend Prof. Dr. Stefan Dech. It is only thanks to his trust and permanent support over more than 20 years that allowed to follow this line of research. A heartfelt 'thank you very much' is not enough, I know, but I fear I just do not have better words.

A big thank you to all colleagues at DLR, and especially to the colleagues within our 'Georisks and Civil Security' department, who create this cordial, constructive and highly motivated atmosphere. Special thanks go to the ever-inspiring critical urban thinkers Prof. Dr. Christian Geiß, Dr. Michael Wurm, Dr. Tobias Leichtle, Dr. Marta Sapena, Dr. Thilo Erbertseder, Dr. Nicolas J. Kraff, Prof. Dr. Thomas Esch, Martin Habermeyer, Dr. Isabel Georg, Dr. Matthias Weigand, Dr. Ines Standfuß, Dr. Jeroen Staab, Johannes Mast, Lorenza Gilardi, Marina Reiter and Christoph Otto, and to Dr. Monika Gähler, Dr. Sandro Martinis, Dr. Tanja Kraus, Dr. Doris Klein, Mihaela Blaskovic and Prof. Dr. Günter Strunz for the continuous trustful co-operation. Special thanks goes to my deputy Dr. Torsten Riedlinger for his tireless efforts taking care of the numerous big and small things to keep the wheel turning. And, special thanks to Dr. Elisabeth Schöpfer: without your invaluable support, this report would never have been completed.

A big thank you to all colleagues at the University Würzburg, and especially to the colleagues at the Earth Observation Research Cluster: To Dr. Martin Wegmann, Dr. Michael Thiel and Prof. Dr. Tobias Ullmann, Christine Linge, Sabine Oppmann and the late Susanne Hedler who take care of everything – and really everything, to the ever-inspiring creative urban minds Dr. Ariane Droin, Henri Debray, Dr. Richard Lemoine-Rodriguez, Andrea Sofia Garcia de Leon, De-Cyuan Jheng, Dr. John Friesen, Konstantin Müller, Prof. Dr. Heiko Paeth, Dr. Lisa Lehnen, Dr. Ninja Schulz and Prof. Dr. Carolin Biewer. And to the too many to mention here at DLR and at the University Würzburg who work in the background and foreground on organizational and content-related issues to keep the wheel turning.

I would particularly thank all colleagues and co-authors all across the world that inspired with an incredible richness of knowledge, ideas and know-how the academic urban debate. Only in collaboration it is possible to develop all these arguments, perspectives, data, methods and geographic findings: Prof. Dr. Thomas Blaschke from the Paris-Lodron University Salzburg, Dr. Tirthankar Chakraborty from the Pacific Northwest National Laboratory in Richland, Prof. Dr. Chengbin Deng from the University of Oklahoma, Dr. Fabian Dosch, Dr. Antonia Milbert and Silas Eichfuss from the Federal Institute for Research on Building, Urban Affairs and Spatial Development, Dr. Gregory Duveiller from the Max Planck Institute for Biogeochemistry in Jena, Dr. Jan Goebel and the late Prof. Dr. Gert G. Wagner from the DIW Berlin, Prof. Dr. Xianjin Huang and Prof. Dr. Taiyang Zhong from the Nanjing University, Dr. habil. Luis Inostroza from Mendel University Brno, Tamara Iungman, Dr. Sasha Khomenko, Evelise Pereira Barboza, Marta Cirach, Dr. Paula Petrone, Dr. Karen Gonçalves, and Prof. Dr. Mark Nieuwenhuijsen from the Barcelona Institute for Global Health, Dr. Emanuele Massaro, Alessandro Cescatti, Luca Caporaso, Matteo Piccardo from the Joint Research Center, Prof. Dr. Steffen Mau and Dr. Fabian Gülzau from the Humboldt University of Berlin, Dr. Steward T.A. Pickett from the Cary Institute of Ecosystem Studies, Dr. Nikola Sander from the Federal Institute for Population Research, Dr. Rossano Schifanella from the University of Turin, Prof. Dr. Andreas Schmitt und Philipp Uth from the University of Applied Sciences Munich, Prof. Dr. Michael Schmitt from the University of the Bundeswehr München, Prof. Dr. Karen C. Seto from Yale University, Prof. Dr. Stefan Siedentop, Dr. Christian Gerten, Prof. Dr. Karsten Rusche, Dr. Bastian Heider from the Research Institute for Regional and Urban Development Dortmund and the Technical University Dortmund, Dr. Lifeng Shi from the China University of Mining and Technology, Dr. Eleanor C. Stokes from NASA, Prof. Dr. Yuyu Zhou from the University Honkong, Prof. Dr. Zhe Zhu from the University of Connecticut and Prof. Dr. Xiao Xiang Zhu, Prof. Dr. Mrinalini Kochupillai, Dr. Chunping Qiu, Dr. Jingliang Hu, Dr. Yilei Shi, Prof. Dr. Yuanyuan Wang and Matthias Kahl from the Technical University Munich, Dr. Justice Nana Inkoom from the University Halle-Wittenberg, Marlene Kühnl from SLU, Dr. John Singer from AlpGeorisk, Dr. Moritz Gamperl from the Technical University of Munich, and Carolina Garcia-Londoño from the Geological Society of Colombia.

And, I would especially like to thank my parents, Hans and Ingrid Taubenböck, and my sister Andrea, who have always supported me in all my meaningful and even nonsensical endeavors and have made it possible for me to pursue my interests. And last, but not least: To my wonderful family – Kerstin, Luisa and Alina – that they put up with my love for the city and show me that happiness is not bound to a particular place.

Table of contents

Preface			i
Acknowledgements			iii
Table of contents			v
List of papers contributing to this research report			xi
SYNTHESIS	**URBAN SCIENCE WITH A VIEW FROM SPACE – DYNAMICS, DIMENSIONS, PATTERNS, EFFECTS, ADAPTATION, ETHICS**		
CHAPTER I	1.	Introduction	3
	2.	Research goals	5
	3.	Structure of the research report	7
	4.	The main findings	8
	4.1.	*Part I*: Strategic directions for urban remote sensing as starting point (papers related to Part I)	8

4.2. *Part II:* What 'urban' is (papers related to Part II) 9

4.3. *Part III:* Urban dynamics (papers related to Part III) 11

4.4. *Part IV:* Urban dimensions and rankings (papers related to Part IV) 12

4.5. *Part V:* Urban patterns (papers related to Part V) 14

4.6. *Part VI:* Effects of urban patterns (papers related to Part VI) 16

4.7. *Part VII:* Urban adaptation (papers related to Part VII) 17

4.8. *Part VIII:* Ethical considerations (papers related to Part VIII) 19

5. Embedding the research report into urban research 19

6. A road map of strategic key research directions for urban remote sensing 22

7. Urban Science with a View from Space: A conclusion 28

PART I **STRATEGIC DIRECTIONS FOR URBAN REMOTE SENSING AS STARTING POINT**

CHAPTER II Understanding an urbanizing planet: Strategic directions for remote sensing 31

PART II **WHAT '*URBAN*' IS**

CHAPTER III To be, or not to be '*urban*'? A multi-modal method for the differentiated measurement of the degree of urbanization 61

CHAPTER IV Spatial delineation of urban corridors in North America: An approach incorporating fuzziness based on multi-source geospatial data 83

CHAPTER V Quantitative assessment and comparison of urban patterns in Germany and the United States 109

CHAPTER VI Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception 139

CHAPTER VII Searching for the DNA of urbanisation. A material perspective 153

PART III **URBAN DYNAMICS**

CHAPTER VIII Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios 171

CHAPTER IX The digital urban frontier: Disparities in social media activity between consolidated and newly urbanized areas in Africa 201

CHAPTER X Border regions across the globe: Analyzing border typologies, economic and political disparities, and development dynamics 227

CHAPTER XI NO_2 air pollution trends and settlement growth in Megacities 253

PART IV	**URBAN DIMENSIONS AND RANKINGS**	
CHAPTER XII	The urban morphology on our planet – global perspectives from space	283
CHAPTER XIII	A new ranking of the world's largest cities – do administrative units obscure morphological realities?	303
CHAPTER XIV	Which city is the greenest? A multi-dimensional deconstruction of city rankings	327
PART V	**URBAN PATTERNS**	
CHAPTER XV	Seven city types representing morphologic configurations of cities across the globe	349
CHAPTER XVI	Planned, unplanned, or in-between? A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe	375
CHAPTER XVII	Patterns of Eastern European urbanisation in the mirror of Western trends – convergent, unique or hybrid??	405
CHAPTER XVIII	Measuring the spatial hierarchical urban system in China in reference to the Central Place Theory	427
PART VI	**EFFECTS OF URBAN PATTERNS**	
CHAPTER XIX	The impact of urban configuration types on urban heat islands, air pollution, CO_2 emissions, and mortality in Europe: A data science approach	447
CHAPTER XX	Land consumption in cities: A comparative study across the globe	469

CHAPTER XXI How we live and what that means – a character study with data from space 495

PART VII URBAN ADAPTATION

CHAPTER XXII Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes 503

CHAPTER XXIII Cost estimation for the monitoring instrumentation of landslide early warning systems 519

CHAPTER XXIV Was global urbanization from 1985 to 2015 efficient in terms of land consumption? 545

PART VIII ETHICAL CONSIDERATIONS

CHAPTER XXV Earth Observation and Artificial Intelligence: Understanding emerging ethical issues and opportunities 569

REFERENCES 621

Table of contents

List of papers contributing to the research report

Part I to Part VIII of the research report consist of 23 SCI papers and 1 reviewed conference paper. Presented below is a list organized in chronological order based on the release dates, with each entry also indicating its respective license. The published papers have been converted to a unified format. Appendices and additional material can be found, where applicable, in the original (online) versions. The work is framed by a synthesis. All references have been compiled and can be found at the end of the report.

1. Mast, J., Sapena, M., Debray, H., Inkoom, J.N., Lemoine-Rodríguez, R., Geiß, C., & Taubenböck, H. (2025). The digital urban frontier: Disparities in social media activity between consolidated and newly urbanized areas in Africa. *Applied Geography*, *181*(103687), 1–15. https://doi.org/10.1016/j.apgeog.2025.103687

   © 2025 German Aerospace Center (DLR). Published by Elsevier Ltd. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

2. Taubenböck, H., Mast, J., Lemoine Rodríguez, R., Debray, H., Wurm, M., & Geiß, C. (2025). Was global urbanization from 1985 to 2015 efficient in terms of land consumption? *Habitat International*, *160*(103397), 1–15. https://doi.org/10.1016/j.habitatint.2025.103397

   © 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

3. Erbertseder, T., Taubenböck, H., Esch, T., Gilardi, L., Paeth, H., & Dech, S. (2024). NO_2 Air Pollution Trends and Settlement Growth in Megacities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 12058–12076. https://doi.org/10.1109/JSTARS.2024.3419573

   © 2024 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

4. Inostroza, L., & Taubenböck, H. (2024). Searching for the DNA of urbanisation. A material perspective. *Cities*, *151*(105079), 1–10. https://doi.org/10.1016/j.cities.2024.105079

   © 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

5. Iungman, T., Khomenko, S., Barboza, E.P., Cirach, M., Gonçalves, K., Petrone, P., Erbertseder, T., Taubenböck, H., Chakraborty, T., & Nieuwenhuijsen, M. (2024). The impact of urban configuration types on urban heat islands, air pollution, CO_2 emissions, and mortality in Europe: A data science approach. *The Lancet Planetary Health*, *8*(7), e489–e505. https://doi.org/10.1016/S2542-5196(24)00120-7

© 2024 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

6. Taubenböck, H., Mast, J., Geiß, C., Wurm, M., Esch, T., & Seto, K.C. (2024). Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios. *Cities*, *151*(105117), 1–31. https://doi.org/10.1016/j.cities.2024.105117

© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

7. Debray, H., Kraff, N.J., Zhu, X.X., & Taubenböck, H. (2023). Planned, unplanned, or in-between? A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe. *Landscape and Urban Planning*, *233*(104711), 1–27. https://doi.org/10.1016/j.landurbplan.2023.104711

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

8. Georg, I., Blaschke, T., & Taubenböck, H. (2023). Spatial delineation of urban corridors in North America: An approach incorporating fuzziness based on multi-source geospatial data. *Cities*, *133*, 1–16. https://doi.org/10.1016/j.cities.2022.104129

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

9. Massaro, E., Schifanella, R., Piccardo, M., Caporaso, L., Taubenböck, H., Cescatti, A., & Duveiller, G. (2023). Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes. *Nature Communications*, *14*(2903), 1–10. https://doi.org/10.1038/s41467-023-38596-1

© The Author(s) 2023, corrected publication 2023. This article is licensed under a Creative Commons Attribution 4.0 International License. For more information, see https://creativecommons.org/licenses/by/4.0/.

10. Sapena, M., Gamperl, M., Kühnl, M., Garcia-Londoño, C., Singer, J., & Taubenböck, H. (2023). Cost estimation for the monitoring instrumentation of landslide early warning systems. *Natural Hazards and Earth System Sciences*, *23*(12), 3913–3930. https://doi.org/10.5194/nhess-23-3913-2023

© Author(s) 2023. This work is distributed under the Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.

11. Schmitt, A., Uth, P., Standfuß, I., Heider, B., Siedentop, S., & Taubenböck, H. (2023). Quantitative assessment and comparison of urban patterns in Germany and the United States. *Computers, Environment and Urban Systems*, *100*(101920), 1–18. https://doi.org/10.1016/j.compenvurbsys.2022.101920

© 2022 Elsevier Ltd. All rights reserved.

Note: The paper has been utilized in this research report in accordance with the reuse rights granted to the authors as outlined at: https://www.elsevier.com/about/policies-and-standards/copyright.

12. Taubenböck, H., Otto, C., Gülzau, F., & Mau, S. (2023). Border regions across the globe: Analyzing border typologies, economic and political disparities, and development dynamics. *Applied Geography*, *151*(102866), 1–15. https://doi.org/10.1016/j.apgeog.2022.102866

    © 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

13. Kochupillai, M., Kahl, M., Schmitt, M., Taubenböck, H., & Zhu, X.X. (2022). Earth Observation and Artificial Intelligence: Understanding emerging ethical issues and opportunities. IEEE Geoscience and Remote Sensing Magazine, 10(4), 90–124. https://doi.org/10.1109/MGRS.2022.3208357

    ©2022 IEEE

    Note: A payment was made to CCC RightsLink to reuse the paper in this report.

14. Taubenböck, H., Droin, A., Standfuß, I., Dosch, F., Sander, N., Milbert, A., Eichfuss, S., & Wurm, M. (2022). To be, or not to be 'urban'? A multi-modal method for the differentiated measurement of the degree of urbanization. *Computers, Environment and Urban Systems*, *95*(101830), 1–12. https://doi.org/10.1016/j.compenvurbsys.2022.101830

    © 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

15. Zhu, X.X., Qiu, C., Hu, J., Shi, Y., Wang, Y., Schmitt, M., & Taubenböck, H. (2022). The urban morphology on our planet – Global perspectives from space. *Remote Sensing of Environment*, *269*(112794), 1–11. https://doi.org/10.1016/j.rse.2021.112794

    © 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

16. Hu, J., Wang, Y., Taubenböck, H., & Zhu, X.X. (2021). Land consumption in cities: A comparative study across the globe. *Cities*, *113*(103163), 1–17. https://doi.org/10.1016/j.cities.2021.103163

    © 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

17. Taubenböck, H. (2021). How We Live and What That Means – A Character Study with Data from Space. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 1190–1193. https://doi.org/10.1109/IGARSS47720.2021.9553084

    © 2021 IEEE.

    Note: A payment was made to CCC RightsLink to reuse the paper in this report.

18. Taubenböck, H., Reiter, M., Dosch, F., Leichtle, T., Weigand, M., & Wurm, M. (2021). Which city is the greenest? A multi-dimensional deconstruction of city rankings. *Computers, Environment and Urban Systems*, *89*, 101687. https://doi.org/10.1016/j.compenvurbsys.2021.101687

    © 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

19. Wurm, M., Goebel, J., Wagner, G.G., Weigand, M., Dech, S., & Taubenböck, H. (2021). Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception. *Environment and Planning B: Urban Analytics and City Science*, *48*(2), 265–279. https://doi.org/10.1177/2399808319869341

    © The Author(s) 2019.

20. Shi, L., Wurm, M., Huang, X., Zhong, T., & Taubenböck, H. (2020). Measuring the spatial hierarchical urban system in China in reference to the Central Place Theory. *Habitat International*, *105*(102264), 1–13. https://doi.org/10.1016/j.habitatint.2020.102264

    © 2020 Elsevier Ltd. All rights reserved.

    Note: The paper has been utilized in this research report in accordance with the reuse rights granted to the authors as outlined at: https://www.elsevier.com/about/policies-and-standards/copyright.

21. Taubenböck, H., Debray, H., Qiu, C., Schmitt, M., Wang, Y., & Zhu, X.X. (2020). Seven city types representing morphologic configurations of cities across the globe. *Cities*, *105*(102814), 1–15. https://doi.org/10.1016/j.cities.2020.102814

    © 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

22. Taubenböck, H., Gerten, C., Rusche, K., Siedentop, S., & Wurm, M. (2019). Patterns of Eastern European urbanisation in the mirror of Western trends – Convergent, unique or hybrid? *Environment and Planning B: Urban Analytics and City Science*, *46*(7), 1206–1225. https://doi.org/10.1177/2399808319846902

    © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

23. Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J., & Dech, S. (2019). A new ranking of the world's largest cities – Do administrative units obscure morphological realities? *Remote Sensing of Environment*, *232*(111353), 1–14. https://doi.org/10.1016/j.rse.2019.111353

    © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

24. Zhu, Z., Zhou, Y., Seto, K.C., Stokes, E.C., Deng, C., Pickett, S.T.A., & Taubenböck, H. (2019). Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sensing of Environment*, *228*, 164–182. https://doi.org/10.1016/j.rse.2019.04.020

    © 2019 Elsevier Inc. All rights reserved.

    Note: The paper has been utilized in this research report in accordance with the reuse rights granted to the authors as outlined at: https://www.elsevier.com/about/policies-and-standards/copyright.

Synthesis

Urban Science with a View from Space

Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics

Chapter I

Urban Science with a View from Space

Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics

1. Introduction

In 2023, every single week an estimated 1.5 million people were added to cities worldwide. This means that the global urban population grew by almost 80 million in just one year (UN, 2017). That is about the entire population of Germany and testifies to the gigantic dimension of urban growth – Saunders (2010) puts it in a nutshell that humankind is within the largest migration ever: out of rural agricultural life into cities.

In 2024, 58 % of the global population was urban (UN, 2019; World Bank, 2025). Cities, whether large or small, are already the primary living space for the majority of humanity. Figure 1 illustrates the development of the quantity and spatial distribution of cities across the globe over 70 years from 1950 to 2020. The figure exemplifies all cities larger than 750,000 inhabitants. In 1950, less than one hundred cities had that size with the majority in Western countries – in 2020, it was already about 1,600 with the majority now in Asia. This rapid trend of urbanization will continue (UN, 2019).

An attempt to explain this trend towards cities could begin with this quote: "*Nobody is so brilliant that they could come up with something really valuable all by themselves. Ingenious personalities and cultures collect, make their own, utilize the greatest achievements*" (Pinker, 2018). Places where people densely cohabitate, as Jane Jacobs has argued, are the "*well-spring of innovation*" as new ideas largely develop from the combination of (many) old ideas (Jacobs, 1970). One of the most important advantages in this regard comes from speeding up the flow of ideas (Glaeser & Joshi-Ghani, 2015). In a very abbreviated form, this explains why the sources of knowledge, culture and development were always there where many people cavort – the advantages of density. These advantages thus have predominantly developed in (larger) cities, mostly trading towns at important crossroads (Glaeser, 2011) and led to innovation (Carlino, Chatterjee & Hunt, 2007) and productivity gains (Anderson, Burgess & Lane, 2007; Florida, 2002). And consequently, on a global scale, the central relevance of cities is demonstrated by the facts that they, for example, account for more than 80 % of the gross domestic product (World Bank, 2022), with added value from scaling effects (West, 2017). Parameters such as the number of patents, the gross domestic product, the average income, the number of highly creative groups scale superlinear with the size of cities (Anderson, Burgess & Lane, 2007; Carlino, Chatterjee & Hunt, 2007; Bettencourt, 2013; West, 2017). The larger cities are, the more innovative social capital is created. This somewhat simplistic basic logic, still plausibly explains why people have always been keen travelers in search of a better life (Mau, 2021) – and for the vast majority, cities credibly promise this (Hollis, 2013): "*Trees put down roots, people have feet*" (Pinker, 2018).

However, the number of '*travelers*' from rural to urban areas is currently higher than the number of people who can enjoy the benefits of civilized urbanity in their lifetime (Sloterdijk, 2018). Two general issues are here in conflict: on the one hand, the hopes and promises that urbanity will bring prosperity and actually a better life (e.g. Bettencourt et al., 2007; Glaeser, 2011; Montgomery, 2013). And on the other hand, that the dreams and desires projected and released into urbanity do not always result in genuine integration, participation and prospertiy (e.g. Davis, 2007; Hollis, 2013). Saunders (2010) has argued very

convincingly that cities can make '*arrival*' possible – the city as a place for economic rise, social integration, and in general, the use of personal opportunities. However, he has also shown that urban areas can also be places of failure, descent and hopelessness.

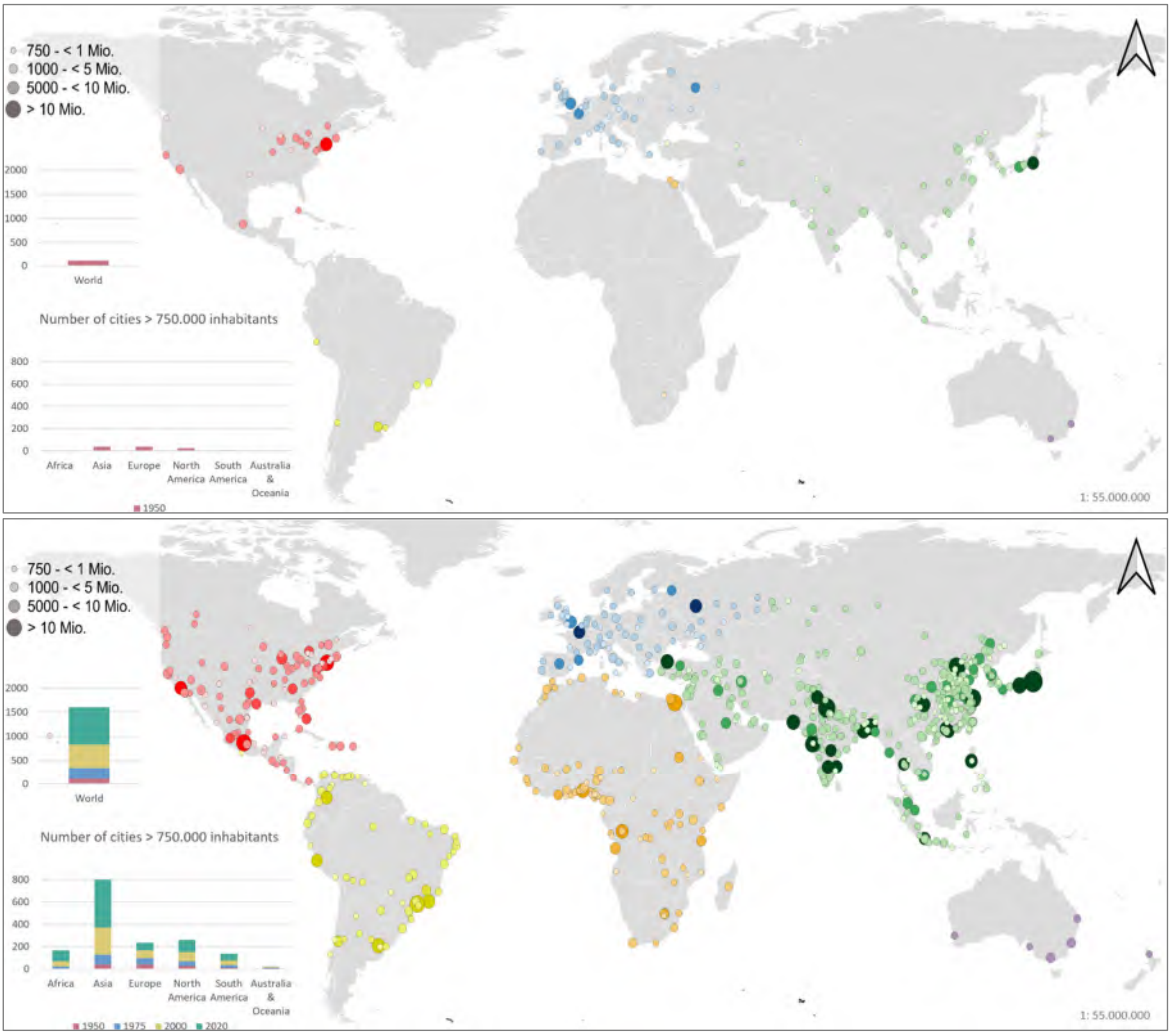


Fig. 1. Quantity and spatial global distribution of cities larger than 750,000 inhabitants in 1950 (top) and in 2020 (bottom) based on the United Nations population statistics (UN, 2017). Source: own elaboration, unpublished.

The relevance of cities on a global scale is also demonstrated by the facts that they account for 66 % of global energy consumption (World Bank, 2022) and contribute two-thirds of global greenhouse gas (GHG) emissions (IPCC, 2022). And, scaling effects that ensure more innovative social capital (as argued above) also show superlinear behavior for undesirable parameters such as environmental pollution, crime rates, risk of disease and much more (Bettencourt, 2013; West, 2017). It is therefore almost a truism to say that the design of the urban living space is of enormous importance, now and even more in the future.

The structures of the built city and the patterns of use determine various issues on an *individual* as well as on a *societal level*: On an *individual level* the structures define our everyday rhythms, our behaviors and our perceptions (Glaeser, 2011; Batty, 2013; Tonkiss, 2013; Sennett, 2018; among many more). Urban form is thus a key element as a social-ecological-technological hybrid (Grimm et al., 2016): It localizes and structures where people live and work, and it influences mobility, interactions, processes, and networks

(Jacobs, 1961; Bramley & Power, 2009; Ewing et al., 2014). From a planning perspective – as the spatial impacts the social (Gehl, 2009; Salat, 2011; Sloterdijk, 2013) – the design of cities allows to influence the usage or perception of space with various forms that have been documented (Jacobs, 1961; Jackson, 1985; Kostof, 1991). As simple examples, low dense urban sprawl strongly relates to private transport in cars (Glaeser & Kahn, 2004) and a higher per person CO_2 emission, high dense urban areas promote walking (Droin et al., 2023) and lower personal CO_2 footprints (Iungman et al., 2024).

On a larger, *societal level*, the relevance of spatial structures and patterns of urban land is given due to their impact on land consumption (Zambon, Serra & Salvati, 2019), land use efficiency (Schiavina et al., 2022), greenhouse gas emissions (IPCC, 2014), the urban micro-climate (Meng et al., 2024), the urban heat island effect (Leichtle et al., 2023), mobility types and patterns (Kang et al., 2012), energy demand (Madlener & Sunak, 2011), crime (Patino et al., 2014), quality of life (Sapena et al., 2021), sustainability (Mobaraki & Oktay Vehbi, 2022), among many other issues.

All these examples show the multi-scale relevance – from local to global and from the individual to society – that cities and their spatial and functional organization have an immense influence on political, economic, social or ecological development goals (e.g. the Sustainable Development Goals by the UN, 2025). This argumentation is only meant here to serve as a basic line of reasoning to bring together more knowledge about the built urban space, its evolution over time, the emerging patterns in different parts of the world, the impact of these patterns and the possibilities for urban adaptation.

However, it is only in recent years that rapid developments in data technology, processing power and methodologies in remote sensing and other geographic domains have made it possible to document and analyze these urban structures both in high spatial resolution and empirically over large areas as well as over time (e.g. Esch et al., 2012; Melchiorri et al., 2018; Marconcini et al., 2018; Zhu et al., 2019; Li et al., 2023; Huang et al., 2021; Yu & Fang, 2023). It is thus a possibility to respond to the United Nation's call for better data (UN, 2025). And this is precisely the general objective of this research report. To create a larger whole from new data sets and the many individual studies on urbanization, and thus provide more and better knowledge using the view from space.

2. Research goals

This research report follows on from a previous research report published in 2019 titled '*Remote Sensing for the Analysis of Global Urbanization*' (Taubenböck, 2019). One of the main outlooks given in this past report was "*the plea towards the urban remote sensing community to no longer neglect research driven by urban geographic questions instead of limiting the research discipline to methodological classification challenges*". It is precisely this plea that is consistently taken up in this current research report. It is an attempt to consistently answer geographical questions with the data and methods of remote sensing for urban areas.

This present research report is a direct follow up titled *"Urban Science with a View from Space – Dynamics, Dimensions, Patterns, Effects, Adaptation, Ethics"*. New methods in the

domains of machine learning and artificial intelligence for pre-processing or classifying remote sensing data have developed rapidly (e.g. Zhu et al., 2017; Reichstein et al., 2019; Li et al., 2023; Tuia et al., 2024; Yu & Fang, 2023; Zhou & Weng, 2024, among many more). They are the basis for the manifold geoinformation products – from global settlement masks to intra-urban classifications to very high-resolution three-dimensional city models – that are primarily relied on in this research report. In addition, manifold works in benchmarking and validation of the resulting geoinformation products have been introduced (e.g. Klotz et al., 2016; Maxwell, Warner & Guillén, 2021; Karasiak et al., 2022). Thus, methodological developments remain of crucial relevance for the research field of urban remote sensing, but they are only marginally touched upon here. The focus here is explicitly on the demands formulated in the predecessor from 2019: "*Studies need to follow the path of multidisciplinary frameworks combining urban concepts, theories and research questions with data-intensive approaches relying on EO and ancillary geodata to develop a clearer, more comprehensive picture of urban form and its role in global urbanization processes*" (Taubenböck, 2019).

In this research report, the higher-value research objective is to document the built urban landscape in its temporal development, to systematize the emerging dynamics and dimensions, to analyze the urban forms and spatial patterns in a global as well as region-specific comparative way, to evaluate the effects of the emerging spatial patterns and to show capabilities for urban adaptation. An underlying theme is the term '*urban*' in itself, how it can be defined theoretically and conceptionally, and how it can be measured and mapped in spatial domains. The input data comes primarily from new geoinformation products from remote sensing.

These topics are seen of particular relevance as urban form is known to have crucial influence on shaping the societies of yesterday, today and tomorrow (Jacobs, 1961; Lynch & Rodwin, 1958; Lynch, 1981; Kostof, 1991, 1992; Anas, Arnott & Small, 1998; Batty, 2008; Burdett & Sudjic, 2007; Glaeser, 2011; Hollis, 2013; Tonkiss, 2013; Sennett, 2018). Whyte (1988) provocatively formulated: "*What attracts people most is other people. Many urban spaces are designed as though the opposite were true*". Thus, the aim is to systematically expand our knowledge of the physical and, to some extent, demographic processes of urbanization and the respective patterns and their effects. New knowledge that allows to re-think, to re-evaluate, to re-discuss urban form and to develop strategies that globally growing urban areas can hopefully be developed as places providing perspectives for people.

From the diverse empirical findings of this compilation of studies in particular, as well as from the urban remote sensing research landscape in general, another goal emerges: Reflections on the key research directions for urban remote sensing as proposed in 2019 (Chapter II; Zhu et al., 2019) and the extent to which these are still relevant today. And beyond that the attempt to derive new key research directions for urban remote sensing and to propose a road map for future paths in this field.

Furthermore, these conceptual, technological and thematic developments in urban remote sensing will be considered from another perspective – with an ethical view. High spatial and temporal resolutions of data and the attempt to provide political and social debates and

decision-makers with better geoinformation and new insights also have implicit ethical consequences – that have widely been neglected. In other words, not everything that is technically feasible is also legitimate or without consequences for groups of people or areas (Kochupillai et al., 2022).

3. Structure of the research report

To systematically approach these research objectives, the research report is structured into *eight parts* which shed light on '*the urban*' from slightly different, yet closely linked conceptual, methodological, thematic and ethical perspectives. In addition to the connecting thematic aspects, the input data from the domain of remote sensing and related disciplines are connecting elements.

Part I – 'Strategic directions for urban remote sensing as starting point' lays the foundation with a large-scale literature study on remote sensing in urban areas for strategically relevant research directions. The four key research directions identified represent the central research strands of this research report. ***Part II – 'What urban is'*** places the fluid concept of '*the urban*' at the center and focuses, on the one hand, on different methodological and thematic variants of understanding this space geographically and, on the other hand, on conceptual ideas of understanding '*the urban*' as a planetary process. ***Part III – 'Urban dynamics'*** focuses on temporal urbanization trends – both on a global level and for specific geographical regions. While urbanization is a global process, the focus here is on documenting the variability in different regions of the world. ***Part IV – 'Urban dimensions and rankings'*** focuses on the emerging new spatial dimensions of urban landscapes as well as on rankings of cities. Rankings are popular in the media but often do not include transparent methods on a scientific basis. Here, methodological approaches are developed and geographic results are presented with regard to global city sizes as well as city hierarchies at national level. ***Part V – 'Urban patterns'*** focusses on the spatial layouts of built cities. Global and regional comparisons are presented here in an attempt to document and categorize structures and patterns of urban space on different scales. ***Part VI – 'Effects of urban patterns'*** focuses on the relationship of spatial patterns on land consumption, utilization intensity or health. It thus shows a variety of dependencies of a socio-demographic or urban planning nature with the physical urban space. ***Part VII – 'Urban adaptation'*** is focusing on the often-necessary reorganization of urban spaces. The possibilities of implementing improvements in such rigid systems for reducing temperatures or by early warning systems are identified and evaluated in order to make cities more resilient to different climatic influences or natural hazards. In addition, the potential and adaptation possibilities of the structures already built for population carrying capabilities of cities is assessed. Finally, ***Part VIII – 'Ethical considerations'*** puts conceptual, technical and thematic developments to the test of ethical challenges and opportunities. A road map is developed that allows scientists to identify ethical issues in their research works.

Structure	Themes	Approach	Results
Synthesis	Remote Sensing for urban geography	**Literature review**, summary of all **contributing papers**, and **development of a road map for urban remote sensing**	**Six key research directions** added to the four key research directions from Part I
Part I	What 'urban' is?	Based on a **literature review**, the development of the **key strategic research directions** for urban remote sensing	Four **key strategic research directions** in the domain of urban remote sensing
Part II	Urban dynamics	Methodological and conceptual approaches to **delineate 'the urban'** from **'the rural'**, respective **effects on urban statistics**, and a **planetary concept** for measuring urbanization	Global differences in **urbanization dynamics**, special foci on **border regions** as well as on national scale on **China**, **air pollution** development in relation to **settlement growth** in **megacities**
Part III	Urban dimensions and rankings	**New data sets** and spatial approaches for **consistent and comparable refence units** and new methods for an appropriate **ranking of cities**	**Global city size ranking**, **urban hierarchies** on national level, and **ranking** based on **green shares**
Part IV	Urban patterns	Methodological approaches to **compare spatial patterns of settlement landscapes** and to **categorize** them globally or regionally	Global structural **typification of cities**, continental **comparison of patterns**, and **intensity of plannedness**
Part V	Effects of urban patterns	Methodical approaches to illuminate the **relations between urban spatial patterns and socio-economic variables**	**Land consumption, and structural effects on air quality, mortality or quality of life**
Part VII	Urban adaptation	**Modelling temperature adaptation** by green spaces, **location analysis for landslide early warning systems**, and evaluation of **population carrying capacities** of cities	Capabilities to **reduce the urban heat island effect, the exposure towards landslides,** and evaluation of **population carrying capacities** of cities
Part VIII	Ethical considerations	Systematization of **ethical challenges and opportunities** and development of a **framework** guiding scientists	A **road map** that allows scientists to systematically **identify ethical issues in their on-going research**

Fig. 2. The main structure of the research report. Source: own elaboration, unpublished.

Figure 2 illustrates the main structure of the research report with respect to the general themes, the approaches and the results.

The ***key findings of Parts I – VIII*** are briefly summarized below (section 4 of Chapter I). Building on this, the research report is embedded in the research landscape (section 5 of Chapter I). The aim is then to reflect and supplement the strategic key research directions formulated for urban remote sensing in 2019 from the various ***Parts II – VIII***. Further key research directions are crystallizing from this, leading to a **road map of key strategic research directions for urban remote sensing** (cf. section 6 of Chapter I). These key research directions are an attempt to work out new fields of development for applied urban remote sensing. And finally, an outlook concludes the summary of the entire research report (section 7 of Chapter I).

4. The main findings

4.1. *Part I*: Strategic directions for urban remote sensing as starting point

Scientific research studies from remote sensing have contributed immensely to our understanding of urban areas over the last fifty years. An increasing number of papers in remote sensing journals, corresponding special issues as well as articles in urban journals bear witness to this (Zhu et al., 2019). Thematic contributions are manifold from the multi-temporal (Seto et al., 2011) or spatial characterization of cities (Taubenböck et al., 2020) in

their built and natural environments (Bechtel et al., 2015) to thermal analysis of urban climates (Li et al., 2017), demographic analysis (Carioli et al., 2023), works in the fields of planning (Debray et al., 2023), health (Machault et al., 2010), equity (Stokes & Seto, 2019), economy (Nordhaus & Chen, 2015), poverty (Kuffer, Pfeffer & Sliuzas, 2016), biochemistry (Imhoff et al., 2004), biodiversity (Qian et al., 2015), hydrology (Ebrahimian, Gulliver & Wilson, 2016), natural hazards (Li, Martinis & Wieland, 2019), risk, exposure and vulnerability (Geiß et al., 2022), or recovery (Román et al., 2019), among many more.

In a synthesis of the diverse research works in the field of urban remote sensing, four key strategic directions were identified from a systematic analysis of the diverse past methodological and thematic developments: *high temporal frequency analysis*, *characterization of urban heterogeneity*, *characterization of urban form and structure in two and three dimensions*, and *linking remote sensing with emerging urban data*. These four key strategic directions are a guideline for this research report, as the following parts are orientated towards these areas (Chapter II; Zhu et al., 2019).

4.2. *Part II*: What 'urban' is

What is '*urban*'? This question seems as banal as it is difficult to answer on closer inspection. Different concepts are devised and implemented depending on the discipline (e.g. Nassehi, 2002; Tonkiss, 2013; Liu et al., 2014; Angel et al., 2018; Balk et al., 2018; Dijkstra et al., 2021). Different types of data (e.g. Malizia, 2013), different methods (e.g. Schabenberger & Gotway, 2005), different scales of aggregation (e.g. Openshaw, 1983) or different indicator(s) (e.g. Shearmur, 2017) have been utilized and applied. However, what can be labelled as '*urban*' cannot be clearly defined, nor are there generally accepted data sets and methods that can generate a universal truth (Deuskar, 2015; French et al., 2022; Taubenböck et al., 2022). Nevertheless, the '*urban*' plays a central role as a reference point in scientific, social and political discourse. As a result, generally accepted truths based on spatial statistics must be constantly scrutinized for their methodological validity.

In *Part II*, this challenge is first approached with a national study that empirically maps urban and rural populations in a spatially and thematically differentiated manner using a multimodal probability-based approach. The results reveal that the delineation of the '*urban*' is in fact very tolerant of ambiguity. The probability-based approach reveals that Germany is by no means as urban as official statistics say. In these official statistics, the German population is measured urban at 80.3 % (UN, 2020). The probability-based approach presented here reveals that Germany can be considered urban by at least 50.0 % up to possibly 68.1 % of the population. The range in between, however, is fraught with uncertainty. These results suggest that figures on the degree of urbanization should be treated with caution and that dichotomous urban/rural classifications do not do justice to the complexity of the situation (Chapter III; Taubenböck et al., 2022). It must therefore be stated that the scientific reference point of '*urban space*' must be chosen very carefully.

The challenge of what can be labelled as '*urban*', thus can logically also be approached from a different angle. Based on the concept of '*urban corridors*' – which are qualitatively defined as massively large, linear urban agglomerations whose dimensions go far beyond individual

cities (cf. Whebell, 1969; Florida, Gulden & Mellander, 2008; UN Habitat, 2008) – a methodological approach is developed to delineate urban corridors in a fuzzy way using multi-source geodata on settlements. Methodologically, a delineation is consistently applied for different input data and the frequency that a pixel is assigned to the respective corridor defines the classification. The results reveal that 14 areas in Northern America show spatial properties of an urban corridor. The approach offers a different methodological approach to the physical delineation of a spatial construct – that of the urban corridor (Chapter IV; Georg, Blaschke & Taubenböck, 2023).

At this point, the possible methodological variability in the definition of '*the urban*' is already apparent. Both studies above have the great advantage that they do not go for a dichotomous delineation, but make the transition differentiable according to probabilities – something that comes close to reality, because transitions between urban and rural areas can, but do not have to be abrupt.

In a further study, this challenge is approached by two very different methodological approaches: a non-parametric approach to evaluate the multi-scale homogeneity of the settlements and one that transfers the settlement patterns into a Gaussian Mixture Model via hierarchical multi-scale clustering. These approaches allow to describe urban land patterns in a numerically stable, consistent and globally comparable manner. From a thematic point of view, it is found that in Germany, many clusters of similar size define '*the urban*' landscape of city regions, whereas in the U.S. one large dominating cluster exceeds most neighboring settlements significantly (Chapter V; Schmitt et al., 2023).

Up to this point, the studies have referred to basic concepts of '*the urban*' and these have been methodically translated into spatial classifications using geodata on built space or population density. The fact that space is also co-defined by perception (e.g. Lynch, 1960, Gibson, 1979) has thus far only been considered indirectly through the conceptual definitions. In a further study, the challenge of how the '*urban*' can be approached is based on the combined analysis of the physical space using remote sensing and other geodata as well as national household survey data. The latter provide location-based statements on the perceptions of individuals (Goebel et al., 2018). The conceptual target objects in this study are city centers. Here, the subjective perception of citizens on what they define as '*the urban center*' is used and related to the urban spatial structure. It is found that a relative threshold for city centers at around 30 % of the built density of the 'central place' is indicative of perceiving the transition of center to suburban areas (Chapter VI; Wurm et al., 2021).

Finally, in *Part II*, this challenge of what can be defined and labelled as '*urban*' is approached by a critical reflection on any of the above categorical understandings of '*the urban*'. All the above and generally used population-based and land-use/land-cover-based categorical understandings of cities are critically questioned since urbanization and the urban functions by now are a planetary process (Brenner et al., 2014; Inostroza, 2014). The urban is networked with places far away from urban centers and their surroundings into the '*last of the wild*'. The conceptual idea is therefore developed that a spatially separate analysis of '*urban*' and '*rural*' areas does not do justice to this planetary interconnectedness. Therefore, a continuous indicator of urbanization which is based on the accumulation of anthropogenic materials is proposed. With it, a physical, rather than a spatial or demographic characteristic

is suggested. This continuous indicator allows the analysis of socio-ecological systems' spatial organization. With it, comparative studies across geographies and times become possible. Such an approach allows to inform on globally generalizable patterns of urbanization processes and gives a material body to address claims for the evaluation of urban development (Chapter VII; Inostroza & Taubenböck, 2024).

4.3. *Part III*: Urban dynamics

Urbanization is truly a process of global impact, even in far off rural landscapes (cf. Chapters VII and X). However, the most visible component of urbanization is the physical expansion as well as the structural transformation of cities. Scholars have – predominantly relying on multitemporal remote sensing data – documented the spatial expansion of cities into hinterlands (e.g., Angel et al., 2011; Seto et al., 2011; Taubenböck et al., 2012; He et al., 2019; Huang & Xu, 2022; among many others) and have also described and compared the regional differences of dynamics (e.g. Güneralp et al., 2020; Huang et al., 2021; Liu et al., 2020; Melchiorri et al., 2018). What has often been missing so far is the combination of a systematic analysis of settlement dynamics for the entire world, combined with a standardized and consistent concept of comparable spatial units for permissible geographical comparisons.

It is precisely this knowledge gap that lies at the heart of the first study on urban dynamics. The expansion of cities observed worldwide is systematized across geographies by their growth rates and their scale: Highest relative urban growth rates are measured in Asia, and specifically in East Asia. In contrast, lowest growth rates are measured in Oceania and Europe. In general, a progressive decline in relative growth rates from 1985 until 2015 is found; however, absolute settlement expansion rates are not declining over time, but are rather constant. In this study the measured urban trends of the past decades are also projected into the future and linked to climate scenarios used by the Panel on Climate Change (IPCC) (IPCC, 2022). Particularly high exposures are expected in South Asia and West Africa in the future: There highly dynamic urban growth rates will intensify exposures due to very high increases in predicted maximum rainfall. Highest increases in extreme temperatures are predicted for Europe and North America; there, however, urban growth rates are expected to be low. Thus, adaptation measures (cf. Chapters XXII and XXVIII) must be brought into focus here (Chapter VIII; Taubenböck et al., 2024).

As outlined in the previous study, high urban dynamics were recorded in Africa. Reason enough to focus an analysis precisely on this continent. As urban space today is both, physical and digital, these two types of '*urban frontiers*' were at focus in this study. The digital divide in social media activity between older and newer urban areas across 135 sites in Africa's most strongly growing cities was analyzed. Findings showed clear evidence of disparities: The '*physical urban frontiers*' often do not run parallel to the '*social media frontiers*'. In the majority of sites, more recently built areas featured significantly lower tweet densities than older ones, while the inverse case was not observed. These disparities were found more frequent in Western Africa and they decreased with increased plannedness and compactness of the built structure. The disparity observed may be due to the fact that the new urban space first has to be inhabited and actively revived until it is represented in

social media data. In general, this study revealed to the urban frontier of settlement expansion a second frontier in digital space. With it a hitherto undocumented form of digital divide was made empirically tangible (Chapter IX; Mast et al., 2025).

Growth analyses can focus on a continent, a region or a country. Another spatially incoherent unit of analysis is places with similar framework conditions. Here, the focus of an analysis of settlement and population dynamics was on border regions worldwide. Borders are intervention in the circulation of goods, information, capital and people (e.g. Mau, 2021; Newman, 2006; Paasi, 2019). The guiding research idea here was the question of the extent to which borders influence the surrounding regions in terms of their dynamics, and if so, which trends can be distilled here. In this study, the impact at border regions was operationalized by border typologies, settlement and population dynamics near borders, and economic and political indicators of neighboring nation-states. The key findings are that higher political or economic differences of neighboring nation-states relate to stronger border fortification, greater economic or political disparities relate to stronger population or settlement accumulation at the poorer or less free side of the border, and stronger fortification of the border hinders settlement and population development to a certain degree. These empirically measured trends, however, are only partially statistically significant and not as strong or unambiguous as assumed (Chapter X; Taubenböck et al., 2023).

Urban dynamics can be related to settlement expansion, but the topic can also be approached in terms of dynamic changes in air quality. It is precisely these dependencies that have been the focus of another research study. The dynamics of settlement growth in all current 38 megacities across the globe have been studied in relation to the trends of tropospheric nitrogen dioxide (NO_2) pollution from 1996 to 2015. The key findings are that on average, an increase in tropospheric NO_2 burden of 5.06 ± 0.83 %/year and a relative settlement growth of 2.87 %/year for the time period is observed. However, there is a wide range of trends across cities: relative NO_2 pollution ranges from an increase of 44.0 ± 5.9 %/year to a decrease of -3.2 ± 0.2 %/year, settlement growth rates vary from 0.3 %/year to 9.5 %/year. Despite this variety, the results exhibit a pronounced relation to the income group following the world's economies classification of the World Bank. In general, NO_2 pollution in megacities in developed, richer countries is at a very high level, bur decreasing over time. In contrast, megacities in recently developing countries show a strong increase in NO_2 pollution (Chapter XI; Erbertseder et al., 2024).

4.4. *Part IV*: Urban dimensions and rankings

How should we understand the term '*dimension*', which is used here in the title of this section as well as in the title of the entire research report? Basically, this term means something like the '*extent of a thing*' or '*in general and simplified terms: the physical size of an object in the space it occupies*'. If this basic understanding is transferred to the urban environment, there are many starting points for research questions: Here, this question is approached from different perspectives. On the one hand, '*the urban*' is heterogeneous in its form (e.g. Bechtel et al., 2015; Seto, Sánchez-Rodríguez & Fragkias, 2010) and the question is first what dimension or share do different structural components take in cities worldwide? This

question, however, is also approached, on the other hand, from another angle. A methodological toolkit is developed to translate the spatial extents (or '*dimensions*') of cities or structural components within them into rankings.

As introduced in Chapter II, the characterization of urban form and structure in two and three dimensions is one of the key strategic research directions (Zhu et al., 2019). However, such a database on intra-urban structures on global scale did not yet exist. To address this data gap, a novel deep learning and big data analytics approach was applied on multi-sensor satellite data to create the first-ever global and quality controlled intra-urban structural classification for all 1,692 cities with a population greater than 300,000. The conceptual foundation relates to the Local Climate Zones classification scheme, i.e. a neutral description of built and natural structural forms of cities (Stewart & Oke, 2012). This makes it possible to quantitatively describe the structural proportions of cities, whether compact or rather low-density, whether with high or low building heights. The key findings reveal that the urban fabric is globally predominantly of low density. The largest type of the built-up area is the '*sparsely built*' LCZ, accounting for about 30 % of areas within the 1,692 cities. In comparison, the compact classes occupy only about 14 %. Beyond this, these dimensions in the shares of structural types reveal inequality when related to population: approximately 40 % of the area defined as compact or light/large low-rise accommodates about 60 % of the total population, whereas approximately 30 % of the area defined as sparsely built accommodates only about 10 % of the total population (Chapter XII; Zhu et al., 2022a).

However, the dimension of cities can also be approached on the spatial unit of the entire city. In order to be able to compare cities in terms of size, a spatial reference unit is needed that is comparable and does not have to rely on random, political or historical city boundaries (cf. Chapter III). Critics claimed that much of existing urban research can be contested due to a rather arbitrary use of geographical boundaries (e.g. Lechner et al., 2013; Riitters et al., 1995; Berry & Okulicz-Kozaryn, 2012; Masucci et al., 2015). A new methodology has been developed to delineate the morphologic extent of cities. By re-territorializing the conventional administrative units, the current global rankings and statistics on the largest cities are scrutinized. With it, such a new ranking is based on a transparent approach, with consistent, and thus less biased reference units – the morphological urban areas (MUAs). The key findings are that the urban agglomeration in the Pearl River Delta in China is today the largest city in the world. It has a calculated population of 42.6 million, and with more than 20 million informal inhabitants, probably close to 65 million people. Tokyo, is now in this new list ranked only number two. It is interesting to note that although many MUAs outperform administrative units by area, and contrary to what was assumed, in most cases MUAs are considerably smaller than administrative units. Only in Europe MUAs are found that are largely outweighing administrative units in extent (Chapter XIII; Taubenböck et al., 2019).

This new ranking on the rather popular topic of global '*city sizes*' shows that with a consistent approach, previous rankings that are generally accepted and used for political, economic, ecologic or social argumentation can be quite mixed up – and with them statistics and our view of things. In this regard, city rankings have, in general, become a marketing tool. One example can be attributed to the practice of '*greenwashing*' (e.g. Seele & Gatti, 2017), i.e. a form of city marketing in which statistics on green shares are used to convince

the public that the city stands out in particular. Against this background, the influence of different reference units was systematically analyzed – on the subject of the proportion of greenery in cities. In addition, a wide range of statistics on urban greenery were used to rule out random effects as far as possible. This was done by using two parameters, i.e. green space shares and green space per capita, by using all green areas as well as only green areas with a minimum size of one hectare, and by referring to administrative spatial units, to morphological spatial units as well as to a standardized buffer unit around the city centers. The diversity of results has been tested among the 80 major cities in Germany. The results allow to discuss the susceptibility of spatial statistics to ambiguities that may arise from the use of different concepts. The key findings reveal that some cities in Germany do indeed practice '*greenwashing*', as they describe themselves as one of the greenest cities in the country, but in empirical terms they end up among the last places of the ranking (Chapter XIV; Taubenböck et al., 2021).

4.5. *Part V*: Urban patterns

What we generalize by the simple terms '*city*' or '*urban*' has, in fact, highly diverse domains: different economies, demographics, ways of living, land uses, built-up morphologies, among other things. In this *Part V*, the focus, however, is on the spatial patterns of the built settlement landscape in cities, both globally and regionally.

The fact that cities are structurally different has been described already more than a hundred years ago by the Chicago School. Various generalized spatial patterns of urban landscapes have been conceptualized by observations and theories, e.g. in concentric zones (Park & Burgess, 1925), sectors (Hoyt, 1939), or multiple nuclei configurations (Harris & Ullman, 1945). Urban researchers have also noticed that cities have regional differences. Spatial models reflected the geographical-cultural diversity between European (e.g. Lichtenberger, 1972), US American (e.g. Hahn, 2014), Latin American (Griffin & Ford, 1980), oriental Islamic (Ehlers, 1993), African (UN, 1973), or Chinese (Gaubatz, 1998) city models, among many others. The basis for such intercultural comparisons of cities is the division of the earth into cultural areas, as it has been done e.g. by Kolb (1962) or Huntington (1997). In this study, an attempt was made to empirically prove these conceptualizations and observations. The morphological-spatial configurations of 110 cities distributed around the globe were empirically examined based on the intra-urban LCZ classification scheme (Stewart & Oke, 2012; cf. Chapter XI; Zhu et al., 2022a). The key findings are that seven city types relate to the global diversity of spatial urban configurations and prove empirically that some of the conceptualizations and observations introduced above hold true. Some types are found to be largely congruent with well-known geographic-cultural spaces such as the Islamic world or Europe. However, in contrast to some theoretical city models, other types reveal an empirically more complex spatial distribution as African-American or Asian-African clusters. In general, the study confirms that similar socio-economic, demographic, cultural, or political conditions do produce similar morphologic-spatial urban configurations. Nevertheless, the world remains more complex than this general trend as there exist also similar morphological configurations across geographic-cultural spaces (Chapter XV; Taubenböck et al., 2020).

The fact that urban structures are different and that structural city types crystallize should be sufficiently proven by the studies above. And the fact that the types of structures indirectly carry the diverse economic, planning, social or political processes '*in stone*' has been discussed and shown many times (e.g. Kostof, 1991, 1992; Salat, 2011). However, the extent to which the planning influence can be mapped via the structures has never been empirically proven on a global level. Thus, the complex relationship between the existing urban morphology and the diversity of planning processes has been the focus of analysis. By the development of an ontology of '*planning intensity*' to conceptualize intermediate categories of whether an urban structure is planned single-handedly or constantly updated by myriads of participants lays the conceptual foundation. With it, the '*planned/unplanned*' dichotomy is overcome by the continuum of the *Intensity of Plannedness (IoP)*. The key findings are an empirical proof that the distinct IoP has indeed significant contrasting structural complexities. At an abstract level, it can be deduced from the empirical analysis that highly planned areas usually have a structurally low complexity and vice versa. With it a relationship between the intensity of the process of planning and the resulting urban morphology is proven (Chapter XVI; Debray et al. 2023).

In the global study on city types introduced above (Chapter XV), Europe appears as a cluster with similarities in terms of urban structure. And yet it is worth taking a look at intra-European differences that are not so easily recognizable in a global comparative setting. The main guiding question in this study was: can the post-socialist transformation of urban spatial structure in Eastern European regions be interpreted as an adaption process to Western-style urbanization and, if yes, how far a process of '*mimicry*' has reached. Based on a cross-sectional analysis of settlement patterns of 230 city regions as well as an in-depth longitudinal analysis of ten selected urban regions, the key findings are: There are indeed significant differences in settlement patterns between Western and Eastern European city regions that can be interpreted as different stages of urbanization. Eastern city regions are less urbanized, less populated, and more dispersed. Post-socialist urban restructuring, however, is also found to have adaptations in land use patterns similar to that of Western Europe. All in all, however, a '*catch-up*' leading to full convergence with Western-style urbanization, is not given nor a realistic future scenario. Eastern European urbanization simultaneously manifest characteristics of convergent adaptation to capitalist urbanization and path-dependency of relics of the socialist past (Chapter XVII; Taubenböck et al., 2019).

We have seen that spatial patterns vary around the world, and even so within one continent even if on global scale the variability there appeared comparatively similar (Chapter XV). And yet there is also significant variability within a country, a region or within a city (Chapter XVI). It is therefore worth taking another look at a specific region.

The focus therefore has been on China –with regard to the hierarchical patterns of city sizes on national level. The spatial distribution, hierarchical structure and the development over time of the urban system in reference to Christaller's Central Place Theory (Christaller, 1966) have been analyzed. The key findings are that on national level the hierarchical spatial urban systems are close to the ideal pattern suggested by the Central Places Theory, in particular in the eastern and central regions of the country. However, the hierarchical urban

systems developed over the last decades in most areas away from this spatial organization (Chapter XVIII; Shi et al. 2020).

4.6. *Part VI*: Effects of urban patterns

Built urban structures and the corresponding functions define everyday rhythms and behaviors to a not inconsiderable degree (Glaeser, 2011; Batty, 2013; Tonkiss, 2013; Sennett, 2018). They define whether we (have to) travel by car (Glaeser, 2011), how easy it is to get our daily necessities on foot (Köberl et al., 2024), or how we perceive or behave in urban space (Gehl, 2009). In addition, these structures have a variety of effects such as exposure to heat (Leichtle et al., 2023), air quality (Pascal et al., 2013), noise (Staab et al., 2021), stress (Hunter, Gillespie & Chen, 2019), and much more. It is thus of crucial importance for urban planning and policy development to understand how different urban configuration patterns affect the environment, human health, our everyday rhythms, among other issues. And thus, an evidence-based weighing up of advantages and disadvantages is a necessity.

To lay an empirical foundation, a large-scale study was carried out to evaluate the relation of structural urban configuration types, i.e. of the built and natural landscape, with Surface Urban Heat Island (SUHI) intensities, tropospheric NO_2, CO_2 per person emissions, motorized traffic flows, and age-standardized mortality. Four distinct types of urban configurations were statistically defined for a sample of 918 cities across entire Europe: '*compact high density*', '*open low-rise medium density*', '*open low-rise low density*', and '*green low density*'. The key findings show an empirical prove that different urban configuration types at city level relate to all the parameters introduced above, of course with varying advantages and disadvantages: On the one hand, the compact high-density city type, to present one example here, is measured with fewer green spaces, higher NO_2 exposure, a higher SUHI effect and higher mortality rates. On the other hand, the compact high-density type allows for a high population density, less land consumption, high accessibility and per person CO_2 emissions were measured significantly lower. The findings indicate that, although the compact city model is considered more sustainable (e.g. Glaeser, 2011; Angel et al., 2021), the European compact cities still feature many disadvantages with respect to poor environmental quality and health (Chapter XIX; Iungman et al., 2024).

The fact that structures have a variety of effects should therefore be sufficiently demonstrated. In the following, the focus is on two studies that deal specifically with one parameter: land consumption.

Land consumption reveals how effectively we use space – whether wastefully in area or efficiently. Here, land consumption analysis relies on settlement layers produced by remote sensing data as well as on population data. The local climate zone classification scheme was again used to map the intra-urban pattern in 40 cities across the globe. In addition, global population grids have been disaggregated to this intra-urban scale. The key findings reveal that urban land consumption is differing immensely within cities but also across the globe. To give examples of the contrasts: Mumbai, India is measured with a population density of 36,741 people/km^2 which is twelve times denser than Melbourne, Australia (2,981

people/km^2). At the intra-urban scale, the high-dense slum structures of Nairobi, Kenya are found to be the densest areas with a population density of 48,241 people/km^2 (20.73 m^2 land per person). A citizen living in Melbourne, however, has 530 m^2 for living. The study shows inequality in land consumption and makes clear that aggregated per-city land consumption studies lose sight of intra-urban variability and statements at city level often lead to distortions (Chapter XX; Hu et al., 2021).

A further study in this regard focused on the relationship between building types and land consumption. The building types of '*urban block developments*', '*terraced houses*', '*semi-detached houses*' and '*detached houses*' have been investigated. In order to make the effects of the statistical correlations between the two variables tangible, model calculations were carried out. For the example of the entire country of Germany it is calculated what it would mean in terms of land consumption if all 83 million inhabitants of the country were to live in one building type or another. The key findings reveal for the scenarios that if everyone lived in detached houses, over 42 % of the land would have to be impervious surfaces in order to cover the land requirements of this type of structure. In contrast, if everyone would live in high-density block developments, only 2.8 % of the land would be needed. This difference with respect to land consumption shows how important forms of housing are and what societal decisions in favor of or against certain building types can achieve when adapting or newly constructing. And, this analysis also allows to take a look at how reasonable and resilient our way of living currently is (Chapter XXI; Taubenböck, 2021).

4.7. *Part VII*: Urban adaptation

The research studies presented so far show, among other things, how dynamically cities grow, what spatial patterns exist and which varying effects these patterns have. On the one hand, this shows that, in principle, we could use this knowledge to develop new urban expansion areas in a more targeted manner. On the other hand, however, it must also be recognized that the already built city is comparatively long-lived and that adaptations require enormous effort. It is undisputed that cities have to adapt to new boundary conditions, whether through growth, climate change or other factors (e.g. Chapters VIII, XI, or XIX).

With reference to the necessary adaptation of cities, three studies are presented below that demonstrate how urban remote sensing can be used for this purpose.

In a first study possible adaptions to the urban heat island effect are at focus. The aim is to model what is necessary so that fewer people are exposed to heat by reducing the temperature. A spatial regression model using remote sensing data was developed to assess the population exposure to Land Surface Temperature (LST) extremes in more than 200 cities across the globe. The model relies on surface properties like vegetation cover and distance to water bodies. The key findings reveal that vegetation plays a considerable role in decreasing the exposure of the urban population to LST extremes. Nevertheless, as the analyses show, there are competing uses: the need for more green space or a shortage of housing, which cannot be solved in a one-dimensional way. In particular, it could be shown that adaptations in high-exposure areas reduces vegetation needed for the same decrease in exposure compared to uniform treatment (Chapter XXII; Massaro et al., 2023).

Another expected consequence of climate change is increasing heavy rainfall events (IPCC, 2022; Chapter VIII). The devastating events in the Ahr valley in Germany in 2021 or in Valencia, Spain in 2024 show how vulnerable exposed areas, structures and societies are in such situations. In addition to flash floods, landslide events can also be a consequence. In Colombia, for example, landslides are the most frequent hazards (Ruiz Peña et al., 2017). The interplay of heavy rainfall events and the mostly informal growth of cities in landslide-prone areas increases the associated risks (World Bank, 2012). To reduce the risks of catastrophic events, Landslide Early Warning Systems (LEWSs) are a possible measure. The city of Medellín, Colombia, is used as an example as it is located in a steep, deep valley with large informal, make-shift, unplanned settlements expanding up the slopes. To minimize the population's exposure, adaptation measures are needed. However, what and where one could adapt to which costs is not easy to answer. For an evidence-based assessment of location-based needs, effects and costs, a data-driven approach has been developed. This approach localized landslide-prone areas suitable for LEWS instrumentation, and allowed to estimate associated costs. An area-wide landslide susceptibility map was derived from a random forest model based on heterogenous data on hazard levels, landslide inventories, and geological and topographic factors. The synergistic analysis of the landslide susceptibility map and remote sensing-derived layers of settlement types and population density allowed to localize highly exposed areas. In an interdisciplinary approach with geologists and urban planners (cf. Werthmann et al., 2024), a cost function was developed to estimate the cost of LEWS monitoring sensors at particular sites. The application enables decision-makers to balance risk levels, financial costs, and the possibly protected exposed populations. The key findings reveal that LEWS monitoring sensors could be installed in several landslide prone areas across Medellín, Colombia. Depending on the location the financial costs vary between 5 to 41 Euros per inhabitant of the city. Possibly the resilience of over 190,000 exposed inhabitants, 81 % of whom are located in precarious informal neighborhoods, could be improved by these measures (Chapter XXIII; Sapena et al., 2023).

With respect to the necessary urban adaption, the topic of '*land consumption*' (cf. Chapter XX and XXI) is again taken up: The aim here was to show what potential the already built structures or adaptations of these structure (e.g. re-densification) of the cities around the world have with respect to potential population carrying capacities. As empirical baseline, the morphologic and demographic urban expansion across 1,567 major cities around the globe from 1985 to 2015 and the changes in the morphological and population densities in the newly built areas in comparison to the status quo in 1985 was analyzed. Theoretical retroactive scenarios of alternative land consumptions were modelled. They allowed to determine possible carrying capacities if denser structured would have been built. The key findings are that between 1985 and 2015 cities across the globe have become structurally less dense than the areas built before 1985. Thus, global urban expansion consumed comparatively more land which means that it has become more inefficient with respect to land consumption. Of course, this varies around the globe, with South Asia and East Asia having higher densities. In contrast Europe and North America reveal lowest densities. In general, however, the scenarios reveal the large unutilized potential in city building and for urban adaptation: If urban expansion would have been structurally denser, e.g. by '*compact mid-rise*' structures, our model calculations show that 1,466 billion more people could live

in the same newly built urban areas. The study reveals that with current built structures a less resource consuming urbanization is possible. (Chapter XXIV; Taubenböck et al., 2024).

4.8. *Part VIII*: Ethical considerations

The presented studies rely on 'big earth data' and methodological advancements or applications in the domains of Artificial Intelligence (AI) and Machine Learning (ML). The applications shown here intend an impact on science, governance, policy, and ultimately, the lives of people by providing better geoinformation for evidence-based decision-making. Ethical issues are thus moving more and more into focus (e.g. Floridi et al., 2018; Coffer et al., 2020).

This research reports closes with a practical overview of the key ethical issues emerging in this field, with concrete examples from the Earth observation and remote sensing domain to explain these issues. Four research fields – the very urban topics of slum classification, and population assessments, as well as the monitoring of small-scale mining and the protection of biodiversity – were used as guiding examples how scientists can redesign research questions to more effectively grab ethical opportunities. It aimed towards a framework to address real-world problems that are otherwise akin to ethical dilemmas with no win-win solution. A road map was developed that allows scientists to systematically identify ethical issues in their own on-going research. With it, it has been intended to sensitize scientists to these issues and the road map allows them to critically reflect that not everything that is technologically doable might be ethical (Chapter XXV; Kochupillai et al., 2022).

5. Embedding the research report into urban research

More and more people are moving the center of their lives to metropolitan areas (UN, 2017). The expansion of built urban landscapes are the most direct result. Meanwhile, it is often unclear where urban space ends and peri-urban or rural space even begins, i.e. what the spatial unit as a reference point actually is (Chapters III – VI), and whether such a differentiation makes sense at all (Chapter VII). The immense urban dynamics in recent decades (Chapters VIII – XI) and the gigantic dimension of the urban landscapes that have been created (Chapter XIII) show that, on the one hand, they provide rising global populations with opportunities to live and protection from nature (Chapter XXIV). At the same time, the process of urbanization has reshaped the planet beyond recognition (Chapter XX and XXI). These urban patterns carry socio-political-economic information and indeed reveal geographic patterns across the globe (Chapters XV – XVIII). On the other hand, the effects of these structural patterns and processes impact individual and societal issues (Chapter XI and XIX), and global change triggers the need for adaptation (Chapter XXII – XXIV). And while an improved understanding of these developments is created – in this research report – by high-resolution and constantly growing remote sensing data, new methods in the field of machine learning and artificial intelligence in combination with the corresponding increasing processing power, ethical questions arise in parallel in the course of these possibilities. These have so far only been given rudimentary consideration in the field of Earth observation (Chapter XXV).

Remote sensing technology enables both, large-scale, even global observations and analyses of the built urban landscapes as well as highly detailed surveys of the variable objects in heterogeneous, small-scale urban areas. It thus offers multi-scale geodata for a wide range of urban geographical studies – the focus of this research report.

Four key research directions – *high temporal frequency analysis*, *characterization of urban heterogeneity*, *characterization of urban form and structure in two and three dimensions*, and *linking remote sensing with emerging urban data* – were established for urban remote sensing in 2019 (*Part I*), which were followed in this research report. The studies presented here have shown that new conceptual and theoretical foundations of '*the urban*' and new methods for implementing these concepts for spatial analyzes are necessary for any geographic (comparative) study and statistical validation of results (*Part II*). The uneven growth of cities across the globe as well as of peripheral settlement areas has been demonstrated empirically and in regionally differentiated manners (*Part III*). To which enormous spatial dimensions cities have grown by now and how comparisons among cities can be carried out transparently was also demonstrated (*Part IV*). How variable built urban structures are worldwide, which codings are embedded in them with regard to planning or which spatial patterns have emerged in different regions was also empirically shown (*Part V*). The effects of these urban structures and patterns could be shown by variables such as land consumption or the urban heat island effect. The value-adding by joint analyses with other geodata allowed to reveal effects related to per person CO_2 emissions or mortality rates in dependence of their location (*Part VI*). The potential of adaptation measures towards climate-induced changes against urban heat islands or landslides was demonstrated as well as theoretical adaptations of urban structures and their relation to potential population carrying capacities of cities have been evaluated (*Part VII*). And last, but not least, ethical challenges and opportunities at the intersection of (urban) remote sensing, big data, methods in artificial intelligence and machine learning and possible effects and consequences in the realworld have been systemized (*Part VIII*).

Of course, the studies presented here are not at least exhaustive, as urban remote sensing contributes to urban geography in a variety of ways: for example, in the fields of socio-economy, disaster management, risk and environmental analyses, urban planning, to name a few or in the subject areas that were the focus here in this research report, such as urban growth, urban structures and patterns, urban definitions, effects of structures or adaptation measures.

Seen in this light, the intrinsic approach of remote sensing to urban space via the built and natural landscape initially appears one-dimensional. However, when you consider that since 2021, the global human-made mass exceeds all living biomass (Elhacham et al., 2020), i.e. the weight of all houses, infrastructures, machines, consumer goods, etc. exceeds the weight of all living animals and plants, the built urban space reveals itself as a crucial starting point. And, when you consider that the built landscape materializes people's relationships to the world, to other living beings, to space and time (von Borries, 2024) – in other words, the built urban landscapes give an indication of what is going on and paint a '*societal psychogram*' of its time – then remote sensing-derived geodata can be understood as an important starting point of a multi-layered urban geography. A geography that, in joint

analysis with other, heterogenous geodata (e.g. Chapters III, VI, IX, X, XI, XIX and XXIII), makes the complex, often non-linear dependencies and simultaneities over space tangible. Then, this approach is anything but one-dimensional.

This naturally raises the question that has preoccupied many people: How should urban space as the center of life for the majority of our planet's inhabitants be designed? A wide range of scientific disciplines have dedicated themselves to the goal of understanding cities better as a basis for making them better. Urban remote sensing is making an increasingly important contribution here with its ever-growing data, derived geoinformation and respective empirical studies: from global comparative studies to highly detailed local analyses. The aim of this research report was to confirm and expand existing knowledge in this domain. The report is thus positioned at the interface between remote sensing data, derived geoinformation and empirical urban research.

However, science alone is not enough to achieve progress. Progress consists of the application of knowledge (Pinker, 2018). A wide range of players have committed themselves to this goal: Supranational organizations such as the United Nations have committed themselves for example in the Sustainable Development Goals (SDGs), which explicitly address cities in *SDG 11: Sustainable cities and communities* (UN, 2025). There is also a wide variety of civil society involvement, e.g. in NGOs that are trying to make the urban environment more sustainable, less poor or more ecological, to name just a few examples. In addition, there are many state organizations in the fields of construction, urban development, planning and much more to pursue these goals. And, there are actors from business and industry, from housing associations to architectural firms, or the like. And finally, there are many citizens who want to have a positive impact on the city with their unorganized commitment (e.g. urban gardening).

The research report here is a contribution from science. And, realizing that science alone is not enough, we have to acknowledge a gap between the general statements of '*the urban*' being of crucial importance to the global society – as just argued in the paragraphs above – and the possibilities of the sub-discipline '*urban remote sensing*' and its scientific data, methods and results. However, I am arguing here that these two subject areas do have each other in mind and yet urgently need to be brought closer together. In the knowledge that the spatial influences the social (e.g. Tonkiss, 2013; Sloterdijk, 2013; Jackson, 1985), in the knowledge that the linking of spatial data with other geodata increases the understanding of processes, behavior and perception (e.g. Wurm et al., 2021; Zhu et al., 2019), in the knowledge that the capability for large-scale analysis allows to learn from comparisons (e.g. Schneider & Woodcock, 2008; Güneralp et al., 2020), it can be concluded that new empirical knowledge can have, maybe must have effects beyond the scientific domain. "*Evidence-based knowledge must support rethinking urban form in a way that these areas can be developed as places providing perspectives for people to win*" (Taubenböck, 2019). This is the final sentence of my habilitation in 2019, which I would still sign today.

The challenges of interdisciplinary and transdisciplinary collaboration from scientific studies to real-world application were demonstrated, for example, by Werthmann et al. (2024). In this example, the scientific contribution of urban remote sensing lies particularly in the localization of highly exposed urban areas with regard to landslide hazards for the

implementation of early warning systems (Chapter XXIII). However, the study also demonstrates the many hurdles between scientific knowledge generation, the physical realization of such an early warning system and ultimately the sustainable operation of a system. This example reveals the complexity between a humanistic research idea of improving urban space and its implementation in reality.

And still, by today, empirical knowledge on urban issues has explosively grown since 2019 and is available in incredible diversity – even if it must be recognized that gaps in knowledge are always greater than new empirical findings. But, I think it is fair to say that urban remote sensing is making a not inconsiderable contribution.

This brings me to the statement I made in the Preface of this research report: the question that stands above everything else is – *how do we adapt and design the cities of the future so that they are worth living in, socially balanced, as well as economically prosperous, ecologically sustainable and resilient to changes?* Or is this question actually already wrong, since it implies an active process of action? And the real question should rather be: Do many small-scale, informal processes create a more open, adaptable city than formally planned ones? And this raises many more questions that I am not able to answer: Should planning represent society or try to change it? How do we deal with ambiguities and contradictions in our scientific findings? How does our remote sensing data actually find its way into the wider urban debate? And, do current urban remote sensing data and research studies really come up with suggestions to adapt cities or do they operate in a theoretical sphere? Are research projects still too far removed from the necessities of reality? And, how can uncertainties be communicated? Which impact does our research field actually have?

Perhaps it is not particularly surprising that there are at least as many questions as there were before all this presented work here. Because I cannot (yet) answer these overarching questions, I can only argue that the discipline '*urban remote sensing*' needs to sit back and think from time to time – where exactly do we need to go from here? And so, in the next section, I will limit myself more to reflecting on and expanding the strategic goals for '*urban remote sensing*'. In the following, I will therefore try to briefly reflect on the four key strategic directions for urban remote sensing which formed the starting point for this research report (Chapter II) and suggest further key strategic directions.

6. A road map of strategic key research directions for urban remote sensing

The starting point for this research report were four key strategic directions that had crystallized from past urban remote sensing studies: *high temporal frequency analysis*, *characterization of urban heterogeneity*, *characterization of urban form and structure in two and three dimensions*, and *linking remote sensing with emerging urban data* (Chapter II: Zhu et al., 2019).

The compilation of the research works in this report followed these strategic directions: the temporal analysis of the urban (Chapters VIII – XI, XXIV), the characterization of urban heterogeneity (Chapters II – VII and XII – XIV), the characterization of urban form and structure in two- and three-dimensions (Chapters XV – XVIII), and the linking of remote sensing with emerging urban data (Chapters VI, IX, X, XIX). At this point, it should be noted

that in my opinion these key strategic directions developed in 2019 remain just as relevant today as they were then. It remains here to expand these key research directions with a few more lines of thought:

- The meaning of '**high temporal frequency analysis**' in remote sensing has changed over time, especially in the last few years since 2019: The analysis of changes in cities was usually carried out over long periods of time, often decades. In this research report, the analyses performed on urban dynamics and changes are done in these time frames (Chapters VIII, IX, X, XI, XVII, XXIV). Now, annual or even intra-annual updates are possible – for the analyses of settlement growth as done in this study higher frequency as years or decades might not be decisive. However, new sensor systems from public and commercial providers now make it possible to take several images per day, if the weather permits. This gives '*high temporal frequency analysis*' a new meaning and opens up a variety of new applications, but also methodological challenges. The dynamics of the urban can now also include short-term, monthly, weekly- or daytime or even near-real-time-dependent changes, e.g. in the intra-annual cycle of vegetation (Cheng et al., 2024), heatwaves (Li et al., 2019), construction activities, traffic situations (Butilă & Boboc, 2022) in general and parking lot utilization (Drouyer & de Franchis, 2020) in particular, but also in domains such as the dynamic changes in conflict regions such as developments in refugee camps (Aravena-Pelizari et al., 2018) or in war situations (Pushkarenko & Zaslavskyi, 2024), among other thematic domains. As Grimm et al. (2017) put it, urban disturbance, i.e. anything that disrupts aspects of the structure of an urban system can now be addressed.
- The **characterization of urban heterogeneity** relates to the urban varying significantly around the world as well as within particular cities or even within neighborhoods, blocks, streets, plots, or buildings (e.g. Brenner, 2014; Kostof, 1991; McHale, 2015). On the one hand, the new availability of geodata at large-scale or even global level (e.g. Zhu et al., 2022a; Demuzere et al., 2022; Li et al., 2022) makes it now more and more possible to describe, categorize and compare these heterogeneities more precisely at intra-urban scale. This has been shown e.g. in Chapters XII, XV, XIX and XXIV, however, there is still no systematization of all forms or regional specifics. And, the material- or object-specific heterogeneities, which can be seen in such variable objects as roof shapes, solar panels, roof gardens, street furniture and the like, can be captured and analyzed with the high-resolution data. However, research on urban heterogeneities still remain far below the possibilities with regard to comparative studies across cities. In addition, this data now also makes it possible to better understand these heterogeneities and dependencies across different scales – spatially and temporally. On the other hand, while many different methodological approaches have crystallized (e.g. Fleischmann, Romice & Porta, 2021; Lemoine-Rodríguez, Inostroza & Zepp, 2020; Liu et al., 2022; Luck & Wu, 2002; Chapters XVI, XVII and XIX), the toolbox remains opaque to the extent that their variability often does not easily allow comparisons and overall assessments across studies.
- The **characterization of urban form and structure in two- and three-dimensions** remains an essential research direction for urban remote sensing as many biophysical and socioeconomic processes are shaped by the horizontal as well as vertical structure

and form of the built and the natural urban landscape (Grafius et al., 2018; Wentz et al., 2018). Data today allow for studies at finer patterns, i.e. at the level of streets and blocks, plots or parcels, buildings, and green- and blue spaces (Oliveira, 2016), and over time. The systematization of patterns and their spatial relations in dependency of locations, time, dimensions and much more is in demand. In more concrete words, the discovery of hidden rules of urban patterns and trends need a systematization (e.g. Chapter XV, XVI, XVII and XVIII). This systematization will also help to investigate many different effects of urban structures in a more targeted and, above all, more comparable way.

- The **linking remote sensing with emerging (urban) data** is a necessity for urban remote sensing as only information on the physical environment can directly be retrieved from remotely sensed imagery. Other complementary or dependent information are in need: This includes the integration of geoinformation from different perspectives than the top-view: Structured data with georeference collected usually by federal agencies such as Census information; the human perspective in general operationalized e.g. by volunteered geographic data (Goodchild, 2007), street level images (e.g. Aravena Pelizari et al., 2021), webcams (Idé et al., 2017), wireless sensor networks (Rashid & Rehmani, 2016), data from measuring stations on the ground (Karges et al., 2025), sensors on people or objects moving through the urban space (Resch et al., 2014), among others. Other, non-image data are powerful to support retrieving information that is crucial like urban functions or usages (Zhou & Zhang, 2016), population (Balk et al., 2018), spatially explicit human activity and mobility (Frias-Martinez et al., 2012), feelings and emotions at particular urban places (Wang & Stewart, 2015) and how things are said at a particular place (Lemoine-Rodríguez et al., 2024) are of relevance. Data sources are e.g. in social media (Zhu et al., 2019), in (online) media news (Senaratne et al., 2023), points-of-interest (Ye et al., 2019), among many other possibilities. Despite sometimes a bias in the data, other spatial scales or incompleteness and errors in the data (Goodchild, 2013), they still allow to compensate for certain data gaps or at least supplement them in a meaningful way.

Following these paths, new and manifold research questions can now be created for each individual key strategic direction reflecting the possibilities in 2025 and the coming years. In my view, with the few extensions introduced above, these key research directions need to be continuously followed. Urban remote sensing should therefore continue along this path and strive to systematize all four of these key research directions to statistically and spatially validate the trends and develop new thematic avenues.

But this continuation is not my main point here. I am more interested in **expanding these four key research directions for urban remote sensing**. Of course, methodological developments in pre-processing, classification, validation, spatial metrics and much more will play a decisive role in future developments. However, I would like to focus more on the applied side here. What now emerges from the synopsis of these various research studies complied in this research report are further aspects that I consider to now also be strategically central for the urban remote sensing community.

- Until now, '*urban remote sensing*' has been understood in a spatially delimited and localized and, in consequence, often dichotomous way. The urban is therefore often

understood in terms of a demarcation to urban hinterlands or rural areas. This research report admittedly also contains approaches in this regard (e.g. Chapters VIII, XIII, XIV, or XV). **Urbanization**, however, **is a planetary process** that extends into "*the last of the wild*" (Brenner, 2014; Inostroza & Zepp, 2021; Chapter VII) through exchanges of all kinds (Castells, 2002; Inostroza et al., 2014). It is too short-sighted to understand the construction and operating of the built urban landscape without the totality of the space remodeled by human intervention (von Borries, 2024).

(Urban) remote sensing with its intrinsic spatially continuous data sets has the capability to deal with the influences of these processes beyond the self-limiting spatial delimitations of cities. Thus, dichotomous analyses and ways of thinking in '*urban*'/'*rural*' categories need to be supplemented by non-space-limiting approaches. Some studies in this report already address this challenge, be it in a conceptual way (Chapter VII), be it in an area-wide assessment of localities according to their urbanity in a continuous rather than dichotomous way (Chapter III and IV), be it that studies analyze settlement areas as a whole rather than just structures of larger agglomerations (Chapter X). This way of thinking is intrinsically part of some analysis. Studies on exogenous pressures on urban populations, such as air pollutions (e.g. Chapter XI) or climate-induced exposures (e.g. Chapter XXIII) basically analyze effects of the global interrelated systems on the particular urban spaces. These are pressures that are not necessarily produced at the particular place. However, the analysis of endogenous pressures of the urban on rural, non-urban or peri-urban environments by anthropization, sprawl, or on the ecosystems have less been the focus of research. For example, activities in the peripheries of this world (be it mining (e.g. Ahmed, Bryant & Edwards, 2020; Ngom et al., 2020), the development of infrastructure (Fiorentini & Losa, 2025), the dumping of waste (Karimi, Ng & Richter, 2022) or the production of food in remote areas (Wu et al., 2023), to name just few examples) are inextricably linked to the urban and must therefore be considered conceptually and analytically in future analyses.

- In the literature, there are many applications on urban issues based on remote sensing data: be it on spatial structure, on population, on urban heat islands, on ecological, social, planning or political topics, among many more. The studies, however, often appear to be isolated in their respective subject areas (cf. McPhearson et al., 2016). Only rarely the many interdependencies with other perspectives are analyzed in an integrative way and treated as complex systems (Batty, 2013) with manifold feedbacks, interdependencies, and non-linearities. It seems that the **topics need to be analyzed in a more multidimensional way,** i.e. in all, or at least more dependencies and interrelations. This also demands for frameworks that allow to harmonize heterogeneous spatial data to tackle these interdependencies.

  The city is defined by the most diverse life phenomena, which are concentrated within a small area and they also take place simultaneously. As one example: The remote sensing-based recording of the surface urban heat island is well known and shown many times (e.g. Stewart, 2011; Chapter XIX), but only a few studies relate the surface temperatures to air temperatures (e.g. Leichtle et al., 2023), and even fewer transfer from these temperatures to heat stress (e.g. Schaefer et al., 2021), and even fewer studies analyze

the effects of heat stress in relation to spatial structures or green areas (Chapter XXII in parts), and even fewer relate this stress to exposed groups of the population and include planning scenarios or adaptation measures, or show conflicts of use with other objectives, such as additional demand for housing or infrastructure, and we could go further. The new multi-thematic data possibilities in urban remote sensing and related fields allow and require to think about the problem in a more multidimensional way through to the end.

- The argument above is closely related to the development of **true interdisciplinary studies**. These will be, in my opinion, central in the future. Remote sensing is brilliant when it comes to describing and quantifying the spatial patterns of the world. Too often, however, the true valorization of the better information generated remains underdeveloped. '*Information is not yet knowledge*' means that research questions are too often orientated towards methodological developments or simple descriptions of the environment. These are often questions that are not based on the basic demands of other disciplines, e.g. in social sciences, economics, ecology, politics, planning, geography, history, or architecture, among many others, but on the relatively superficial knowledge of urban remote sensing specialists on the respective topic. And many thematic avenues are perhaps still untrodden because the specific domain knowledge is lacking. This also means that the transfer of a lot of generated knowledge to support governance remains underdeveloped. True interdisciplinarity is exhausting – different terminology, different schools of thought, different concepts, data and methods, and much more. However, I believe that travelling this long road is crucial and rewarding if urban remote sensing is to make an even more important contribution to the urban geography debate.

  There are certainly many good examples of true interdisciplinary studies. I would like to highlight one example here: At first glance, the disciplines of '*remote sensing*' and '*linguistics*' have little in common – in terms of their concepts, methods and thematic approaches and the respective scientific communities seldomly interact. However, the "*Geolingual Studies*" initiative (https://www.phil.uni-wuerzburg.de/gls) reveals the potential that is made possible by combining these disciplines. Understanding what, how and why something is said or reported at a certain point in time at or about a certain place and how this place can be described in its spatial characteristics and in its embedding in the larger spatial context allows a deeper understanding of the manifold dependencies of space, time, thematic fields of interest, discourses, subjective perception, and much more (e.g. Häberle, Hoffmann & Zhu, 2022; Lemoine-Rodriguez et al., 2025; Jheng et al., 2025; Mast et al., 2025).

- The data and images in remote sensing often make it possible to identify similarities and dissimilarities (e.g. of urban structures), which are sometimes visually obvious and sometimes not. Remote sensing is particularly good at making the invisible visible. **Language is needed to convey these spatial patterns beyond the pictorial**. Text and image often complement each other and yet **we often lack terminologies that make spatial (urban) patterns, geometries, the spatial configuration of objects truly describable**. Urban pattern language is still far from complete, even if there are attempts to do so (Alexander, Ishikawa & Silverstein, 1977; Fleischmann, Romice & Porta, 2021;

Wu et al., 2025; Debray et al., 2025). There is not yet sufficient and coherent terminology to do unambiguous justice to the diverse scales, objects, spatial patterns of the urban. It is in demand to develop these terminologies and concepts further, to systemize and standardize a new, more comprehensive visual language in order to make the visual semantically describable and tangible.

- Remote sensing image data has become part of everyday life. We constantly encounter them in the media or when surfing the internet. Derivatives from them, be they image manipulations or classifications, are also becoming more common. Not only has image production become highly complex – from the sensor on satellites to the digital image to a derivative – the image content is also highly complex. **Remote sensing images thus become a medium of communication.** In the urban domain, these images allow to perceive complex urban landscapes or districts or neighborhoods with a top-view perspective and usually as a visual whole. They thus allow a kind of interpretation of reality. And results presented in maps or complex images are a language of their own for communicating knowledge. Promising approaches can be mentioned here, such as the representation of settlement structures and cities in satellite data (Seto & Reba, 2018; Eisl et al., 2016; Strubelt, Dosch & Meinel, 2021) or the representation of processed information in infographics (Burdett & Rode, 2018; Rendgen, 2020). Images thus become a medium for conveying truth (Sloterdijk, 2013). **The associated creation of reality (or distorted images, AI-generated synthetic data or statistics that reflect a one-sided or even distorted view of reality) is a subject area that should also be explored in greater depth in research**. We need to be aware that the way data and images are presented defines how it changes our perception and understanding. A truly scientific approach to visual language is still not sufficiently taken into account in urban remote sensing.
- Urban remote sensing has so far been driven by new data availability, technical possibilities and knowledge-based research – usually linked to a humanistic narrative. Better data for better decisions, as the UN also prominently argues in its SDGs (UN, 2016). Big (remote sensing) data in combination with artificial intelligence or machine learning in image analysis with ever-increasing image resolutions as well as new use cases reveal that even with a humanistic research idea, many of the methods and applications can also be used in other ways and/or are dual-use applications – with consequences that have so far been little considered. This can be, for example, the bias in training data, the underrepresentation of studies for certain geographic areas, the stigmatization of a place or its inhabitants through remote sensing-based classification (e.g. by the labelling of an area as ‘slum’ and thus with a term that carries negative connotations), the backdrop raising awareness of an informally populated neighborhood by mapping, including possible eviction, among other issues. Remote sensing in general, and urban remote sensing in particular, is slowly beginning to address the manifold intrinsic (unintended and yet often thoughtless) issues (e.g. Chapter XXV; Kochupillai et al., 2022; Zhu et al., 2022b; Kochupillai & Taubenböck, 2023; Šćepanović et al., 2023; Tuia et al., 2024; Ghamisi et al., 2025; Šćepanović, Bogucka & Quercia 2025) realizing that these developments can have a direct impact on governance, policy, and the lives of

people. However**, ethical opportunities and challenges** are still poorly anchored in current research work and often little reflected upon. Technical possibility and scientific curiosity do not seem to me to be enough to tread any path, but in the future the possible negative consequences of one's own actions must be thought through and act accordingly.

These additionally identified six key research directions are certainly not complete. And yet, together with the other four key research directions, they can open up research avenues for the future or at least stimulate contradiction and discussion. In summary, Fig. 3 provides an overview of these ten research directions. Some identified and suggested in 2019, and some that crystallized from this research report to complement the other objectives.

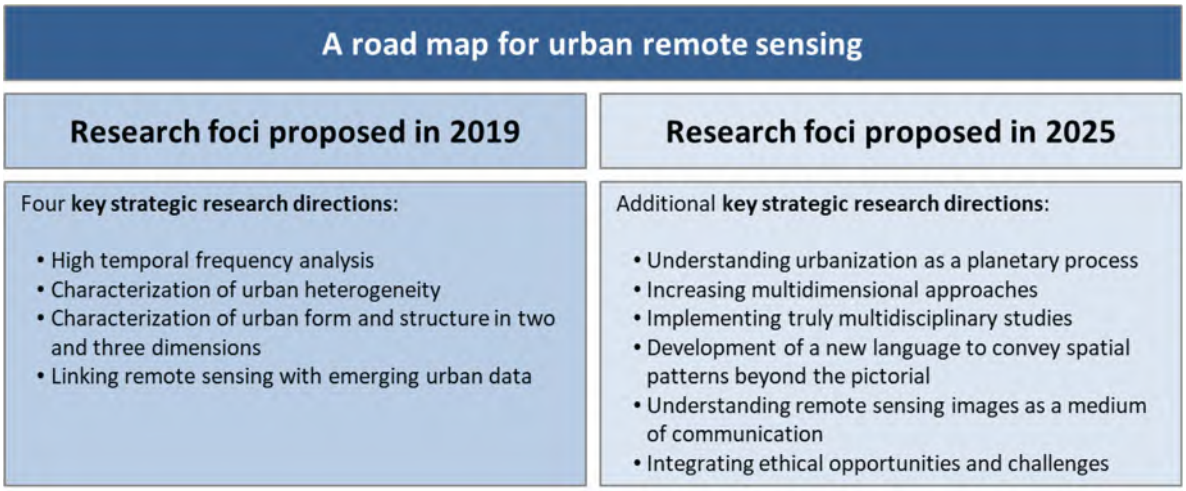


Fig. 3. A road map for future research directions of urban remote sensing based on overall observations from this research report. Source: own elaboration, unpublished.

7. Urban Science with a View from Space: A conclusion

Urban remote sensing is great: The increasing availability of data and the methodological developments of recent years have made completely new layers of geoinformation possible – on the one hand with a high spatial and temporal resolution, which does justice to the small-scale complexity of the city, and on the other hand covering large areas, often even globally, in order to capture the big trends. We now know so much more than we did five or ten years ago and this report attempts to add further empirical knowledge to the topics of urban dynamics, dimensions, patterns, effects, adaptations as well as on ethical perspectives. And yet we should be clear: Even if we know everything about building structures, we might still know nothing about social life there. What I want to say is that the complexity, diversity, simultaneity and perception of the city can never be fully grasped.

I would therefore like to take this opportunity to make a reference to the '*urban remote sensing*' community – with its intrinsic distant view on cities. Yes, sometimes '*distance creates clarity*' – and remote sensing is particularly good at this. But a reference to perhaps the most famous urbanist of the 20th century is permitted – who said: "*you need to walk the city's streets to see its soul*" (Jacobs, 1961). All the informative maps, figures and statistics

urban remote sensing produces are invaluable. But nothing in the world can replace being on site.

The privilege of repeatedly discussing the results of urban remote sensing research with colleagues from the scientific community and with the interested public often reveals the following for me: On the one hand, there is a great openness towards new findings, amazement at the dimensions and patterns of global and local urban processes that are not on everybody's radar. On the other hand, however, there is often also a great perplexity – perplexity because people want to improve social processes and yet do not want to bear the consequences of this; perplexity because the ambiguity and contradictions of many scientific results do not allow for simple solutions. This is how I would like to conclude this research report: We should bear in mind that more and better data, information and knowledge does not necessarily lead to simplification or a consensus on how cities should be developed. We have to accept that our world in general and the urban environment in particular are highly complex. Here we have to endure the ambiguities, simultaneities, contradictions and also compromises and break new ground in exchange and discourse.

Part I

Strategic directions for urban remote sensing as starting point

CHAPTER II	Understanding an urbanizing planet: Strategic directions for remote sensing	31

Chapter II

Understanding an urbanizing planet

Strategic directions for remote sensing

Remote Sensing of Environment, 2019, vol. 228

Zhe Zhu, Yuyu Zhou, Karen C. Seto, Eleanor C. Stokes, Chengbin Deng, Steward T.A. Pickett & Hannes Taubenböck

© 2019 Elsevier Inc. All rights reserved.

Note: The paper has been utilized in this research report in accordance with the reuse rights granted to the authors as outlined at: https://www.elsevier.com/about/policies-and-standards/copyright.

https://doi.org/10.1016/j.rse.2019.04.020

Received 29 Oct 2018; Accepted 16 April 2019; Published 02 May 2019

Abstract

Scientific contributions from remote sensing over the last fifty years have significantly advanced our understanding of urban areas. Key contributions of urban remote sensing include but are not limited to characterization of urban areas, urban land cover changes and thermal remote sensing of urban climates. Today, the proliferation of new sensors, long time series of the satellite record, joint analysis of Earth observation data with ancillary data sets, widespread availability of high-performance computing facilities, and slow but increasing use of remote sensing data and methods in non-remote sensing fields together offer new opportunities to generate scientific knowledge for an urbanizing planet. Simultaneously, the scale and pace of contemporary urbanization require new information about urban areas from both the science and policy communities. This paper examines remote sensing contributions to the scientific understanding of urban areas over the last 50 years until today. Based on this assessment and current needs by user communities, we identify four strategic directions for future urban remote sensing research: high temporal frequency analysis, characterization of urban heterogeneity, characterization of urban form and structure in two and three dimensions, and linking remote sensing with emerging urban data. Advances in these four areas are likely to generate significant new insights that will be useful to both science and policy.

Keywords

Urbanization; UN Sustainable Development Goals (SDGs); review; urban; human dimensions; environmental change

1. Introduction

Urban areas are central to sustainability. They generate more than 75% of global GDP, contribute to about 75% of carbon emissions from global final energy use, and produce approximately 2 billion tons of waste per year (Hoornweg and Bhada-Tata, 2012; Seto et al., 2014). Every week, the global urban population increases by about 1 million. Every day urban areas expand by an area equivalent in size to 20,000 American football fields (Seto et al., 2011; United Nation, 2018). The resources required to build and operate the cities of tomorrow will be enormous, social and governance challenges aside. The world needs scientific knowledge that can help transition society into a more sustainable urban future.

The need for this knowledge has become apparent in the last few years to governments and decision-makers, at the local, national, and even international level. Indeed, several international agreements and frameworks have highlighted the need for more information about cities and urban areas. For example, in 2015, more than 150 world leaders adopted the 2030 United Nations Sustainable Development Agenda, including a stand-alone Sustainable Development Goal (SDG) to "make cities and human settlements inclusive, safe, resilient and sustainable". One year later, 170 countries agreed to the UN New Urban Agenda (NUA). A central part of the NUA is recognition of the importance of National Urban Policies (NUPs) as a key component of achieving national economic, social, and environmental goals. In response to these efforts, decision-makers are asking for more urban knowledge from scientists that can inform policies to help create a more sustainable urban future.

Concurrently, the evidence has been mounting on how urban systems significantly impact key components of the Earth system: the atmosphere, biosphere, geosphere, and pedosphere (Zhang et al., 2013a; Zhang et al., 2013b). Beyond the environment, urban areas and the socio-technological systems within them also have been shown to affect many human and social issues, including health and well-being, economic development, and social cohesion (Eckert and Kohler, 2014). As such, a number of science communities from global change to the health sciences are calling for more science-based knowledge about urban areas in order to better understand a changing planet.

It is against the backdrop of these recent developments that we assess how remote sensing can contribute to creating knowledge of our urbanizing planet that is needed by both science and policy communities. The primary goal of this paper is to assess the key types of Earth Observation (EO) based information that informs urban research, and at what spatial and temporal scales. We review the literature and provide a broad overview of the types of information that the remote sensing community has produced to date, and suggest directions for future urban remote sensing research that could generate significant insights useful for science and policy communities.

2. The science we need for an urbanizing planet

Who are the research and policy communities that want urban information and what type of information do they need? How frequently do they need this information and at what spatial scale? The United Nations, the Intergovernmental Panel on Climate Change (IPCC), Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services

(IPBES), the UN Convention on Biodiversity (CBD), and the World Health Organization (WHO) are just some examples of the supra-national organizations that use urban information. The CBD recently initiated the second Cities and Biodiversity Outlook, an international assessment of the threats to biodiversity from urbanization. The IPCC recently convened its first Cities and Climate Change conference, the goal being to inspire new research on how cities can mitigate and adapt to climate change. The conference produced a Research and Action Agenda on Cities and Climate Science, which includes a call for more observational data at the urban scale (Prieur-Richard et al., 2018). The recent Global Urban Observation and Information Implementation Plan for 2020–2022 identifies a range of global scale and regional stakeholders for satellite-based urban information, including UN-Habitat, World Bank, and the EU Directorate for Regional Policy.

Although there is much need for information that will support these scientific and policy efforts, the majority of the scientific literature on urban areas tends to be single case studies, with significantly fewer studies that are comparative. A 2017 special issue of the Proceedings of the National Academy of Sciences of the USA highlighted the lack of comparative empirical work on urban sustainability: "*…cross-comparative empirical work on sustainability crossing typologies of urban areas and across different geographic regions is sparse*" (Seto et al., 2017). The special issue also highlighted the need for more solutions-oriented fundamental science that is both place-based and spans multiple geographic and administrative scales. In a similar vein, the urban science-policy Expert Panel for Nature Sustainability concluded that there is a need "*to forge new knowledge that corresponds to complex urban challenges*" and "*to accelerate the uptake of urban science by practitioners*" (Acuto et al., 2018).

In order for urban science to be used by practitioners, however, remote sensing scholars need to produce scientific findings that have practical utility. Currently, the majority of remote sensing studies are focused on methodological advancement, especially algorithm development. Figure 1 illustrates the number of peer-reviewed articles since 1991 in the urban remote sensing domain in the two remote sensing journals with the highest current impact factors (Remote Sensing of Environment and ISPRS Journal of Photogrammetry and Remote Sensing), and a common non-remote sensing journal (Applied Geography). We searched the archive since 1991 according to the same keywords: 'urban' and 'city', among others. In addition, in Applied Geography, we added method-specific terms like "remote sensing", "Earth observation", etc. All papers were divided into two categories: 1) 'Methods', if the main focus of the paper was on data (pre-) processing algorithms, and 2) 'Applications', if the main focus of the paper was on geographical findings. In general, we found an uptake of urban remote sensing studies. However, only recent studies were focused on applications, and most were published in Applied Geography. It is worth noting that the general increase in urban scientific remote sensing studies since 1999 was likely triggered by the advent and proliferation of very high spatial resolution satellite sensors.

There is a growing consensus among decision-makers that more data does not necessarily lead to better decision-making. The Nature Sustainability Expert Panel as well as the cities IPCC research agenda makes that clear. A lot of data being generated are not accessible either tangibly or conceptually to the decision-making or practitioner communities. Hence,

it is not just a matter of providing more data points. There is a need to provide different types of information that can be used, interpreted or digested.

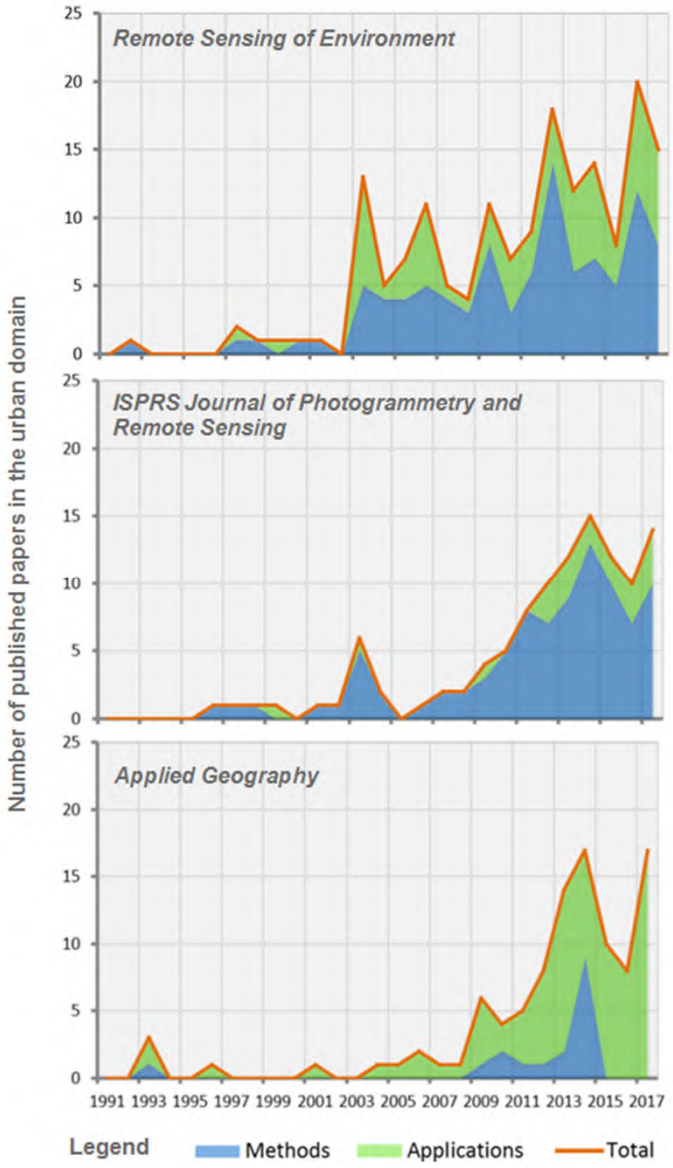


Fig. 1. Number of peer-reviewed articles since 1991 in the urban remote sensing domain for three selected journals.

Science and policy communities are calling for more and new information about variation within urban areas and comparative analysis across urban areas. In order to understand the impact of policy, we need benchmark data across spatial and temporal scales. This requires information about change, changing conditions and changing outcomes that are comparable. Although a snapshot about conditions is more valuable than no information at all, it is clear that change information is essential to know if we are making progress towards sustainable urban systems and to evaluate whether policies are effective.

Moreover, urban remote sensing studies need to be focused on a higher diversity of regions and urban area sizes. The majority of urban remote sensing studies focus on either large cities or those located in China, the U.S., or Europe. As one example, although data from the UN Urbanization Projections show that most of the future urbanization will occur in medium-sized cities of half a million to two million, 54% of urban change detection

algorithms developed since 1981 have been developed for cities of greater than five million (Reba & Seto, 2020). Nearly 40% of all urban change detection algorithms were developed for China, despite the country having 19% of the world's urban population (Reba & Seto, 2020). To better inform science and policy, urban remote sensing research must be more representative, that is, focused on a higher diversity of regions and urban area sizes.

The science and policy communities are also calling for more harmonized data collection and data interpretation efforts. There is no consensus of what is urban, or at which scales comparisons are meaningful, and remote sensing science can fill this gap. New methodologies predominantly in the machine learning domain (e.g., deep learning) are incredibly powerful ways to identify patterns of physical urban structure, but in many cases, there is limited understanding of what the outcomes mean. An open question is whether machine learning should be embraced as central to remote sensing applications or used cautiously because of its black-box characteristics (Zhu et al., 2017).

Practitioners who are engaged in design, planning, and architecture in urban areas also have expressed a need for more, relevant, and spatially explicit scientific knowledge (Felson et al., 2013; McGrath, 2018; Steiner, 2014). Urban planners and designers often think, imagine, create and communicate ideas and strategies for cities through spatial visualizations (Marshall et al., 2019). The intrinsic spatiality of EO data and their classification products contain a visual language that can relate to the workflow of these practitioners and their concerns, such as sustainability in cities, neighborhood revitalization, and installation of green and blue infrastructure to promote health and manage stormwater.

In short, urban data needs to be measured across several dimensions: within urban heterogeneity and across time. Urban data needs to be consistent and harmonized for boundaries, comparable across cities and over time. Urban data also needs to be more relevant to inform solutions such as policy, planning and design solutions and user communities such as practitioners, policymakers, and the public and private sectors.

3. Historical developments of urban remote sensing

Urban remote sensing has been evolving consistently over the last fifty years. With the first urban land cover maps derived from color infrared film on hand-held camera from the Gemini and Apollo mission in 1965 (Thrower, 1970), urban remote sensing was mostly based on aerial photo interpretation in the early days. On July 23, 1972, the launch of Landsat-1 (originally named Earth Resources Technology Satellite 1), revolutionized urban remote sensing and transitioned it from airborne to satellite remote sensing. The series of Landsat satellites have been the “gold standard” for urban remote sensing with myriads of applications (Donnay et al., 2014; Forster, 1983; Lo & Welch, 1977), because of their moderate spatial resolution that is fine enough to capture various kinds of urban developments and their long-term continuity. Since 1982, the thermal band included on Landsats 4–8 started to provide consistent measurements of land surface temperature that are ideal for urban heat island studies (Voogt & Oke, 2003; Roy et al., 2014). In 1999, the launch of commercial high spatial resolution satellites (5–0.5 m), such as IKONOS and QuickBird has further stimulated urban remote sensing research (Bhaskaran et al., 2010;

Myint et al., 2011). These commercial satellites can provide imagery at a similar spatial resolution to aerial photos, but they provide this data routinely and with synoptic coverage of Earth's surface. Global, wall-to-wall mapping of urban areas was not possible until the launch of the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor in 1999 (Friedl et al., 2002; Schneider et al., 2009).

Though a majority of urban remote sensing studies are based on daytime optical and thermal sensors, there are other sensors that also provide unique observations of urban characteristics, such as nighttime lights sensors, LiDAR, and RADAR. Nighttime light sensors can measure anthropogenic light at night, which corresponds with urban extents and certain human activities (Elvidge et al., 1997). The Defense Meteorological Satellite Program Operational LineScan System (DMSP-OLS) and VIIRS Day/Night Band are the two major sensors, and the latter provides daily observations with improved capabilities for detecting city lights at night, especially in small and less developed settlements (Román et al., 2018).

Though a majority of urban remote sensing studies are based on daytime optical and thermal sensors, there are other sensors that also provide unique observations of urban characteristics, such as nighttime lights sensors, LiDAR, and RADAR. Nighttime light sensors can measure anthropogenic light at night, which corresponds with urban extents and certain human activities (Elvidge et al., 1997). The Defense Meteorological Satellite Program Operational LineScan System (DMSP-OLS) and VIIRS Day/Night Band are the two major sensors, and the latter provides daily observations with improved capabilities for detecting city lights at night, especially in small and less developed settlements (Román et al., 2018). Additionally, LiDAR system provides new measurements of the three-dimension features of urban infrastructure, but most of the studies are only based on airborne LiDAR system at the city scale (Yu et al., 2010), with very limited studies on space-borne LiDAR at the global scale (Gong et al., 2011). RADAR data has been used to map urban growth since the launch of the European Remote Sensing Satellite 1 (ERS 1) in 1991, but most of the studies are small scale applications. Nevertheless, with more RADAR satellites launched in the last decades (e.g., TerraSAR-X, TanDEM-X, and Sentinel-1), continental and global scale urban mapping is emerging (Esch et al., 2012; Lisini et al., 2018). Moreover, recent studies suggest RADAR and optical data are complementary with each other, and the combined use of the two can provide more accurate urban mapping results (Zhu et al., 2012). The detailed list of civilian satellites that have been discussed in this review for urban remote sensing is shown in Table 1.

Table 1. A list of civilian satellites discussed in this review for urban remote sensing. From is launch date, to is end of (imaging) life, (Active means the satellite or at least one of the satellite series is still actively collecting data) Pan (Panchromatic), Optical (Optical multispectral sensor that is capable of recording in at least red and near-infrared), Thermal, ighttime Light, LiDAR, and SAR provide highest resolution (in meters) for sensors carried by each mission. MSS: Multispectral Scanner System; DMSP: Defense Meteorological Satellite Program; OLS: Operational Linescan System; NOAA: National Oceanic and Atmospheric Administration; AVHRR: Advanced Very High Resolution Radiometer; TM: Thematic Mapper; European Remote Sensing (ERS); SPOT: Satellite Pour l'Observation de la Terre; SAR: Synthetic Aperture RADAR; ETM+: Enhanced Thematic Mapper Plus; MODIS: Moderate Resolution Imaging Spectroradiometer;

ASTER: Advanced Spaceborne Thermal Emission and Reflection Radiometer; Envisat: Environmental Satellite; ASAR: Advanced Synthetic Aperture RADAR; MERIS: MEdium Resolution Imaging Spectrometer; ICESat: Ice, Cloud, and land Elevation Satellite; GLAS: Geoscience Laser Altimeter System; HJ: Huanjing; TanDEM-X: TerraSAR-X add-on for Digital Elevation Measurement; TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurement); Suomi-NPP: Suomi National Polar-orbiting Partnership; VIIRS: Visible Infrared Imaging Radiometer Suite; OLI: Operational Land Imager; TIRS: Thermal Infrared Sensor; SLSTR: Sea and Land Surface Temperature Radiometer. HyspIRI: Hyperspectral Infrared Imager.

Satellite	Sensor	From	To	Pan	Optical	Thermal	Nighttime light	LiDAR	SAR
Landsat 1–3	MSS	1972	1983		79 × 57				
DMSP F1–F19	OLS	1976	Active				560		
NOAA 6–19	AVHRR	1979	2019		1100	1100			
Landsat 4–5	MSS/TM	1982	2013		30	120			
SPOT 1–4		1986	2013	10	20				
ERS 1–2	SAR	1991	2011						30
Landsat 7	ETM+	1999	Active	15	30	60			
IKONOS		1999	2015	1	3.3				
Terra	MODIS	1999	Active		250	1000			
Terra	ASTER	1999	Active		15	90			
QuickBird		2001	2014	0.6	2.4				
Envisat	ASAR	2002	2012						30
Envisat	MERIS	2002	2012		300				
Aqua	MODIS	2002	Active		250				
SPOT 5		2002	2015	2.5	10				
ICESat	GLAS	2003	2010					70	
WorldView 1		2007	Active	0.5					
TerraSAR X		2007	Active						1
HJ-1A		2008	Active		30				
HJ-1B		2008	Active		30	300			
WorldView 2		2009	Active	0.46	1.84				
TanDEM-X		2010	Active						1
Suomi-NPP	VIIRS	2011	Active		375	375	750		
ZiYuan3 01-02		2012	Active	2.1	30	100			
SPOT 6–7		2012	Active	1.5	8				
Landsat 8	OLI/TIRS	2013	Active	15	30	100			
Sentinel-1		2014	Active						5 × 20
Sentinel-2		2015	Active		10				
Sentinel-3	SLSTR	2016	Active		500	1000			
Luojia 1-01		2018	Active				130		
HyspIRI		Expected in 2022	3–5 years		60	60			

4. Urban remote sensing contributions to environmental change and human dimensions research

In this section, we briefly summarize major contributions from remote sensing studies to knowledge about urban areas for environmental change and human dimensions research. Because methodological advances in urban remote sensing have been covered extensively in the literature elsewhere (Voogt & Oke, 2003; Weng, 2012, 2009; Wentz et al., 2014), we limit this discussion to contributions to non-algorithmic topical knowledge that provide measurable advances in the understanding of key urban processes.

Urban areas are multi-spatial coupled human-environment systems. By their very nature, urban systems link ecological, physical, and socioeconomic systems across spatial scales (Pickett et al., 2001). As such, efforts to separate them into various components or categories will always be limited, and absolute delineations of purely social or biophysical are not possible. Nevertheless, for this analysis, we distinguish between remote sensing contributions that focus more on issues of environmental change versus those that are more human in scope, recognizing that some topics, such as the built environment, span both categories. Furthermore, we analyze various applications of urban remote sensing in environmental change and human dimension research based on the relationship of spatial resolution versus spatial scale, and temporal frequency versus time scale, respectively (Fig. 2).

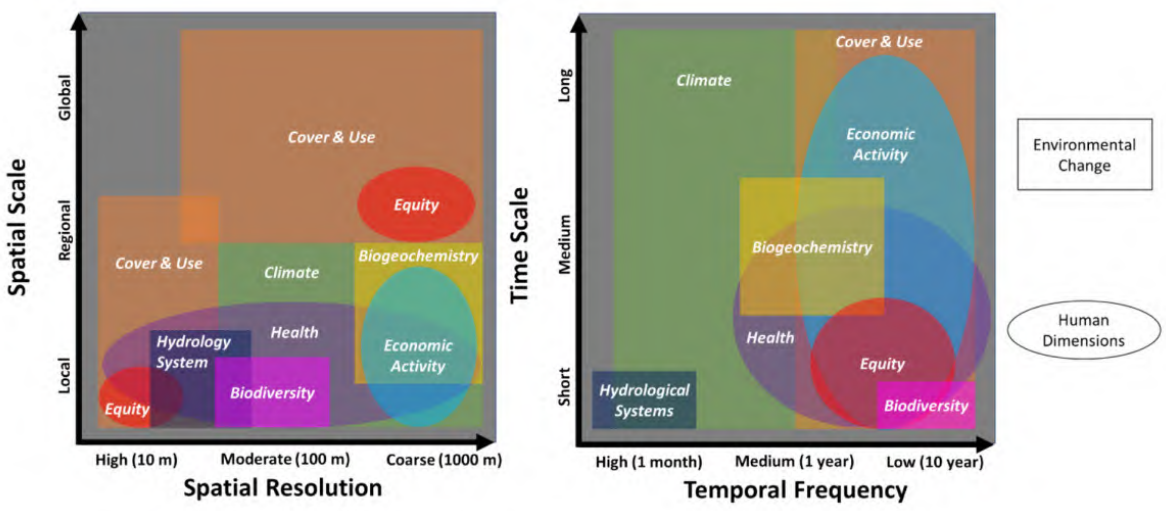


Fig. 2. Spatial resolution vs spatial scale and temporal frequency vs time scale for the remotely sensed imagery that have been applied in urban areas for environmental change and human dimension research. The sizes and locations of each rectangle or ellipse are approximate values estimated based on all the papers reviewed in this study.

4.1. Contributions to environmental change research

Although there is a myriad of ecological processes to which urban areas contribute, we focus here only on environmental change, defined as change or disturbances of the environment, most often caused by human influences and natural ecological processes. Urban areas cause environmental change primarily through five mechanisms: 1) land use and cover, 2) biodiversity, 3) climate, 4) hydrological systems, and 5) biogeochemistry (Grimm et al., 2008). We briefly describe how urban remote sensing contributes to knowledge in these five domains.

4.1.1 Land use and cover

4.1.1.1 Defining and mapping urban areas

There are nearly 200 United Nations member states and almost every country has a different definition of urban. Urban areas are principally a type of human settlement, and about one hundred UN member states use countryspecific minimum population thresholds to differentiate “urban” from other types of settlements. In addition to population thresholds, many UN member states define urban areas as human settlements which have a mix of built infrastructure (e.g. road, permanent dwellings, and municipal services such as electricity), and where a significant portion of the population engages in the non-agricultural economy. An examination of the definitions suggests that aggregating them into a single estimate is problematic for many reasons, including significant differences in scale and criteria. For example, urban is defined in Norway as localities with more than 200 inhabitants, while in Japan, it is defined as 60,000 inhabitants plus “urban facilities” and nonagricultural economic activities. In Nicaragua, localities need to have at least 2000 inhabitants as well as streets and electric lights in order to be deemed urban. Since there is no universal definition of urban, statistics on global urban population are developed by aggregating these vastly differing definitions, which is problematic.

Hence, one of the principal contributions of remote sensing to urban knowledge to date has been to characterize, measure, and map urban areas in a consistent fashion. Instead of relying on population size or density, urban areas are mapped based on the physical characteristics of land surface (e.g., the amount of impervious surface area or built-up area). From satellites, some of the earliest urban insights were the use of multispectral data to differentiate urban land cover from agriculture and forests (Haack, 1983; Quattrochi, 1985), to discriminate urban textures from surrounding rural areas using RADAR data (Henderson et al., 1980; Bryan, 1975), and to estimate urban population and densities (Kraus et al., 1974; Murai, 1974).

Today, with the routine collection of Earth observation data from a variety of satellites, more studies are focused on mapping urban areas at large scales (e.g., global urban area mapping), something that was unimaginable thirty years ago. Most of the global urban maps are derived from coarse spatial resolution images (300 m–10 km), such as MODIS (Schneider et al., 2009), Medium Resolution Imaging Spectrometer (MERIS) (Arino et al., 2008), and DMSP-OLS (Small et al., 2005; Zhou et al., 2015). Recently, a few global urban maps have been created with moderate to high resolution (10 m–50 m), including the Global Land Cover product (GlobeLand30) using Landsat and China Environmental Disaster Alleviation Satellite (HJ-1) (Chen et al., 2014), the Global Human Settlement Layer (GHSL) based on Landsat images (Pesaresi et al., 2016b), and the Global Urban Footprint (GUF) generated from RADAR satellite constellation of TerraSAR-X and TanDEM-X (Esch et al., 2012). Although estimates of global urban areas vary significantly, in part due to differences in the spatial resolution of the imagery and the varying accuracies across the globe (Klotz et al., 2016), these maps have been essential in building a scientific knowledge of the direct urban imprint on Earth.

While it is easier to map urban land cover as the physical evidence of the environment, it is more difficult to map urban land use – the human activities on land, particularly in modern cities where buildings often host multiple purposes and can be renovated and repurposed. The distinction of land cover and land use is important because land cover information relates to the natural environment of the urban system, while land use data relates to the human dimensions. Notably, some researchers argue for a joint human-natural structural characterization of urban (Rademacher et al., 2019). Texture, contexture, and proportional land cover data have been used in conjunction with spectral data to extract urban land use information (Gong & Howarth, 1990, 1989; Gong & Howarth, 1992a, 1992b).

4.1.1.2 Urban land cover and land use change

In addition to mapping a snapshot of urban areas at a single time point, the notion of using

satellite data to identify urban change, such as from non-urban to urban land cover, has a long history (Howarth & Boasson, 1983; Maxwell & Riordan, 1980). Even before the launch of Landsat 1, geographers proposed urban change detection methods using aerial photography (Dueker & Horton, 1972). Thus, mapping urban areas and urban land cover change have been a fundamental contribution of remote sensing to urban knowledge.

The international research community has been instrumental in developing knowledge about urban areas. In particular, the inception of the NASA Land-Cover/Land-Use Change (LCLUC) program in 1996, the International Human Dimensions Programme

(IHDP)/International Geosphere-Biosphere Programme (IGBP)'s Land Use/Cover Change international science project that existed from 1994 to 2005, and the IHDP Urbanization and Global Environmental Change Project that ran from 2005 to 2016, all helped to foster an international research community that examines land dynamics, including urban land change, and associated drivers and environmental and human impacts (Lambin et al., 2006; Seto et al., 2014, 2015). Urban land systems are now an integral part of the land system science research, a legacy that continues with the Future Earth Programme.

Remote sensing data have been used extensively to track urban expansion (Angel & Sheppard, 2005; Taubenböck et al., 2012; Wang et al., 2012; Zhou et al., 2018), to determine the drivers of urban land demand (Burchfield et al., 2006), and to forecast future urban growth (Seto et al., 2012). Since 1988, over 1080 peer-reviewed papers in English language journals covering over 800 locations have used remote sensing to detect, characterize and map urban land cover and land use change (Seto et al., 2011). Mapping urban land cover change continues to be a dominant component of urban remote sensing research and has been essential to understanding the drivers of urban development patterns. To date, satellite-based studies have shown that urban expansion is driven by population and economic growth, transportation infrastructure, governance and planning controls, and characteristics of the natural environment (Angel & Sheppard, 2005; Burchfield et al., 2006; Christensen & McCord, 2016; Seto et al., 2011).

4.1.2 Biodiversity

Urbanization impacts biodiversity directly through urban expansion, and indirectly through supply chains and consumption. Here we limit our discussion to direct urban impacts on biodiversity through land use. We discuss the insights from remote sensing of vegetation and green space within urban areas as well as urban expansion and habitat and biodiversity loss.

Though urbanization usually results in a reduction in biodiversity, urban vegetation, and green space can also play a vital role in supporting biodiversity by providing wildlife migration corridors or by acting as refuges for native biodiversity (Goddard et al., 2010). Advanced information on urban vegetation condition and green space distribution can assist planners in designing strategies for the optimization of urban ecosystem services and biodiversity conservation. Remote sensing is of great value in monitoring the condition of urban vegetation and mapping the spatial distribution of urban green space. Such studies are most often carried out at moderate scales, with a spatial resolution of 30 m (Small, 2001; Zhu et al., 2016). It has been demonstrated that finer resolutions are required to assess urban green space more comprehensively, since coarse imagery would miss small patches of vegetation (Qian et al., 2015). Habitat fragmentation is one of the most important concerns about urban expansion. Agricultural and pastoral lands, as well as production and native forests, remnant grasslands, and deserts, are threatened by urban land conversion. At the coarse scale, fragmentation in these productive, biodiverse, and native habits is readily detected by remote sensing (Skole & Tucker, 1993). Additionally, biodiversity as a concept has several dimensions. It addresses heterogeneity from the genetic to the landscape levels of organization. Most often in discussions of urban biodiversity, it is the species level that is of concern. Although the current satellite remote sensing systems are still too coarse to discriminate among species, they are increasingly feasible for mapping certain aspects of

biodiversity, such as distinguishing species assemblages or diversity patterns (e.g., richness) (Alonzo et al., 2014; Bino et al., 2008; Goetz et al., 2007; Lawes & Wallace, 2008; Pu & Landry, 2012; Turner et al., 2003; Xiao et al., 2004).

4.1.3 Climate

Urbanization may profoundly affect climate locally, and even globally, by land cover conversion, increasing impervious surface, anthropogenic heat discharge, and greenhouse gases emissions (Chen et al., 2006; Wakode et al., 2018; Xiao et al., 2007; Zhou et al., 2012).

These land surface processes that occur at the interface between the Earth's land and atmosphere are complex and influence the climate system at different scales (Gluch et al., 2006). One of the most important effects in climate induced by urbanization is the Urban Heat Island (UHI), referring to the temperature differences between urban and surrounding rural areas (Oke, 1982, 1973; Roth et al., 1989). Biogeophysical parameters of urban land surface cover and temperature from satellite observations are of great value for investigating urban surface energy budgets, and therefore improving the understanding of urbanization impacts on local and global climate.

With the advent of thermal remote sensing satellites, we are able to observe UHIs at large spatial scales and understand their primary causal factors (Voogt & Oke, 2003; Weng, 2009). Since the first UHI study by Rao (1972), a variety of remote sensing platforms and sensors have been used, of which Advanced Very High Resolution Radiometer (AVHRR) (Gallo & Owen, 1998; Kidder & Wu, 1987; Streutker, 2002), MODIS (Li et al., 2017; Meng et al., 2018; Schwarz et al., 2011), Landsat (Carnahan & Larson, 1990; Kim, 1992; Weng et al., 2004), and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (Kato & Yamaguchi, 2007; Nichol et al., 2009; Tiangco et al., 2008) are the four most important sensors. Airborne data have also been widely explored in UHI studies (Ben-Dor & Saaroni, 1997; Jenerette et al., 2016). Although the coarse resolution data from AVHRR and MODIS are particularly useful for mapping urban temperature over large areas, they are not ideal for establishing meaningful relationships with ground measurement. Rather, moderate resolution data with spatial resolutions ranging from 60 to 120 m (e.g. Landsats 4–8), are more suitable to study UHI at regional scales. Studies that have used data at these scales have found that the Land Surface Temperature (LST) products derived from Landsat images can better reflect the spatial variation of LST, which is of great potential for study areas with high spatial heterogeneities. High spatial resolution imagery from airborne remote sensing systems have also been used to extract temperature from specific urban surfaces for analysis (Ben-Dor & Saaroni, 1997; Eliasson, 1992; Gaitani et al., 2017).

There is well-developed literature that uses large scale thermal remote sensing data to explore how urban biophysical factors relate to LST. These biophysical parameters extracted from satellite observations include land use and land cover type (Amiri et al., 2009; Li et al., 2009), vegetation indices (Gallo et al., 1995), subpixel vegetation abundance (Weng et al., 2004), subpixel impervious surface fraction (Imhoff et al., 2010; Yuan & Bauer, 2007), and landscape compositions and combinations (Li et al., 2011; Zhou et al., 2011). Studies have found that LST is usually negatively related to vegetation indices (e.g. NDVI), and positively correlated with urban land cover indices (e.g. impervious surface area). Instead of the

traditional binary interplay between LST and a single spectral indicator, UHI is better modeled and explained as multiple interplays between LST and various land biophysical compositions with different thermal properties (Deng & Wu, 2013a). From a modeling perspective, LST is a product of the surface energy balance, and the physical basis of spatial variation in LST is the response to the factors that affect surface energy balance. These small spatial scales measurements can be made using ground or airborne thermal remote sensing. Therefore, the issue at the satellite scale is that much of the spatial variability of surface temperature is still at subpixel scale (mixed with issues related to emissivity and sensor viewing geometry). In addition to examining the direct association between LST and urban biophysical factors, surface UHI intensity, magnitude, and spatial extent have also been modeled as continuous functions (Rajasekar & Weng, 2009a, 2009b; Streutker, 2003, 2002).

New insights on urban climate are emerging in the remote sensing community, enabled by new data, new sensors, and new algorithms. First, remotely sensed big data (especially time series) has created the potential to depict and analyze UHI over time. Using these time series, the relationship between LST increases and urbanization can be determined (Fu & Weng, 2016). Also, because of the sensitivity of urban vegetation to urban climate change, long term time series satellite observations can be used to compare the annual vegetation phenological transition dates in urban areas to quantify the impacts of urbanization on urban climate (Zhang et al., 2004). Other new data, such as volunteer geographic data on weather websites, have been collectively used with remote sensing data for urban climate studies (Ho et al., 2014). Second, new sensors are emerging that can provide more spectral features or more frequent time series of urban climate variables. Two or more thermal bands are known to be able to provide more accurate estimation of land surface temperature and emissivity in urban areas (Wan & Dozier, 1996). However, most of the current satellite sensors are either without any thermal band or only have a single thermal band. The Sentinel-3 Sea Land Surface Temperature Radiometer (SLSTR) that was launched in 2016 has three thermal bands for better estimates of land surface temperature (Donlon et al., 2012). The Hyperspectral Infrared Imager (HyspIRI) which will be launched in 2022 has eight thermal bands between 3 and 12 μm (Lee et al., 2015). Moreover, recent studies have explored the use of new geostationary satellites to create hourly land surface temperature at a similar spatial resolution to that of Landsat, which greatly improves our understanding of the spatiotemporal dynamics of urban thermal environments (Quan et al., 2018). Finally, new algorithms are being developed that improve on the low temporal frequency of existing moderate resolution thermal data by modeling of diurnal and annual time series that couples equations with remotely sensed data. For example, algorithms that are able to simulate daily or even hourly Landsat thermal images by using data fusion have been developed (Gao et al., 2006; Liu & Weng, 2012; Quan et al., 2018; Zhu et al., 2010).

4.1.4 Hydrological systems

Urban land cover and land use, especially urban form, have extensive impacts on hydrological processes such as surface runoff and infiltration in urban systems. The geometry of road networks, the height of buildings or infrastructure, and the configuration of pervious and impervious surfaces, have notable impacts on the flow vector and water

depth during the surface runoff process in urban areas (Vojinovic et al., 2011; Yu & Coulthard, 2015), particularly during flooding events.

Recent innovations in this area include the use of SAR data to map flood dynamics and hydrologic characteristics such as channel and floodplain connectivity (Schumann et al., 2011). Satellite data can provide near real-time and cost-effective data for estimating stormwater runoff, urban floods, and other hydrological changes. SAR data can provide near real-time flood water levels that are closely correlated with data collected from gauges (Mason et al., 2012). The causes and propensity of urban-induced rainfall have also been explored, using remotely sensed imagery (Shepherd, 2005). These insights have important implications for post-disaster response, hazards management, and disaster risk reduction. Moreover, satellite-derived estimates of effective impervious surface area is an essential indicator for urban runoff and floods (Ebrahimian et al., 2016).

In addition to SAR, high resolution satellite data can provide finely resolved terrain information that is invaluable for modeling urban flood inundation. Although high resolution terrain data have been generated for rural and large floodplain areas for some time, urban areas present unique challenges due to the built-up infrastructure and small spatial dimensions of roads. Airborne LiDAR data have been used to estimate urban vertical structure with height accuracies in the range of 10–15 cm Root Mean Square Error (RMSE), which is sufficient for urban flood modeling (Schubert et al., 2008).

There is a strong correlation between urban development, urban subsidence and urban flooding (Abidin et al., 2011). As such, having information about subsidence can assist in the modeling of urban floods. InSAR data have also been used to quantify subsidence in urban areas (Crosetto et al., 2003; Raucoules et al., 2013).

4.1.5 Biogeochemistry

Although biogeochemistry is a key component of urban ecosystems (Kaye et al., 2006), assessments by remote sensing are largely absent. Most remote sensing assessments of biogeochemical parameters such as productivity and productivity potential, or concentrations of nutrients in the system of interest have not addressed urban systems. Rather they have been applied in large, contiguous areas of non-urban aquatic marine, or terrestrial habitats (Asner & Vitousek, 2005; Schimel, 1995). A few studies have tried to model the carbon cycle of urban systems based on light use efficiency model and coarse resolution remotely sensed data from AVHRR (Imhoff et al., 2004; Zhao et al., 2007) and MODIS (Milesi et al., 2003). Different urbanization impacts on the carbon cycle have been reported, where both Imhoff et al. (2004) and Milesi et al. (2003) suggested that urbanization had a large negative impact on net primary productivity, while Zhao et al. (2007) indicated

that the region-wide increase in tree cover in urban areas actually increased regional gross primary production. A more integrated evaluation of the most prominent impacts of urbanization and finer spatial resolution of satellite images are needed for better understanding of urbanization impacts on the carbon cycle.

Considering that the carbon cycle is just one of the many processes in urban biogeochemistry and the relevant studies using remote sensing data are very limited, there is a great need for

remote analysis of urban biogeochemistry. We think future research should use higher spatial resolution remotely sensed data and include fine and medium scales (e.g., Landsat and Sentinel-2 data) at which ecosystem process are measured in watersheds and in the various terrestrial habitat types found in and around cities.

4.2. Contributions to human dimensions research

Remote sensing is a powerful tool in a wide variety of human dimensions research. Due to its high spatial and temporal resolution and global availability, remote sensing data is a complement to demographic, social, economic, and health surveys that are sparse and difficult to obtain, and often limited in geographic coverage. Here we discuss the key insights of urban remote sensing data to human dimensions research, focusing on health, economic activity, and equity. We organize the discussion around the main topics where satellite data have been applied and the key remote sensing contributions to each area.

4.2.1 Health

Urbanization impacts health directly both through the natural environments that are altered and the urban environments that are built, and indirectly through the way urban environments shape human behaviors. Insights from remote sensing on monitoring these impacts are organized around three types of pathways: how urban environments affect the prevalence and spread of infectious diseases, how they influence environmental exposures that lead to non-communicable disease or mortality risks, and how they shape human behaviors and lifestyles.

4.2.1.1 Infectious diseases

Infectious diseases can be caused by various biological agents, such as bacteria, viruses, parasites or fungi, and can be spread between people either directly or indirectly. Though the agents that cause many infectious diseases are not unique to urban areas, urban environments can shape the breeding ground of disease carrying vectors and can influence interactions between vectors and human hosts (Goodman et al., 2018; LaDeau et al., 2015), or human hosts with each other. Urban remote sensing of the study of infectious diseases has therefore revolved around three topics: (1) understanding the distribution of disease vectors; (2) mapping urban environments that increase disease risk; and (3) understanding disease transmission within urban areas.

While remote sensing has been used extensively to map the environmental attributes that influence the reproductive rates and distribution of a wide variety of disease vectors, studies with an urban focus are almost entirely limited to research on mosquitos. Mosquitos are the primary vector of dengue, yellow fever, Zika, and malaria, and controlling them within urban areas is a policy priority for many developing cities. Recent research in Argentina (Albrieu-Llinás et al., 2018), the US (Hartfield et al., 2011; Liu & Weng, 2012), and parts of Africa (De Castro et al., 2004; Kabaria et al., 2016) have used high resolution imagery, and sometimes LIDAR, to monitor environmental variables for prediction of mosquito distribution and creation of probability maps. These types of studies have also shown how the same environmental conditions that cause urban heat islands can cause an increased incidence of mosquito-borne diseases. In Sao Paulo, for example, urban heat islands, mapped

with land surface temperature data, were shown to have more dengue cases (Araujo et al., 2015). Risk distribution studies aim to provide urban decision makers with the information they need to try to control mosquito populations and mitigate outbreaks.

Other studies have focused on disease risk associated with the quality of urban built environments. It is estimated that about one out of three urban residents live in a slum or informal settlement, where the combination of high levels of poverty, dense housing, inadequate sanitation and waste collection, and poor drinking water supply contribute to high levels and rapid transmission of infectious disease. Understanding the urban mechanisms underlying infectious disease transmission is essential for early warning systems and control. Remote sensing has shown promise for slum mapping by differentiating between textures of formal and informal built environments (Kuffer et al., 2016; Wurm et al., 2017). These contributions can help identify communities for targeted interventions.

In addition to mapping the distribution of risk, urban remote sensing has been combined with demographic data to understand how human populations have interacted with disease vectors, and with each other to influence epidemic progression. One study tracked human seasonal population migrations in and out of urban areas in Niger using a time series of night-time light satellite images and found that these spatiotemporal changes in population density were associated with measles outbreaks (Bharti et al., 2011). Another study that mapped the temporal changes of malaria transmission risk in Dakar, Senegal found that urbanization decreased the risk of malaria transmission (Machault et al., 2010).

4.2.1.2 Environmental exposure and non-communicable diseases

Though less developed, there is a growing literature on the impact of environmental factors, measured with remote sensing data, on noncommunicable diseases such as diabetes, obesity, heart health, cancers, and asthma, which are prevalent in urban environments. Though areas with high urbanization levels have been linked to a higher prevalence of chronic disease in recent decades (Li et al., 2012), the causal mechanisms are often less clear.

One mechanism, air pollution, is a common problem in many urbanizing areas that has been linked to several chronic diseases. Because of the absence of networks of ground-based indicators of particulates in some parts of the world (Engel-Cox et al., 2004; Schaap et al., 2009), the demonstrated relationship between aerosol optical depth (derived from MODIS) and standard particulate measures has proven a useful in epidemiological exposure assessment studies (Brauer et al., 2012; Jerrett et al., 2017; Kloog et al., 2011; Kumar et al., 2007; van Donkelaar et al., 2010; Wang & Christopher, 2003). However, because MODIS satellite observations only date back to 1999, satellite-based studies are limited in their characterizations of long-term changes in PM 2.5 (Butt et al., 2017).

Along with air pollution, the greenness of urban neighborhoods (measured in remote sensing as NDVI) has been linked to cardiovascular disease (Pereira et al., 2012). Urban light pollution, measured with night-time light satellites data has been shown to correlate with breast cancer (Kloog et al., 2010) and cardiovascular disease (Lane et al., 2017). Pollen released by vegetation in the urban and suburban area (measured in remote sensing as vegetation phenology) is the most important factor for respiratory allergies (Li et al., 2019a).

In addition, the health impacts of climate change are likely to be severe in urban areas because of the high concentration of people and infrastructure in one place. In the Carbon Disclosure Project's 2017 survey of leaders from 478 global cities, over half of respondents expected climate change to seriously compromise their public health infrastructure (Watts et al., 2018). Remote sensing is increasingly being used to understand the exposure of urban residents to hazards like extreme weather events, sea level rise, heat waves, flooding and landslides, associated with climate change. For instance, high resolution AVHRR and SPOT imagery were used to analyze the spatial variations in land surface temperature during the 2003 Paris heat wave that caused 4867 deaths (Dousset et al., 2011). Satellite-based emergency mapping, using SAR, VHR optical imagery, and thermal imagery was critical in major disaster events like the 2004 Indian Ocean Tsunami (Voigt et al., 2007), the Wenchuan Earthquake of 2008 (Tong et al., 2012), the Haiti earthquake in 2010 (Duda & Jones, 2011), and the Pakistan flood of 2010 (Gaurav et al., 2011). Recent disaster assessments have also included the use of night-time light sensors–tracking power outages from earthquakes, floods, and storms (Zhao et al., 2018), as well as long-term power restoration protocols after storms (Román et al., 2019).

4.2.1.3 Behaviors and lifestyles

There is additional epidemiological literature that has linked characteristics of urban environments with daily health habits, like physical activity (Sallis et al., 2016; Feng et al., 2010), dietary behaviors (Seto & Ramankutty, 2016), sleep (Carta et al., 2018), and access to health care (Sibley & Weiner, 2011; Caldwell et al., 2016). Measures of urban structure, such as street intersection density, land use mix, population, and housing unit density have all been positively associated with physical activity (Sallis et al., 2016; Feng et al., 2010; Saelens et al., 2003; Troped et al., 2010). However, few studies have used remote sensing data to measure these attributes of the urban built environment, more often relying on GIS or survey data.

4.2.2 Economic activity

Scholars have long been interested in the ability of remote sensing data to proxy economic activity. To this end, night-time light captured by the DMSP-OLS, has become commonly used proxies for local economic activity, since they reflect variation in population density (e.g. Sutton et al., 2001), energy use (e.g. Amaral et al., 2005), and infrastructure development (e.g. Imhoff et al., 1997). In cross-sectional studies, night-time light has been shown to correlate strongly with economic activity (Doll et al., 2000; Sutton & Costanza, 2002; Xin et al., 2017). As a thirty-year panel dataset, change in lights has also been used to approximate change in GDP growth, with a lights-GDP elasticity of 0.28–0.32 (Henderson et al., 2011).

Nightlights data have two main advantages for approximating economic activity and economic growth that are complementary to traditional measures. First, night-time light data is global in scope, whereas traditional economic data is difficult to assess for countries with low-quality statistical systems in place. Night-time lights have been used to estimate economic activity primarily in countries where data is sparse (Lee, 2018; Michalopoulos & Papaioannou, 2014). When included with traditional measures of GDP, nightlights improved

the measurement of economic output in areas that had inferior statistical collection systems but added little information for areas with high-quality economic data available (Deville et al., 2014; Nordhaus & Chen, 2015). The second advantage is the fine spatial resolution of night-time light data, compared to conventional economic data, which is most often reported at the national level. Since sub-national data is typically unavailable, night-time light can be used to create a disaggregated map of economic activity (Ghosh et al., 2010), with lights as the means of proportionally allocating the national GDP. Downscaled

proxies of national economic activity have been used in research about the determinants of variation in economic activity across cities (Florida et al., 2012), subnational administrative regions (Hodler & Raschky, 2014), and grid cells (Henderson et al., 2017). The launch of a new Day/Night Band on the Visible Infrared Imaging Radiometer Suite (VIIRS) with higher spatial, radiometric, and temporal resolution (Román et al., 2018; Román et al., 2019) presents a new opportunity to expand on previous efforts for using night-time light to understand economic activity.

Beyond night-time lights, the concentration of the built environment derived from passive or active sensors has been used to map and understand economic activity. At the continental scale city networks measured by settlement density are shown as a well-suited proxy for revealing economic disparities (Taubenböck et al., 2017a, 2017b). At the city scale, the specific characteristics of settlement structures, such as morphological slums, are also a proven proxy for the economic capabilities of a social group (Wurm & Taubenböck, 2018).

4.2.3 Equity

Though urban areas are sites of concentrated innovation and opportunity, they are also increasingly sites of concentrated poverty and inequality. Very high-resolution optical satellite images have been used successfully to identify qualitative aspects of the built environment that identify informal housing (Kohli et al., 2012; Kuffer et al., 2016; Owen & Wong, 2013; Taubenböck & Kraff, 2014) and concentrated poverty. However, since morphologic forms of poverty vary significantly across the world (Taubenböck et al., 2018), detecting slums based on built-up characteristics in EO-data capture only one aspect of poverty. A combination of DMSP-OLS nighttime imagery with high-resolution Google Static Maps imagery were used to estimate household consumption and assets in five African countries—Nigeria, Tanzania, Uganda, Malawi, and Rwanda—all of which lack high resolution income data (Jean et al., 2016). In addition to mapping informal settlements, nightlights images have provided a new dataset for assessing energy poverty. Studies using DMSP have demonstrated the capability of night-time light to track rural electrification efforts in developing countries (Doll & Pachauri, 2010; Min et al., 2013; Min & Gaba, 2014). However, at present, these studies have only considered electricity access as a binary—whether there are lights or not. At higher resolution, Kuffer et al. (2018) have used night-time light data taken from ISS Astronauts to highlight the lower light emission in slum areas compared to formal settlement areas. In addition, electrification and infrastructure-poor development within urban areas have been identified by considering time series data from the VIIRS Day Night Band, land cover records, and population datasets together (Stokes & Seto, 2019). Capturing a more nuanced picture of the gradients of energy poverty and access could be a promising future direction of remote sensing research.

5. Strategic directions for urban remote sensing research

Improving the scientific understanding of urban areas in a rapidly changing and urbanizing world requires four key advancements (Fig. 3). First, the rapid rates of urban change require high frequency observations as well as a long temporal record. Second, urban environments vary significantly around the world as well as within particular cities, and there is a rich array of urban features and dynamics that might be better assessed by remote sensing. Third, the details of urban form and structure, the urban surface, and across the vertical dimension must be better and more widely known. Finally, linking remote sensing with emerging urban data that reflects the ecological, social, and economic processes of urban areas is urgently needed. We discuss these distinct, but related topics in this section.

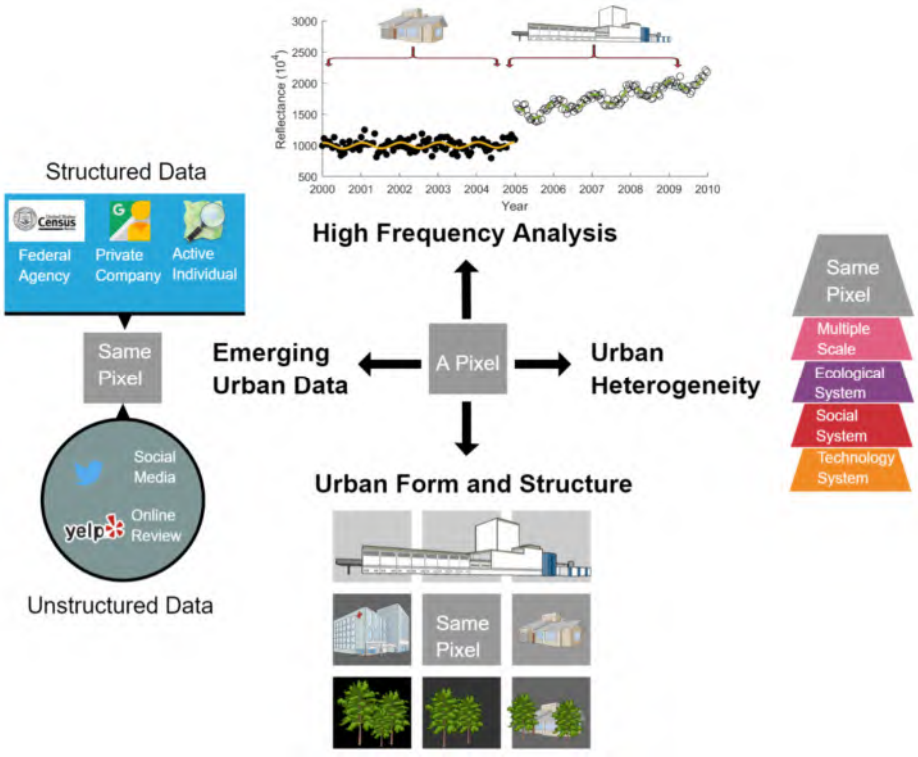


Fig. 3. Four major strategic directions for urban remote sensing research.

5.1. High frequency analysis of urban environments

Unlike forests that can change frequently, urban environments are generally assumed to be a relatively stable land type (Mertes et al., 2015). Therefore, historically, urban areas were mapped at a time interval of 5 to 10 years apart (Seto et al., 2002; Xian & Homer, 2010; Yuan et al., 2005). However, in reality, urban environments are highly dynamic and can undergo many qualitative and subtle changes for which high frequency observations would generate significant new insights (Li et al., 2018). High temporal frequency analysis for urban remote sensing has become an emerging trend with the improvements in satellite technology (Schott et al., 2016), the open data policy (Zhu et al., 2019a), and application of methods from allied fields, such as geostatistics (Boucher et al., 2006), time series methods (Kaufmann & Seto, 2001; Zhu, 2017), among others. Nevertheless, the use of high frequency satellite data did not become the frontier of research until the Landsat data were provided

free of charge after 2008 (Woodcock et al., 2008; Wulder et al., 2012; Wulder et al., 2019; Zhu, 2017; Zhu et al., 2019a), enabling new capabilities for mapping urban impervious surface area at annual time scales (Gao et al., 2012; Li et al., 2015; Sexton et al., 2013; Zhang & Weng, 2016). These newly developed approaches reveal the full continuum of change intensity of urban areas, including acceleration of urban growth and other nonlinearities in time, that the former approaches with low temporal frequency observations could not provide.

Mapping urban or urban impervious surface annually has improved knowledge about rates and patterns of urbanization. Nevertheless, considering the extremely dynamic character of urban environments, the capability of detecting urban change at sub-annual scale could illuminate different insights on urban development. This requires satellite observations with even higher temporal frequency (e.g., weekly or monthly), which will enable remote sensing applications for monitoring urban disturbance events that are usually transient in time, detecting subtle or gradual urban changes, and providing relevant near real-time information for urban response and management.

Unlike ecosystem disturbance, urban disturbance has been rarely studied. Grimm et al. (2017) defined urban disturbance as an event that disrupts any aspects of the structure of an urban system as specified in an explicit model, such as land cover and land use change, human activity, soil disturbance, and hazards/disasters (e.g., fire, heatwaves, flood, windstorms, and earthquakes). Most of these urban disturbances will only occur for a relatively short time, and this signal can be easily buried in annual time series. Though there are studies that use high temporal frequency time series for urban remote sensing, they either select the best observation per year or create an annual time series trajectory to map annual urban change, which would miss many urban disturbances or signals that occur seasonally. Moreover, most urban disturbances do not cause changes in urban area or impervious surface area, and are thus "invisible" to algorithms designed for mapping impervious surface area. For example, urban land surfaces can be highly dynamic, with frequent surface modifications caused by human activities, but with stable aggregate amounts of impervious surface area (Deng & Zhu, 2018). If we only focus on impervious surface area, all this disturbance information will be lost.

In addition, subtle and gradual qualitative changes, such as upgrading of the built environment, frequently occur in urban environments. Due to their small change magnitude, methods that select one image per year during the peak growing season are not able to capture these changes, since the magnitude of change can be smaller than the environmental and systematic noise in the imagery. Using high temporal frequency observations, remote sensing data can isolate "noise" in the time series, such as differences caused by urban vegetation phenology, solar angle change, inter-annual condition trends, and other predictable periodical and long-term changes (Zhu & Woodcock, 2014; Zhu et al., 2019b). With all those sources of "noise" excluded, subtle and gradual changes in urban environments can be identified more easily (Zhu et al., 2016). Moreover, with high temporal frequency observations, seasonal changes, which have caused dramatic changes in impervious surface area estimation and urban land cover classification (Deng et al., 2017; Zhu et al., 2012), can be estimated, modeled, and corrected (Deng & Zhu, 2018).

Near real-time monitoring of urban change is critical for raising quick awareness and reducing negative impacts in urban environments. At present, most urban change detection algorithms and products can provide accurate urban change maps, but usually, there is a substantial lag (a few years) between the time when the urban change occurred and the time when the change is detected. This delay in detection will greatly reduce the impact of change maps as they lose relevancy for response and management efforts. High temporal frequency observations are essential for near real-time response applications, such as mapping and evaluating urban fires, heatwaves, floods, windstorms, earthquakes, etc. For example, by combining Sentinel-2A/2B and Landsats 7–8, near real-time monitoring (3-day delay) of urban change at sub-30-meter resolution is possible.

Near real-time monitoring also requires change detection algorithms that are online. COntinuous monitoring of Land Disturbance (COLD: Zhu et al., 2019b), Continuous Change Detection and Classification (CCDC; Zhu & Woodcock, 2014), and Break detection For Additive Season and Trend (BFAST; Verbesselt et al., 2012) algorithms are three examples of methods. Note that COLD, CCDC, and BFAST are not specifically designed for monitoring urban change. COLD is designed for monitoring all land surface disturbances, CCDC is designed for monitoring all land cover types, and BFAST is designed for monitoring vegetation change. Similar methods that focus on monitoring urban change in near real-time are needed. The Continuous Subpixel Monitoring (CSM) algorithm is one of the few examples that integrate CCDC and subpixel mapping for continuous monitoring of urban impervious surface area (Deng & Zhu, 2018).

5.2. Urban heterogeneity

Urban environments vary significantly around the world as well as within particular cities or urban regions. Although urbanization is often presented as a uniform process, reflecting the northern hemisphere experience with industrial development, colonial metropolises, or post World War II suburbanization, there are many kinds and styles of urbanization underway today (Brenner, 2014; McGee, 2014; McHale et al., 2015; Taubenböck et al., 2018). Contemporary cities emerge or transform based on many rationales and drivers, including the familiar driver of industrialization, but also due to consumerism, the flight of refugees, dispossession in the countryside, shifts in government policy, and unfettered real estate speculation. Many cities, of course, combine several of these drivers. All of these drivers are themselves shaped by cultural, historical, climatic, regional environmental, geomorphic, economic, and governance contexts, among others. This variety of drivers creates the condition for a large heterogeneity of urban development patterns.

Despite this heterogeneity, the vast majority of urban remote sensing studies focus on a single “urban class”, particularly for studies based on coarse spatial resolution satellite data. In other words, the world is categorized as a binary — urban versus non-urban and the richness of urban form and heterogeneous ways in which it can change are not usually taken into account. Of course, the rampant patterns of urbanization in the world today do lend some value to such binary depictions, in which urban land conversion is illustrated by a garish red splotch overspreading the calm earth colors representing agricultural, village,

pastoral, or wild lands. Such coarse scale changes have for a long time been of concern as pressing global changes (Berry, 1990; Vitousek et al., 1997).

However, such a coarse scale, aggregated approaches to quantifying urban change can be problematic. For example, “urban” in Lagos (Myers, 2011) is not equivalent to “urban” in Los Angeles (Dear & Dahmann, 2008). Furthermore, neighborhoods within Lagos and Los

Angeles may be strikingly different in structure and form. Details relevant to the social, ecological, and economic aspects of urbanization, that is, the elements of sustainability, are hidden in such aggregated, generalized approaches to urban assessment. For example, shrinking cities in post-industrial situations (Haase, 2008), infill or replacement in older suburban zones (Wilson et al., 2013), and the dynamic creation and occupancy of vacant lands in even established city cores (Johnson et al., 2014; Li et al., 2018), are all invisible to the single class approach to urban mapping. There is a great opportunity to focus urban remote sensing on the within city scales that would better document and follow such patchiness and change. A further shortcoming of the coarse scale, binary approach to urban delineation, is that ecological structures, composition, and processes are essentially omitted by definition. If urban is defined as only built elements, then very likely only the largest green and blue spaces will be recognized, if any are recognized at all. Such conceptually coarse (as opposed to merely coarse in terms of pixel size) examinations will miss the small, interstitial green spaces, or the fact that many urban patches are to some extent hybrids of natural, engineered, and constructed components.

In contrast to the mixed pixel issue originated from per-pixel classification with coarse resolution remote sensing imagery, a variety of attempts have been made to alleviate the concerns of urban heterogeneity by characterizing inter-urban variation. These methods can be grouped into two categories. The first category is to develop new spectral indices specifically designed for depicting urban environments. Comparing to the binary urban/non-urban classification, these urban indices are usually easy to implement and convenient in practical applications (Deng & Wu, 2012; He et al., 2010; Zha et al., 2003; Zhang et al., 2013a; Zhang et al., 2013b). The second category is to improve information from the spatial domain. This can be achieved by mapping urban compositions at the subpixel level as a continuous field (Deng & Wu, 2013b; Powell et al., 2007; Roberts et al., 2012; Wetherley et al., 2017; Xian & Homer, 2010; Yang et al., 2003). Both approaches are designed to provide more heterogeneous information within each urban pixel by using continuous variables in the spectral or spatial domains, which is more beneficial and realistic for mapping the heterogeneous urban landscapes. In particular, by following Ridd's urban morphology model (Ridd, 1995) — the Vegetation-Impervious surfaceSoil (VIS) model — multiple subpixel urban attributes can be simultaneously considered to characterize urban heterogeneity, which is already an improvement when compared with traditional binary classes from the per-pixel classification approach. Nevertheless, simply using three elements may be still far from enough to characterize the mixed nature of urban environment for applied practices.

Two recent classification frameworks have characterized interurban variation using more than three remotely-sensed based, or remotely-sensed derived elements. The first from urban climate/meteorology, the World Urban Databased and Portal Tool (WUDAPT) framework, uses multiple remotely-sensed variables to classify areas of the urban fabric according to

how these areas shape Local Climate Zones (LCZs). Landscapes in this framework can be divided into 17 LCZs, in which 15 of them are defined by land cover and surface structure combinations, and 2 of them are defined by construction materials and anthropogenic heat emissions (Stewart et al., 2014; Stewart & Oke, 2012). A second example, the urban stands classification system, characterizes neighborhoods within urban areas based on their building layouts, street-network configuration, and job-housing balance. The stand classification concept has been used to characterize urban heterogeneity as it relates to the human and biophysical processes that occur (e.g. travel behavior) (Stokes & Seto, 2019). These novel representations of urban heterogeneity help to better describe the complexity of urban areas.

The points reviewed in this section suggest a rich array of urban features and dynamics that might be better assessed by remote sensing. Given the current and increasing sophistication of remote sensing, the field is poised to contribute to the understanding of urban systems at a wider range of scales, with sensitivity to a more inclusive range of urban drivers and resultant forms. If remote sensing can be more closely linked with other disciplines involved in understanding and predicting urban form and dynamics, it is easy to envision significant advances. Perhaps most obvious would be a processing system that takes into account the multiple scales of urban processes, the fact that cities are lived places as well as physical forms (Marcotullio & Solecki, 2013), and the fact that urban systems are — often cryptically — ecological as well as social and technological systems (Grimm et al., 2016).

5.3. Urban form and structure in two- and three-dimensions, emphasizing spatial patterns and connectivity

Many biophysical and socioeconomic processes are shaped by the vertical structure and form of urban areas (Grafius et al., 2018). Thus, in addition to heterogeneity, mapping urban form (e.g., built-up structures) is an essential dimension of future urban remote sensing research.

Urban form (or structure) plays a crucial role in how the urban biophysical environment functions and interacts with human activities. Urban form — the pattern and spatial configuration of land use, transportation, and urban designed elements (Hamin & Gurran, 2009; Seto & Dhakal, 2014) — is crucial for sustainable and smart urban development (Ramaswami et al., 2016). Urban form can be characterized by the physical urban extents, the layout of streets and buildings, as well as the internal configuration of settlements and greenspace (Seto & Dhakal, 2014). Moreover, the dynamics of urban form in the horizontal and vertical dimensions can shape many associated biophysical processes in the urban system (Grafius et al., 2018). It is worth noting that 3D urban models are not yet available for large areas due to data costs or data scarcity. However, recent efforts using digital surface models with lower resolutions (such as from stereo sensors like Cartosat or RADAR sensors such as TanDEM-X) than the commonly used VHR optical stereo sensors (such as WorldView) or airborne LIDAR data have proven their capability of significant reduction of data costs and overcoming data availability (Wurm et al., 2014; Geiß et al., 2017). This section provides insights on research opportunities for urban form in a variety of biophysical processes.

Urban form often emphasizes the variety of urban spatial arrangement in the form of two-dimensional maps. Even from this "flat" perspective, a rich typology emerges. Simple contrasts like spider webs, stars, corridors, infill, that describe the layout of urban land cover, are common, but urban form also includes finer patterns — the configuration of streets and blocks, plots or parcels, buildings, and greenspace (Oliveira, 2016). Urban form has been extensively studied in a variety of disciplines such as architecture, geography, history, and planning (Moudon, 1997; Conzen, 2001). When the third dimension is added, urban form appears as still more complex.

Urbanization does not just increase urban land cover but can also dramatically alter urban form (Dupras et al., 2016; Park et al., 2014). There has been a growing interest in capturing form, with an extensive body of work using landscape metrics regarding the size, density, shape, and distribution of urban patches to characterize and quantitatively describe urban form (Larondelle et al., 2014; Seto & Dhakal, 2014; She et al., 2017; Wang et al., 2017). For example, one study found that small urban patches will self-grow and/or merge with surrounding large urban clusters during the urbanization process (Li and Gong, 2016). As a consequence, the overall fragmentation of urban clusters increases simultaneously with urban expansion. Recent work in this area has used three simple dimension—human constructed, water, and soil-plant continuum—to develop a multi-dimensional conceptualization and mapping of urban form (Wentz et al., 2018).

The effect of urban form requires further investigation to advance our understanding of the urbanization impacts on all dimension of urban sustainability. For example, recent work has characterized the urban landscape with an explicit aim towards understanding how urban form relates to transport behavior and contributes to sustainability (Stokes & Seto, 2019). Another example is how urban form is related to biodiversity. Urban form influences the biodiversity in the urban ecosystem through changing its habitat environments and accessibility to surrounding ecosystems. In general, the influence of urban form on biodiversity can be depicted as the compositions of relevant elements such as residential area, pavements, road network, and green spaces (Nielsen & Jensen, 2015; Zhong et al., 2014). For example, the bird richness responds negatively to the fraction of impervious surface area (Silva et al., 2015). Compared to apartment buildings in a housing neighborhood, detached houses with interspersed trees among them have more insectivores (Andersson & Colding, 2014). Moreover, the connectivity of ecological corridors is crucial to maintaining the biodiversity and species abundance of urban ecosystem (Park et al., 2014; Silva et al., 2015).

Urban form assessed at the microscale could also inform future UHI studies. Urban form can intensify or mitigate UHI effects through changing relevant biophysical factors. The quantity and distribution of urban green spaces in the residential or commercial area can alter urban form through adjusting the spatial configuration of built-up areas, which will further affect the cooling effect offered by plants to mitigate UHI effects (Kong et al., 2014; Zhou et al., 2017). Urbanization encroaches upon natural areas with green spaces, generally leading to a more compact urban form, in turn, the UHI effect could be intensified. In addition to the biophysical environment, urban forms (e.g., sparse or dense residential areas) are also associated with human activities. Thus, building energy use and resulting anthropogenic heat

discharge vary with different urban forms, resulting in different UHI effects (Güneralp et al., 2017; Rodríguez-Álvarez, 2016). Moreover, urban form induces changes in the surface energy balance and further affect the spatial pattern of UHI in the urban system (She et al., 2017; Tratalos et al., 2007). Change in urban form, which can foster natural ventilation in building, has been proposed to improve the thermal environment with appropriate distribution of different height buildings (Middel et al., 2014).

Table 2. Comparisons between structured and unstructured data as a supplement data for urban remote sensing.

	Structured data	Unstructured data
Data source	Federal agencies, government departments, private sectors, individuals	Private sectors, individuals
Location/geographic information	Specific, and clear	Some are specific, while others are only implied in the text (or hashtag)
Data cleaning necessity	A few but limited	Considerable, and could be labor intensive
Data quality	Well organized	Unknown/poor
Data easiness to researchers	Easy access/downloadable	Usually by web scrapping
Supplement contents	Demographic and socioeconomic information, human activity	Human activity, human perception

The consideration of urban form, particularly its vertical dimension (i.e., building height and volume), will bring new insights in projecting future emissions under urbanization and climate change. Studies that compare the vertical dimension of urban areas to night lights or horizontal expansion show that cities with similar extent can vary significantly in terms of the building stock and vertical urban structure (Frolking et al., 2013). Greenhouse gases emissions, a primary source influencing future climate change, are associated with urban form (Hamin & Gurran, 2009). A compact urban form can reduce anthropogenic CO_2 emissions by enabling low-carbon transportation modes and travel behaviors (Rodríguez-Álvarez, 2016; Wang et al., 2017). With a spatially aggregated urban form having many high buildings, e.g., dense distributed residential and commercial areas (Zhong et al., 2014), per capita CO_2 emissions from transportation can be significantly reduced (Creutzig et al., 2016, 2015).

Understanding of the role of urban form in shaping the biogeochemical cycles of water, land, and atmosphere is also a significant gap. Nutrients such as atmospheric depositions, pet waste, and fertilizers, which have become a crucial threat to water quality in streams or rivers, are influenced by urban form during their transportation through the rainfall and surface runoff processes (Yang & Lusk, 2018). Soil sealing, defined as the covering of land by impervious surface areas, is a common phenomenon in suburban regions, due to

expansive urban development and the prevalence of parking lots (Tombolini et al., 2016). Moreover, the geometry and distribution of buildings are closely related to the atmospheric environment (e.g., air quality), which is a rising concern to public health for people living in megacities. For example, fragmented cities are associated with higher pollution concentration (e.g., NOx and PM 2.5) (Fan et al., 2018; Rodríguez-Álvarez, 2016; She et al., 2017).

5.4. Linking remote sensing with emerging urban data

Remote sensing has long been used to extract urban land cover and use information and to describe different aspects of urban environments, such as urban infrastructure, urban impervious surface, urban vegetation, urban tree canopy cover, and surface urban heat island. Despite the long history of success, the mere usage of remote sensing data in the literature may not be appropriate for future urban studies. The reason is twofold. First, only

the physical environment information of urban can be retrieved from remotely sensed imagery, while other types of urban information cannot be directly obtained. Examples include land use (Huang et al., 2018c), sentiment and emotion in a city (Wang & Stewart, 2015), spatially explicit human activity and mobility (Frias-Martinez et al., 2012; Soliman et al., 2017). Such information is essential to understanding the interaction between human beings and their living environment, and analyzing how urban inhabitants view, sense, and change their living environment. Second, local knowledge is always required as an important input for per-pixel or sub-pixel image classification of urban environments. Such information, however, relies heavily on human inputs and the expertise of the researchers. In support of urban remote sensing for better understanding of urban environments, one of the solutions in the future is to take advantage of emerging urban information from a variety of data sources, especially those reported by humans as a new type of sensor which is called human sensors (Goodchild, 2007). More importantly, most, if not all, of these multi-source datasets are publicly available, and free of charge.

Two major types of open access urban data can be used to supplement remote sensing imagery – including structured and unstructured data. Structured data is preprocessed, well-organized, and generally, do not require substantial data cleaning. Structured data are usually collected and processed by federal agencies (e.g., U.S. Census and U.S. city open data), and more recently, private companies (e.g., Google street view, mobile phone data, subway smart card, and taxi GPS data), and active individuals (e.g., OpenStreet Map). Structured data usually contain a specific geographic location that can be directly and conveniently linked to remote sensing imagery. In contrast, unstructured datasets are more likely to be generated by numerous internet users and retrieved by researchers from the Internet. Such information is loosely distributed online without an appropriate organization. Examples include social media data (e.g., Twitter, and Flickr photos), and online text reviews (e.g., check-in data of FourSquare, and Yelp). A comparison of the characteristics of structured and unstructured data as a complementary dataset to remote sensing are described in Table 2.

Despite the complementary features between urban remote sensing imagery and open urban data, several data issues exist that make their integration challenging. First, open data and remote sensing data are collected at different scales. Social scientists, demographers, sociologists, urban planners, and policy makers prefer to employ predefined administrative boundaries, such as Census geographical unit, postal unit, school district, and congressional district boundaries. In addition, because open data is not collected evenly across space, using too fine a scale might lead to the data sparsity issue (i.e., few or no data in some areas), especially for social media data.

Comparatively, remote sensing scientists are used to employing uniform grids at a scale either inherently determined by the spatial resolution of remote sensing images or the image segment level. In most of the attempts to combine remote sensing and open data (most of which are point data), researchers still follow the tradition of remote sensing studies and select grid as the study scale, in which point-based open urban data are aggregated into each grid. These grids and scales, however, are not frequently used by domain experts. To make remote sensing research more useful and actionable, researchers can further spatially

aggregate the per-pixel results in a scale or unit most commonly used in the relevant discipline. Another type of scale difference exists between Google street view and remote sensing imagery. The former data provide a 360-degree view of the streetscape of a city, while the latter only offers a vertical aerial view. For these cases, domain experts should play a major role on constructing a big picture of the application question, while remote sensing scientists may need to spatially aggregate the per-pixel results to a scale or unit most commonly used in the relevant discipline.

Second, open access urban data can be biased. For unstructured data, there is uncertainty around the data's representativeness, since the population of social media users is not reflective of the whole population. Additional sources of uncertainty include missing data from web crawling, repetitive posts from one single highly active user, underrepresentative samples (e.g., younger generations tend to dominate the contributing group). Prior to the collective usages of remote sensing imagery and open data for urban applications, considerable automated data collection, cleaning and matching, and compilation are necessary. To address this, remote sensing scientists may need to collaborate with urban data scientists who have expertise in addressing the challenges of unstructured data.

6. Concluding remarks and perspectives

Our assessment of the research literature and discussions between the policy and science communities leads us to conclude that urban remote sensing cannot continue to focus primarily on mapping urban cover and use. While continued assessments of urban land cover and land use will be necessary for a number of user communities, other types of analyses are needed. Strategic directions, including *high frequency analysis of urban environments*, *urban heterogeneity*, *urban form and structure in 2- and 3-dimensions*, and *improving the performance with other types of data* can expand the set of urban research question answerable with EO data.

Though this review focuses only on the major scientific advances in urban remote sensing, we cannot ignore the new technique and technology improvements that push urban remote sensing forward aggressively. For example, new machine-learning techniques, such as deep learning, are changing the game. Deep learning is characterized by neural networks (NNs), involving more than two layers (this is where the term “deep” originates). As a “black-box” solution, deep neural networks learn and explore exclusively from the data, without any domain-specific knowledge. The Convolutional Neural Networks (CNNs) approach has proven extremely effective for image analysis, due to its capability of considering the spatial patterns by interleaving convolutional and pooling layers. Though there are contrasting opinions in remote sensing community regarding the use of deep learning techniques, this technique has attracted much attention due to its superior performances (Zhang et al., 2016; Zhu et al., 2017). A recent study suggests that deep learning is particularly good at capturing the fine features of complex urban areas, and performs better than other traditional classification methods (Huang et al., 2018a; Huang et al., 2018b).

New remotely sensed big data streams are also revolutionizing urban remote sensing. Historically, the urban remote sensing community has relied on data from NASA assets.

However, many new opportunities have emerged. The Sentinel mission from the European Earth monitoring program has started to launch a constellation of satellites since 2012 (Aschbacher & Milagro-Pérez, 2012). With a span of 13 spectral bands at a spatial resolution as high as 10 m, Sentinel-2 has great potential for urban remote sensing (Drusch et al., 2012). Studies have suggested that Sentinel-2 adds value to the mapping of built-up areas, over predecessor Landsat-based products (Lefebvre et al., 2016; Pesaresi et al., 2016a). China has also launched a few satellites that are unique for urban remote sensing. For example, high spatial resolution nigh-time light imagery (130-meter) are now available from Luojia-1 (Li et al., 2019b), and the multi-view imagery provided by ZiYuan-3 satellite can map urban in 3D at the global scale (Liu et al., 2019). At the same time, an almost unmanageable number of new commercial satellites, particularly smaller satellites, provide data with both high spatial resolution and high temporal frequencies. For example, a total of more than one hundred small satellites from Planetscope can monitor our planet at approximately 4 m every single day. These new datasets stimulate a variety of new urban remote sensing algorithms and applications. Over the past five years, the number of operational satellites has increased 40%, and nearly 1400 now orbit the Earth. The data collection capability of each new satellite iteration is also increasing at an astonishing rate. For example, Landsat 5 collects 40 GB of remotely sensed data per day, and Landsat 7, Landsat 8, and Sentinel-2 collect 260 GB, 750 GB, and 1.6 TB data per day respectively (Wulder et al., 2016). Those exponentially growing open and free satellite data create a new era for urban remote sensing, but also challenge the established urban remote sensing science and applications, as researchers are easily buried in this enormous volume of data. As this “data tsunami” arrives, the question becomes whether we are ready to derive meaningful new knowledge about cities and sustainable urbanization from it.

The rapid development of modern information technology, hosted computing platforms, such as cloud computing and grid computing, offer a promising solution (Esch et al., 2018). Google Earth Engine is one example of how big data and cloud computing can be seamlessly integrated, facilitating tens of thousands of researchers who may have limited remote sensing knowledge but need to process huge amount of remotely sensed data, requiring extensive computing resources (Gorelick et al., 2017).

Last but not least, remote sensing scientists need to collaborate more with other scholars and practitioners from other urban and allied fields. Urban scholars from other disciplines ask different types of questions from urban remote sensing. As a community, we need to think about not only how to derive the information (e.g., algorithms development), but *what* to do with it and *who* will use it. It is not sufficient to have one sentence at the end of a paper indicating “this will be useful” without a clear understanding of who will use it and how to use it. Urban remote sensing community should always keep science and policy in perspective to facilitate the transition towards a smart, sustainable, and healthy urban environment.

Acknowledgments

We would like to thank both the Editor and Reviewers for the valuable insights and constructive suggestions made to improve this manuscript.

Part II

What '*urban*' is

CHAPTER III	To be, or not to be '*urban*'? A multi-modal method for the differentiated measurement of the degree of urbanization	61
CHAPTER IV	Spatial delineation of urban corridors in North America: An approach incorporating fuzziness based on multi-source geospatial data	83
CHAPTER V	Quantitative assessment and comparison of urban patterns in Germany and the United States	109
CHAPTER VI	Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception	139
CHAPTER VII	Searching for the DNA of urbanisation. A material perspective	153

Chapter III

To be, or not to be '*urban*'?

A multi-modal method for the differentiated measurement of the degree of urbanization

Computers, Environment and Urban Systems, 2022, vol. 95, 101830

Hannes Taubenböck, Ariane Droin, Ines Standfuß, Fabian Dosch, Nikola Sander, Antonia Milbert, Silas Eichfuss & Michael Wurm

© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.compenvurbsys.2022.101830

Received 22 Feb 2022; Accepted 17 May 2022; Published 26 May 2022

Abstract

Today, 56.6% of the world's population is urban and the trend is rising; in Germany, the urbanization process is almost complete at 80.3%. This is the ubiquitously used narrative. Surprisingly, the figures behind it are rarely questioned. The spatial statistics responsible for this meta-narrative are, however, prone to ambiguity. They suffer from a lack of systematic empirical justification, and from cross-national differences in cut-off values used to differentiate between urban and rural populations. In this study, we present an empirical approach that allows to systematically map urban and rural populations in a spatially and thematically differentiated manner using a multimodal method. On the one hand, we resort to the common approach of presenting the degree of urbanization in terms of population figures for administrative units. However, we do not only use the common national threshold value, but we project various national thresholds applied in different countries across the globe to classify multiple degrees of urbanization onto our study site Germany. On the other hand, we also calculate various degrees of urbanization at a higher spatial resolution using a regular grid. Beyond the common approach of calculating the degree of urbanization by population figures, we also apply at grid-level two additional variables: building density and the share of a certain building type. By systematically applying thresholds between minimum and maximum per variable, we trace the effects on the resulting degree of urbanization. These multiple perspectives lead us to propose that a range rather than a singular threshold allows us to estimate the degree of urbanization in a more differentiated way. To do so, we estimate the degree of urbanization for Germany on a probability-based basis. Therefore, we combine possible variants from the administrative approach using population figures and the grid-based approach using thresholds of population, building density and the share of a certain building type. Our results show that Germany can be considered urban by at least 50.0% up to possibly 68.1% of the population, which by no means comes near the reported 80.3%. We conclude that the results of the commonly used approach to quantify urban populations are not tenable in their clarity and should therefore be used only with great caution for political and societal decision-making.

Keywords

Degree of urbanization; urban; rural; spatial statistics; building density; population density

1. Introduction

In the best case, we base decisions – political, social, economic, ecologic, or even personal – on facts. Science is a source of profound facts. Nevertheless, scientific knowledge is subject to uncertainties due to applied concepts, data or methods that are not always easy to interpret. There is the danger that mathematically unambiguous calculable 'facts', nevertheless lead to ambiguity or even misinformation. Therefore, appropriate interpretation and a deeper understanding of the situation is a necessity.

In spatial sciences or urban geography, to take an example, there are often gaps between computable results and their interpretation. Results based on precise mathematical calculations may lead to biased statements due to inconsistencies or inaccuracies in the input data (e.g. Malizia, 2013), the preference for a particular method (e.g. Schabenberger & Gotway, 2005) or indicator (e.g. Shearmur, 2017), or due to the modifiable areal unit problem (MAUP) (e.g. Openshaw, 1983). With respect to the latter, the shape and scale of aggregation units, for example, can produce a statistical bias (e.g. Taubenböck et al., 2016). On the other hand, certain statistics of spatial science remain fuzzy by nature, as results depend on the defined conceptual framework, the data, the scale or the methods applied. Let's take 'polycentricity' for an illustration – a concept that grasps a hierarchical division of centers in urban space (e.g. Batty et al., 2004). Until today, there are no generally accepted indicators (used are e.g. job concentrations (Krehl, 2016), built densities (Taubenböck et al., 2017), among others) or methods (used are e.g. threshold approaches (Kim, Yeo & Kwon, 2014), locally weighted regression (Roca Cladera, Marmolejo Duarte & Moix, 2009), among others) to unambiguously distinguish these centers and whether those centers distinguished define poly- or monocentricity or the state in between (e.g. Bartosiewicz & Marcinczak, 2020).

1.1. Questioning the common logic of urban population

And yet, science in the urban domain has produced seemingly unambiguous facts we all – in science, in media, in politics, among other branches – take for granted: what the largest city in the world is or what the proportion of urban population on our globe is. These are generally accepted facts. Incidentally, it is usually stated that the largest city in the world is Tokyo (UN, 2018) and, that 55.6% of the global population today is urban (UN, 2018). These figures, with few exceptions (cf. e.g. Bocquier, 2004; Buettner, 2014), are the basis of argumentation in most contributions. In this study, we question if we are already subject to a fallacy on these 'facts'?

Recent studies challenge these two types of 'facts' by trying to generate estimates with their own concepts and their own empirical investigations aiming at more reliable figures. On the city sizes, official statistics are challenged e.g. for Germany (Budde & Neumann, 2018) or on a global scale (Taubenböck et al., 2019). The latter study reveals that not Tokyo, as the UN (2018) has stated, is the largest city in the world, but the urban agglomeration in the Pearl River Delta in China. It is demonstrated that long-established administrative boundaries can become obsolete in times of highly dynamic urbanization. And beyond, these units are inconsistent across municipalities and thus are not acceptable as spatial comparators due to a distortion of statistics. They, in contrast, develop a method to distinguish urbanized

from rural areas in a consistent manner for all cities across the globe. The advantage of this analysis lies in its spatial consistency.

1.2. Arbitrary classifications of 'the urban' and 'the rural'

On the *degree of urbanization*, scientific studies, reports, media, political statements argue that more than half of the world's population now lives in urban areas. These statements are predominately based on the most authoritative and cited source of global urbanization levels published by the UN (2018). Melchiori et al. (2018), however, claim that the widely accepted UN numbers are incorrect. Using remote sensing and population data and a consistent methodology to define the degree of urbanization, they claim that a significantly higher amount is urbanized than stated by the UN: they state that already 84% of the world is urban. This assessment, in turn, has been challenged by Angel et al. (2018). They claim that what we commonly understand as 'urban' runs counter to this analysis. Based on four arguments relating to low-density areas classified as urban, the migration aspect towards cities, the share of the labor force in agriculture and village non-farm occupations and regularities in settlement hierarchies, they argue that the world is still nowhere near as urbanized as Melchiori et al. (2018) assert.

These few examples elaborated here, together with studies e.g. by Balk et al. (2018), Dorelien, Balk & Todd (2013) or Liu et al. (2014) as well as by initiatives such as the Global Rural-Urban Mapping Project (GRUMP) (Socio-economic Data and Applications Center (SEDAC), 2021) or the GHS-SMOD product of the Global Human Settlement Layer (GHSL) (JRC, 2021), show that in research studies on the degree of urbanization in particular, and in spatial sciences in general, there are often no clear, incontrovertible facts, but only different ways of measuring something and interpreting it accordingly. In the urban/rural domain, this complexity is often encountered by conceptualizations such as 'urban-rural continuum' (e.g. Champion & Hugo, 2004; National Academies of Sciences, Engineering, and Medicine, 2016; Pahl, 1966), urban-rural-interface (e.g. Lopez-Goyburu & García-Montero, 2018) or 'urban-rural gradient' (e.g. du Toit Marie & Cilliers, 2011). Thus, the truth even viewed from a scientific lens is not easily determined and it remains tolerant of ambiguity. Brenner & Schmid (2013) name the above discussed statistics empirically untenable (a statistical artifact) and theoretically incoherent (a chaotic conception).

In this paper, we aim to take an empirical turn by applying the 'chaotic conceptions' for calculating the *degree of urbanization*. We do this using the categories '*urban*' and '*rural*' – artificial as they may be – as they are nevertheless constitutive of our society (Löw, 2008; Redepenning & Hefner, 2018), often providing an argumentative basis for political decisions. However, we believe that the widely accepted statistic that today (i.e. in 2021) 56.6% of the world's population resides in urban environments (UN, 2018) paints a superficial picture of a highly complex situation. Like the studies mentioned above, we want to challenge these statistics, or at least question the validity of these generally accepted and used facts.

In contrast to other studies, we are in no way concerned with eliciting what constitute meaningful variables or thresholds that can be used to measure '*the urban*'. Our aim is to show how arbitrary the classification of 'the urban' and '*the rural*' can be in such statistics due to the variation of concepts, variables and thresholds. With it, we follow the call from Angel et al. (2018) to approach a more definite method by a systematization of the effects of the various methods. We do so by the example of Germany as the very good geodata situation allows us to grasp the complex situation at very high and unprecedented resolution in an empirical manner. We want to approach the statistics, i.e. the degrees of urbanization, from different spatial and thematic angles and show how susceptible these numbers are to different concepts and methods. As we believe that a single threshold value always falls short, we propose to specify the degree of urbanization in a more differentiated, probability-based way which allows a range between the clear poles 'urban' and 'rural'. With it, we want to add to the discussion that we need better data, clearer and maybe other concepts and accepted methods to make interpretations of geodata not arbitrarily reinterpretable for every desired point of view.

2. Conceptualization of '*the urban*' and '*the rural*'

In this study, as mentioned, we are not primarily interested in developing our own conception of '*the urban*' and '*the rural*'. Rather, we are interested in developing an approach that maps these two thematic classes in a more nuanced way. Thus, it is important to have a basic conceptual understanding of '*the urban*' and '*the rural*'. Generally, the core idea of the urban in distinction from the rural are various forms of higher density: Concentrations of people, of everyday' s routines, of social and economic structures, of differentiations and heterogeneity, or of built structures in one place, among other things (e.g. Nassehi, 2002; Tonkiss, 2013).

Urban as a term thereby conceptualizes two different aspects: On the one hand, it testifies to a certain way of life. Tolerance, freedom, social distance, networking, diversity, interculturalism, cosmopolitanism, open-mindedness, intellectuality, and creativity are often attributed to this lifestyle (Florida, Gulden & Mellander, 2008; Glaeser, 2010; Tonkiss, 2013; among others). Even though these attributes are probably most constitutive in larger urban agglomerations, these are non-territorial characteristics, or at least they are very difficult to locate in a very concrete way. The urban lifestyle is spreading spatially unspecifically in a world of increasing mobility and better communicative connectivity. In consequence, it is not what is defined as a city in administrative law that appears as urban, but what is perceived as an urban unit in everyday life (Löw, 2008). On the other hand, the term 'urban' points to the quantitative aspect of density in one place: of populations, of built structures, of infrastructure, and architecture, as well as of physical growth and structural change (Löw, 2008). In this line of thought, urban is, according to Berking (2008), first and foremost a spatial form with territorial characteristics. Even though, as discussed above, our lives are becoming more and more globally networked, we still live locally (Castells, 1996). Social life continues to be organized into coherently bounded spatial envelopes ('human settlements') whose demographic properties can be defined within spatial boundaries (Wirth,

1969). In this line of argumentation, the place of residence and its territorial design still have a considerable relevance and thus, territorial approaches remain important.

In this study, we resort to the basic concept that the *urban is a form of higher density at a place*. It is thus our basic assumption that 'the urban' can be territorialized. Although it remains conceptually difficult to determine, we assume that in principle spatial boundaries between urban and rural exist (Sievers, 1997). However, for our approach we disregard forms of density in terms of social dimensions, lifestyle, heterogeneity, etc. and focus on the density of population and built space. And fundamentally, we do not want to commit ourselves a priori to which thresholds or which definitions constitute *the urban*, since we are aware that, at the margins, the characteristics of the urban may be ambiguous or can also be understood as transition zones, urban-rural gradients or a continuum. Rather, we aim to explore the question of what influences these ambiguous distinctions have on the assessment of the degree of urbanization at national scale. And finally, we propose a method that allows these ambiguities to be presented in a more differentiated way.

3. Experimental set-up

3.1. Classification of '*the urban*' versus '*the rural*'

We experiment with different characterizations of '*the urban*' and its complement '*the rural*'. For the classification the spatial unit referred to is important. Administrative units are of jurisdictional, administrative and political relevance and therefore remain a central element in any discourse. At the same time, these administrative demarcations result from historical-political processes, they are heterogenous in size, feature spatial outlines which are somehow artificial, and they may feature within their area both types, i.e. dense areas that can be attributed more to '*the urban*' as well as low dense areas that can be attributed more to '*the rural*'. A gridbased approach, in contrast, allows the variables to be positioned more in the local, i.e. they allow for a less spatially cohesive, more porous type of mapping.

Against this background, we use *two distinct spatial approaches*: (1) a territorial approach where we focus on classifying the official administrative units; (2) a grid-based approach where we focus on a standardized spatial unit determined by a grid.

On this basis, we set up our experiments as follows:

1) For the *territorial approach*, we quantify '*the urban*' vs. '*the rural*' by using threshold values to assign the administrative units of municipalities to one of the two types. We base this on the commonly used variable of population numbers. For the selection of thresholds, we use the wide variety of thresholds defining the urban (ranging according to UN World Urbanization Prospects from 200 to 50,000 people) as applied by different nations across the globe (UN, 2020) as well as the threshold of 100,000 people suggested by the study of Angel et al. (2018). Then, we calculate the proportion of the cumulative population values classified as urban in relation to the total population at national scale.

2) For the *grid-based approach*, we determine '*the urban*' vs. '*the rural*' at a spatially continuous and consistent grid of 100 × 100 m. Based on our available data, we focus here

on *three variables: building density, population density and the share of a certain building type*.

Building density is measured as the percentage of the sum of all ground areas of buildings per grid, i.e. per hectare. *Population density* is the number of inhabitants per grid. For this specific variable, we relate the population density relative to the maximum value of population density at a grid cell in Germany to scale the variable accordingly be tween 0 and 100. The *share of a certain building type* is measured as the share of a particular building type per grid expressed as a percentage. Here, we assume that certain building types are characteristic of urban living. For this purpose, we make use of the following argument: When we imagine urban housing, we usually think of high-rise buildings, block development, large housing estates, or multi-family houses. In this conception, these structural appearances constitute '*the urban*', while single-family or semi-detached houses or the like do constitute a more rural way of structure.

All these variables are clearly spatially locatable. Landscape always consists of a complex web of structural types even on the here applied grid level. As we do not want to determine a priori which grid cell we categorize as 'urban' and which one we categorize as 'rural', we systematically calculate for all three variables thresholds from the mini mum to the maximum in order to quantify the effects on the degree of urbanization. In the experiment, we apply the entire possible range of thresholds: in 1% steps from a minimum greater than 0% share of the respective variable to a maximum of a 100% per grid. So, as an example, we assume that in the two most extreme cases, we count a grid cell as 'urban' even if there is not a single building at all, i.e. 0% share, or we count a grid cell as 'urban' only if 100% of the area is occupied by buildings. In doing so, we determine how the degree of urbanization varies as a function of thresholds. The result is a non-contiguous spatial grid constituting '*the urban*'.

We present the variability of our results on the degree of urbanization in the form of progress plots over the thresholds. In addition, we also map the spatial effects of the various thresholds per variables on the degrees of urbanization. Of course, since a myriad of thresholds exist, we select representative examples for comparison in the cartographic representations.

3) Finally, we bring the different manifestations of the approaches together and propose a *combined, probability-based approach*. Each approach introduced above has its justification. And yet, due to the large number of variants, the evaluation of *'the urban'* versus *'the rural'* in spatial statistics falls short of expectations: a definitive statement about the degree of urbanization in Germany. We approach this definitive statement by combining the different variants without leaving aside the statistical ambiguity.

The thresholds from the various national approaches on adminis trative units as well as the systematically calculated thresholds of the different variables on the grid-based approach, produce different quantitative results and cartographic formations of *'the urban'* vs. *'the rural'*. Since we do not define a priori whether a certain approach is more appropriate than others, we give the national thresholds at the administrative units the same weight as the grid level variants. As there are 14 variants for the territorial approach based on national definition thresholds, we aim to match these with grid-based ones. As we have three different

grid-based approaches, we select five variants each that add up to 15 variants. Of course, due to the systematic test in 1% steps of the threshold values for the grid-based approaches, there are significantly more variants per variable available than we can use if we weight both approaches equally. To guarantee a generally bias-free calculation at best, the chosen urbanization thresholds for the grid-based approaches are evenly distributed across the entire range: 1, 25, 50, 75 and 99% respectively. The sum of all 29 variants (14 administrative, 15 grid-based) allows a probability-based classification of the degree of urbanization at grid level.

In the thought process that the delineation of *'the urban'* from *'the rural'* constitutes itself not at the extremes, but somewhere around the middle, we generate for the combined, probability-based approach a three-part classification: high probability of being '*urban*' or '*rural*', as well as a '*range in between'*. By the latter class, we define the uncertainty for the quantification of the degree of urbanization. For classification, we apply the standard deviation around the median to narrow down the range. This range of uncertainty naturally has grid cells along the continuous range of probabilities that are closer to *'the urban'* or to *'the rural'*. We subclassify these again into three classes based on quantiles of population for these designated areas.

3.2. Study area and data

Our study site is Germany. The degree of urbanization for the country is given at 80.3% (UN, 2020). These statistics refer to the spatial level of municipalities with more than 5000 inhabitants (DeStatis, 2020), i.e. a territorial approach is used. Despite this apparent clarity, other approaches go beyond the simple urban-rural dichotomy. Eurostat (2021), as an example, differentiates the degree of urbanization ac cording to densely populated (cities), medium population density (towns and suburbs), and sparsely populated (rural). At a spatial grid level of 1km2, *'the urban'* is classified here as clusters of neighboring grid cells with a density of at least 300 inhabitants per km2 and at least 5000 inhabitants suggesting the degree of urbanization for Germany at 69.6%. These two examples show how different approaches applying different spatial levels and different variables and corresponding thresholds cause different results. The latter conceptualization indicates that a dichotomous distinction could fall short with regard to the complexity of urban-rural gradients or continua. With our proposed classification of the degree of urbanization based on probabilities and with a range of uncertainty, we believe that the susceptibility of statistics can be mastered in a more differentiated way than with a dichotomous characterization of the territory.

We apply the different constitutions of '*the urban*' to entire Germany. We do this because here, in comparison to the continental or global input data used in the cited studies above, the spatial and thematic resolutions of the available data is particularly good and yet area-wide and consistent.

For our analysis, we use as basic data sets (1) *administrative units of Germany at the municipality level* as provided by German Federal Agency for Cartography and Geodesy (BKG) (BKG, 2020), (2) *census data* on population. Here, we rely on population data projected to 2018 based on the 2011 census from the German Federal Statistical Office. Population information is available at 100x100m grid cells (DeStatis, 2020), 3) *level-of-*

detail-1 (LoD-1) building models. Here, we rely on the official LoD-1 building model (BKG, 2021) which provides the building ground floors and the height of the building, and (4) data from *publications on the degree of urbanization*. We rely on the data published in the World Urbanization Prospects (UN, 2020).

Based on these raw data, we further process the data to obtain the following indicators: (5) *building structural type.* As discussed above, we assume that multi-family houses (including high-rise buildings, block development, etc.) constitute structurally urban forms of living. Thus, we specify two different classes, namely *urban* and *rural structural building types*. Using the designated function of the buildings based on the official real estate cadaster (ADV, 2008), buildings are differentiated into residential and non-residential buildings in a first step. Buildings in federal states of Germany that do not include a differentiation of the respective functions are classified using a Random Forest approach (Droin, Wurm, & Sulzer 2020; Wurm, Schmitt & Taubenböck, 2016). Furthermore, Droin et al. (2020) differentiated the designated residential buildings into four thematic classes (detached, semidetached and terraced houses are understood as rural and multifamily houses (MFH) are understood as urban, cf. Section 3.1.). We use this classification to derive an (6) *estimation of population numbers*. We disaggregate the population data from the 100 × 100 meter grid onto the available living space in the residential buildings from (5). From the LoD-1 building heights, we calculate the number of floors from building height based on established correlations between the two variables (Wurm et al., 2019; Wurm et al., 2011). Hence, we derive the living area per building and disaggregate the population proportionally by a linear disaggregation model. With it, we achieve people per residential building.

4. Results

As already mentioned, the story is told that in principle the urbanization process is largely completed in Western countries (e.g. Zukunftsinstitut, 2019). The benchmark for Germany is 80.3% of the population already lives in urban areas and the trend is rising (UN, 2020). In the following, we show variants of such statistics that question these general narratives and we propose a procedure to present these statistics in a more nuanced way.

4.1. The territorial approach

What is classified as *'urban'* is highly heterogeneous between countries. If one applies these multitudes of used thresholds across the globe to the administrative areas of Germany, one obtains very heterogeneous degrees of urbanization indeed. In this range of thresholds, one extreme is 200 inhabitants per municipality (applied in Iceland, Greenland, Australia, Denmark). Applied to Germany means that 99.9% - almost the entire population - would be classified as 'urban' (Fig. 1). The other extremes are, if the Japanese threshold of 50,000 inhabitants per municipality is applied, that only around 41.0% of Germany's population would be considered 'urban'. Angel et al. (2018) even propose 100,000 people or more for the classification of the 'urban' to ensure a focus on cities. In this case, even only 32.1% would be classified 'urban' in Germany. Figure 1 provides an overview of the quantitative share of the population classified as urban versus rural in a progression diagram using the

different national thresholds. It becomes clear that the range between 99.9% and 32.1% is extremely wide.

These applied thresholds and the related variable, of course, have spatial implications on the degrees of urbanization. If we again focus on the extremes, then according to the Icelandic definition, virtually all of Germany is mapped as urban (Fig. 2a). Both, in the context of the conceptual differentiation between *'the urban'* and *'the rural'* in general, as well as in the sense of what we understand as urban, i.e. a dense form of living, this threshold value seems nonsensical for Germany. The threshold applied in Japan (>50,000) or the even more restrictive one of Angel et al. (2018) (>100,000), in turn, identify only large cities and few metropolitan areas such as Berlin, Hamburg, Munich or the Rhein-Ruhr area as urban. Small and medium-sized towns, however, are classified as rural (Fig. 2a). From these extremes, one can conclude that a meaningful threshold must lie somewhere in between. However, there can be no unequivocal true or false. It is evident in the gradations in between, that setting a threshold will always lead to discussions.

To add to these variants, some countries apply not just an absolute threshold, but a second condition for the classification of '*the urban*'. The combination of absolute population numbers with a population density threshold adds to the complexity. Classifying the urban at e.g. 1000 inhabitants and a population density at 400 inhabitants per km2, as Canada does, makes an immense difference compared to the threshold of 1000. Figure 2b illustrates the enormous spatial impact. Here, too, it is predominantly the large cities and metropolitan areas that remain.

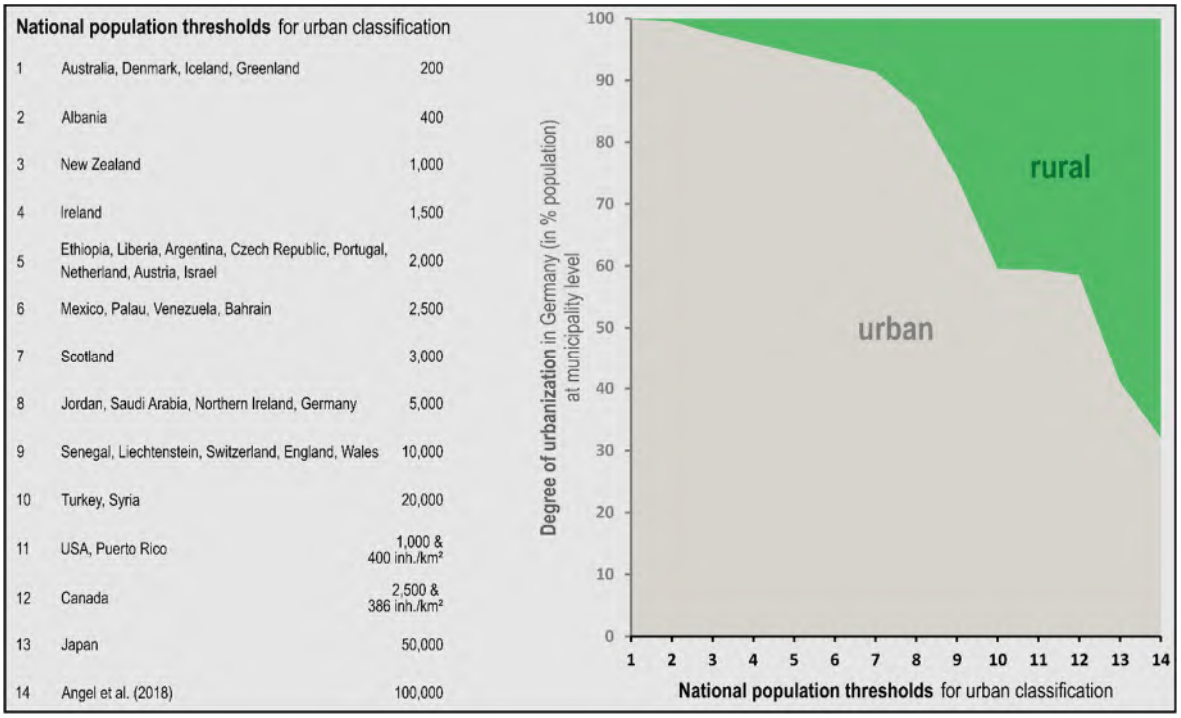


Fig. 1. Degree of urbanization for Germany depending on the population thresholds applied (thresholds as in UN, 2020 as well as from Angel et al., 2018) at administrative units.

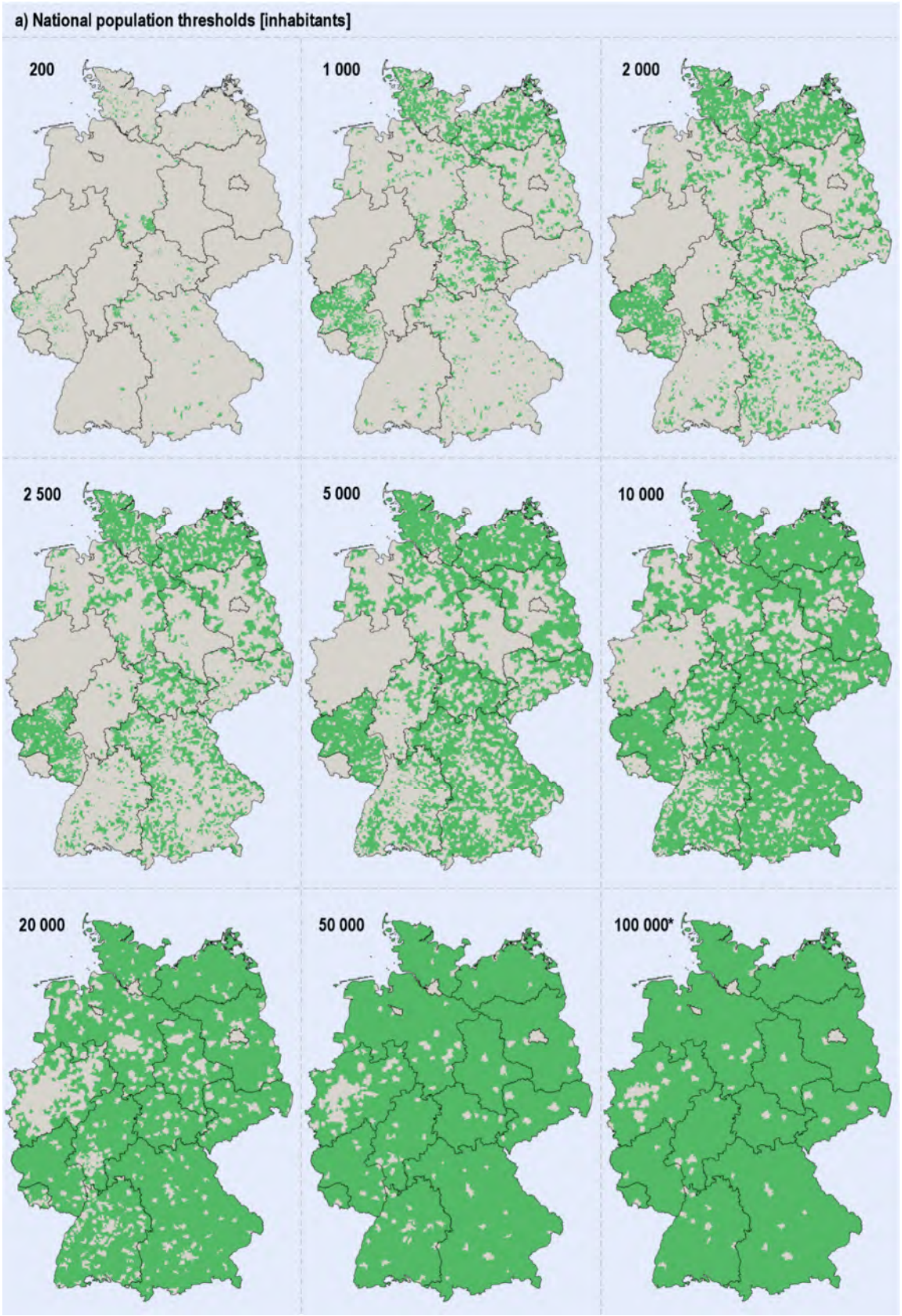


Fig. 2. Degrees of urbanization for Germany: at administrative units depending on a) national population single thresholds, and b) on national thresholds combined from absolute population and density (thresholds as in UN, 2020 as well as in Angel et al., 2018); at grid-level depending on c) a selected threshold on the share of multi-family homes per grid cell, d) a selection of thresholds on the building density, and e) a selection of thresholds on the population density relative to the measured maximum value of population density in Germany.

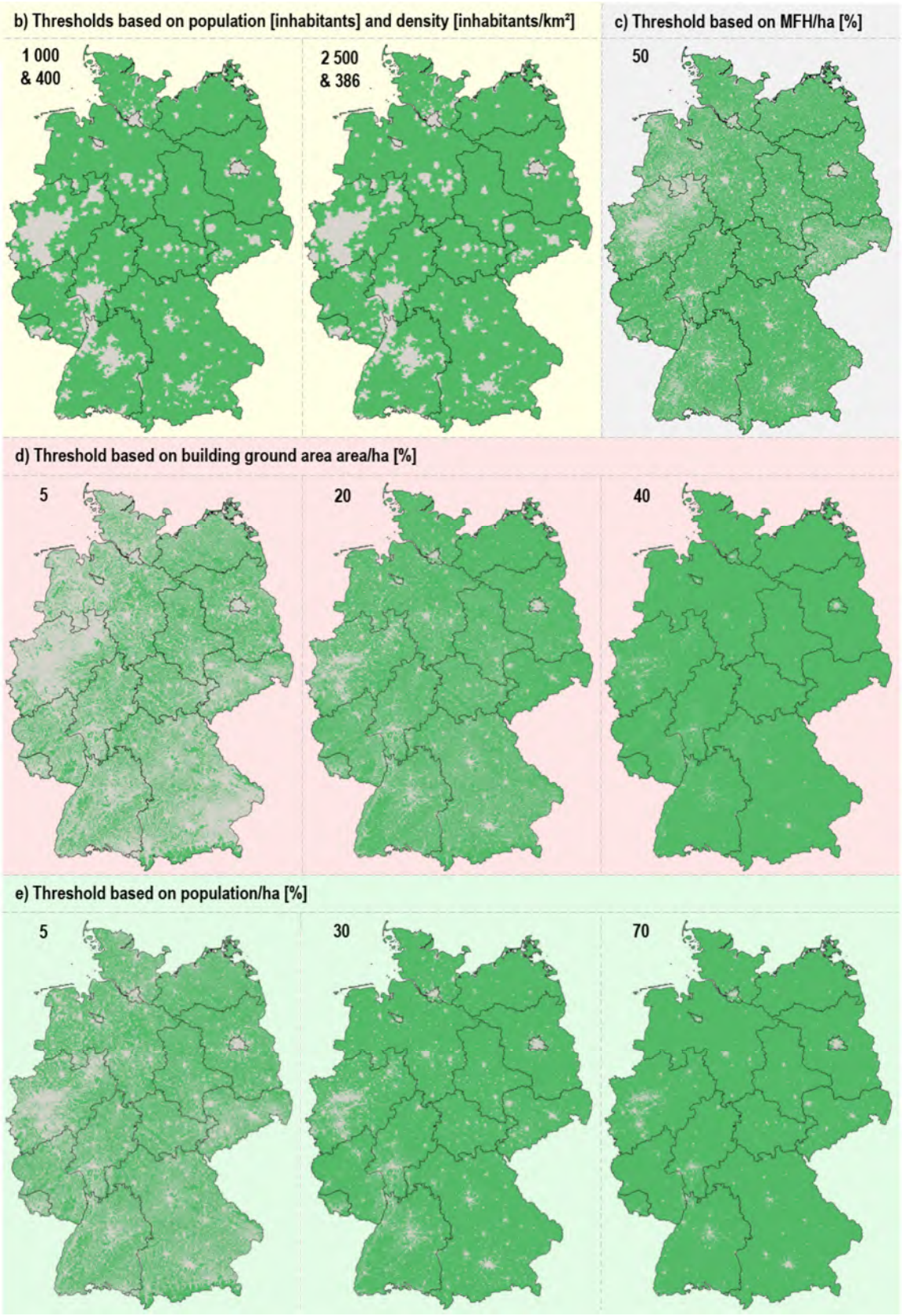


Fig. 2: (*continued*).

4.2. The grid-based approach

For the grid-based approach, we apply three different variables for mapping: the building density, the population density and the share of a certain building type. The mapping results are generally not as continuous as the administrative level apparently; rather, a porous picture of the settlement area emerges (Fig. 2c, d, e). The threshold values illustrated, as examples, also show a very variable, cartographic picture here: From the very low threshold of 5% of building density, which then classifies '*the urban*' widely beyond the area of what is regarded as large, medium and small-sized cities to the threshold of 40% of building density, where even within the metropolises only the centers remain urban. It shows again that almost any image can be created.

If we systematically calculate thresholds for all variables from the minimum to the maximum, all of Germany is classified as 'urban' at the minimum and at the maximum it is classified as completely 'rural'. The course between the extremes shows us the respective degree of urbanization in relation to the particular threshold applied (Fig. 3).

For interpretation purposes, let us compare the results in Fig. 3 with the United Nations benchmark that Germany is 80.3% urbanized: If we use the variable 'building density' to match these 80.3%, we find that this would be the case from a threshold value of 4.7% (see Fig. 3a1). If we reference the urbanization rate of 80.3% to the variable population instead of building density, then this would be the case for a threshold of 11.2% (see Fig. 3a2). And if we use the share to the maximum value of population density in Germany per grid cell, then this would be the case for a threshold of 6.9% (see Fig. 3b). Finally, if we reference the 80.3% degree of urbanization on the share of MFH, this would be the case from a threshold of 1.4% (see Fig. 3c). All of these threshold values derived to achieve the 80.3% degree of urbanization as suggested by the UN (2018) appear to have comparatively low densities, both in terms of the built landscape as well as for population. Studies that deal with suburban structures, i.e. in peri-urban areas that are generally assumed to be of low-density, show densities in the medians that are 15% in the lowest range, but around 20% appears to be much more characteristic (e.g. Taubenböck, Kehrer & Wurm, 2015). Thus, the built structures that are counted as urban using the 80.3% degree of urbanization do not seem to correspond to what we perceive as urban.

4.3. A probability-based approach for specifying the degree of urbanization in a differentiated way

Which variable for the classification of '*the urban*' versus '*the rural*' is best, which threshold is suitable, which spatial entity makes sense? This cannot be said with any certainty. To nevertheless arrive at a statement of how urbanized Germany is, we propose the combination of all the different conceptual approaches introduced above. In our example, we choose the combination of 29 variants (cf. Section 3.1.). The result is a classification at grid level, which allows a probability-based allocation to 'urban' or 'rural' (Fig. 4). The illustration shows the complexity of the problem in that the graded probabilities are a product of conceptual, parametric, and threshold-based variations.

The suggested three-part classification, i.e. high probability of being '*urban*' or '*rural*', as well as a '*range in between*', allows a differentiated quantification: We can safely say that 1.9% of Germany's land area has a clearly urban land cover. It is remarkable that on this comparatively small area, almost 50% (41.59 million) of the population resides. In other words, Germany is at least 50.0% urbanized. 32% of the area, in turn, is clearly assigned to a rural land cover. However, only 3.95 million people (4.8%) live there. The 66.1% of the land area or the 45.3% (37.65 million) of the population between these clear classifications could of course also be classified dichotomously into urban and rural according to a methodological specification. The probabilities for one of the two classes, however, are not so clear-cut; they represent a kind of transition.

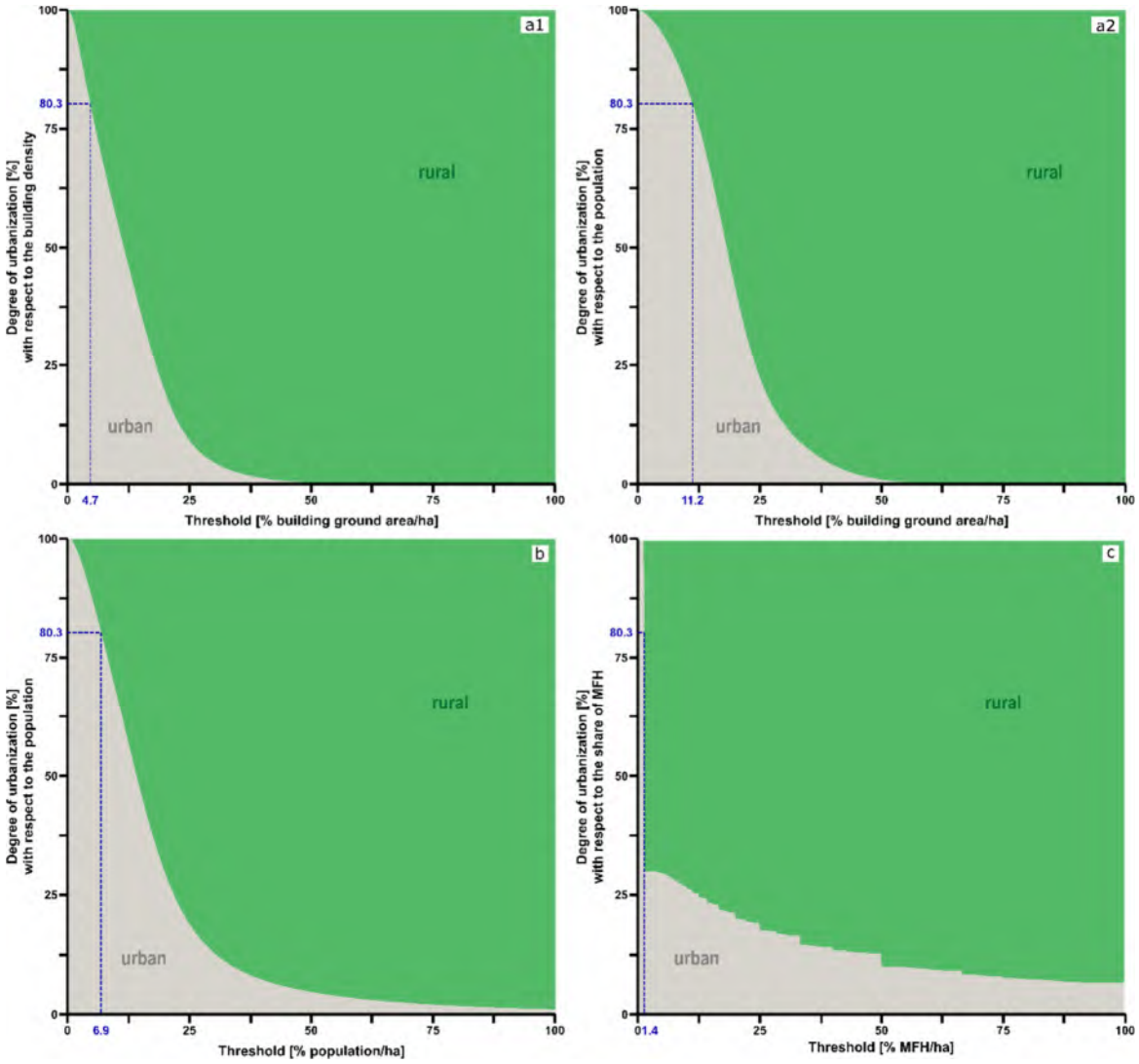


Fig. 3. Degree of urbanization for Germany depending on the thresholds applied at grid level; a1) thresholds for % building ground area/ha with respect to the building density, and a2) with respect to the population; b) thresholds for % population/ha with respect to the population; c) thresholds for % MFH/ha with respect to the share of MFH. The horizontal dotted blue lines indicate the 80.3% degree of urbanization as suggested by the UN (2018). The vertical dotted blue lines indicate the particular threshold to achieve the 80.3% degree of urbanization for Germany.

To present this transition in a more differentiated way, we have subclassified this range of uncertainty based on quantiles. With it, we count another 18.1% of the population closer to

'*the urban*', and 12.6% correspondingly closer to '*the rural*' (Fig. 4). In other words, in this conceptualization Germany is certainly at least 50.0% urbanized, probably as much as 68.1% of the population. And, Germany is certainly rural to at least 4.8% of the population, probably as much as 17.3%. The missing 14.6% feature a high level of uncertainty according to our probability-based approach. We thus do not assign these to a type and these remain in the *"range in between"* class. This type of land cover is sometimes regarded as suburbia, inter mediate city (Zwischenstadt) or settlement mash. We believe that this more nuanced, multi-part specification allows for a more differentiated statement about the degree of urbanization.

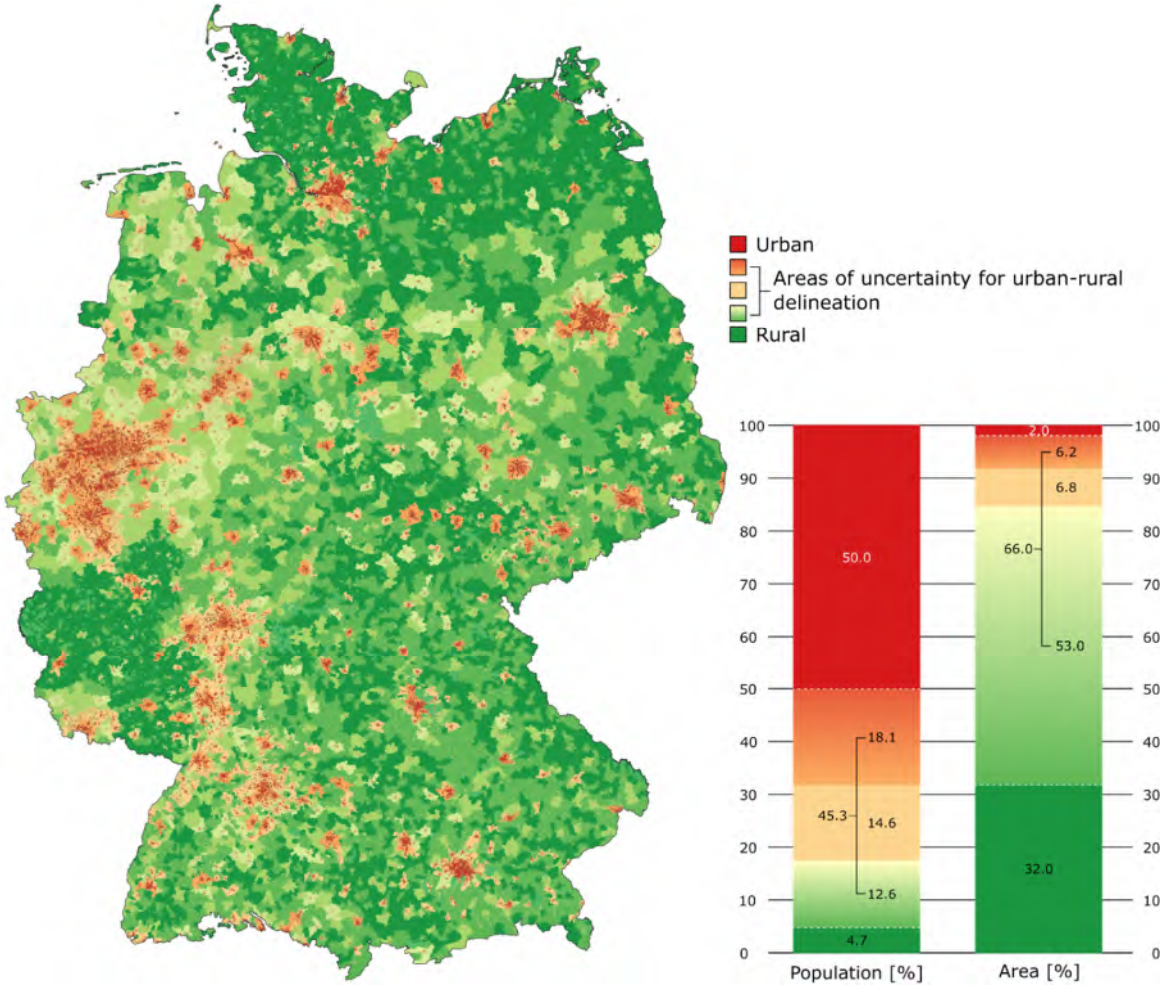


Fig. 4: Degree of urbanization for Germany based on the proposed probability-based approach: Cartographic illustration and statistics on population and area shares.

5. Discussion

"The world is becoming an urban one" has been embraced as metanarrative by influential thinkers, researchers, practitioners or politicians (Burdett & Sudjic, 2011; Glaeser, 2010; Ash, Jasny, Robert, Stone, & Sugden, 2008; Sennett, 2018; UN, 2018, among many others). This narrative is repeated with monotonous regularity across diverse discursive, scientific, social, economic, environmental, institutional and political terrains (Brenner & Schmid, 2013). The referenced urbanization trend is certainly correct to describe one of the main processes of global change. However, our study shows that this statement is based on an

inconsistent conception and that the numbers can easily be mis interpreted or manipulated. So, what does this mean?

First of all, we have to constitute: The figure that 56.6% of the world's population lives in urbanized areas (UN, 2018) arises from the synopsis of national statistics. As a compendium of national realities, it has therefore absolute justification. However, the global figure cites results from a highly inconsistent approach. In our view, it will only remain valid if the inconsistency of the methodological calculation is clearly stated and the uncertainties are thus expressed. If one projects, as we have done in this study, e.g. the extremes of the thresholds used across the globe to Germany in a consistent manner, the range of results is from 32.1% to 99.9% urbanized. This range shows how fragile the statistic is and it does not really allow any truly justified statement about the degree of urbanization on our planet. Thus, we should always critically question the accepted and often used statistics so that they cannot be reinterpreted depending on the particular direction of any ideological argument. While the trend towards higher population shares in urban areas may certainly be correct, any absolute share given by any study must be viewed with caution. Since facts in the spatial sciences, as shown, contain ambiguities, we tried to ensure (decisional) transparency (Beerbaum & Puaschunder, 2019), and explainability (Benke & Benke, 2018), i.e. we made our conceptual thinking, the specific sources of data used and the ones not used or not available, the applied methods and the interpretation of variable results transparent.

Secondly, it becomes obvious that it is difficult to determine how urbanized our empirical example of Germany de facto is. With a closer look to our results, however, from our point of view, Germany is by far not as 'urban' as commonly assumed. The UN (2020) set the degree of urbanization to 80.3% based on the threshold value 5000 inhabitants per administrative unit. Eurostat (2021) defines the degree of urbanization for Germany at 69.6% using a 1km^2 grid and clusters of neigh boring grid cells with a density of at least 300 inhabitants per km^2 and at least 5000 inhabitants. It is clear to us that any specification of methodological approaches and thresholds must lead to opposition. Nevertheless, we would like to make the following arguments here: In our experience, the self-perception of German citizens in places with 5000 or only slightly more inhabitants is by no means an entirely or unambiguously urban one. If we additionally allow for the structural perspective at grid-level, i.e. that urban also means a dense form of living in multifamily buildings, then we even measure much lower figures. And, to achieve e.g. the suggested 80.3% at grid-level, building density values need to be set at 11.2% or even only to 4.7%. These are building densities that are lower than typical densities of areas of single-detached houses which are often located at the transition between urban and rural. We therefore propose a different way of presenting the degree of urbanization: A differentiated representation over a range of un certainties. By combining different approaches, we can definitely say that 50.0% of the population in Germany is urbanized, 4.8% is rural. The range in between of 45.3% is fraught with uncertainty. But this can be differentiated as well, further 18.1% tend to be closer to '*the urban*', further 12.6% tend to be closer to '*the rural*', and 14.6% feature such a high level of uncertainty that we classify them to the *'range in between'* class. This 'range in between' is a spatial location and empirical description of an urban-rural transition zone, which correspond to conceptual terms such as urban-rural continuum or the urban-rural

gradient. The combined, probability-based approach allows us, through the high spatial resolution of the grid-level, to differentiate the degree of urbanization within administrative areas: for example, the probability of being classified 'urban' for a single-family residential area in a metropolitan community is higher than for a comparable single-family residential area in a low-density, village unit. We think that there is more clarity in this complication of the classification. In consequence, we believe from these figures that the degree of urbanization in territorial and structural terms is far from the indicated 80.3% or 69.6%. Thus, we conclude that we are far from the meta-narrative that urbanization is nearly complete in Germany. It is obvious that a higher proportion of the degree of urbanization should be in line with efficient land use, less land consumption and higher degrees of settlement density. However, it remains to be remembered, that this statement of lower degrees of densities in spatio-quantitative numbers could also be calculated with other variants which might change the numbers. And, this particular approach does not necessarily apply to an urban lifestyle which possibly make up higher proportions of the population.

This latter argument leads us, *thirdly,* based on the data available and applied here, to the remark that our approach is reduced to the variables building density, population density, and the share of certain building types. The 'urban' does also have other manifestations: The indicated urban lifestyle, which increasingly applies not only to city dwellers but is becoming the model of lifestyle in general, is not considered in our study. It manifests itself by a porous spatial form that applies more to people than to territorial forms of society. It is the formation of different spatial logics that produce their own logic of socialization through density in space (Held, 2005). Multi-criteria approaches using physical, demographic as well as socio-economic variables try to take that into account (e.g. Küpper, 2016). The data specifically for such multi-criteria approaches beyond physical and demographic domains, however, are unavailable at the very high spatial grid-resolution applied in our study and therefore cannot be integrated here. It is perfectly clear to us that our assumptions are abstracting reality: single-family homes, as example, framed by dense, urban block developments may have to do more with the urban (life) than isolated block developments in a rural, village setting. We account for this by showing what different shares of MFHs mean with respect to being classified as 'urban' or 'rural' by systematically testing thresholds over the fraction of MFHs at grid level. In our combined, probability-based approach, we apply a distribution of thresholds of 1, 25, 50,75 and 99% along the entire range, i.e. we consider that there can be no clear right or wrong for the classification. This approach, for all its systematic nature, is of course only one conceptual approach among many possible ones. It is the attempt to combine the most diverse conceptual approaches as the categorization into urban versus rural remains a socially constituting feature. Thus, we think the spatial-quantitative delimitation of these categories, as done in our study, remains necessary: it combines several perspectives and thus approaches the topic in a more multi-faceted way than with one or two variables, but it is clear that it is not a holistic perspective.

Fourthly, our input data is comprehensive, largely consistent and features, compared to data sets in other large area studies on the degree of urbanization, spatially and thematically unprecedented resolutions. And, since the data sets come from official databases, we are convinced that the errors are comparatively small and our result is generally reliable.

Nevertheless, they contain errors, in completeness, in consistency, in accuracy. We cannot quantify this error here due to the lack of better data sets, and thus the impact on our analyses cannot be quantified.

Fifth, there is now a need for a consistent approach beyond our sample of Germany to be able to determine the degree of urbanization better and make it more comparable. However, geodata on the individual building level comparable to our cadastral data are not available globally. And yet, there are high resolution and even freely available cadastral building data, e.g. in France (IGN, 2022), the Netherlands (PDOK, 2022) or Luxembourg (Le Governement du Grand-Duche de Luxembourg, 2022). Other data sets from remote sensing or community-based approaches such as Open Street Map (OSM) are also possible input data at the individual building level: Remote sensing-based continental building data, e.g. for Africa (Sirko et al., 2021), South America and Australia (Microsoft, 2022), and national building data for the US, Canada, Uganda, Tanzania, Nigeria, Kenya (Microsoft, 2022) have been derived; however, yet with lower accuracies than cadastral data. OSM data with complete and accurate building footprint coverage exist especially in urbanized areas; however, country-wide coverage is often not given (as e.g. researched by Salvucci & Salvati (2022) for Italy). At the global level, more and more geodata on settlement areas (especially from earth observation, e.g. the Global Urban Footprint (Esch et al., 2012), the Global Human Settlement Layer (Pesaresi et al., 2013), or the Global Impervious Surface Area (GISA) (Huang et al., 2021), among others, become available. For global studies, we still have to make compromises in terms of spatial and thematic resolution – one reason why we have focused on Germany only for this study to experiment with the spatially and thematically unprecedented resolutions. A promising approach without the need for high resolution geospatial data on building footprints or settlement areas is machine learning. The probabilities for being urban, as classified in our study for Germany, can serve as a training dataset. The capabilities applying such classifiers on areawide remote sensing input data have yet to be systematically explored. We are fully aware that there cannot be one and best method or threshold. We even believe that the systematic test of thresholds and the elaboration of a probability-based range, as we have done for Germany, is needed to be able to grasp the status quo more clearly. By this example, we are more concerned with reigniting the discourse. The point is that we want to make clear: we must not accept the widely unquestioned meta-narrative as the only truth per se. We want to call for constantly questioning ourselves about concepts, data, methods and perspectives. Here we plead for approaches that are more interdisciplinary and pragmatic, i.e. what constitutes the 'urban' has to be illuminated by demographers, geographers, spatial scientists as well as by sociologists, economists, psychologists or lawyers, among others. In order to then generate consistent approaches, one must pragmatically fall back anyway on existing, accurate, area-wide data sets. In this field of tension, it is worth getting closer to the truth.

And *sixth and last*, the places of life, *'the urban'* vs. *'the rural'* are sometimes ideologically, sometimes emotionally charged. A factual discourse, a social, ecological, economic, demographic, or political discussion of direction, must make room for a more differentiated approach. This study attempts to do just that, to put the status quo on the topic of the degree of urbanization on better, more differentiated empirical facts. It has been shown that urban

and rural life have highly different implications: Higher density, for instance, is resource-saving (e.g. Bettencourt, 2007), it increases creativity and productivity (e.g. Glaeser, 2010), among many other effects. At the same time, this is offset by higher exposures to air (e.g. Müller, Erbertseder, & Taubenböck, 2022) or noise pollution (Staab, Schady, Weigand, Lakes, & Taubenböck, 2021), by higher living costs, among other effects. Only both, the reliable knowledge of the effects, and the reliable knowledge of how the distribution actually is, allow a discourse. In this paper, we tackled the latter for better empirical evidence in the hope to benefit policy makers on the global and national levels working on agendas.

6. Conclusion

It can be frustrating to answer the guiding question of this paper *'how urbanized is Germany?'*, with *'it depends'* in the conclusion. It depends on the concept, the data, the spatial units of measurement, the variables or the thresholds applied. In consequence, we conclude that there is a lack of that one and best concept, set of data or methods and just accept that it does not exist. Here, we proposed that the combination of different approaches resulting in a probability-based range of the degree of urbanization providing uncertainties seems to fit the question better. We suggest that further research is needed that could complement this method with other appropriate approaches. Therefore, we advocate reconsidering simple and often superficial dichotomous classifications and we propose in this contribution a more differentiated way of specifying the degree of urbanization.

We believe based on our empirical set-up that Germany's degree of urbanization is between certainly 50.0% up to possibly 68.1% of the population. We must allow this ambiguity, we must allow more narratives instead of the intended single truth, because in spatial science, rarely anything is unambiguous. Nevertheless, this means that the degree of urbanization in Germany is, in our estimation, lower than presented by other statistics.

While this is a clear result, we believe that still more and different concepts, data and methods from various disciplines are needed – concurrently and side by side. This corresponds to the call of Acuto, Parnell & Seto (2018) that bringing together scholars from disparate fields and reorganizing existing knowledge domains for a transformation of current science – policy interfaces is in demand for a better understanding of '*the urban'*. For the comparably data-rich situation in Germany, as applied in this study, a complement to the applied physical approaches on building density and types and to the population density would be, e.g., data on lifestyles, behavior pattern or attitudes; at the global level, physical or population related approaches cannot rely on such high-resolution data and accuracies yet, and much less so in socioeconomic or other data. Here this presented approach needs to be adapted to the data circumstances. We suggest that it takes both, more different approaches based on other and different data as well as the systematic tests of thresholds to draw a more reliable, more holistic picture of the degree of urbanization. The different results, when viewed together, as we have done here for Germany, promise a clearer, more differentiated picture of the degree of urbanization, without concealing statistical ambiguity. The acceptance that the world is more complex than mathematical correct results is the basis for this. And, the acceptance of lower urbanization degrees than commonly used, demands

for a re-consideration of political argumentation, policy development and re-doing empirical research on different implications of '*the urban*' vs. '*the rural*'.

We believe the extension of these multiple approaches to a multidimensional perspective or to larger areas, even the whole world, is necessary to really test the meta-narratives. From our perspective, the German meta-narrative needs to be adapted, and we would not be surprised if it were similar for the entire world.

Acknowledgments

This work is funded by the German Federal Ministry of the Interior and Community under the project "*Fe4ErsiGG_IF-Bund - Fernerkundung für die Ermittlung siedlungsstruktureller Grenzen und Gebietseinheiten*".

Chapter IV

Spatial delineation of urban corridors in North America:

An approach incorporating fuzziness based on multi-source geospatial data

Cities, 2023, vol. 133

Isabel Georg, Thomas Blaschke & Hannes Taubenböck

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.1016/j.cities.2022.104129

Received 8 Aug 2021; Accepted 27 Nov 2022; Published 15 Dec 2022

Abstract

Urban corridors are – from a spatial perspective – massively large, linear urban agglomerations consisting of a number of big cities or clusters aligned along high-speed road or rail lines. Fixed administrative boundaries are commonly used to define such urban areas. However, this does not usually reflect the actual extent of the built-up space in today's changing, multi-faceted urban landscape. Earth observation data provide the means to identify urban space in its spatial dimension, disregarding preconceived boundaries. We therefore use multi-source geodata including night-time lights, settlement patterns and population density to spatially delineate large urban corridors in the United States and parts of Canada and Mexico. Using pre-classified input layers, we identify and present varying outlines of 14 urban corridors through geospatial methods. With this approach, we address spatial ambiguities of such concepts and show fuzziness at the edges of these corridors.

Keywords

Urban corridor; large urban areas; urban extent; spatial delineation

1. Introduction

It is commonly accepted that urban areas are rapidly growing, absorbing the largest share of the world's population increase (e.g., UNDESA, 2019; UN-Habitat, 2013). Urban expansion happens not only on a horizontal (e.g., Angel et al., 2005; Soja & Kanai, 2008; Taubenböck et al., 2014), but also on a vertical scale (e.g., Lin et al., 2014; Qin et al., 2015). While the vertical dimension of cities can be measured through 3D methods – such as LiDAR point clouds or Digital Surface Models (DSM) – down to individual building heights, their horizontal expansion is less distinct. For a long time, the measurement of horizontal urban extent most commonly relied on administrative boundaries and related population densities (e.g., UN-DESA, 2019). However, these artificial spatial entities are only an inadequate approximation of the actual urban dimension using a physical perspective on the settlement pattern and, due to technical developments and data availability, a shift towards urban area delineation based on morphological features or functional connectedness can be observed (e.g., Davoudi, 2007; Taubenböck et al., 2019). In ever-changing, expanding urban landscapes, the underlying spatial structures may not reveal themselves immediately. Understanding spatial contiguities and dimensionalities outside their administrative constraints, thus, is a valuable tool for the governance of urban areas beyond existing spatial units and spatial contexts and a worthwhile concern for urban studies.

1.1. Urban corridors: a spatial view

Urbanization dynamics are mostly measured and monitored for individual cities. However, nowadays spatial constellations emerge beyond the administrative spatial units that form the typical spatial reference unit, in which individual cities merge and become a larger conglomerate (Taubenböck et al., 2019). This emergence of new shapes and sizes of large urban constellations with their ambiguous definitions, non-apparent spatial forms and often overlapping spatial extents (cf. Georg, Blaschke, Taubenböck 2016a) calls for new methods to empirically describe and map them.

According to UN-Habitat (2008, 2013), three main forms of these novel configurations are megaregions (or mega-regions), urban corridors and city regions (or city-regions); in addition, a plethora of terms describing large urban areas of various shapes and sizes exists. Of these, megaregions in particular have been extensively studied (e.g., Banerjee, 2012; Florida et al., 2008; Growe et al., 2015; Hall, 2009), but their spatial delineation remains subjective: There is no common agreement with respect to which indicators (e.g. economic output, population, physical dimensions or functional linkages) or spatial entities (e.g., administrative boundaries or natural regions such as watersheds) allow a meaningful spatial delimitation. Similarly, other concepts of large urban areas are also not universally defined (see for example Fang & Yu, 2017, who aim to provide a definition of ‘urban agglomeration’ based on a literature review). This includes the construct of urban corridors which, although the term was introduced over 50 years ago (Whebell, 1969) to describe the urban development in Southern Ontario (Canada), has not gained a large amount of influence or empirical evidence.

From a purely spatial perspective, an urban corridor is a linearly shaped and massively large polycentric agglomeration consisting of several big cities (often megacities) or city clusters and their hinterlands, connected through high-speed road or rail infrastructure (Li & Cao, 2005; Neuman, 2000; Priemus & Zonneveld, 2003; Trip, 2003; UN Habitat, 2008 & 2013). Other interpretations imply the transportation network itself – as is the case with road corridor (Forman & Alexander, 1998), transport corridor (Curtis, 2006; Jain et al., 2021; Trip, 2003; Van Houtum & Lagendijk, 2001) or freight corridor (Rodrigue, 2004). Examples include *BESETO* (Beijing–Seoul–Tokyo) (Choe, 1996), GILA (Greater Ibadan–Lagos–Accra) or Lagos–Abidjan (Hertzog, 2020; UN-Habitat, 2008), the *Blue Banana* from Birmingham to Milan (Hospers, 2002), the Delhi-Mumbai Industrial Corridor (DMIC) (Chandrashekar & Aithal, 2021; Jain et al., 2014) or – one of the most typical examples – the *Boswash* area from Boston to Washington (Florida et al., 2008; Gottmann, 1957).

To date, not many studies actually focus on urban corridor delimitation or mapping and a commonly accepted definition of the term is still missing. This lack of research may be due to the fact that a number of corridors have only emerged in recent times and the connectivity within the linear extent frequently does not immediately show (as is the case, for example, with BESETO, which crosses international borders and water bodies). The technology and high-speed transport to connect people over large distances has only emerged in the last decades (e.g. the Shinkansen train in Japan, TGV in France or ICE in Germany) (Glocker, 2018). In their massive dimensions, most corridors are also mega regions, however, their linear development along high-speed transport routes create spatial characteristics which make them stand out in the urbanized world of today (Kar et al., 2021). Their impact on land use and environment can be significant due to changes in land use and land cover, particularly through an increase in buildings and impervious surfaces, resulting in higher surface temperatures (Jain et al., 2014).

To describe large urban agglomerations such as megaregions or urban corridors, morphological, functional or network approaches can be applied (Glocker, 2018; Yang & Hillier, 2007). Morphological approaches use, for example, settlement densities (or the indirect derivative of population densities) or proxy activity by night-time lights derived from multi-sensoral satellite imagery, e.g. Taubenböck et al. (2017a, b) who use settlement density patterns derived from Earth observation data to categorize urban areas and their spatial connectivity in Europe and China respectively, or Wei et al. (2017b) who implement a multi-temporal estimate of population density in the Pearl River Delta. An increasing number of studies delineate large urban areas without administrative boundaries (e.g., Taubenböck et al., 2014; Florida et al., 2008; Arcaute et al., 2015). Functional or network approaches are less tangible and often hampered by the lack of data availability or data consistency. They describe material flows such as commutes or immaterial flows in the form of observable (email, telephone) and non-observable (knowledge) exchanges (Glocker, 2018). For example, Nelson & Rae (2016) describe megaregions in the US based on commuter flows, showing a new perspective of functional economic geography. Crowd-sourced and social media data (such as OpenStreetMap, Foursquare or Twitter) have been utilized to identify urban centers or urban forms (e.g. by Cai et al., 2017, who use social media data in combination with night-time light imagery to locate main centers in

polycentric Chinese cities, or by Crooks et al., 2014, who present a typology of urban space from crowd-sourced data).

For urban corridors in particular, not many studies to map and describe them exist and these vary in their concepts, input data and methods. Khanna (2016) outlines 17 urban corridors in the US with a variety of shapes and extents based on the Regional Plan Association's (RPA, 2007) megaregions, Taubenböck et al. (2017b) map corridors across Europe using settlement patterns derived from Earth Observation data, and UN-Habitat (2008) sketch out several on a global map, and Georg, Blaschke & Taubenböck (2016b) categorize 67 urban corridors worldwide with respect to their length, width, area and connection using night-time light imagery in combination with questionnaires. Although the actual extents remain somewhat uncertain, urban corridors have been described to be between 400 and 1200 km long and 70 to 200 km wide, indicating a lengthy shape with a typical length-to-width ratio between four and ten. The urbanized area has been suggested to generally be at least 10,000 km^2 (Georg, Blaschke & Taubenböck, 2016b).

Over recent years, the term has been used to label some large urban areas in various countries. Li and Cao (2005) describe the evolution of the Guangzhou to Hong Kong urban corridor in China's Pearl River Delta, while Taubenböck et al. (2017a) identify an urban corridor along the Eastern seaboard of China from Beijing to Shanghai, reaching inland as far as Zhengzhou, through calculating spatial connectivity using Earth observation and population data. Several megaregions and megapolitans in the US have been termed urban corridors, such as the I-35, I-95 or Arizona Sun Corridor (RPA, 2007; Lang & Dhavale, 2005; Miller, 2021). In Indonesia, the Jakarata–Bandung urban corridor has grown along highways essentially through conversion of rice fields, affecting the rice supplies for Western Java (Rustiadi et al., 2021). The role of labor migration for the development and growth of the Lagos–Abidjan corridor is analyzed by Hertzog (2020). Furthermore, new corridors have emerged in India, a number of which driven by industrial activities, e.g. Ahmadabad-Pune or the Delhi-Mumbai Industrial Corridor (DMIC) (Kar et al., 2021; Mukhopadhyay, 2018).

The spatial coherence of urban corridors is frequently not acknowledged sufficiently in urban planning and administration: their governance is complex and involves decision-making on multiple levels and across a large number of local and regional agencies (such as water, electricity, transport, police and fire departments, health care etc.), most of which work within administrative boundaries (Innes et al., 2010; Ross et al., 2008). This results in urban corridors being governed and managed by a plethora of administrative bodies, while no agencies exist for certain types (e.g., bioregions, commutersheds, cultural regions) (Wheeler, 2009).

This study aims to provide a spatial outline of urban corridors using population, settlement and night time data for a multi-dimensional representation. Without pre-defining spatial units, we use a fuzzy approach to describe urban corridors in North America.

2. Data, study area and method

In this paper, we explore the spatial phenomenon of urban corridors with their specific large, linear shape, aiming to identify these in a data-driven, consistent, and thus comparable approach. The basic concept is as follows:

- Without a priori knowledge about the spatial properties of urban corridors and without predefining them, we base our approach on a 'master' corridor that is generally accepted by scholars to be a prototype: the *Boswash* area on the northeastern seaboard of the US (e.g. Florida et al., 2008; Regional Plan Association, 2007; Lang & Dhavale, 2005; Gottmann, 1957). We identify individual thresholds for several spatial input layers to describe the *Boswash* 'master' corridor as one coalescent patch. This basic threshold is then transferred to the whole study area to delineate coherent large urban corridors. Our initial focus is on 13 regions which we identified through a combination of a literature review and questionnaire (Georg, Blaschke & Taubenböck, 2016b). From these, we derive 'urban corridor candidates', calculate their dimensions and check for conformity with criteria describing urban corridors. The Method section below further describes the generation of probabilistic extents of urban corridors and grouping into four categories according to their spatial dimensions.
- Spatial analyses are inherently ambiguous and contain uncertainties. These include distortions and errors of data scaling and aggregation described as the Modifiable Aerial Unit Problem (MAUP) (Fotheringham & Rogerson, 1993): Even at the same scale, there are numerous possible combinations of spatial units which can make a difference to analysis results (Openshaw, 1984; Pietrzak, 2014; Marceau, 1999). In this study, we aggregate all input data sets to a 1 km^2 grid which we believe provides sufficient detail for a large spatial extent such as our study area. 1 km^2 grids have been used elsewhere for similar-scale projects, such as the INSPIRE Geographical Grid System (INSPIRE, 2014) or Hansen et al. (2000).
- Another issue in spatial analysis are boundary effects (Fotheringham & Rogerson, 1993). Hence, rather than relying on administrative boundaries, we use a homogeneous base without pre-defined borders and edges. Furthermore, our results show variable, fuzzy extents of urban corridors, which we believe is a good fit for the concept. Fuzzy boundaries are more effective for a functional connectedness and flexible cooperation between stakeholders (Nelson & Rae, 2016; Yang & Hillier, 2007). Fuzziness is frequently expressed in Boolean terms through a class membership (such as 'artificial surfaces', 'vegetation', 'water' etc.) of either 0 or 1. Fuzzy logic, on the other hand, deals with the possibility of partial membership to a specific class, without clear boundaries (Zadeh, 1965). Membership to a class is between 0 and 1, with the degree described as numerical information (Malczewski, 2004). For the delimitation of urban corridors, a distinction between "corridor" and "not corridor" is diffuse and a single criterion is not likely to provide adequate information; rather, a combination of characteristics is necessary. This is addressed in this manuscript through the use of different datasets representing different aspects of urbanity. The actual membership of an area to an urban corridor is described through membership values obtained through different input datasets.

- The input data used are from a variety of sources to ensure a multidimensional and therefore better empirical robustness of our approach. By this we avoid data related coincidences and smooth out inaccuracies of the datasets. Furthermore, urban corridors are fundamentally based on several indicators (rather than just one) which means that our data approach approximates the actual concept as closely as possible.
- Our approach describes urban corridors from a spatial perspective using physically measurable parameters. Functional or relational aspects are not taken into account due to data availability.

2.1. Input data

For our multidimensional analysis, we use a set of different input data, consisting of earth observation imagery and infrastructure data. Although these datasets vary in their resolution and quality, we chose them with regard to their global and mostly free availability from a similar timeframe (around the year 2010). Table 1 provides a summary of the input layers used in this study.

- *Night-time light data (NTL)*: We used data from the Defense Meteo rological Satellite Program Operational Linescan System (DMSPOLS) cloud-free stable lights composite from 2010 (Elvidge, n.d.), with a resolution of 30 arcsec and gas flares removed. Night-time imagery has been used for urban studies since the 1990s (Elvidge et al., 1997; Imhoff et al., 1997; Small et al., 2005). Apart from urban area extent, these night-time lights are an indicator for energy consumption, population density and GDP (gross domestic product) (e.g. Florida et al., 2008; Loveland et al., 2000).
- The availability of consistent global records for human settlements is quite limited and includes datasets such as the Global Human Set tlements Layer (GHSL) (Pesaresi et al., 2015) and *Global Urban Footprint (GUF)* (Esch et al., 2012) which has been used in this study. The GUF is a data product from the German Aerospace Centre (DLR) derived from TerraSAR-X imagery. Based on 2010–2012 data, the 12m product shows man-made settlement structures with a vertical dimension (particularly buildings) (Esch et al., 2012).
- *Global Land Cover (GLC-share)*: The dataset provided by the United Nation's Food and Agriculture Organization (FAO) consists of eleven different land use classes from different sources such as medium resolution satellite sensors – e.g., MERIS (Medium Resolution Imaging Spectrometer) or MODIS (Moderate-Resolution Imaging Spectroradiometer) – or national mapping schemes (Latham et al., 2014). For this study, we only use the artificial surfaces layer.
- *UCLouvain's GlobCover*: The land cover dataset from ESA is obtained from MERIS data (Bontemps et al., 2011; European Space Agency (ESA), n.d.). It consists of 22 land cover classes and similar to the GLC data, we only select the artificial surfaces class. This is defined as having an 'urban' percentage of more than 50 % for each pixel.
- *Global Impervious Surface Area (ISA) dataset*: It is based on DMSP-OLS night-time imagery and population count with a resolution of 1000 m. It includes man-made surfaces including buildings and roads (NOAA (National Oceanic and Atmospheric Administration), n.d.).

- *Population density data*: these are available from the Center for International Earth Science Information Network at Columbia University (CIESIN)'s Socioeconomic Data and Applications Center (SEDAC) (CIESIN, 2017). The version 4 (GPWv4) datasets consists of population counts from national censuses and population registers.

Table 1. Input datasets used in for the delimitation of urban corridors in this study.

Dataset	Source	Original resolution	Release year	Category	Content
NTL (Night-time lights - DMSP-OLS)	NOAA	30 arcsec (~1 km)	2010	Built environment	Night-time light emissions (proxy for economic activity)
GUF (Global Urban Footprint)	DLR	12 m	2012	Built environment	Artificial structures with a vertical dimension only
GLC-share (Global Land Cover) - artificial surfaces	FAO	30 arcsec (~1 km)	2013	Built environment	Artificial structures including buildings and roads
GlobCover - artificial surfaces	ESA / UC Louvain	300 m	2010	Built environment	Artificial structures including buildings and roads
ISA (Estimated Impervious Surface Area)	NOAA	1000 m	2010	Built environment	Artificial structures including buildings and roads
GPWv4 (Population Density)	CIESIN - SEDAC	30 arcsec (~1 km)	2017	Socio-economic	Population density

Each of the datasets offers a different interpretation of urban extents and their spatial relationships. High *NTL emissions* are a proxy for eco nomic activity (Mellander et al., 2015) and thus used here for the spatial delineation of contiguous areas of increased economic activity. An increased *settlement density* as shown in the GUF implies coherence and connectedness of vertical built structures, i.e. buildings: The spatial distribution of settlements is less fragmented with higher degrees of urbanization (Angel et al., 2012). Hence, settlement density serves as an indication of where human activity is concentrated. *Artificial surfaces layers* in land cover datasets usually include urban features such as roads, carparks, parkland or sports facilities, i.e. those without a vertical dimension (as is the case with GLC-share) (Latham et al., 2014). Thus, these datasets differ in their urban extent compared to settlement datasets such as the GUF which do not include those features. *Population density* describes how many people live within an areal unit (here: 1 km^2) and is closely related to infrastructure, particularly roads (Cervero & Murakami, 2010). In this way, we illuminate spatial contexts from a variety of perspectives.

For a consistent spatial analysis, we aggregate all input data to a consistent 1 km^2 grid and re-project to Albers Equal Area coordinate system. To reduce artificial noise, we remove all single-pixel urban areas (i.e., all areas with just 1 km^2).

Because of their global scale and their intrinsic aggregation effects due to data origins, each of the datasets used has its limitations. Commonly, night-time imagery will overestimate urban areas because of blooming effects (Henderson et al., 2003), while the GlobCover dataset underestimates or omits cities (Potere et al., 2009). Hence, our use of a collection of datasets from various sources helps to describe urban corridors from different perspectives.

2.2. Study area

Our study area covers the whole connected United States territory plus several hundred km to the north and south. With an extent ranging north of Edmonton, the most densely populated parts of Canada are included. To the south, the study area extends south of Aguascalientes in Mexico and comprises the northern part of Cuba and most of the Bahamas.

The northern and southern boundaries were chosen to include possible urban corridor areas which transgress the administrative boundaries of the US.

For most parts of this region, large urban areas have been extensively studied with some of the input data we also use (see above). Our findings complement these studies by focusing on the urban corridor phenome non, by disregarding international boundaries, allowing for consistent, empirical delimitations, and by taking fuzziness at the boundaries into account.

2.3. Method

In order to identify urban corridors in a consistent and comparable way, we use a two-step approach: First, we spatially isolate and describe large urban areas in general. In a second step, we derive our particular urban corridors from these. Figure 1 shows the general workflow which relies on the same procedure for each individual layer: defining appropriate thresholds for the classification of areas potentially belonging to an urban corridor, identifying non-coalescing urban patches belonging to the respective urban corridor and ultimately mapping coherent regions. The combination of all individual spatial results allows the generation of probabilistic extents for possible urban corridors with fuzzy extents.

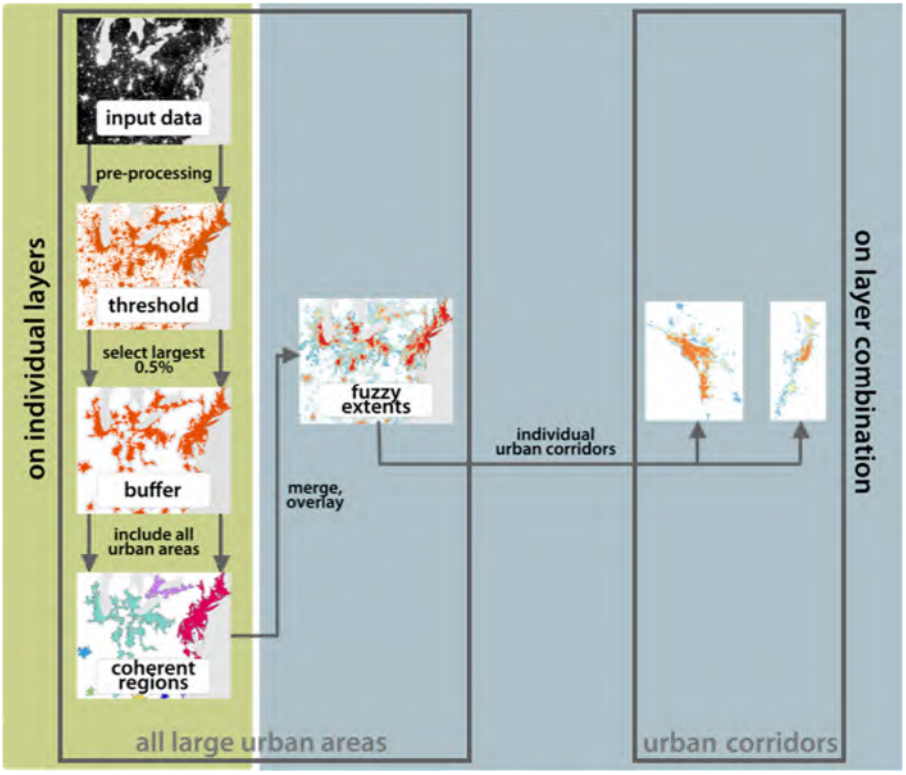


Fig. 1. Workflow of the urban corridor delineation. The green part shows the analysis performed on each individual input layer separately (pre-processing, se lection based on a threshold, buffer and identification of coherent regions). The steps in the blue part are completed on all layers combined. The first steps including the fuzzy extents (left half of the image) are applied to all large urban areas; individual urban corridors are identified at the end. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.3.1 Identification of urban corridor candidates

Assuming that high densities and spatial coherence are indicators to distinguish between urban centers and their rural hinterlands, we base our spatial identification of large urban areas and specifically of urban corridors on a high concentration of settlements, infrastructure or population over large distances. First, we derive spatial delineations using the various input datasets separately to account for the different perspectives (i.e., settlement, population etc.). We then combine these individual spatial representations of large urban areas for a multi-perspective, fuzzy delineation independent of administrative boundaries.

As an initial step, we define how large urban areas can be best described uniformly. Since we chose not to make a priori any predetermination of threshold values for delineation, we model this on an area generally accepted in scientific studies as a prime example of an urban corridor: the well-documented *Boswash* area from Boston, MA, to Washington, DC, which we use as a 'master'. Also known as *Megalopolis* (Gottmann, 1957) or *Northeast Megaregion* (Lang & Dhavale, 2005), the area covers around a fifth of both US population and GDP, but only around 2 % of the land area (Lang & Dhavale, 2005). Because it can be separated from surrounding, less urbanized space, the area is suitable as a model from which to derive a spatial delimitation of large urban areas. Georg et al. (2018) developed a method for a fuzzy delimitation of Boswash, using multi-source earth observation, infrastructure and socioeconomic data including night-time lights, settlement classifications, road density and income data.

Having established Boswash as our 'master', we apply a threshold on each individual input layer which best describes this area as being spatially coherent, allowing for the greatest spatial contiguity. Threshold approaches have been used to describe urban extents and allow to explore the variability of possible urban forms (e.g., Florida et al., 2008; Pesaresi et al., 2008; Rozenfeld et al., 2011; Xie & Weng, 2016; Xu, 2008; Zhou et al., 2014). Since our aim is to spatially map very large urban areas, particularly urban corridors, we only select the largest urban patches derived from these thresholds for each input layer. To avoid absolute figures, we choose a percentage value (the largest 0.5 %, see Discussion). This returns a different total amount of urban areas for each input layer, but with a distinct, traceable and clearly visible spatial separation between patches.

For other regions, the spatial coherence is not as strong as in Boswash, resulting in smaller individual patches. The connectedness can only approximated through 'nearness', i.e. how close urban areas are to each other. In order to establish this proximity, we apply a buffer around the urban patches on each layer. This buffer is determined through the prerequisite that Boswash remains a spatially separate urban region, i.e. does not merge with any of the large neighboring areas (such as the *Great Lakes* megaregion or *Golden Horseshoe* around Toronto). This way, we connect all urban areas within each buffer, allowing us to merge them while at the same time separating them from other connected clusters (see also Fig. 2). Through this procedure, we obtain the main urban centers and their spatially close urbanized surroundings of the previously selected largest 0.5 % of all urban patches larger than 1 km^2. We then add the omitted 99.5 % into the buffers to account for all urban areas. Figure 2 illustrates this approach for the night-time light imagery in the Chicago area.

Performing this for every single input layer, we obtain six different implementations of connected urban areas where a grid value of 1 represents 'urban' and 0 denotes 'non-urban' space. To show how robustly an area is characterized as urban, we sum up the individual grid values of the six different implementations. Each output value, thus, represents the sum of the values in the corresponding cells of the six input layers. Higher values (up to a maximum of six) mean that a cell has been classified as urban by more input layers. Through this method, we can identify and map urban space as a representation of varying degrees of agreement between differing input layers.

Fig. 2. The Chicago urban area identified by the largest 0.5 % of urban patches (a) in the NTL dataset. A buffer is applied (b) and all omitted smaller patches within this buffer are included again (c). The bottom right patch is an example of the buffer size (b, c): it is not part of the buffer around Chicago but belongs to a different cluster.

2.3.2 From candidates to corridors

In the first part of our analysis, we established the general urban extent of possible urban corridors. Whether an empirical baseline study using multiple spatial layers confirms these as urban corridors - our main focus in this research - based on the identified spatial indicators is being examined in this second part.

Within our North American study area, 13 regions have been pro posed as urban corridors in a global inventory by Georg, Blaschke & Taubenböck (2016b). These are, in alphabetical order:

Arizona Sun Corridor, Boswash, Calgary–Edmonton, Cascadia, Chicago, Colorado Front Range, Florida, Gulf Coast, I-35, Northern California, Piedmont Atlantic, Quebec–Windsor (Southern Ontario), South California.

We investigate how these possible corridor areas can be described using the following approach: For the 13 areas, we determine the spatial coherence of urban patches for each input layer and – analogous to the identification of urban corridor candidates – establish the agreement between the individual results, i.e. we add up the individual grid values ('urban' = 1, 'non-urban' = 0). To allow for overlap and different main urban centers, we perform this for each corridor separately. More precisely, for each of the potential urban corridors, we select the main urban centers and all urban areas which are spatially amalgamated with this center (as established through the individual thresholds and buffers). We then calculate the number of matching input layers to obtain the fuzzy extent for each cluster. The result ranges from 0 to 6 for each pixel, where, like for the corridor candidates, 0 means that none of the input layers identify a pixel as urban, while 6 means that all input layers do. In two cases, *Chicago* and *Piedmont Atlantic*, we cross-check the extents by using alternative cities as core (Chicago/Detroit and Atlanta/ Charlotte - see also Fig. 9) since the resulting urban corridors covered a similar region for a low number of matching input layers.

To describe these areas further, we measure their length, width, length-to-width ratio and area. Length and width are measured manually to allow for specific curved shapes and to exclude water bodies. Although water bodies can be part of an urban corridor (Georg, Blaschke & Taubenböck, 2016b), we believe that the length of an area such as the 'Golden Horseshoe' with Toronto as main urban center is better expressed by not cutting through the water body since it represents an obstacle that can only be traversed with significantly more difficulty than a high-speed road. The area calculations include both the total area and the specific urbanized space within. The total area is approximated through a convex hull, the urbanized space is the area of all urban patches combined.

Finally, we eliminate areas identified through this approach which do not show characteristics typical for urban corridors (a specific minimum length, width, length-to-width ratio or area) as suggested by Georg, Blaschke & Taubenböck (2016b). The remaining areas represent those regions best classified as urban corridors. To account for the varying extents of these, we group them into four categories based on the area measurements for three agreeing input layers.

For this analysis, we used two software packages, namely ArcGIS and QGIS. ArcGIS was used for data aggregation, reprojection, removal of artificial noise, selection of the largest 0.5 % of urban patches, setting of thresholds, buffering and the spatial joining of the buffered urban areas to individual corridors. The classification of urban corridors and raster calculations (summing up pixel values for several input layers) were performed with QGIS.

3. Results

3.1. Mapping of urban corridor candidates

Applying the method described above, a map of large urban areas with a graduation from weak (teal) to strong (red) membership derived through combining all six input layers is generated in consistent and area-wide manner for the entire North-American study site (Fig. 3). The more layers classify an area as urban, the stronger the membership. In principle, the possible urban corridors of the memberships achieved become visible in this overview.

We detect here the bulk of possible corridors in the east, with a particular accumulation in the north-east. Clearly visible is Boswash, the 'master corridor' in the Northeastern US as the largest region with a strong membership to urban areas. Other prominent large extents include parts of Florida and Northern California as well as major cities of the *Great Lakes* megaregion (e.g., Chicago, Detroit, Pittsburgh, or Cincinnati).

The result at this stage implies a membership to any large urban area without taking into account which specific spatial cluster an urban pixel can be assigned to. While the red areas show more or less clearly visible clusters, it is not always unanimously obvious – especially in the eastern part of the study region – to which particular stronger (orange to red) cluster a lower-membership (teal) area belongs to. The delineation of specific spatially connected, linear clusters – urban corridors – is detailed in the next section.

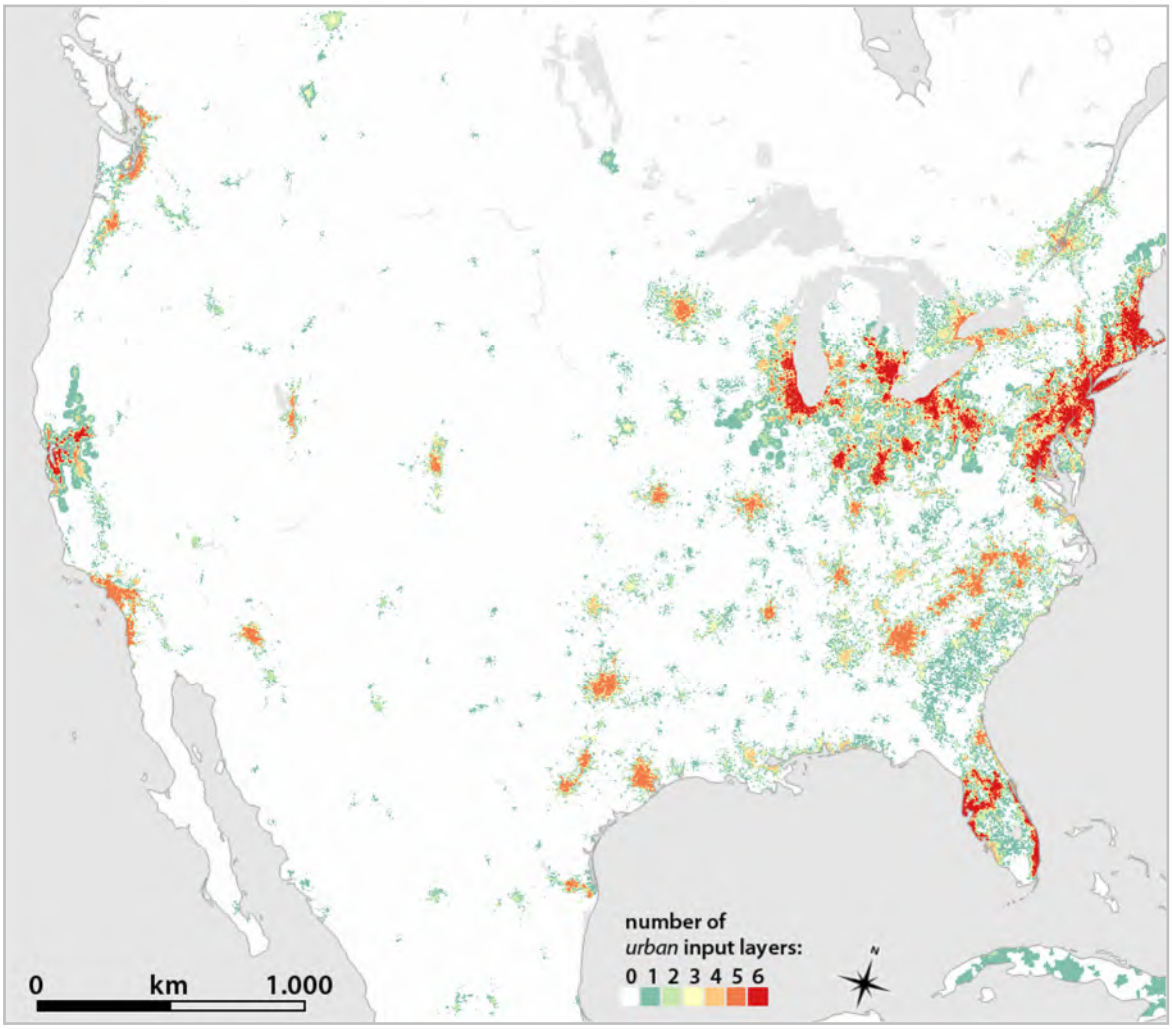


Fig. 3. Overlay of top 0.5 % of large urban areas, color-coded according to the number of input layers defining 'urban'.

3.2. Identification and characterization of urban corridors within the study area

For the delineation of our specific urban corridors, we perform a grouping of urban patches for each input layer. Figure 4 shows, for the north-eastern part of the study area, how heterogeneously these clusters are represented in the different datasets. The Boswash area, as specified through our requirements in the thresholding/buffering approach, is a separate, connected patch in all six input layers. Other patches are less concordant: For example, Chicago and Detroit form separate patches in three datasets (GUF, GLC Land Cover and GlobCover) but are connected in the other datasets.

From these separate large urban area clusters, we select for each layer the areas identified as urban corridors (cf. Georg, Blaschke & Taubenböck, 2016b), overlay the results and establish how many of the input layers agree. Although this method mirrors the identification of all large urban areas with their respective corridor candidates, the attribution to a specific potential urban corridor yields different results. Fig. 5 shows, for Cascadia with Seattle as the main center, the difference between all urban areas from the initial large urban area map (Fig. 3) and the coherent, specific urban corridor.

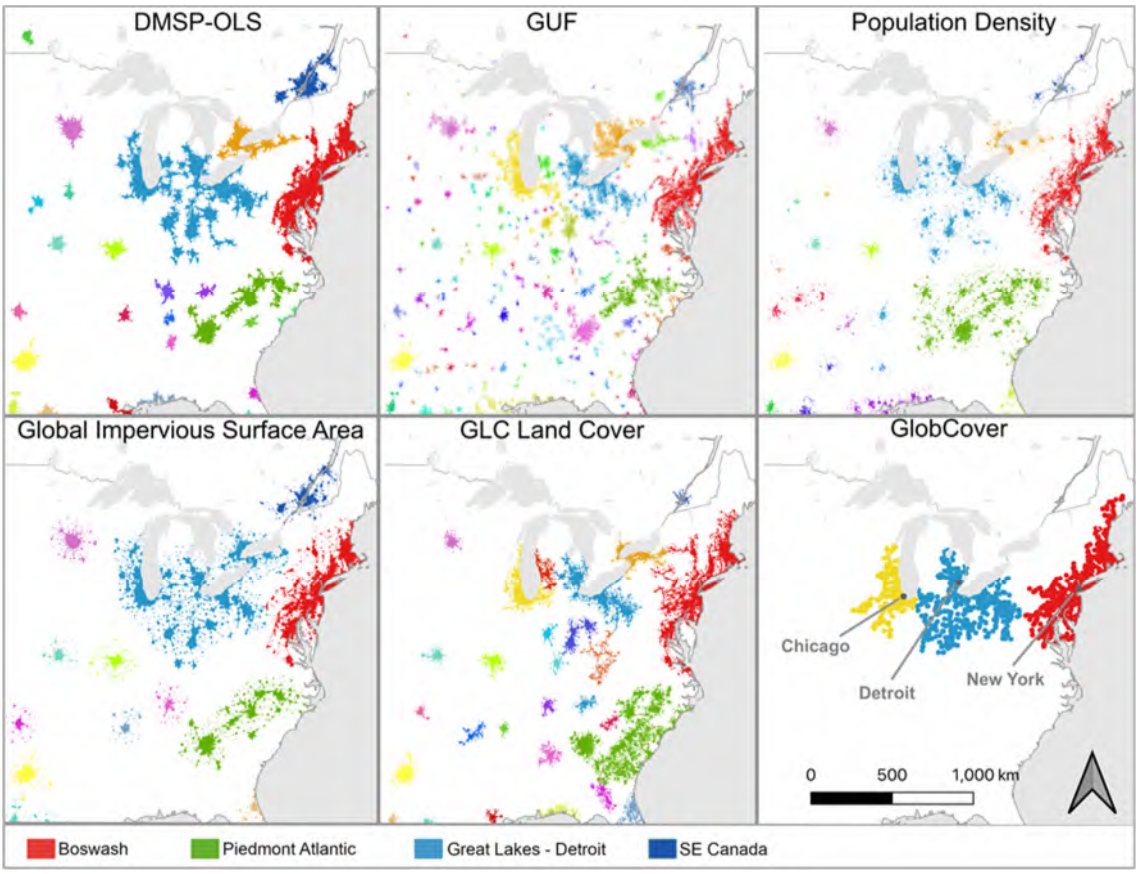


Fig. 4. Grouped urban clusters for each input layer (north-eastern part of the study area). While Boswash is a separate connected patch in each input layer, other patches are less concordant. Main colors: red (Boswash), green (Piedmont Atlantic), blue (Great Lakes - Detroit), dark blue (SE Canada), random colors for the remaining patches. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

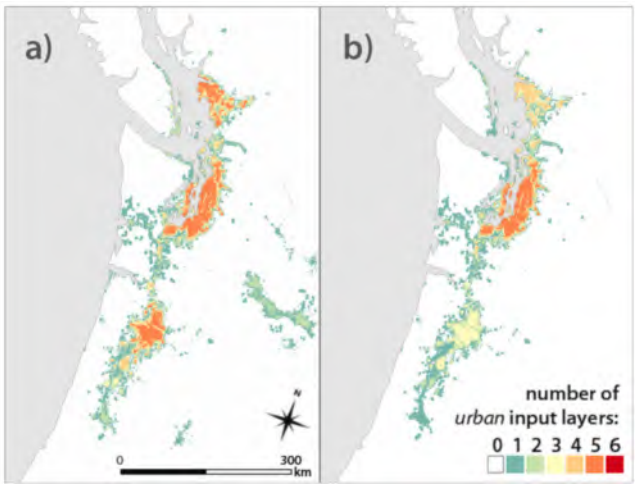


Fig. 5. Cascadia: a) all urban areas classified, b) urban areas belonging to the urban corridor based on our methodology.

Applying the methodology, we find that ambiguous extents are measured for several areas due to the selection of the core cities. In particular, the areas identified in Georg, Blaschke & Taubenböck (2016b) as *Chicago* and *Piedmont Atlantic* were indistinct with respect to which urban area could be regarded as the most prominent one. Therefore, we cross-check the results by using different cities as core. For *Chicago*, we also used Detroit, and for

Piedmont Atlantic, we use both Atlanta and Charlotte. The *Quebec–Windsor* (Southern Ontario) area turns out to split in a Quebec/Montreal and a Toronto section; coherence cannot be established for the total region. Consequently, we examine both parts independently. Similarly, the *Gulf Coast*, as outlined in Georg, Blaschke & Taubenböck (2016b), includes Houston which shows no connection with New Orleans through our approach; thus, we describe the two areas separately. For *Calgary–Edmonton*, we restrict the analysis to Calgary only since Edmonton does not show up as a large enough urban patch through our method. Also, the *I-35 Corridor* cannot be identified as a connected region; rather, the main areas are spatially separate and fragmented patches. Hence, we single out the largest of these (Oklahoma, Dallas–Fort Worth, and Austin–San Antonio) for an approximation. Through these additions and omissions, we obtain areas we preliminarily define as urban corridor candidates. For each of these, we determine length, width, length-to-width ratio and different area extent in relation to the number of 'urban' input layers (Table 2). In total, we identified 19 possible urban corridors with a length between 120 and 1500 km, width between 75 and 630 km, length-to-width ratio between 1.3 and 5 and an area between 3500 and 290,000 km^2.

Table 2. 19 corridor candidates with their respective extents.

Corridor	Corridor candidates				Initial corridors			
	Length	Width	Length-to-width ratio	Area in km^2 (rounded)	Length	Width	Length-to-width ratio	Area in km^2 (rounded)
Arizona Sun Corridor	400	140	2.86	9000	350	80	4.38	12,000
Boswash	1400	330	4.24	214,000	950	275	3.45	142,000
Calgary	120	75	1.6	3500	380	90	4.22	23,000
Cascadia	700	150	4.67	33,000	665	80	8.31	27,000
Colorado Front Range	310	175	1.77	10,000	890	80	11.13	16,000
Florida	730	225	3.24	86,000	615	240	2.56	49,000
Great Lakes – Chicago	1500	470	3.19	259,000	300	125	2.4	34,000
Great Lakes – Detroit	1500	470	3.19	290,000				
Gulf Coast – Houston	550	370	1.49	25,000	1530	160	9.56	52,000
Gulf Coast – New Orleans	700	170	4.12	22,000				
I-35 – Austin–San Antonio	320	110	2.91	13,000	1400	180	7.78	51,000
I-35 – Dallas–Fort Worth	320	250	1.28	20,000				
I-35 – Oklahoma	460	200	2.3	10,000				
North California	760	190	4	160,000	650	140	4.64	24,000
Piedmont Atlantic – Atlanta	950	630	1.51	160,000	1050	160	6.56	73,000
Piedmont Atlantic – Charlotte	950	630	1.51	169,000				
SE Canada – Montreal	600	180	3.33	43,000	1350	230	5.87	108,000
SE Canada – Toronto	800	160	5	54,000				
South California	525	130	4.04	27,000	350	100	3.5	23,000

In a next step, we check if those 19 candidates meet the criteria to be classified 'urban corridors'. The typical dimensions of such corridors have been identified as 400–1200 km in length, 70–200 km in width and exceeding 10,000 km^2 in size, with a high length-to-width ratio (which indicates a linear rather than compact shape) (Georg, Blaschke & Taubenböck 2016b). Calgary as a smallish, round patch fails to fulfill either of these criteria due to the fact that, as mentioned above, our method does not recognize Edmonton as sufficiently large to form a connected region with Calgary. Houston, Dallas and the I-35 in Oklahoma show a very low length-to-width ratio (generally under two) and can therefore be regarded as too compact to be classified as an urban corridor. The same applies to the Arizona Sun Corridor which only exhibits a lengthier shape for one match which incorporates Tucson. We therefore disregard these five clusters (Fig. 6) which are better described as metropolitan areas or urban fields (see also Georg, Blaschke & Taubenböck, 2016a).

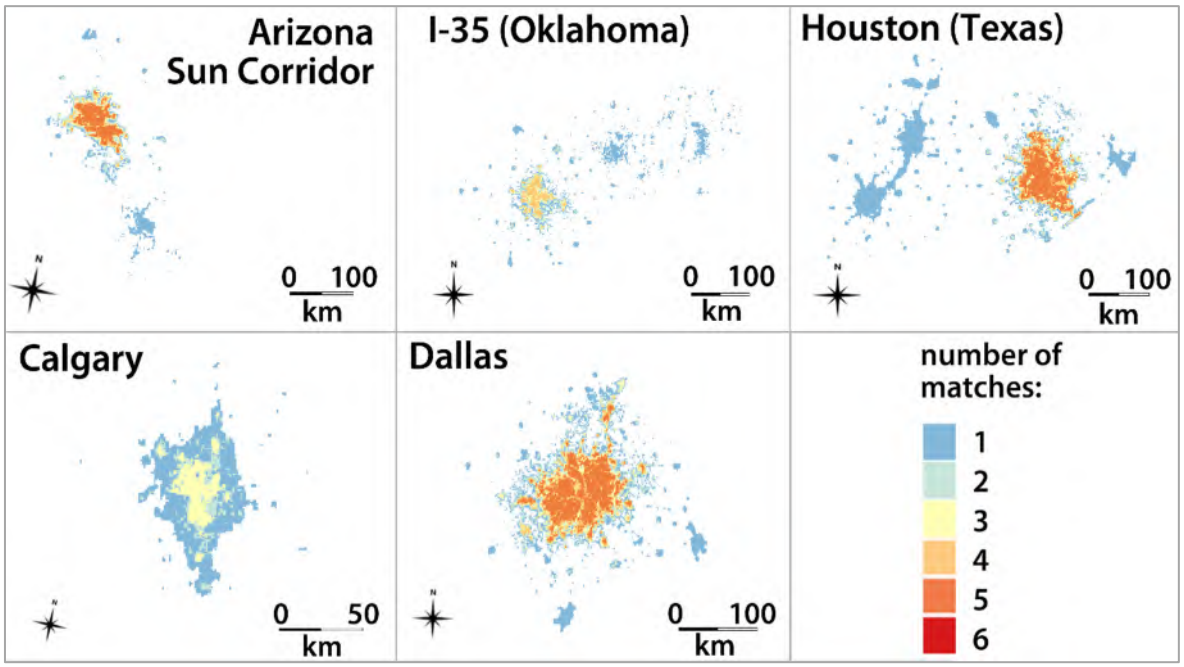


Fig. 6. Omitted corridor candidates. These five areas do not match the criteria to qualify as urban corridors due to their compact shape and/or small size and are therefore excluded from further analysis.

In order to answer our main research question of which urban corridors we can empirically identify and verify, we present the 14 areas in Fig. 7 and Fig. 8. Figure 7 shows the spatial outline of the urban corridors. While some of them show distinct hotspots (such as Detroit, Chicago, Boswash and Florida), others are less distinct. In particular, the Gulf Coast around New Orleans shows a very low agreement in the input datasets with no more than three matches. Fig. 8 shows the decline in length and width with an increasing number of matching input layers as well as the changing length-to-width ratio.

For a clearer visual picture of the varying extents of a sample region, Fig. 9 shows the extents for the Piedmont Atlantic area for one to five (which is the maximum in this case) matching input layers. As mentioned above, there are two possible urban cores for a clustering. Both versions (Atlanta and Charlotte) are very compact for one match, more linear for 3–4 matches, and center around the respective 'seed' city for five matches.

As a trend for both Chicago and Detroit we can establish a more compact shape for a small number of matches, turning lengthier for an increasing number of matching input layers. For one match, the Golden Horseshoe (around Toronto) is connected with both Chicago and Detroit. The same coherence, consequently, is given for Toronto described through one match. Remarkable for Toronto is the massive extent for a single match (due to the influence of the Global Impervious Surface input layer), returning more or less the same area as Detroit and Chicago. However, Toronto shows a significant decrease in size with a higher length-to-width ratio for a larger number of matches and is not detected by all input layers. Because of this, the actual urbanized space is significantly smaller than in the Chicago/Detroit urban corridors (as can be seen in Table 2).

Boswash and South California display a relatively consistent length-to-width ratio, Florida and North California a slight decrease. Quebec-Montreal and – despite a peak in between –

Cascadia show a significant ratio decrease, indicating a stronger focus on one central urban area (here: Montreal and Seattle) with a higher compactness.

Some areas are only identified through a maximum of three (Gulf Coast) or four (Austin-San Antonio, I-35 Oklahoma) matching layers. With the exception of Oklahoma (which has been excluded above), all show a relatively lengthy shape, indicating a possible urban corridor. In Texas, it is noticeable that a connection between Houston and AustinSan Antonio can be seen for one match, while for several matches, this coherence is no longer given and Austin and San Antonio are no longer part of the Houston cluster.

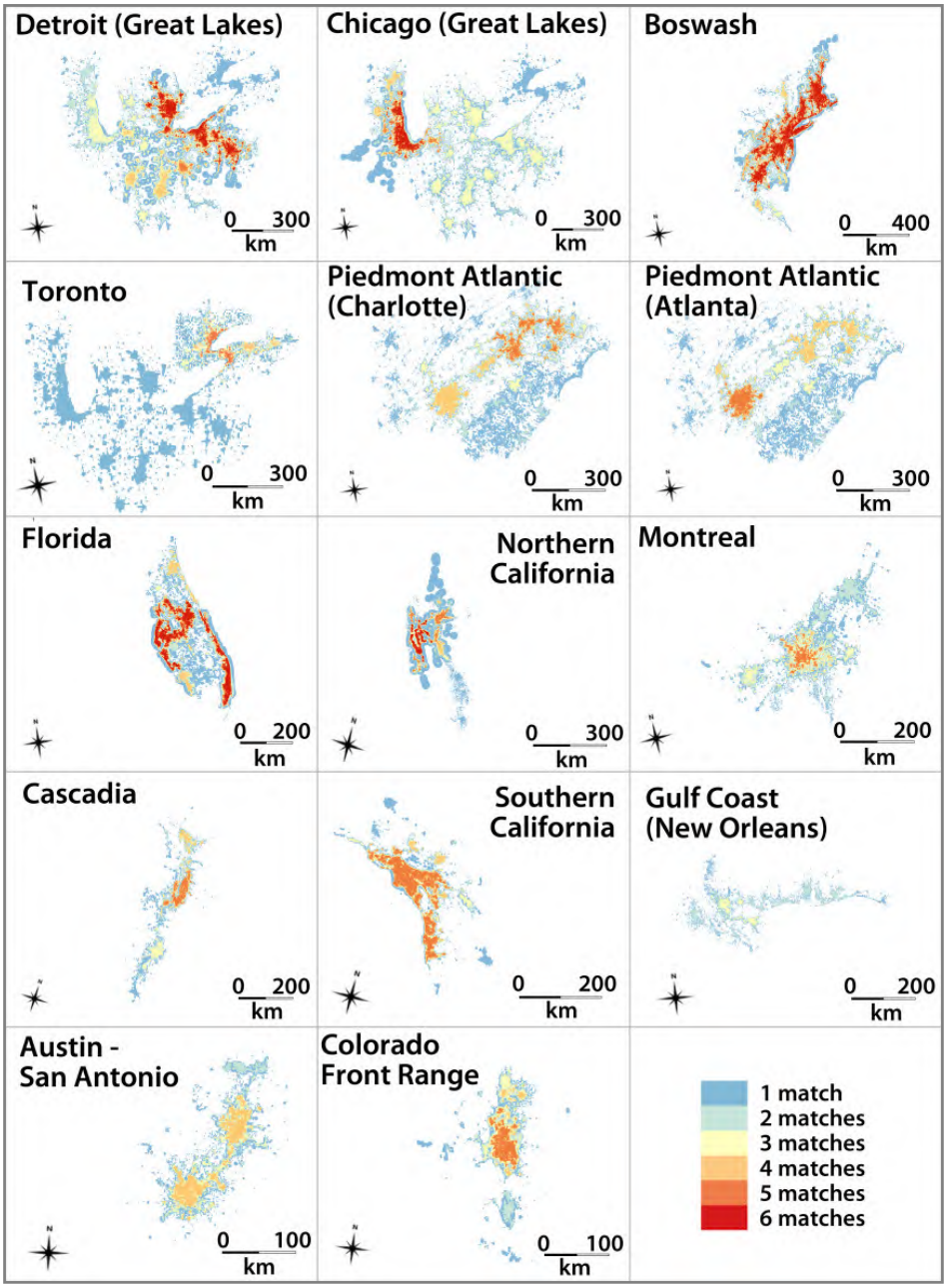


Fig. 7. Urban corridors identified through overlay of input layers.

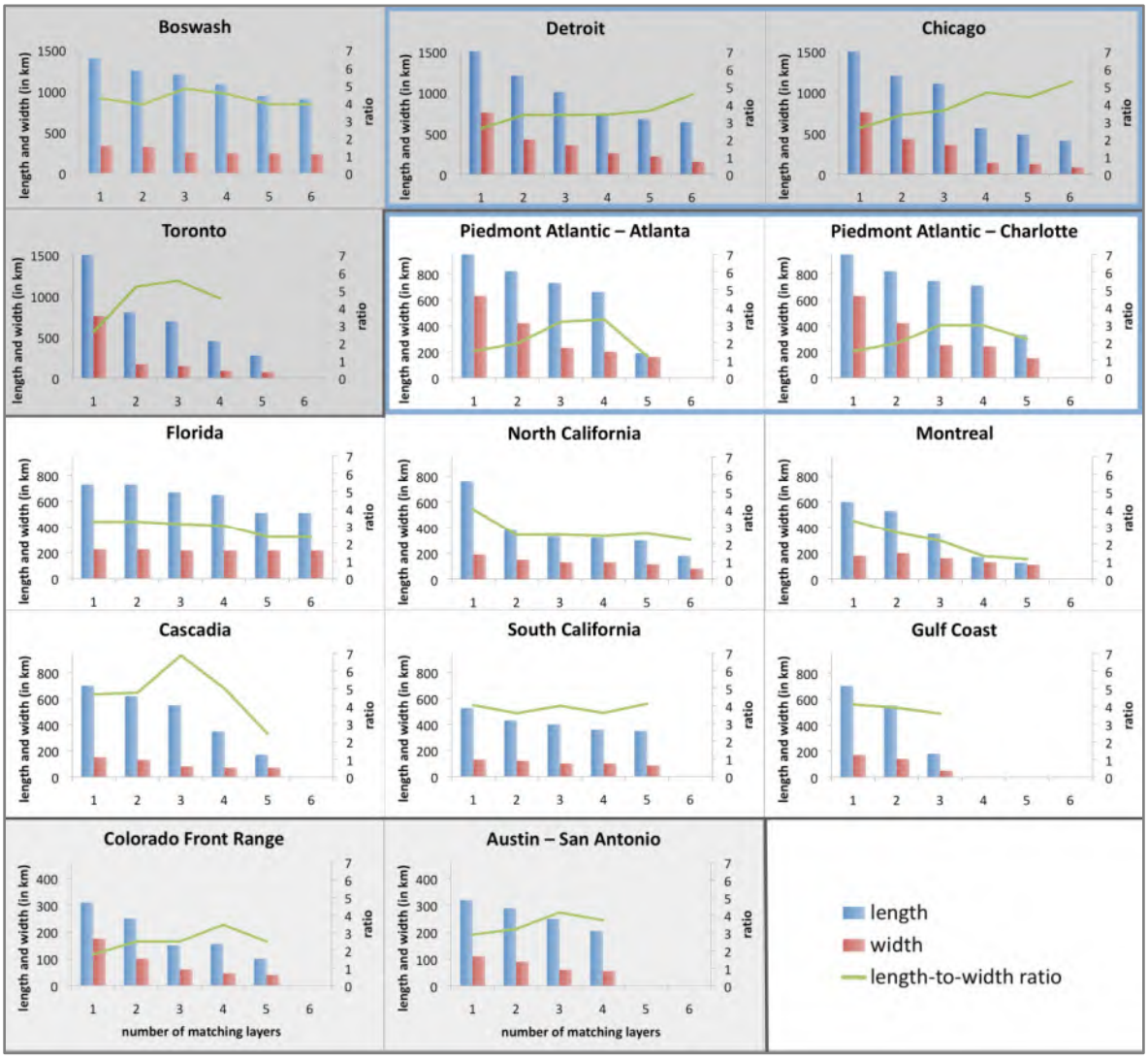


Fig. 8. The 14 corridors identified and their physical characteristics with respect to the number of matching layers. Note the different scales (marked by back ground color).

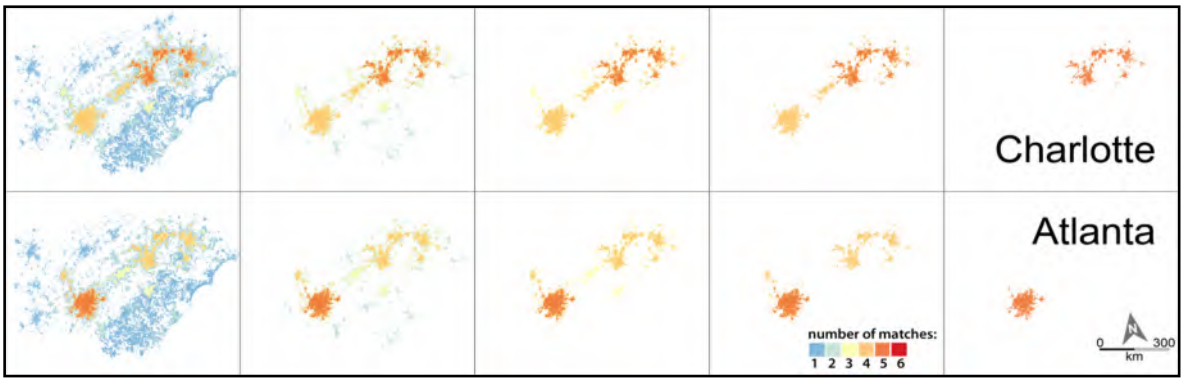


Fig. 9. Piedmont Atlantic: varying extents for Atlanta (top row) and Charlotte (bottom row) as main centers.

Table 3 displays the corridor candidates and the resulting urban corridors with their respective length, width and area measurements for the examples of one and three matching input layers. It is notable that, for one match, all corridors exceed the originally derived dimensions by far, due to the influence of single individual input layers. An exception is the

I-35 (Austin–San Antonio) region because of the subdivision into several smaller areas. In order to reduce the influence of a single input layer, we re-examine the resulting areas for 3 or more matches. That way, we disregard the extents where only one or two (i.e., one third) of the input layers return a positive result. The resulting measurements show three different sets of characteristics:

(1) Urban corridors smaller than the candidate estimate. This mainly includes areas we had to split into smaller individual parts (Montreal/Toronto, the Gulf Coast or the I-35 sections).

(2) Urban corridors moderately larger than the candidate estimate which is the case for a large number of areas. Due to a different approach in measuring urban corridor extent, most areas identified through three overlapping input layers are about two to three times larger than previously measured.

(3) Urban corridors significantly larger than the candidate estimate. This comprises the massive extents of the Great Lakes region with Chicago and Detroit which have been greatly underestimated before. This is mainly due to the night-time lights, impervious surface area and population density input layers which generate a very large coherent patch for the Great Lakes area.

As a last step, we group the urban corridors into four categories to account for the varying extents. The categories are based on the area for three matching input layers (see Fig. 10).

- **Category A** includes the largest urban corridors: Boswash, Detroit, Chicago
- **Category B** contains all areas between 100,000 and 200,000 km^2: Charlotte, Atlanta, Florida
- **Category C** consists of Toronto, Cascadia, Northern and Southern California, Montreal
- **Category D** with the smallest areas (under 25,000 km2): Austin–San Antonio, New Orleans, Colorado Front Range

4. Discussion

In today's fast-grown urban space, actual spatial dimensions do not always directly manifest themselves at first glance. Connectivity may be revealed between seemingly distant places, with the reverse being just as applicable (see for example Taubenböck & Wiesner, 2015). The spatial delimitation of urban areas is a complex task and subject to the under lying concept, data and methods and the related ambiguity in spatial statistics. The significance of urban corridors, in this respect, may not be immediately obvious since they share a number of properties with other conceptual approaches to describe large urban regions such as mega regions. However, their sometimes vast, linear extent is not always easily recognized, meaning that any planning or governance efforts lag behind the dynamic expansion of such an area. Our effort to identify urban corridors can only represent one of many possible approaches, but allows one conceptually and methodologically consistent perspective. In this section we discuss the use of some of the steps of our analysis including its limitations and potential future work.

The novelty of our study is the presentation of urban corridors and their variable extents in a large study area, Northern America. In contrast, most other studies are either case studies on smaller regions (such as individual agglomerations but not necessarily identified as urban corridors) or global urban maps. Our approach uses multiple freely available input layers from different sources and with different content (including built environment with or without vertical structures, night light emissions and population data). All input layers are available on a global scale and therefore our method can be applied on other parts of the world without much modification.

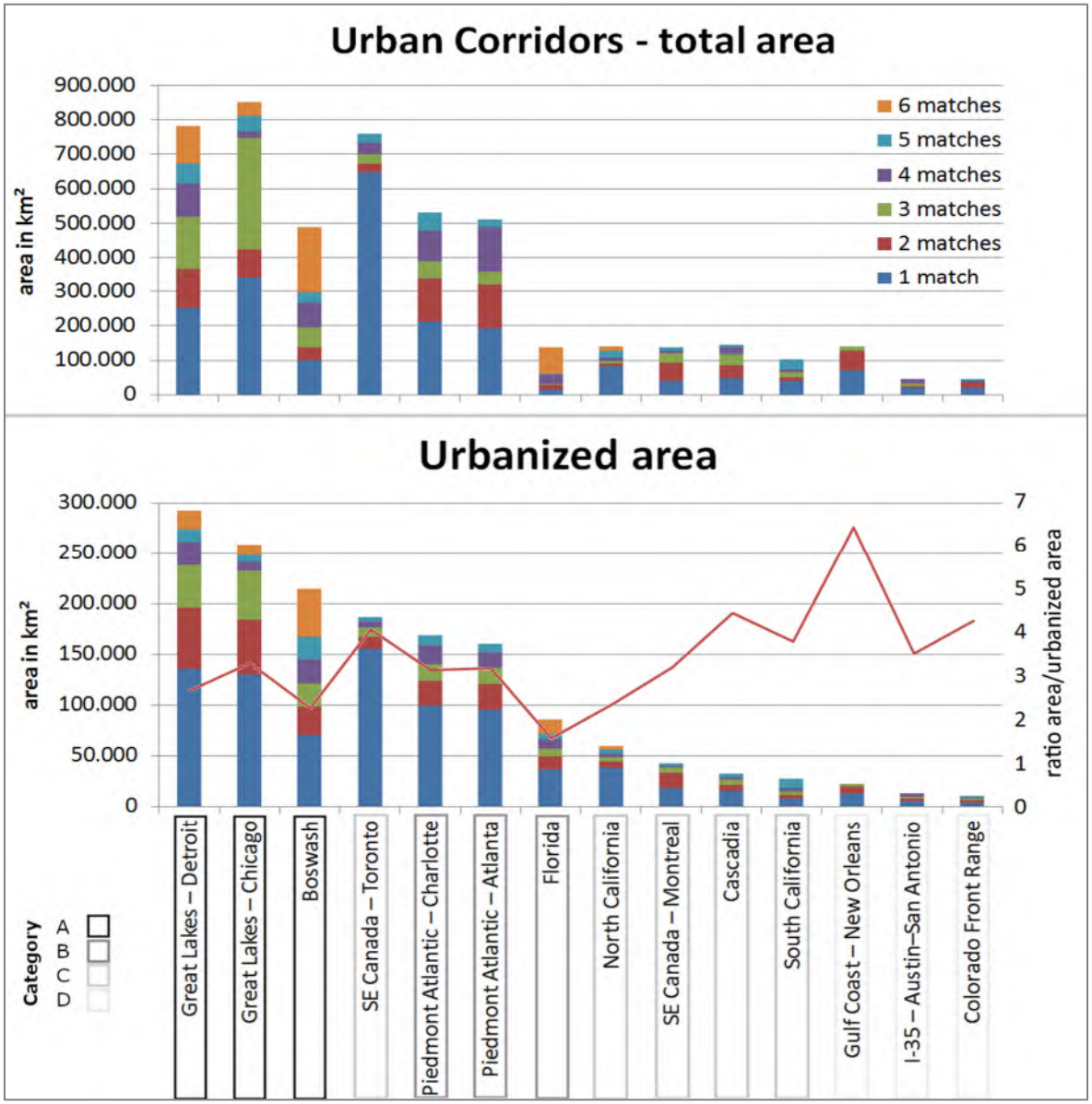


Fig. 10. Urban corridors: total area (top) and urbanized space within that area plus the ratio between total and urbanized space (bottom). Note the different scales. The boxes around the corridor names indicate the respective urban corridor category.

As scattered and small urbanized areas are seen as not decisive in forming an urban corridor, we selected, for each input layer, the largest 0.5 % of all urban patches. While this seems a subjective and arbitrary number, a closer inspection reveals that a percentage less than 0.5

% would have omitted some areas which have been categorized as possible urban corridors (Georg, Blaschke & Taubenböck 2016b), while a larger share would have included a considerable number of relatively small areas which are not considered relevant for our study. Also, using a percentage is, in contrast to an absolute number like for example 'the 25 biggest patches', or a manual selection of only certain areas, transferable across scales and input data. We have not systematically tested the effects of alternating thresholds, but believe that the strength of this approach is in its consistency.

Table 3. Urban corridor dimensions (candidates and results) for one and three matches. Corridor candidates are marked with 'cand.' and highlighted in light grey. No highlight and italics: excluded corridor candidates. Dark grey and bold: final corridors.

Corridor name	Minimum: 1 match					Minimum: 3 matches				
	Length (km)	Width (km)	Length-to-width ratio	Area (km²)	corridor/ candidate area	Length (km)	Width (km)	Length-to-width ratio	Area (km²)	corridor/ candidate area
Arizona Sun Corridor (cand.)	350	80	4.38	12,418						
Arizona Sun Corridor	*400*	*140*	*2.86*	*52,684*	*4.24*	*115*	*65*	*1.77*	*8,519*	*0.69*
Boswash (cand.)	950	275	3.45	142,452						
Boswash	**1400**	**330**	**4.24**	**489,383**	**3.44**	**1200**	**250**	**4.80**	**351,394**	**2.47**
Calgary–Edmonton (cand.)	380	90	4.22	22,760						
Calgary	*120*	*75*	*1.6*	*13,782*	*0.61*	*85*	*35*	*2.43*	*2,210*	*0.10*
Cascadia (cand.)	665	80	8.31	26,665						
Cascadia	**700**	**150**	**4.67**	**145,696**	**5.46**	**550**	**80**	**6.88**	**61,627**	**2.31**
Colorado Front Range (cand.)	890	80	11.13	16,271	0.11					
Colorado Front Range	**310**	**175**	**1.77**	**44,166**	**2.71**	**150**	**60**	**2.50**	**9,779**	**0.60**
Florida (cand.)	615	240	2.56	48,973	1.11					
Florida	**730**	**225**	**3.24**	**136,149**	**2.78**	**670**	**215**	**3.12**	**108,984**	**2.23**
Chicago (cand.)	300	125	2.4	34,125						
Great Lakes – Chicago	**1500**	**750**	**2**	**852,012**	**24.97**	**1100**	**350**	**3.14**	**427,549**	**12.53**
Great Lakes – Detroit	**1500**	**750**	**2**	**783,100**	**22.95**	**1000**	**350**	**2.86**	**415,762**	**12.2**
Gulf Coast (cand.)	1530	160	9.56	51,995						
Gulf Coast – Houston	*550*	*370*	*1.49*	*123,207*	*2.37*	*160*	*130*	*1.23*	*17,123*	*0.33*
Gulf Coast – New Orleans	**700**	**170**	**4.12**	**140,276**	**2.7**	**180**	**50**	**3.60**	**12,791**	**0.25**
I-35 (cand.)	1400	180	7.78	51,031						
I-35 – Austin–San Antonio	**320**	**110**	**2.91**	**46,098**	**0.9**	**250**	**60**	**4.17**	**20,113**	**0.39**
I-35 – Dallas–Fort Worth	*320*	*250*	*1.28*	*64,426*	*1.26*	*250*	*150*	*1.67*	*27,329*	*0.54*
I-35 – Oklahoma	*460*	*200*	*2.3*	*93,893*	*1.84*	*100*	*75*	*1.33*	*6,518*	*0.13*
Northern California (cand.)	650	140	4.64	23,854						
Northern California	**760**	**190**	**4**	**139,380**	**5.84**	**330**	**130**	**2.54**	**49,541**	**2.08**
Piedmont Atlantic (cand.)	1050	160	6.56	73,394						
Piedmont Atlantic – Atlanta	**950**	**630**	**1.51**	**511,060**	**6.96**	**730**	**230**	**3.17**	**191,937**	**2.62**
Piedmont Atlantic – Charlotte	**950**	**630**	**1.51**	**531,024**	**7.24**	**745**	**250**	**2.98**	**194,558**	**2.65**
Québec–Windsor (Southern Ontario) (cand.)	1350	230	5.87	108,071						
SE Canada – Montreal	**600**	**180**	**3.33**	**136,771**	**1.27**	**350**	**160**	**2.19**	**43,208**	**0.40**
SE Canada – Toronto	**1500**	**750**	**2**	**758,631**	**7.02**	**430**	**250**	**1.72**	**85,381**	**0.79**
Southern California (cand.)	350	100	3.5	23,481						
Southern California	**525**	**130**	**4.04**	**102,951**	**4.38**	**400**	**100**	**4.00**	**52,126**	**2.22**

All input layers used here are excerpts from global datasets, allowing for a worldwide application of the presented concept and applied method. These layers are from a variety of sources and with different foci: built environment, population density and economic activity.

We are aware that these show some limitations (Klotz et al., 2016). This is particularly obvious for the GlobCover dataset which underestimates urban areas and only identifies five of the urban corridors examined. Also, night-time imagery is only an indirect measure of urban areas, quantifying human economic activity rather than built-up extent (Florida et al., 2008). Because of a blooming effect and sensor proper ties, DMSP-OLS imagery tends to overestimate urban areas (Henderson et al., 2003). Many studies use DMSP-OLS imagery, but lately, the newer VIIRS (Visible Infrared Imaging Radiometer Suite) sensor has proven superior in both resolution (spatial, radiometric and spectral) (Elvidge et al., 2013) and accuracy (Shi et al., 2014). However, since it has only been available since 2012, we utilized DMSP-OLS imagery for a better temporal match with the remaining input layers.

All input data were aggregated to a 1 km^2 grid. This may seem quite coarse, particularly when considering that there is a variety of high-resolution, often free imagery available. However, for large spatial ex tents such as urban corridors, we believe that a higher resolution is not required for our purposes. A 1 km^2 grid is also used by transnational programs such as the European INSPIRE initiative (INSPIRE, 2014) or Hansen et al. (2000). Furthermore, when using densities (such as road or rail densities as used by Georg et al., 2018), data aggregation is necessary. As Wei et al. (2017a) describe, this aggregation does not necessarily imply less appropriate outcomes. It has to be kept in mind that spatial statistics are hardly ever unambiguous and results depend on input data and parameters used.

While the analysis strives to provide objectively measurable outlines of urban corridors, there is some bias in the use of thresholds and buffers. The use of thresholds has its weaknesses since small adjustments can lead to obvious changes in the outcome (Jing et al., 2015). As can be the case with thresholds, a global applicability is often not given due to economic and physical differences (Small et al., 2005). To avoid using different thresholds for different areas – a method used, for example, by Henderson et al. (2003) –, we applied a buffer (a technique frequently used in landscape metrics, e.g. Schneider & Woodcock, 2008 or Seto & Fragkias, 2005) around each area of interest to establish a coherence that way.

A major criterion for describing urban corridors is their length-to-width-ratio which, on a global scale, was suggested to range from four to ten (Georg, Blaschke & Taubenböck 2016b). Our method reveals that for only one match, the linear character of an urban corridor is less prevalent than for more matches (see Table 2 and 3).

The dimensions with which these areas change in relation to the number of matching input layers vary significantly and the resulting urban corridors generally exceed the dimensions previously suggested based on less input data (Georg, Blaschke & Taubenböck 2016b): While Florida, largely enclosed by water, only shows a relatively small change in length and width, the Golden Horseshoe with Toronto, on the other extreme, covers a massive area for one match (around 1500 km in length) but is reduced to a rather small-sized urban area less about a fifth of the length for a maximum number of matches. The Great Lakes megaregion with Chicago or Detroit as main center shows a much larger area through our approach than previously identified. However, because of the interspersion of lakes, the length and width measurements are prone to subjectivity.

Our method describes the fuzzy concept of ‘urban corridors’ in a traceable and thus spatially objective way. With our fuzzy view of urban corridors, we argue that this fits the concept and that it contributes to wards complementing existing classifications. This serves to go beyond a dichotomous delimitation towards a definition where a classification as ‘part of an urban region’ is probable but not necessarily unambiguous. The extents derived through this can serve as recommendations for decision-making in spatial planning or urban administration.

We acknowledge that this approach is following a ‘space of place’ logic (Castells, 2009). This is of relevance as these areas define the places where living and working is established. However, this physicalism does not account for the ‘spaces of flows' (Castells, 2009). Thus, some areas integrated in these national or continental to global networks which might not have the related physicalism cannot be identified using these data and approaches. Functional aspects are of importance when describing large urban areas, particularly concepts such as ‘Functional Urban Region (FUR)’. However, these functional connections cannot be measured globally and consistently – this can only be achieved for the spatial patterns. We are aware that, while spatial patterns help ‘describe’ urban areas, they cannot ‘explain’ them. We acknowledge that for a comprehensive understanding of the dynamics behind urban change, economic and other functional data would have to be taken into account.

Our analysis only reflects one point of time, with datasets generated around the same date. While this can provide an accurate description of urban corridors at a given time, it only generates a static picture without acknowledging the development into such a large urban area. The long-term driving forces of corridor growth, their history and projected future growth cannot be explained this way, but a multi-temporal application of the approach is suggested for monitoring development.

5. Summary and Outlook

In ever-expanding urban landscapes, new physical forms of cities and connected cities evolve. However, what constitutes the urban in comparison to the rural is fuzzy and far from any unambiguity (Taubenböck et al., 2022), and the spatial delineation of conceptual terms such as urban corridors or the like is just as challenging due to different conceptual approaches, varying datasets available and variables used or methods applied. Beyond, the delineation usually relies on administrative boundaries or other, fixed extents. The reality, though, is more complex, as urbanization processes have, over time and in many locations, outgrown those fixed boundaries. This work aimed to take into account the underlying fuzziness of a concept such as an ‘urban corridor’, the variability of different input data and a specific thresholding approach resulting in a probability-based, malleable definition for each area investigated here.

This paper examined the extents of large urban areas within the study area covering the connected United States territory plus several hundred km to the north and south, allowing to disregard administrative boundaries. Particular focus was on ‘urban corridors’, i.e. massively large, lengthy and spatially coherent urbanized areas. The input data for our

analysis consist of six globally available datasets (night-time lights, population density and different artificial surfaces layers) – thus allowing for transferability to other parts of the world. For each input layer, we derived an individual extent of each region and combined the individual results for a fuzzy delineation of the area independent of administrative boundaries: the likelihood that an area is part of an urban corridor is a factor of how many input layers define an area as urban. Our results reveal that 14 areas in Northern America show spatial properties of an urban corridor as a specific type of large urban areas.

The method in our research was designed to be easily transferable both in terms of geographic regions as well as datasets used. Future work would therefore include applying this method on a global scale and compare the empirical results to the urban corridor inventory established in a previous study (Georg, Blaschke & Taubenböck, 2016b) which was based on a literature review and perceptions. Regular updates to this inventory will help contribute towards an understanding of global urban dynamics, with the current population trends indicating an in crease in both number and dimensions of urban corridors. Furthermore, the robustness of our approach will need to be verified using different datasets, which would also provide insight into other aspects of urbanization (such as environmental factors like land surface temperatures or air quality, commuting patterns, health aspects). A sensitivity analysis could account for any variations in input data or parameters used. Lastly, a time-series analysis could provide information on the development of urban corridors over the last decades and show differences and similarities in their emergence for different regions as well as help predict future growth of such massive areas. On a more local level, the question is what kind of urban living is being generated within urban corridors – will the living standard increase or does the extended shape lead to more uneven development, particularly along the high-speed transportation infrastructure?

We believe that this research contributes to the understanding of the spatial extent of urban corridors, thus helping stakeholders for governance on different levels of interest. The dynamics within urban environments are a multi-disciplinary challenge – the information gained will be useful for urban governance and planning institutions including transport and utilities. The implications of this research are therefore interdisciplinary and include the prediction of the growth patterns of cities (e.g., hot-spots of urban expansion or growth directions) (see for example Dewan & Yamaguchi, 2009), land use planning, resource management, and inter-urban administration.

Chapter V

Quantitative assessment and comparison of urban patterns in Germany and the United States

Computers, Environment and Urban Systems, 2023, vol. 100, 101920

Andreas Schmitt, Philipp Uth, Ines Standfuß, Bastian Heider, Stefan Siedentop & Hannes Taubenböck

© 2022 Elsevier Ltd. All rights reserved.

Note: The paper has been utilized in this research report in accordance with the reuse rights granted to the authors as outlined at: https://www.elsevier.com/about/policies-and-standards/copyright.

https://doi.org/10.1016/j.compenvurbsys.2022.101920

Received 14 Aug 2022; Accepted 19 Nov 2022; Published 30 Nov 2022

Abstract

In this paper, we present methods to assess the homogeneity of the settlement landscape and to identify settlement clusters in terms of location and size. Due to various input data, methodologies, and spatial concepts, there are only a few international comparative research studies with quantitative, spatial approaches to date. Existing studies are mostly qualitative and empirically hardly replicable. Here, we introduce two novel methodological approaches to describe urban land patterns in a numerically stable, consistent and globally comparable manner. The first method evaluates the multi-scale homogeneity of the settlements and the respective density distributions in a purely non-parametric approach. The second method transfers the settlement patterns into a Gaussian Mixture Model via hierarchical multi-scale clustering in order to describe each cluster by the parameters of a 2D Gaussian distribution. In order to proof the robustness of both approaches, two different reference units are considered for the interpretation of the results: OECD functional urban areas and standardized subsets of 200 km × 200 km around the same central cities. We show that German urban areas are characterized by a higher heterogeneity within a smaller neighborhood with an almost symmetric density distribution in comparison to urban areas in the U.S. Furthermore, German urban areas possess an increasing homogeneity with increasing size, i.e., the larger the urban agglomeration, the more uniform it appears. It is the contrary in the U.S. These findings are substantiated by the analysis of individual city centers. In Germany, many clusters of similar size define the urban region, whereas in the U.S. one large dominating cluster exceeds most neighboring settlements significantly. The study is performed on the Global Urban Footprint Density, which assigns a density value to each urban pixel of about 30 m by 30 m. In this way, the approach is a blueprint to be extended to urban patterns of any spatial unit worldwide.

Keywords

Urban patterns; remote sensing; global urban footprint; built-up density; homogeneity; Gaussian mixture model

1. Introduction

In recent decades, urbanization processes in the global North are characterized by a simultaneity of concentration and deconcentration. Higher economic functions find themselves increasingly concentrated in metropolitan areas (Hall & Pain, 2006) and, at the same time, deconcentration processes lead to the emergence of new suburban centers (Garreau, 1992; Krehl, 2016; Taubenböck et al., 2013) forming polycentric-dispersed, multi-nuclei spatial patterns (Anas, Arnott & Small, 1998; Garcia-Lopez & Muniz, 2010). Formerly dominant core cities with monocentric patterns more and more transform into spatially extended, polycentric urban regions (Angel & Blei, 2016; Romero, Solís & Urena, 2014).

These restructuring processes of metropolitan regions are considered more or less universal in the Western world. And this structural change, driven by global economic processes, results in changing functional, political-institutional, and morphological patterns. Aspects studied include the suburbanization of residents (e.g. Kim, 2007) and workplaces (e.g., Heider & Siedentop, 2020), the decline in settlement density (e.g. Angel et al., 2011), urban sprawl (e.g., Siedentop & Fina, 2012) or the emergence of polycentric-dispersed patterns of companies, employees and built urban masses (Angel & Blei, 2016; Meijers, 2013; Taubenböck et al., 2017).

Within the literature on urban spatial structure, there are debates about the "Americanization" of the European or the "Europeanization" of the American city. It exemplifies the assumption of converging developments of urban regions in the global North (Gordon & Cox, 2012; Häußermann, 1997; Le Gales & Zagrodzki, 2006). Other perspectives, however, emphasize the importance of the national, regional and local conditions, which have led to considerable variance in settlement structure developments (Angel et al., 2011; Bertaud & Richardson, 2004). In this context, reference is made above all to varying degrees of deconcentration and dispersion of jobs (Heider & Siedentop, 2020; Shearmur & Coffey, 2002; Siedentop et al., 2016), urban density (Angel et al., 2011), and differences in the size and location of suburban centers (Bontje & Burdack, 2005; Taubenböck et al., 2013). In Europe, for instance, the development of large office agglomerations in peri-urban suburban locations, so-called '*edge cities*' (Garreau, 1992) has been relatively scarce. New service locations and related concentrations of large buildings have frequently emerged in spatial proximity to historic city centers or in the periphery of core cities (e.g. Bontje & Burdack, 2005; Krehl & Siedentop, 2018; Romero et al., 2014). Moreover, the process of polycentric development in European metropolitan areas has not been driven by market-forces, but was rather influenced by regional planning frameworks aiming at concentrating urban growth in pre-designated subcentres (Shearmur & Alvergne, 2003).

Such assessments of convergent and/or divergent development pat terns are still – primarily – based on rather qualitative, empirically hardly replicable findings. There are only few international comparative research studies with quantitative, spatial approaches (Batty et al., 2004; Hall & Pain, 2006; Heider & Siedentop, 2020; Mahendra & Seto, 2019; Taubenböck et al., 2019). The widespread absence of international comparative research or the predominance of qualitative-descriptive (and mostly single) studies on the development of urban spatial patterns can on the one hand be attributed to the limited availability of appropriate data across countries. Equivalent data in terms of population or jobs often suffer

from inconsistent spatial reference units, challenges regarding reporting requirements, or outdated data. On the other hand, it can be attributed to methodological challenges of scale, units of measurement, thematic variables used or the applied metrics (e.g. Openshaw, 1983; Taubenböck et al., 2019; Bartosiewicz & Marcinczak, 2020). Approaches targeting urban spatial structures vary from supra-regional to local spatial levels, with measurements at administrative city to intra-urban block level and thematic variables such as residents (Garcia-Lopez, 2010), jobs (Giuliano & Small, 1991), companies (Arauzo- Carod & Villadecans-Marsal, 2009), commuter flows (Roca Cladera, Marmolejo Duarte & Moix, 2009) or the built urban landscape (Taubenböck et al., 2017). Furthermore, there is a great variety of different methods applied for the identification of urban centers such as density thresholds (Giuliano & Small, 1991), locally weighted regression (McMillen, 2001), or exploratory spatial data analysis (Arribas-Bel & Sanz-Gracia, 2014) and a cornucopia of metrics for the determination of urban form such as rank-size distributions (Arribas-Bel, Ramos & Sanz-Gracia, 2015), concentration indices (Heider & Siedentop, 2020), density functions (Kim, Yeo & Kwon, 2014), measures of spatial auto correlation (Tsai, 2005), landscape metrics (Fragkias & Seto, 2009) or model-based measurements (Taubenböck et al., 2019). There is no consensus on suitable approaches. Rather it has been shown that results strongly depend on the selected units of analysis or – statistically spoken – the reference unit and the method for the measurement of urban form (e.g. Bartosiewicz & Marcinczak, 2020). Thus, it can be stated that there is a methodological ambiguity as to how urban patterns can be quantified and compared optimally. Against this background, our goal is to present a set of methods that can robustly and comparably measure the settlement and center structure of an urban region given the same input data in a quantitative manner. Spatial structures define to a not inconsiderable extent how we behave, and accordingly they have an influence on many factors such as traffic volume (e.g., Glaeser, 2010), land consumption (e.g., Taubenböck 2021), or economic prosperity (e.g., Bettencourt, 2007). To know which structural types occur where on our planet and to be able to relate them in a comparable way thus is crucial.

2. Conceptualization and research questions

Our basic objective is to consistently measure and compare urban patterns worldwide by robust and reliable spatial approaches. We aim to add to the reviewed literature by introducing two methodological approaches that have not yet been applied in this thematic domain. The approaches shall allow transferability to be ubiquitously applicable for comparative urban research. From a geographic point of view, we like to contribute to the debate of transnational studies between U.S. and European (German) urban patterns.

We consider the U.S. and Germany to be suitable for comparison because both countries differ in many relevant framework conditions (e. g., national policy systems, path-dependent local-regional planning institutions, divergent market dynamics, and cultural preferences), but share similarities in spatial developments. These include the ongoing decentralization of urban functions and the emergence of polycentric dispersed spatial structures. A comparative research strategy can help to reflect the equivalence of universal concepts (such as “polycentricity” or “sprawl”) and to scrutinize more closely the spatially specific

manifestations of universal development factors in times of globalization, digitalization, and urbanization.

One data source that is and will be increasingly and consistently available for large areas or even globally is remote sensing from space. Thus, we rely on a remotely sensed mapping product providing consistent land-cover information on location and density of settlements. We explicitly use consistent data on the built landscape instead of the often applied distributions of jobs, companies or other functions. This information can be seen as stable at least over the period of a few years while other data sets are subject to high temporal variations.

Methodologically, we apply a parameter-free approach of discrete statistics as introduced by Schmitt et al. (2018) and refined by Schmitt et al. (2020). This can be understood as a measure of multi-scale heterogeneity of settlements. In other words, this measure is intended to determine the size of the patch needed to cover the complete heterogeneity of existing settlement structures for a study site (in our case a (functional) urban area). One can imagine this as the distance one citizen has to travel in order to reach the full spectrum of existent structural types such as central business district, large housing estates, low density suburban structures or parks. This interpretation presumes that the built-up density provided by the built-up density layer directly refers to the urban morphology, e.g., low density means sub urban living whereas high density indicates the central business district. In fact, maximum values in the GUF-DenS - used in this study - only denote that the land surface is impervious over the complete area of interest, which can also point to large industrial zones or airports.

Secondly, we develop a hierarchical multi-scale clustering approach based on the least square adjustment of point clouds (commonly used in geodetic network adjustment) which results in a so-called Gaussian Mixture Model (GMM) known from continuous, parametric statistics. With it, a varying number of centers with an individual extension and orientation can be identified in the respective metropolitan areas. Number and size of the clusters are variables to define the type of urban patterns from poly- to monocentric regions. As both approaches utilize the image pyramid technology, they do not depend on external reference units. The reference unit is adapted gradually by the patch size according to the pyramid level. In contrast to that, the interpretation of the outcomes requires the introduction of comparable reference units for which the derived information is aggregated. We decided to rely on two different definitions: functional areas of a larger city and for the same central city subsets of a standardized extent. Based on these considerations, we aim to answer the following research questions:

1. Does the non-parametric multi-scale heterogeneity approach deliver robust and comparable results independent of the reference unit?

2. Does the parametric hierarchical clustering approach deliver robust and comparable results independent of the reference unit?

3. Do these approaches reveal distinct differences between urban pat terns in the United States of America and Germany?

The remainder of this paper is organized as follows. *Section 3* introduces the data basis, the study sites and the spatial concept. *Section 4* introduces the applied methods in detail. In

Section 5, we apply the parametric and non-parametric methods onto urban patterns of study sites in the U.S. as well as Germany and we present the measured similarities and differences across the regions. This is followed by a discussion in *Section 6*, concluding with a perspective on the implications of the results in *Section 7*.

3. Study design

We rely on globally available classifications of settlement areas and related attributes derived from remote sensing data. The mapping products studied are subsets of the Global Urban Footprint suite: the Global Urban Footprint (GUF) and the GUF Density (GUF-DenS). The GUF captures vertical built-up structures based on TerraSAR-X and TanDEM-X StripMap radar images as binary classification of settlement vs. non-settlement areas with a spacing of 0.4 arcsec (about 12 m) (Esch et al., 2012; Esch et al., 2013). The experimental GUF-DenS layer adds to the settlement areas the built-up density as attribute. Landsat images were classified to derive the percentage of impervious surfaces within the settlement areas for 30 m by 30 m patches (Esch et al., 2018). The GUF layer then was aggregated to the 30 m so that both settlement classification and settlement density were uniform. For settlement pixels, the built-up density is specified as percentage of built-up structures covering a reference pixel. All image subsets are processed in geographic coordinates (latitude and longitude in WGS84) as archived at the German Aerospace Center (DLR) in order to prevent unavoidable re-projection and interpolation errors and to enable seamless image tiling.

We aim at describing the urban land cover patterns in and around selected urban areas and subject them to a geographical comparison. The non-parametric method evaluates the multi-scale homogeneity of the built-up density around settlement pixels. For the scale of the minimum homogeneity and therewith, the maximum heterogeneity, the characteristic density distribution is studied likewise. This results in an estimate which spatial extent is required in order to cover the whole spectrum of sub-urban environments from low to high built-up density. The respective distribution denotes the dominant type of built-up density at this specific scale, which might be a more or less pronounced peak in the histogram. In a second step, the number of settlement clusters is determined by a hierarchical multi-scale clustering that decomposes the settlement pattern into a GMM. The necessary parameter settings are derived from the non-parametric approach before. This parametric approach delivers the number, the individual location, and the size of the settlement clusters in the region of interest.

Comparing cities across different countries is often challenging, because national definitions of cities can strongly vary and may rely on administrative boundaries that do not reflect the functional and eco nomic area of a city, which typically consists of at least one densified urban core and a commuter belt of lower density. An international comparative analysis relying on administrative definitions of cities (e.g. municipalities) might not only be biased due to the modifiable areal unit problem, but also has a high risk that the spatial units of observation do not reflect the real extent of urban settlement systems.

Hence, in order to obtain consistent delineations of functionally autonomous urban areas, the regions of interest are chosen from the OECD dataset of functional urban areas (FUAs). FUAs were jointly developed by the OECD and the European Commission in order to delineate urban areas in a consistent way across countries (Dijkstra, Poelman & Veneri, 2019). The delineation of FUAs basically involves two steps: First, an urban core is defined based on population density, and second, a commuter zone is defined based on commuter flows be tween the urban core and less densely populated local units in its sur rounding area. While FUAs have some analytical limitations since they still consist of local administrative units, their delineation is based on urban theory and is consistent across countries. The FUAs in this collection are visualized in Fig. 1. Additionally, the areas of interest are fixed to a standardized image size of 200 km by 200 km around the central city of the functional areas. Widely overlapping areas like Berlin and Potsdam (e.g.) are restricted to one region of interest. Hence, we get a list of 68 German and 162 U.S. urban areas (see Appendix for their names). For each area (functional and standardized), a homogeneity diagram and a map of the identified clusters is produced. This results in 230 areas with two reference units and two methodologies, i.e. 920 output diagrams. In order to simplify their interpretation, the main findings are summed-up in one diagram per reference unit and methodology: (1) scale of minimum homogeneity, maximum heterogeneity, skewness (3rd moment) of the respective density distribution for the non-parametric approach as well as (2) the number and size of the urban clusters for the parametric approach. These diagrams are supposed to highlight similarities and differences between the studied urban areas in the U.S. and in Germany. They thus gain new insights into the convergence and divergence of urbanization patterns across continental and national boundaries.

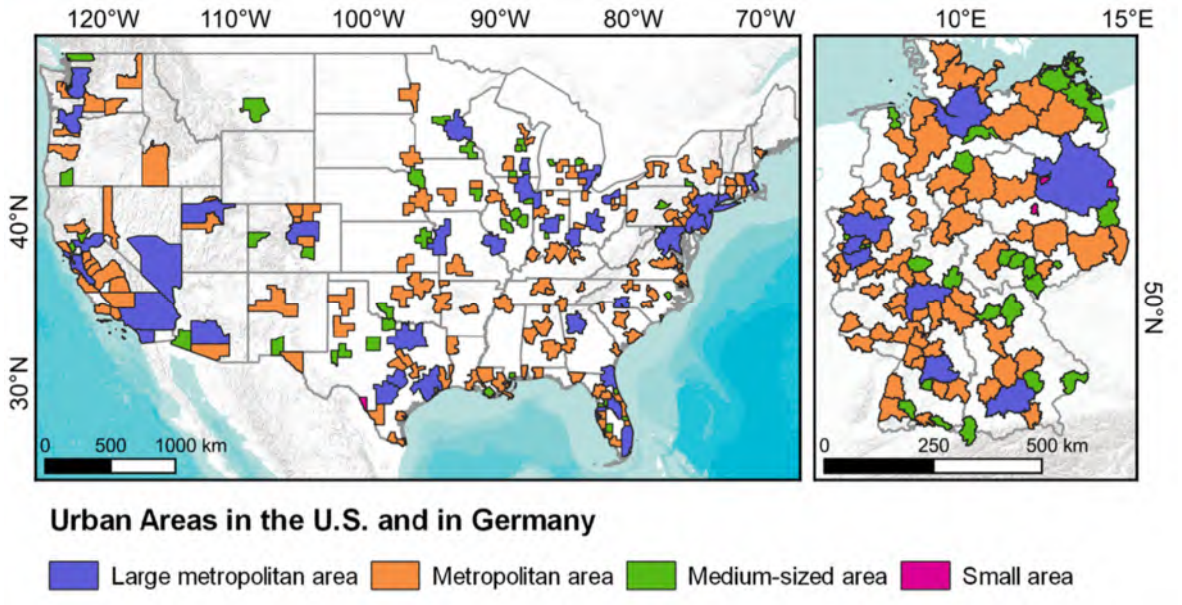


Fig. 1. Map of the OECD functional urban areas in the U.S. and Germany.

4. Methodology

In this section, we explain the design of the two methodological approaches for measuring urban patterns, i.e., (1) a non-parametric from discrete, model-free statistics and (2) a

parametric approach from continuous, model-based statistics. Furthermore, we introduce the geostatistical evaluation of the results. It is well-known that the scale is playing a key role in the spatial analysis (Schmitt et al., 2018; Schmitt, Wendleder & Hinz, 2014). In the following, we consider variations in the scale by the size of patches around one settlement pixel of interest that is included in the calculation of the desired feature. It is typically referred to as level of the Gaussian image pyramid, cf. Fig. 2 and Fig. 3.

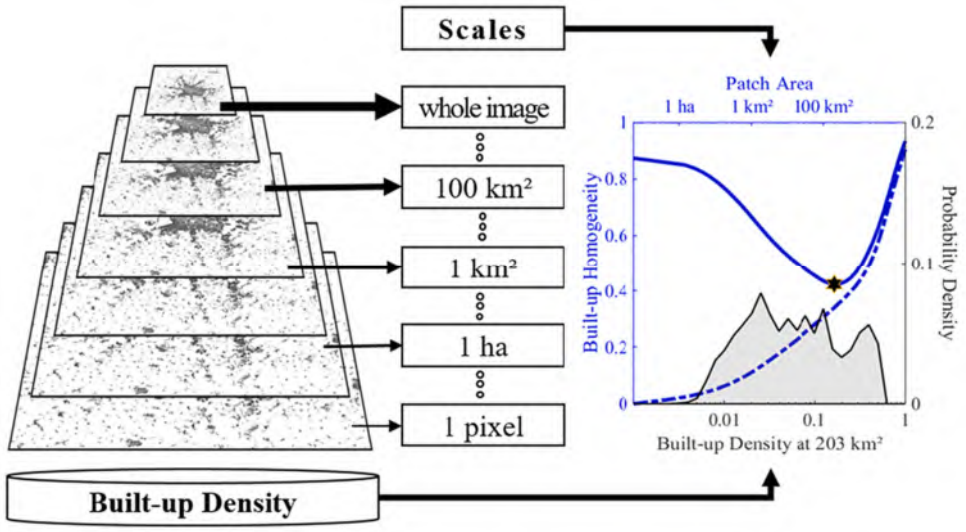


Fig. 2. Workflow of the non-parametric approach: multi-scale built-up homogeneity around settled (solid) and non-settled (dash-dotted) areas as well as the density distribution (grey histogram) at the scale of minimum homogeneity (star) which is denoted numerically in the bottom label by an area of 203 km² in the visualized case. The blue lines show the homogeneity of built-up areas around built-up pixels (solid line) and around non-built-up pixels (dash-dotted line). The latter indicates if natural areas are clearly separated from settlements (low values) or commonly mixed-up with settlement nearby (high values). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

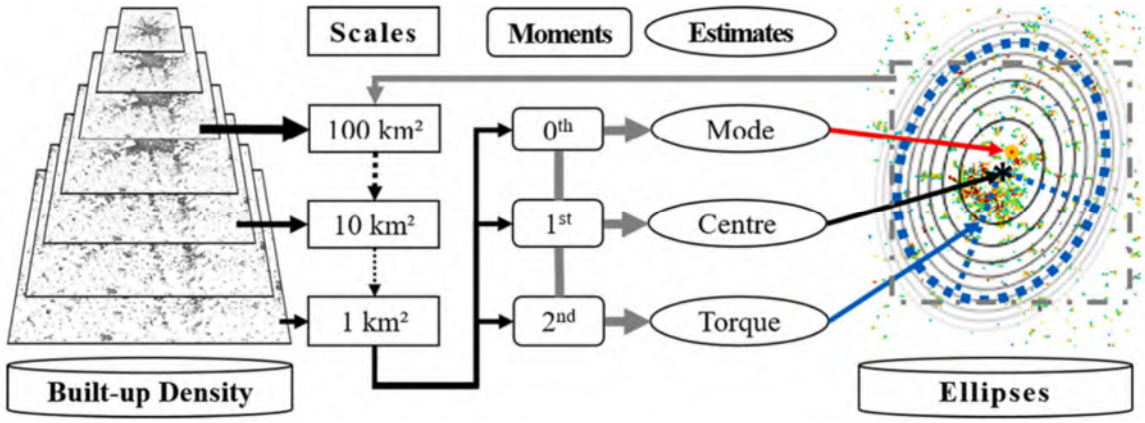


Fig. 3. Workflow of the parametric approach: from coarser to finer scales, the local mode, gravity center and torque are calculated in order to model a two-dimensional Gaussian distribution via least square estimation and to assign the single pixels to a specific cluster. For each image a certain number of ellipses (blue dotted closed line) is derived, which are described by their centers, their major and minor half axes (blue dotted straight lines), and their orientation in the coordinate system. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The finest scale (smallest patch) of the pyramid starts with one isolated pixel and increases by factor √4 2 (in our implementation) until the coarsest scale (largest patch) is reached, which resembles almost the whole image. The respective patch area in square kilometers can be calculated by the help of the ground sampling distance which slightly varies from image to image according to the latitude of the city in the geographic reference system used.

4.1. Non-parametric approach

The homogeneity of the non-parametric approach represents a further development of the similarity measure once developed for the histogram classification of multi-modal SAR images (Schmitt et al., 2018). It unfolds to the similarity of a considered discrete distribution with an equal distribution (Schmitt et al., 2020; Uth, 2019) and is now extended to arbitrary real numbers in ratio scale in this study. Applied to a discrete probability density function, it can be interpreted as entropy measure, i.e. high entropy means an almost equal distribution. Applied to the density values directly, it resembles a homogeneity measure, i.e. high values stand for a uniform density all over the image and low values for a variety of different density values. Hence, the minimum homogeneity can be seen as scale of the maximum heterogeneity or mathematically spoken: *heterogeneity* = 1 - *homogeneity*. The scale of maximum heterogeneity defines the size of the patch in which the most different built-up density values are observed. A narrow scale of maximum heterogeneity stands for a great variety in a rather small neighborhood, which can be seen as a high-quality living environment. The homogeneity bases on the similarity measure in relation to a uniform distribution that unfolds to

$$H = \frac{1}{n} \bullet \frac{\langle density \rangle^2}{\langle density^2 \rangle} = \frac{1}{\sum m} \bullet \frac{(m*density)^2}{m*density^2}$$

The sum over a certain patch containing n pixels is indicated by the angle brackets $\langle\rangle$. It can be implemented as convolution (denoted by the asterisk *) with the 0th order filter kernel from Fig. 4. The homogeneity has a closed value range [0,1] as it has been shown by Schmitt et al. (2018): it reaches its maximum, when the density is equal all over the image, and drops down, when the density becomes heterogeneous, i.e., the minimum of the homogeneity denotes the scale of maximum het erogeneity, see Fig. 2. The study includes both, the built-up density around settlement pixels (solid line) and, for the sake of completeness, also the built-up density around non-built-up pixels (dash-dotted line). Additionally, the density distribution at the scale of maximum heterogeneity is provided (grey histogram). The skewness of this characteristic distribution will allow us to discriminate urban patterns in Germany from the ones in the U.S. later on.

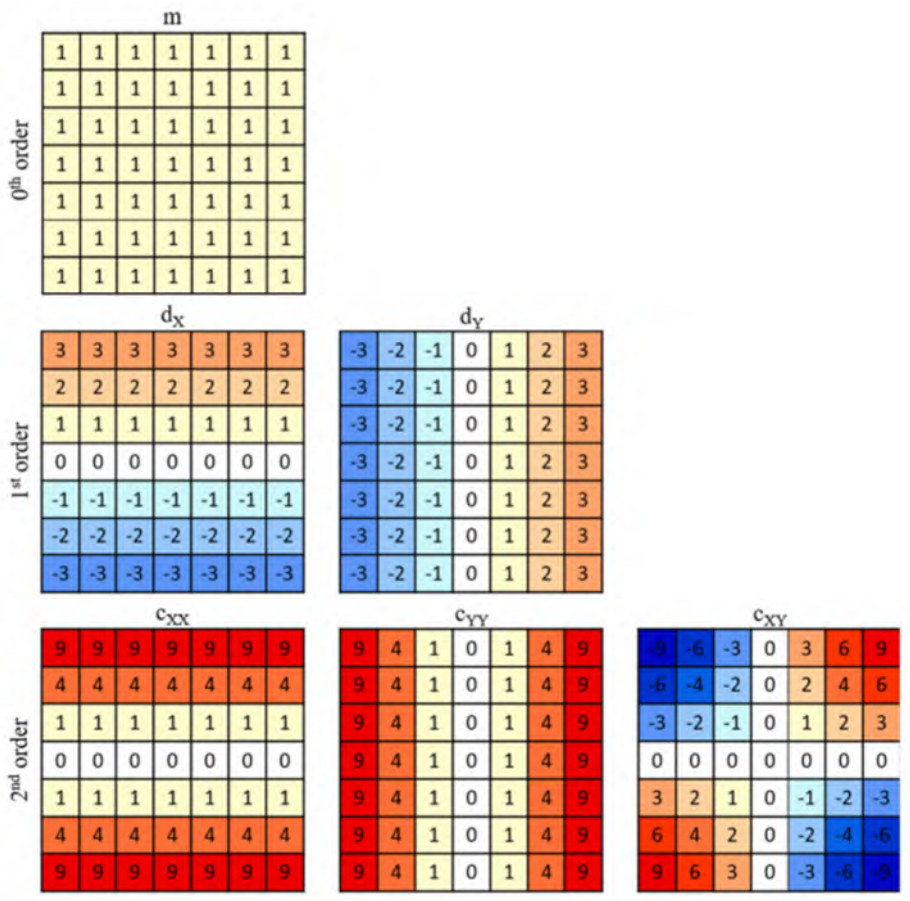


Fig. 4. Exemplary filter kernels of the fourth scale covering a 7 × 7-pixel neighborhood as used in the parametric approach. The 0th order delivers the total sum which is used as normalization factor on the one hand and whose maximum defines the local mode on the other hand. The 1st order calculates the mean shift after normalization by the summed-up weights, i.e., the vector from the center pixel to the gravity center. The 2nd order (normalized by the mean) estimates the elements of the covariance matrix with respect to the center pixel by applying the squared distance from the center in the respective direction as weight. The covariance matrix can then be referred to the identified gravity center by Steiner's law or by choosing the nearest center pixel. Although only the 7 × 7-pixel neighborhood is visualized here, the implementation covers varying filter sizes according to the image pyramid, see Section 6.2.

4.2. Parametric approach

For the distinction of individual settlement clusters, we apply a parametric approach. This approach allows us to estimate the center of clusters by their local modes (coarse, discrete, but robust towards outliers) and gravity centers (fine and continuous via least square estimation) as well as their size by their torques (error ellipses from highly sensitive second order statistics). To do so, we reconstruct the urban landscape as GMM in a hierarchical multi-scale clustering process, which is visualized in Fig. 3. The basic idea of the multi-scale hierarchical clustering refers to the Multi-Scale Multi-Looking for the radio metric enhancement of radar intensity images (Schmitt et al., 2014). This approach also runs through the image pyramid from coarser to finer scales and selects the appropriate scale for the local image content. The selection of the appropriate scale is now replaced by the adaptation of a local Gaussian distribution. While the Multi-scale Multi-Looking only knows about circular filter kernels, its multi-directional advancement – called Schmittlets – also

takes elongated structures in different directions into account (Schmitt, 2016). Instead of searching for Gaussian or TANH-distributions based on a predefined set of parameters as done by the Schmittlets, the approach presented here individually estimates the parameters of the two-dimensional Gaussian Distribution, i.e., mean vector and covariance matrix. Hence, the image is described by a GMM consisting of a varying number of centers with an individual location, extension and orientation. Additionally, each pixel has a certain prob ability with respect to each Gaussian distribution, i.e., a certain degree of belonging to the respective cluster. In this way, each pixel can be attributed to the cluster showing the highest probability similar to the Maximum Likelihood method. Furthermore, the total probability of belonging to any urban cluster instead of being located in a rural environment, i.e., the probability provided by the resulting GMM, is given by the summed-up probabilities of all clusters found in the image subset.

The algorithm starts at the coarsest level and calculates the local mode as rough location of the maximum density after the convolution with the filtering kernel (Fig. 4) of 0th order, which resembles a simple boxcar smoothing filter:

$$mode = max(m*density)$$

The mode is used to identify urban clusters at the respective scale. As soon as the mode of an urban cluster is detected, the gravity center as relative shift from the center pixel (i.e., the mode with maximum built-up density) in both directions is determined. This step bases on the 1st order filtering kernels (Fig. 4), which contain the oriented distance in the respective direction and apply a boxcar smoothing in the orthogonal direction for stability reasons:

$$shift_X = \frac{d_X*density}{m*density}$$

$$shift_Y = \frac{d_Y*density}{m*density}$$

The convolution delivers the refined relative coordinates according to the weighted least square estimation of the detected cluster center. In order to derive the size of the cluster, the elements of the local covariance matrix are calculated based on the 2nd order filtering kernels as approximation of the torque. The second order kernels (Fig. 4) contain the squared distance and again a smoothing component in the case of the variance and the cross-correlation coefficients in both directions in the case of the covariance.

$$variance_{XX} = \frac{c_{XX}*density}{m*density}$$

$$variance_{YY} = \frac{c_{YY}*density}{m*density}$$

$$covariance_{XY} = \frac{c_{XY}*density}{m*density}$$

All operations are performed as simple convolutions in spatial domain using the multi-scale and multi-directional filter bank described above and visualized in Fig. 4. The element-wise division by the convolution with the 0th order kernel leads to a weighted mean with the

density values as weights. The correction of the variance estimation by the reduced redundancy - because the mean value is estimated from the same sample - can be regarded as negligible due to the immense number of measurements in one patch. The Gaussian bell curve N of one urban center then unfolds to

$$N = \begin{bmatrix} mode_X + shift_X \\ mode_Y + shift_Y \end{bmatrix} + Z \bullet \sqrt{\begin{bmatrix} variance_{XX} & covariance_{XY} \\ covariance_{XY} & variance_{YY} \end{bmatrix}},$$

which is the sum of the mean vector (mode plus mean shift) pointing at the gravity center and a standardized normally distributed variable Z stretched by the matrix root of the covariance matrix denoting the spatial extension. The decomposition into several urban centers – maybe even at varying scales – leads to several Gaussian bell curves and therewith, in combination to a GMM.

Centre and torque are the common parameters to describe a two-dimensional Gaussian Distribution with its typical error ellipse. All pixels showing more than 1 ppm probability of belonging to the detected cluster are assigned to this cluster and removed from the further processing. If no more clusters with their mode greater than 10% density are found, the algorithm switches to finer scales until a grid of 5 km by 5 km is reached, i.e., agglomerations covering less than 25 km^2 are not considered in this study in order to simplify the further evaluation and not to become too fragmented. The choice of the minimum size will be justified by the results of the non-parametric analysis: the least average minimum homogeneity is found at a patch size of 26 km^2 in FUAs of Germany (Table 1). Doing so, the algorithm gradually approximates the parameters of the Gaussian distribution in a hierarchical (top-down the image pyramid) and robust (from mode via gravity center to torque) way. For each image, consequently a varying number of centers with varying sizes, extents and orientations is estimated. Furthermore, each pixel can now be assigned to the cluster center showing the highest probability. The sum of all probabilities per pixel denotes the probability of belonging to an urban agglomeration instead of being located in a rural landscape. The summed-up probability is plotted by the help of contour lines with 0.1, 0.01, 0.001, etc. probability in the corresponding maps, see Fig. 5 and Fig. 6 for instance. Overlays of several urban clusters are possible, but of low influence because clusters in close vicinity are quite rare.

Table 1. Average values for the scale (area), minimum homogeneity, and skewness with respect to the functional regions and the standardized patches in Germany and in the U.S.

Functional	Scale [km^2]		Homogeneity [%]		Skewness	
Germany	26	19	58	56	−0.43	−0.28
U.S.	105	49	69	70	−1.59	−1.53
	Mean	Median	Mean	Median	Mean	Median
Standardized	**Scale [km^2]**		**Homogeneity [%]**		**Skewness**	
Germany	44	27	54	55	−0.27	−0.31
U.S.	362	243	61	62	−1.11	−1.00
	Mean	Median	Mean	Median	Mean	Median

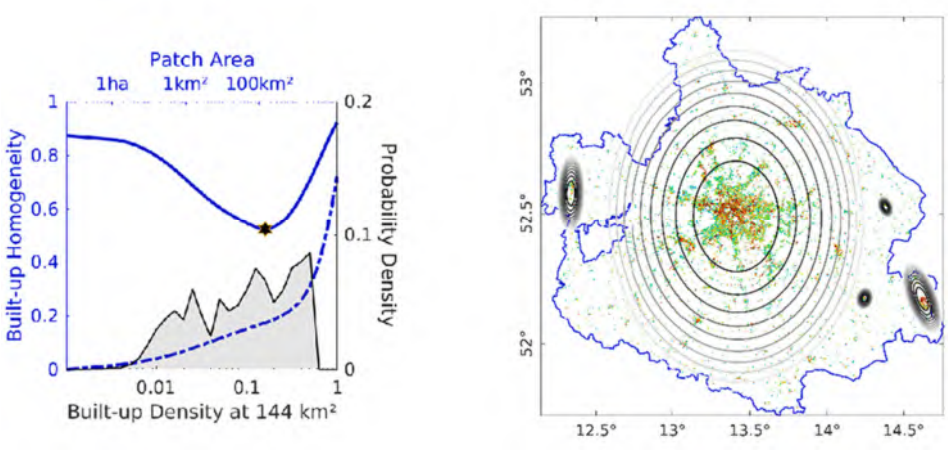


Fig. 5. Berlin – functional.

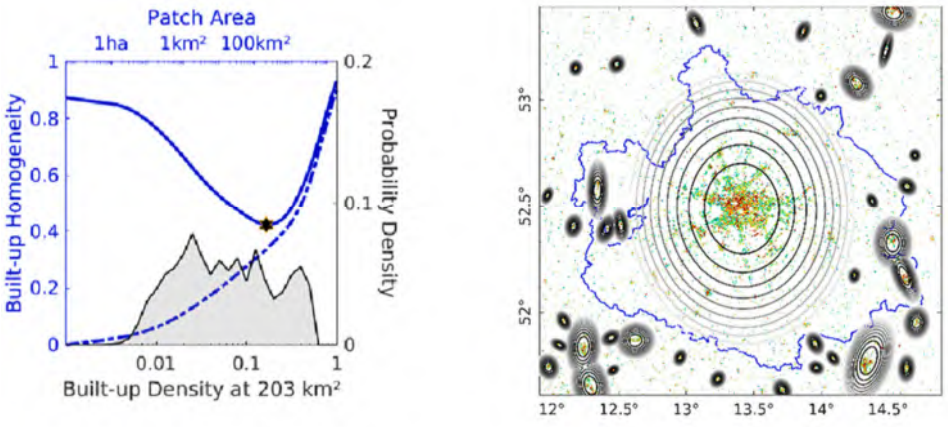


Fig. 6. Berlin – standardized.

4.3. Geographic analysis

The two methodological approaches result in quantifying the spatial patterns and center structures of settlement areas. In the following, we will show how we draw geographical conclusions from this.

The non-parametric approach reports on the homogeneity and the local distribution of the built-up density in relation to the scale indicated by the patch area. The results thus reveal the scale of minimum homogeneity and therewith, the maximum heterogeneity which seems to be characteristic for an urban environment. The indicator shows the patch size at which the built-up density varies the most. At this scale, we find pixels of low, middle, and high built-up density next to each other. With finer or coarser scales, the built-up density becomes more homogeneous either because the patch covers only one homogeneous district (finer scales) or because more or less the complete urban area is considered (coarser scales).

We present the results as a 3D scatter plot (with 2D marginal distributions) of the minimum homogeneity value and the 3rd moment (skewness) of the distribution vs. the patch area including all urban regions under study. We also focus explicitly on the similarities and differences between US and German cities (see Fig. 21 and Fig. 22).

The parametric approach delivers centers of urban agglomerations within the considered urban areas and ellipses defining their orientation and size. It is thus a measure of the mono- or polycentricity of the respective urban areas. A previous study already showed that the flattening and the orientation of the ellipses are effects of the local topography and further specific conditions (e.g., infrastructure like highways and railroads) and thus, not necessarily meaningful for the transnational comparison (Uth, 2019). The center coordinates are also more or less insignificant because they depend on the chosen reference system. Consequently, only the number and size of urban clusters are comparable measures. An area plot illustrates the settlement area of the single ellipses in the sorted list, i.e., the smallest cluster area appears on the left leading to the largest on the right. All ellipse areas sum up to the total settlement area in the region of interest. The monotonous increase al lows for a clear and distinct visualization of all urban regions under study in one single diagram, see Fig. 23 and Fig. 24. The more centers were found, the wider is the extent of the area plot towards the right. The higher the peak, the larger is the dominant cluster in comparison to the others.

5. Results

This section summarizes the results of the algorithms presented above applied to the GUF-DenS. First, the influence of varying reference units for interpretation is studied for selected urban areas 1) in Germany and 2) in the U.S. And then, 3) the results of all German and U.S. cities are joined in order to highlight similarities and differences across these two countries and their urban systems.

5.1. Selected results in Germany for reference unit comparison

The following figures illustrate the results of the non-parametric analysis on the left and the detected centers on the right. The functional area is drawn by the blue polygon and the full sections show the 200 km by 200 km regions in Figs. 6, 8, 10 and 12. The results of the functional and the standardized reference units are grouped together for comparison reasons. We selected Berlin (Fig. 5 and Fig. 6), Munich (Fig. 7 and Fig. 8), Cologne (Fig. 9 and Fig. 10), and Dresden (Fig. 11 and Fig. 12) as illustrative examples. Obviously, the functional areas restrict to one dominant cluster in the center and only few adjacent significantly smaller clusters whereas the standardized regions cover numerous smaller or mid-sized clusters beyond the borders of the functional areas. Regarding Berlin in Fig. 5 and Fig. 6, the metropolitan area of Potsdam (small polygon on the left) was excluded from the functional area, but is contained in the standardized area. The algorithm behaves most robust as it delivers almost the identical results for the overlapping areas in dependent of the chosen patch size except for Cologne. The standardized patch size of the region reveals that the cluster identified in the functional area of Cologne is embedded in a larger cluster covering the Ruhr metropolitan region also. The results of the non-parametric analysis are similar throughout the varying regions of interest though the influence of the adjacent smaller cluster is clearly visible. It expresses in the increasing influence of settlement pixels with lower density and there with, a reduction of the peak at the right. This goes along with a decrease of the absolute skewness.

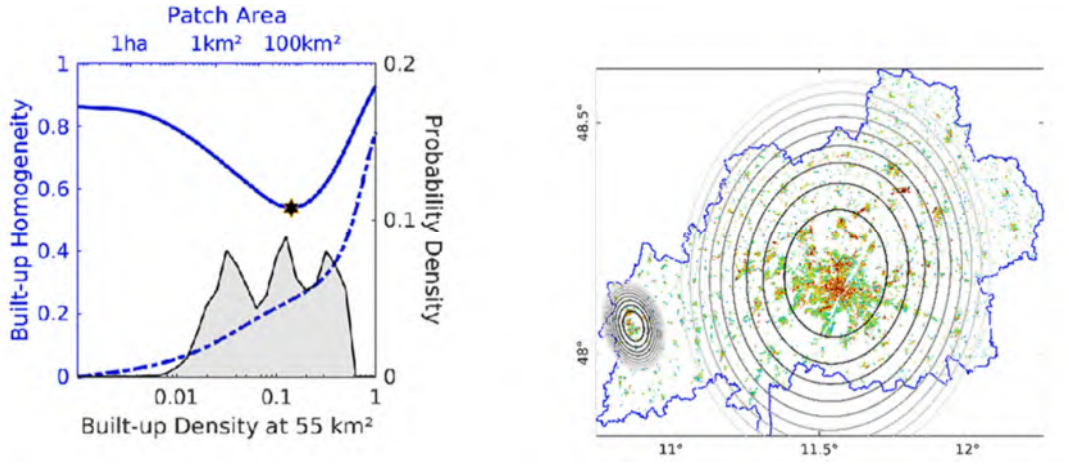


Fig. 7. Munich – functional.

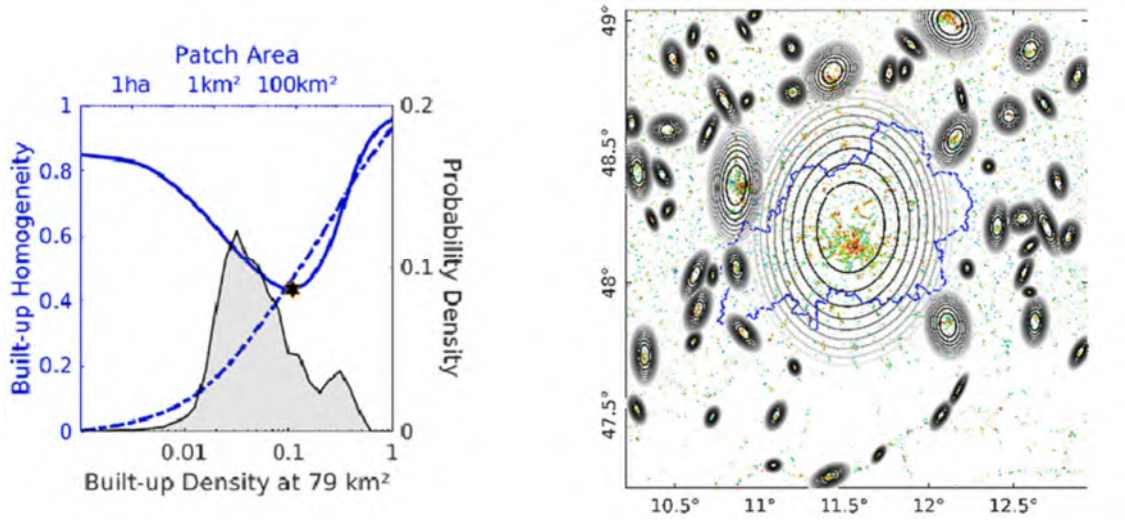


Fig. 8. Munich – standardized.

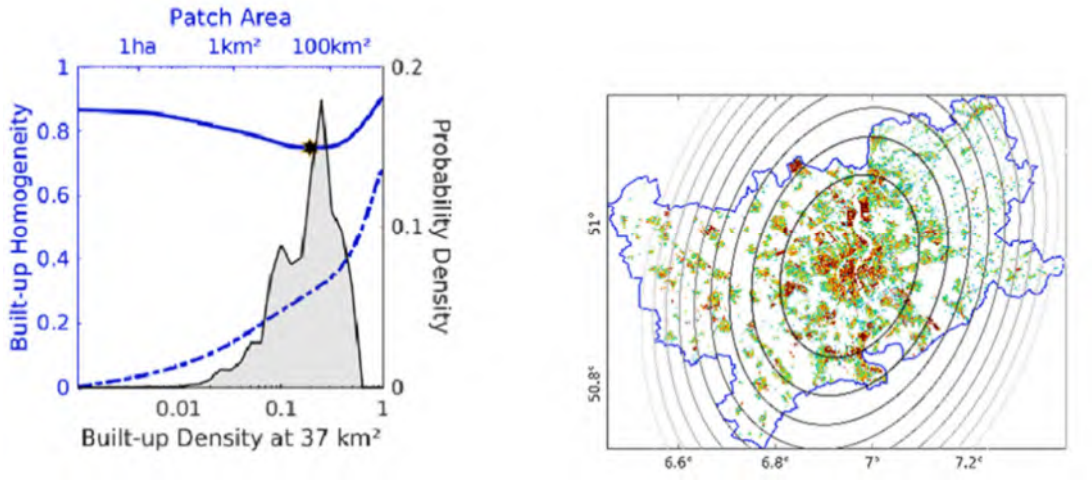


Fig. 9. Cologne – functional.

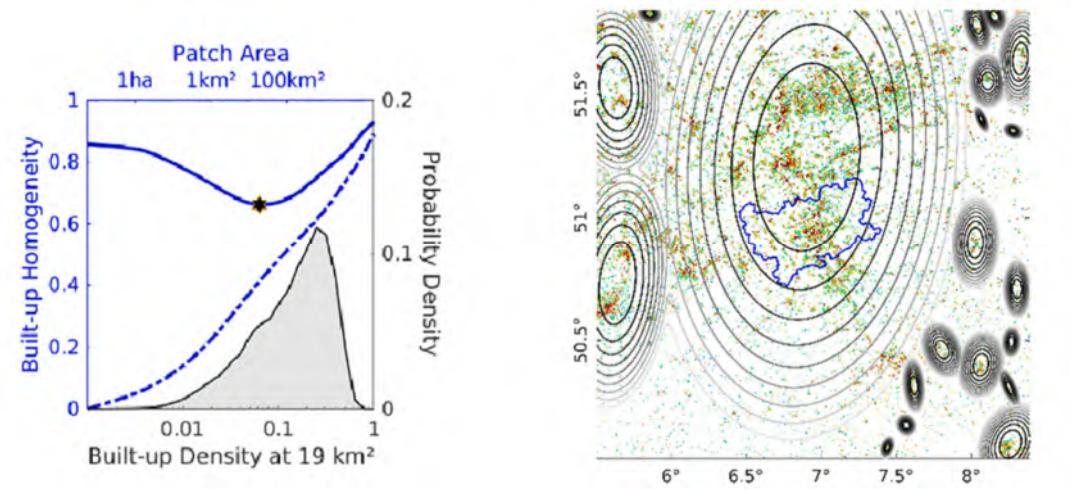


Fig. 10. Cologne – standardized.

5.2. Selected results in the U.S. for reference unit comparison

While for the German urban areas the standardized areas were al ways larger than the functional areas, we find both, smaller and larger functional areas in the U.S. From the sample cities, we selected Los Angeles (Fig. 13 and Fig. 14), Seattle (Fig. 15 and Fig. 16), Providence (Fig. 17 and Fig. 18), and Hartford (Fig. 19 and Fig. 20) as illustrative examples. In the case of Los Angeles, the functional area obviously exceeds the standardized area. Anyways, the shape of the core with two adjacent Gaussian bell curves is identical. All four U.S. cities are characterized by an extremely high, negative skew ness, i.e., a dominance of very dense built-up areas. They show one most dominating cluster and only few neighboring clusters independent of the choice of the reference unit. Hartford and its environment show a pattern similar to the ones observed in Germany with numerous smaller clusters around the functional unit. The functional unit of Providence is embedded in a larger urban cluster similar to the situation regarding Cologne in Germany (Fig. 9 and Fig. 10).

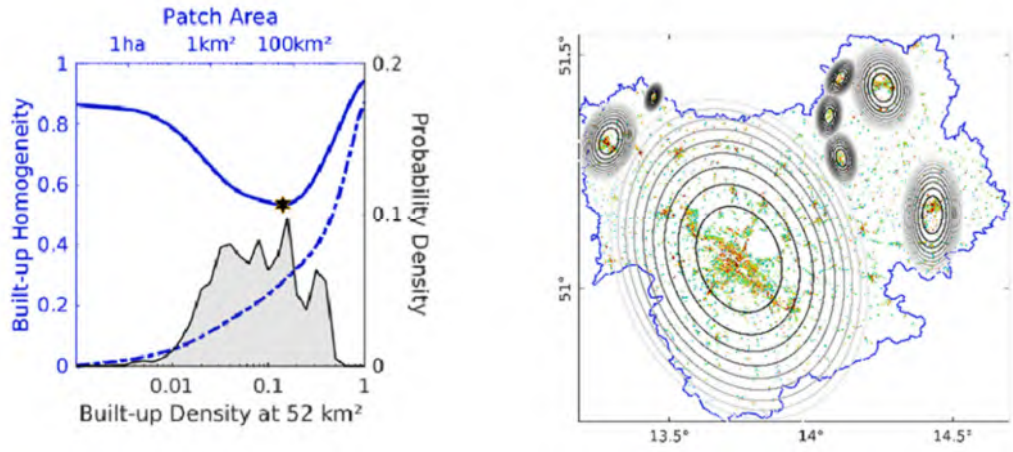


Fig. 11. Dresden – functional.

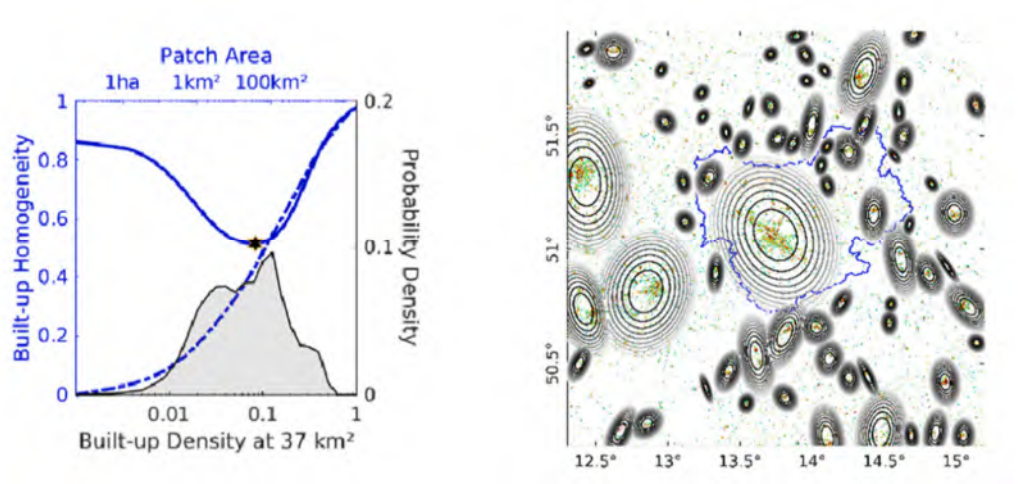


Fig. 12. Dresden – standardized.

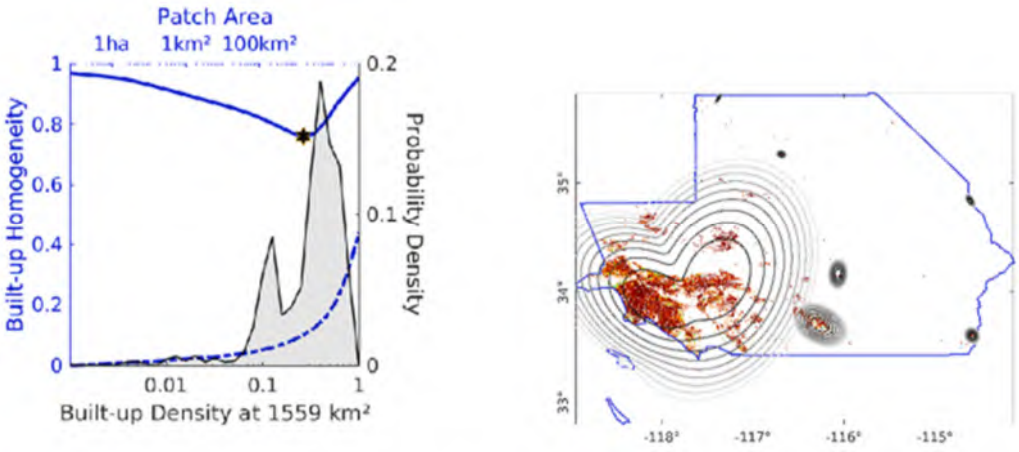


Fig. 13. Los Angeles – functional.

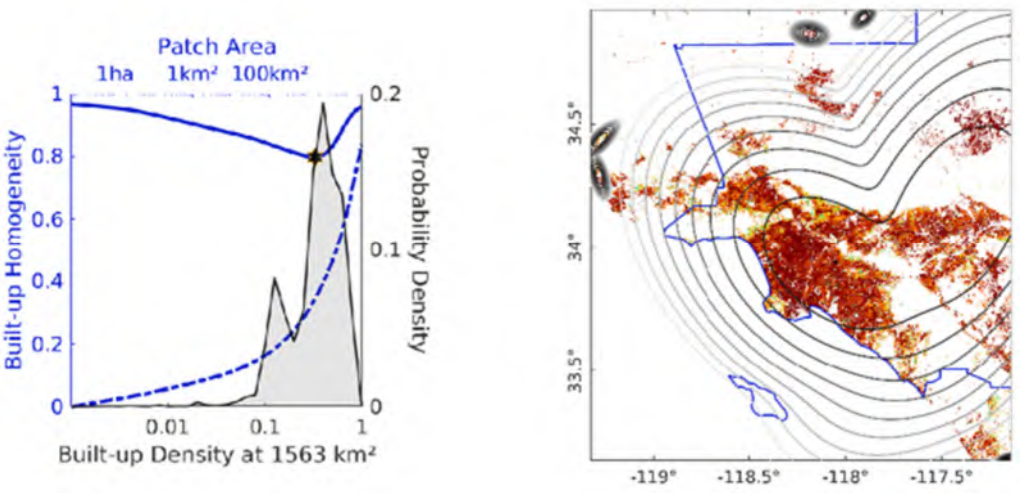


Fig. 14. Los Angeles – standardized.

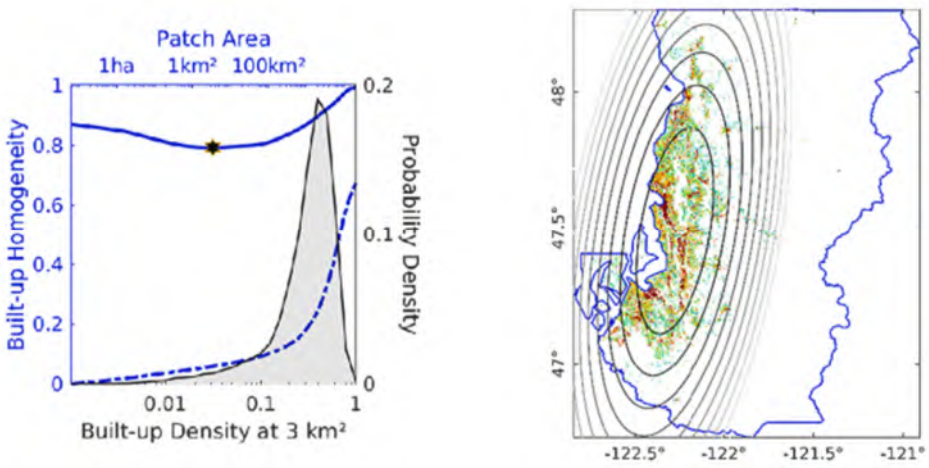


Fig. 15. Seattle – functional.

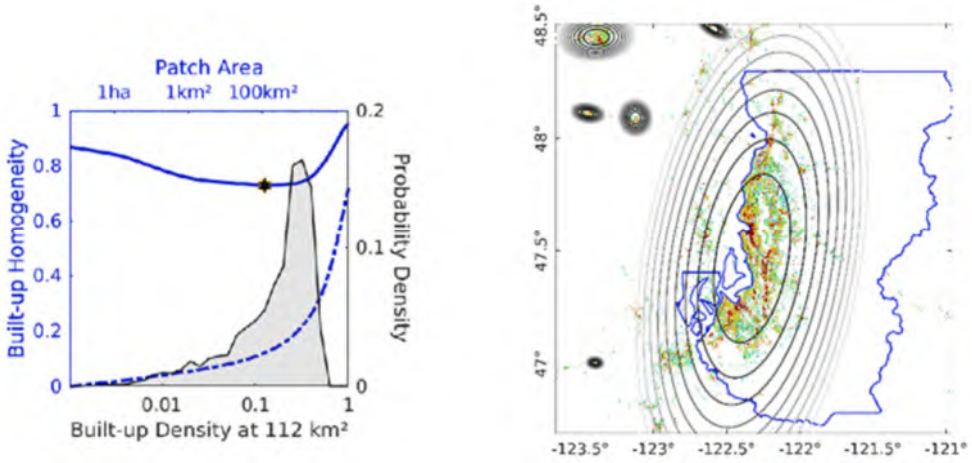


Fig. 16. Seattle – standardized.

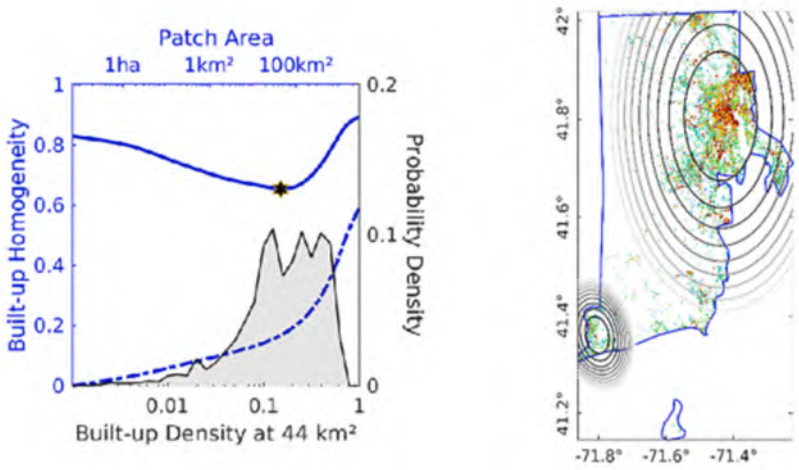


Fig. 17. Providence – functional.

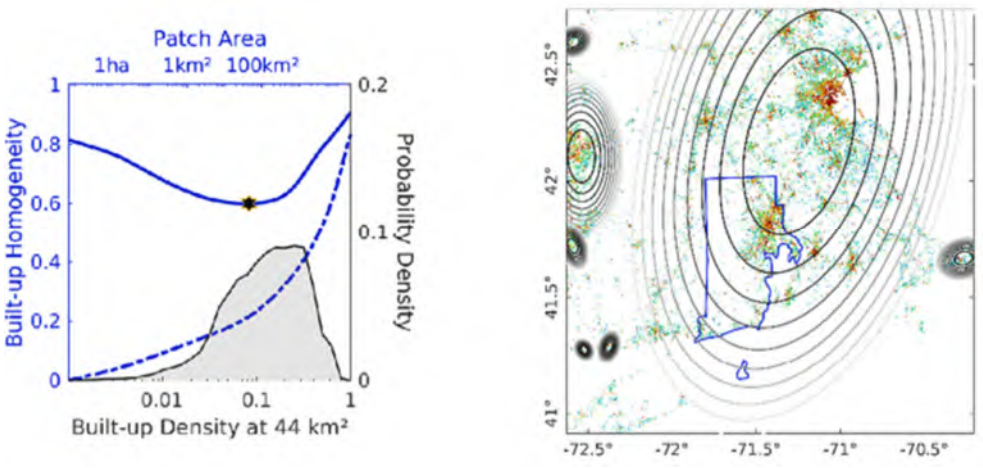


Fig. 18. Providence – standardized.

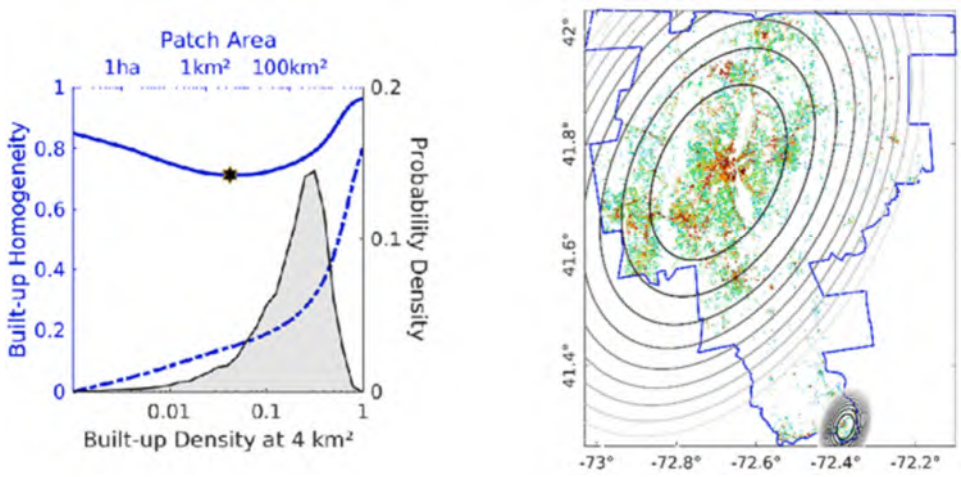


Fig. 19. Hartford – functional.

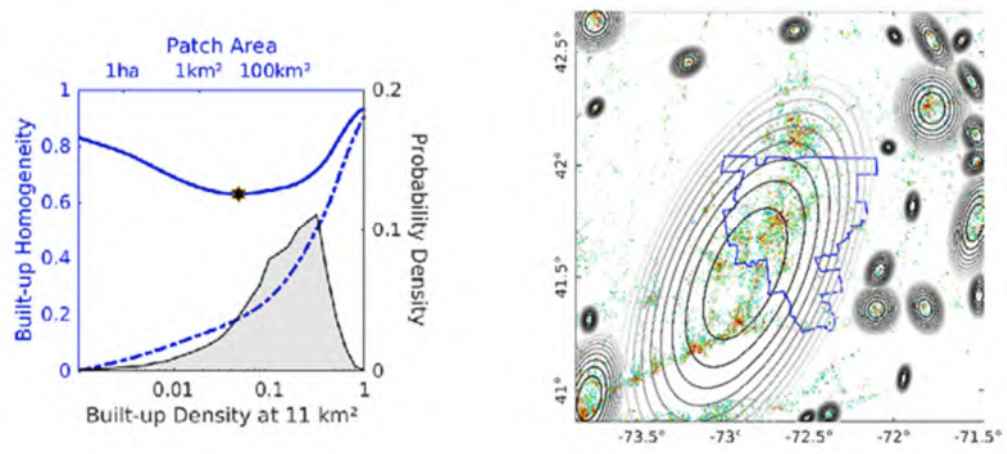


Fig. 20. Hartford – standardized.

5.3. Geographic analysis for intercontinental comparison

To sum up, the figures above indicate that German metropolitan areas are characterized by a steeper minimum of the homogeneity at lower scales than in the U.S. The respective density distribution is almost symmetric while the density distribution of U.S. metropolitan

areas shows a high negative skewness. This indicates the polycentric pattern of urbanization in Germany whereas the U.S. urban system is rather characterized by huge conurbations in larger distances to each other. The parametric approach therefore delivers more numerous, but smaller clusters in Germany than in the U.S. In order to proof these assumptions, the results are joined together in one diagram for the functional metropolitan areas in Fig. 21 and for the standardized patches in Fig. 22.

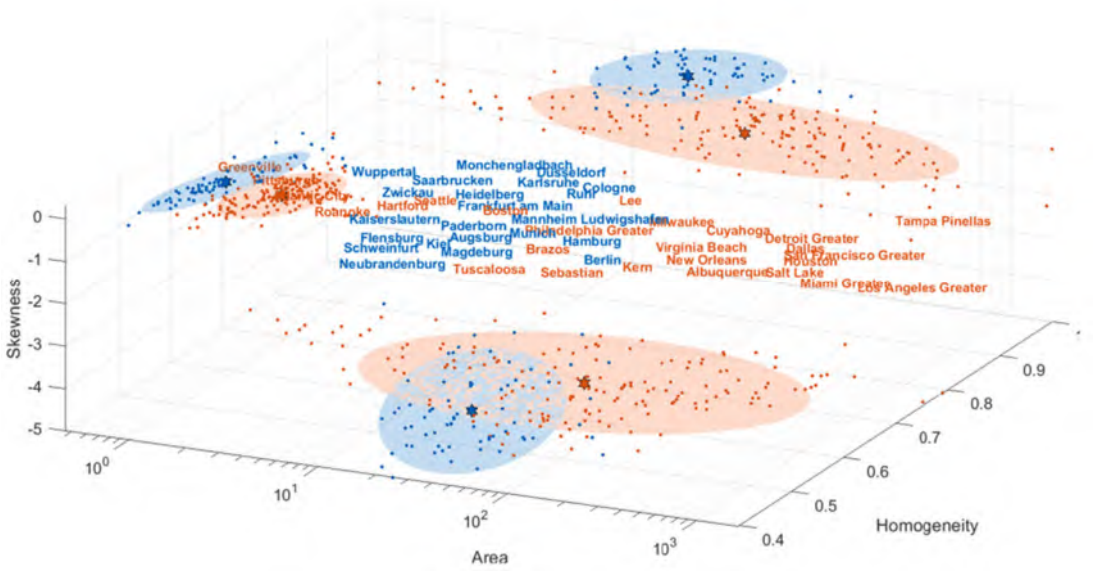


Fig. 21. Homogeneity, scale (area), and skewness of the functional units derived by the non-parametric approach: German metropolitan areas are plotted in Blue, U. S. metropolitan areas in Orange. The stars denote the corresponding mean values. The ellipses visualize the theoretical 90%-quantile in the respective plane, i.e., the marginal 2D normal distribution. The 3D plot shows only the names of notable cities for better readability, but the points in the marginal 2D distributions represent all cities under study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

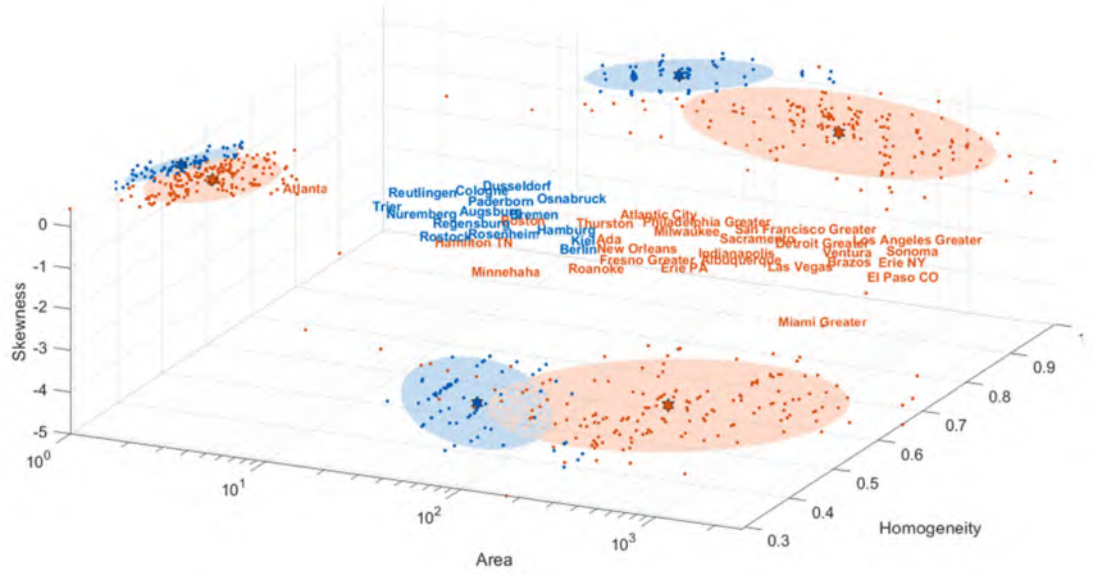


Fig. 22. Homogeneity, scale (area), and skewness of the standardized patches derived by the non-parametric approach: German metropolitan areas are plotted in Blue, U.S. metropolitan areas in Orange. The stars denote the corresponding mean values. The ellipses visualize the theoretical 90%-quantile in the respective plane, i.e., the marginal 2D normal distribution.

The 3D plot shows only the names of notable cities for better readability, but the points in the marginal 2D distributions represent all cities under study. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Based on the functional areas, the different characteristics of German and U.S. metropolitan region become visible. The mean values clearly underline a lower homogeneity at a lower scale in combination with a less skew distribution for German metropolitan regions. The ellipses drawing the 90% quantile overlay partially mainly in the area homogeneity plane. Switching to the standardized regions, the general location and size of the ellipses is similar, but the overlay reduces in such a way, that the German and U.S. metropolitan areas can clearly be distinguished from each other in the area-skewness plane.

The numerical values mean and median are listed in Table 1. The minimum homogeneity is throughout lower in Germany, so is the scale and even the absolute skewness. Though the influence of the reference unit is clearly visible, these conclusions stay robust. Obviously, with larger areas of interest, the scale also increases which is reasonable. The minimum homogeneity slightly reduces with an increasing size of the regions of interest, because more rural environment is captured in the image. The absolute skewness reduces likewise because the influence of high-density areas decreases, which was also apparent in the individual diagrams above.

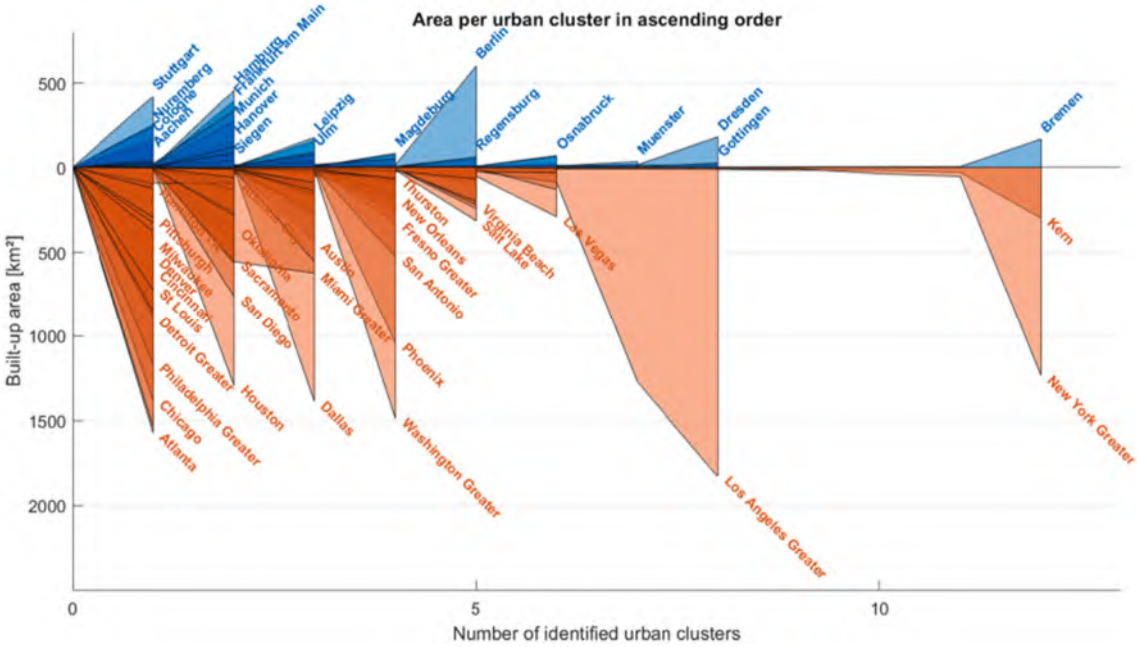


Fig. 23. Number of identified clusters and corresponding area of the ellipses in ascending order from the left to the right for the functional areas. German urban areas are marked in Blue upwards, U.S. urban areas in Orange downwards. The plot shows only the names of notable cities for better readability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The parametric analysis delivers the number of urban centers and the size of the corresponding clusters. For better visualization, the sorted list of clusters by number is presented as area plot in Fig. 23 for the functional areas and in Fig. 24 for the standardized regions. There is only one dominant cluster in one functional area except for Los Angeles

where two clusters of similar size in close vicinity were detected (see Fig. 13 and Fig. 14). The dominant clusters are much larger in the U.S. with almost 2000 km^2 in the case of Los Angeles. The largest dominant clusters in Germany range around 500 km^2. The number of clusters restricts to twelve at maximum in a few cases whereas the vast majority has less than five clusters both in Germany and in the U.S.

The measured values reveal a different result when we switch to standardized areas of interest in Fig. 24. The German samples are typically characterized by up to 80 small-sized clusters. Some sample cities in the Ruhr metropolitan area, however, contain only few, but very large clusters that also reach nearly 2000 km^2 like the dominant clusters in the U.S. The standardized regions in the U.S. still reveal only one dominant cluster with a large extension in most cases. This underlines again a higher diversity in German urban environments.

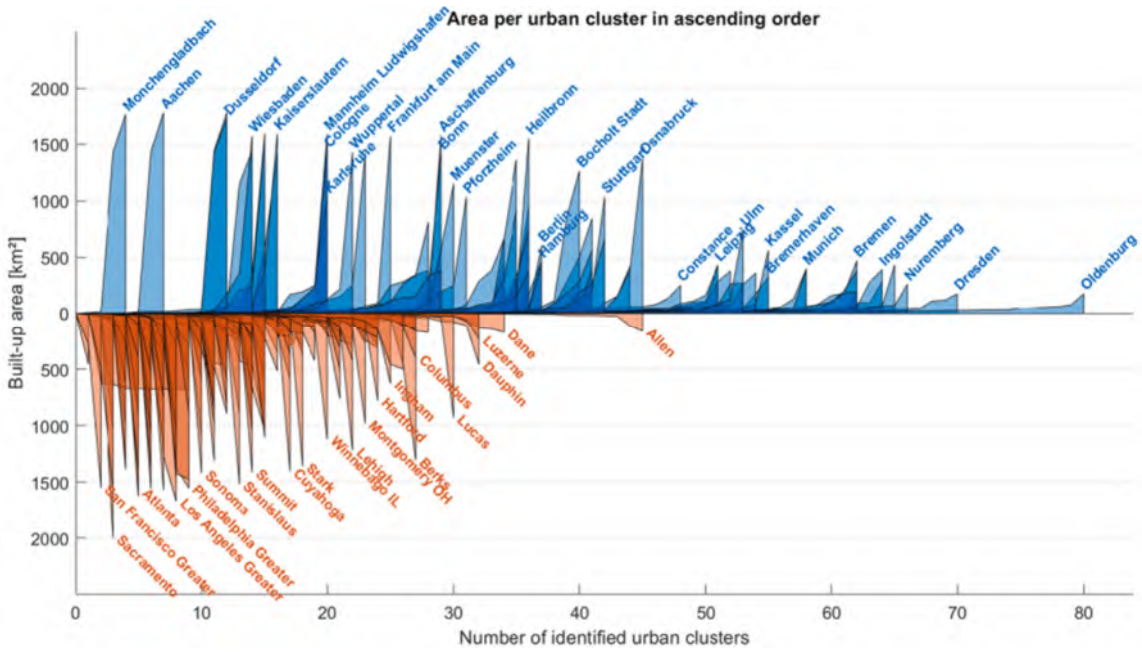


Fig. 24. Number of identified clusters and corresponding area of the ellipses in ascending order from the left to the right for the standardized patches. German urban areas are marked in Blue upwards, U.S. urban areas in Orange downwards. The plot shows only the names of notable cities for better readability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The numerical results of the parametric approach are summed-up in Table 2 and substantiate the above statements. In the functional areas, the expected number of clusters ranges around two in both cases. In the standardized areas, the number rises to around 13 in the U.S. and even to 38 in Germany. The settlement area according to the GMM also rises when switching from functional to standardized areas. Interestingly, the U.S. sites exceed the German sites in the functional areas which indicates that functional areas in Germany are generally smaller. In contrast to that, the settlement areas are almost twice as high in Germany compared to the U.S. which can be referred to the 200 km by 200 km environment which covers mostly rural land in the U.S., but includes many small towns in Germany. The relative standard deviation in the size of the clusters with around 30% is similar for the functional regions both in the U.S. and in Germany. These values reflect the high gradient between the one dominant and the few adjacent clusters. Switching to standardized regions, the relative

standard deviation drops down to around 20% for the U.S. and even 10% for Germany. This fact proofs that the standardized areas in Germany are populated with numerous more or less equal small-sized clusters with a mean deviation in size of only 10% whereas the standardized area in the U.S. have a lower number of clusters, but with a higher deviation in size with up to 20%.

Table 2. Average values for the number and size as well as the respective relative standard deviation in size of settlement clusters within the functional areas and the standardized patches in Germany and in the U.S.

Functional	Number		Area [km^2]		Areal Deviation [%]	
Germany	3	2	113	76	30	32
U.S.	2	2	316	163	30	31
	Mean	Median	Mean	Median	Mean	Median
Standardized	Number		Area [km^2]		Areal Deviation [%]	
Germany	38	37	1541	1471	10	8
U.S.	14	12	880	714	21	18
	Mean	Median	Mean	Median	Mean	Median

6. Discussion

This section interprets the results presented above from a methodological, implementational, statistical, and geographical point of view. Future research needs are discussed.

6.1. Methodological aspects

Our approaches base on sophisticated multi-scale image processing using spatial convolution with finite response filter masks. This procedure is fast with today's multi-core processing and most robust with respect to sparse signals such as the GUF-DenS. The non-parametric approach further develops measures from discrete statistics like the in formation entropy to the more flexible homogeneity measure. The parametric approach applies a step-wise robust estimation using the least square estimation theory with weighted direct measurements. Both approaches circumvent the sticking point of each spatial evaluation of urban environments, i.e., the choice of the reference unit as stated in the beginning, by introducing a variable patch size via the image pyramid. Furthermore, due to the variable reference entity, the identification of potential outliers – which is essential in least square estimation to remove samples belonging to other entities – becomes superfluous. In this way, a most robust and transferrable algorithm is assured.

Nevertheless, the choice of two different reference units for the subsequent interpretation points out challenges for geographical comparisons: the functional areas on the one hand and the standardized areas on the other hand, reveal certain, but explainable differences. The fewer number of only two clusters in the functional areas supports the original choice of these regions as catchment areas for workers that commute to the central city more or less daily. Regarding the urban patterns of a whole country, our study shows, that the analysis of individual catchment areas might fall short as each city is embedded in a larger metropolitan territory, at least in Germany. The situation in the U. S. is different because of the long distances in-between neighboring cities. This becomes obvious in the quite low variation of

the expectance values in Table 1 and Table 2 when comparing functional and standardizes regions in contrast to the high variation regarding Germany.

One drawback might be that the process solely relies on the density in the GUF-DenS product, which does neither directly reflect the number of inhabitants nor any other geographic urban parameter. From a morphological point of view, it is important to emphasize that settlement density can only be approximated in two dimensions and that the third dimension of cities is excluded so far. Beyond, complexes of buildings like airports, exhibition halls or even larger factories show up with an extremely high density over extended areas and may shift the observed gravity center away from the center of the inhabited city. This fact becomes obvious in Fig. 8 where the ellipse of Munich is not centered around the visible shape of the residential urban footprint, but shifted to the north with the airport and its extensive industrial zone whereas the south is characterized by extensive forested areas and therewith, a much lower built-up density. Apart from the fact, that the third dimension and functional variations are not included in the dataset, the GUF-DenS is consistently available at global scale and suitable for transnational comparisons as envisaged in this study.

6.2. Implementational aspects

The parametric approach requires a certain configuration in order to work efficiently: The minimum patch size is set to 25 km^2, the minimum density for a cluster is set to 10%, and the minimum probability of belonging to the considered cluster is set to 1 ppm. To justify the choice of these three variables, the non-parametric approach is consulted: the mean scale of maximum heterogeneity is always above 25 km^2, see Table 1. The density threshold of 10% regarding the mode results from the typical site occupancy index in Germany of 20%. Having in mind, that U.S. settlements might be less dense, we reduced this value to the half. The probability threshold is chosen in order to capture almost the whole set of built-up pixels belonging to the respective cluster. We are aware that there is some subjectivity here in the calibration. However, we argue that this is not the crucial point: what is most important is that it has been carried out consistently for all cities and city regions and thus, becomes comparable. Furthermore, the plausibility of the results confirms that this calibration is appropriate.

The above considerations consequently lead to following kernel sizes in the image pyramid: 1, 3, 5, 7, 9, 11, 13, 15, 19, 21, 25, 31, 37, 45, 53, 63, 75, 89, 107, 127, 151 for the non-parametric approach exclusively, and 181, 215, 255, 303, 361, 429, 511, 607, 723, 861, 1023, 1217, 1447, 1721, 2047, 2435, 2895, 3443 for both approaches. The para metric approach starts at a pyramid level with a patch size of 181 by 181 pixels, which corresponds to an areal coverage of about 20 km^2 and is slightly below the envisaged 25 km^2. Though the number of pixels per dimension is typically doubled with each level of the pyramid, we decided in favor of a much finer sampling of about 1.2 instead of factor 2, which results in four intermediate levels until the number of pixels per dimension is doubled and still two intermediate levels until the covered area is doubled. Additionally, it is necessary to have an odd integer number of pixels in the patch for implementation reasons. Therefore, several adjacent levels of the image pyramid coincide for small patch sizes below ten pixels.

Regarding higher pyramid levels, each level has its own patch size that diverges less than one pixel in each direction from the intended size. The fine scale sampling ensures very stable results along the scale dimension, which is clearly visible in the diagrams (e.g., Fig. 18) because the homogeneity curves are very smooth even without any further smoothing algorithm applied for display purposes (e.g.).

Another question arising at this point is the robustness of the approach in terms of how reproducible the results are. Most clustering approaches depend extremely on the image subset, i.e., a slightly varying subset leads to completely different results. This is not the case for our approaches. The images comparing functional and standardized areas clearly show that almost the identical ellipses are drawn despite of varying subsets, see Fig. 12 and Fig. 14. This fact underlines the robustness of this novel approach and assures its suitability for transnational comparison. The robustness towards varying subsets even enables image tiling in order to cover whole countries or continents. When image tiles are chosen with an overlap, the clusters of the over lapping parts can easily be fused by location and size. The final map would be the settlement clustering of a whole country or continent. The non-parametric approach could be broken down to a scale, homogeneity, and skewness map using a fixed standardized tile size.

6.3. Statistical aspects

The approaches are designed in order to evaluate urban patterns over large areas. Indeed, we only studied the functional urban areas and standardized patches of 200 km by 200 km. The single evaluations are robust and reliable as noticed above, but the mutual evaluation of all cities under study might be affected by unwished correlations coming from overlapping parts. For instance, the overlap of the image subsets regarding the standardized areas, e.g., in the Ruhr metropolitan area, might cause some cities to appear multiple times in the statistics. The effect is low, but not negligible. Two strategies are feasible in order to assure independent measurements: First, extent the image subsets to a whole country with image tiling as described in the preceding section, and second, to evaluate closed, non-overlapping polygons like the functional regions with restriction that a possible embedding in a larger metropolitan region might not be recognized. The coincidence of the results both from the functional as well as from the standardized regions underline, that misbalance introduced by overlaps does not question the basic statements.

The archive of DLR holds the GUF-DenS as well as most other global products in geographic coordinates, i.e. in a grid defined by latitude and longitude. This is reasonable to store a global coverage, but complicates a proper spatial analysis because neither the areas nor the angles are preserved in this most simple projection. A possible solution could be the projection to appropriate map coordinates like UTM. In return, this in duces a necessary resampling, bent image borders, and varying reference systems for the single cities of interest. Especially with view to image tiling, the mosaicking of adjacent UTM zones is not an easy task. In summary, the use of a conformal projection would cause more problems than it solves. Therefore, we decided to adopt a mean pixel area for each image individually that is considered in the statistic evaluation. The unavoidable distortion of the error ellipse is tolerated. A better approximation to the local situation could be realized by the inclusion

of the different sampling distances in latitude and longitude in the filtering kernels, i.e., replacing the distance in pixels by distances in a metric system. To be sure that our approximation is already sufficient, the study on standardized regions was performed in UTM coordinates in parallel with only negligible changes to the final results.

6.4. Geographical aspects

In general, there is a widespread agreement in the urban geographic discourse that American urban regions are characterized by a more polycentric-dispersed structure than European ones (Guillain, Le Gallo, & Boiteux-Orain, 2006; Heider & Siedentop, 2020; Krehl, 2016; Riguelle, Thomas, & Verhetsel, 2007). These findings, however, mostly refer to spatial units similar to the functional urban regions considered in our study, i.e., more sub-centers (or "edge cities") are measured within the urban catchment area, and a loss of importance of the core city as an employment hub. Based on our input data, however, referring to centers as density clusters of settlements without a differentiation based on urban functions, jobs or the like, the quantity of centers and subcenters does not differ systematically between Germany and the USA. What is different, however, is that identified centers in the USA are much larger. However, we must acknowledge that the centers identified by the parametric approach can hardly be compared to the centers and subcenters in the above-mentioned literature. Thus, our method seems to be better suited for the analysis of settlement structures at a larger

spatial scale such as inter-urban polycentricity (Meijers, 2008) than for the measurement of intra-urban polycentricity. When considering the larger, more regional spatial units of 200 km × 200 km around each core city, we measured very different urban patterns between U.S. and German regions. These regional urban patterns in the U.S. have a much stronger dominance of the core cities (including their extended suburban fringe). These have, on the one hand, a larger extent and, on the other hand, fewer sub-centers in the surrounding area than it is measured for German regions. This may be reflective of the much denser settlement system in Germany, which is characterized by a large number of small and medium-sized cities outside of the metropolitan areas. This again shows how vulnerable spatial statistics are due to their specific approaches such as spatial extent, input data, methods, etc. (Thomas, Schmidt, & Siedentop, 2021). Another key finding of this study is that the homogeneity of American cities is measured as much higher. In order to reach the entire urbanized territory, one has to travel much longer distances in the U.S. cities than in German cities. This finding can be interpreted as the result of different planning cultures and trajectories of urban growth. In the U.S., unregulated sprawl has been the dominant form of urban development, resulting in large and homogenous low-density suburbs and long commuting distances. In stark contrast, urban growth in Germany has always been strongly regulated by land-use planning and a hierarchical system of so-called '*central places*', which aims to ensure equal living conditions in all parts of the country (Schmidt & Buehler, 2007).

While our study does not treat the multi-temporal aspect, but focuses on the comparison of urban patterns across space at one moment in time, it does not contribute to the debates about the '*Americanization*' of European or the '*Europeanization*' of American urban regions. How ever, for the specific case of our large spatial units for measurement, at least, it can be

stated that the urbanized landscapes around the respective core cities as of today differ significantly between Germany and the U.S.

These measured differences in urban patterns in Germany versus the U.S. apply to almost all urban areas. But, of course, there are also outliers. In Germany, Monchengladbach, Dusseldorf and Aachen are measured with an atypical urban pattern (Fig. 24). This results from the Rhine-Ruhr metropolitan region, where many larger core cities cluster in comparatively close proximity. This fact becomes clearly visible in the upper half of Fig. 10 showing the standardized patch around Cologne. Thus, there are less small sub-centers as typically found in the sur rounding of core cities for the other German cities. In addition, the specific local situation of the Eifel National Park (lower left part of Fig. 10) with its very sparse population and the absence of urban centers adds to this pattern. But, also for the American examples there are outliers: Allen, a comparatively small city in our sample, has an un usually large number of centers in its vicinity and even the core city does not possess the typical dominance in terms of a large ellipse area.

6.5. Future aspects

Some tasks for future studies were already mentioned like the extension of the image subsets to the coverage of complete countries, e. g., by image tiling and mosaicking. Another idea is the inclusion of or transfer to other data sets. Possible candidates could be the composites of Night Time Lights of VIIRS, which map the areas of human activity by their night-time light emission. Alternatively, the Sentinel-1 coverages could also be analyzed, as settlements are characterized by extremely high reflectance values due to the double bounce effect and low speckle impact. With respect to Sentinel-2, the diverse built-up indices like the newly published LEAIndex (Schollerer et al. 2022) could also provide a suitable data base in addition to the ones already mentioned. With respect to the parameter estimation in the GMM, even information on infrastructure, e.g., traffic routes that establish fast connections between urban clusters could be integrated as done in a recent study on the mobility network in Germany (Feicht & Schmitt, 2022). With the additional information on inter-urban connections, the assignment of the single built-up pixels to distinct urban clusters gains more importance. The current model delivers a pixel-wise probability of belonging to each of the clusters. Thus, the highest probability decides on the assignment to a specific cluster. With the inclusion of quick inter-urban connections, a place of interest might have a better transport link to another cluster than to the one it belongs to spatially according to the GMM. Thus, a multiple assignment including direct distance measurements becomes most interesting to be studied in the future. Geodetic network adjustment could be one suitable mathematic tool for this evaluation.

A more methodological refinement is the extension from the two-dimensional image space (with latitude and longitude) to a multidimensional feature space. Both, the parametric and non-parametric approach can cope with n-dimensional feature spaces in principle. With an increasing number of features, the number of indistinct features unfortunately increases likewise. So far, the approaches only consider concise descriptors. One solution might be the orthogonal trans formation of the feature space to hyper-complex bases in order to select distinct features (Schmitt et al., 2020) or the use of machine learning methods that are known

to be resilient towards outliers and indistinct features. Having solved this problem, both approaches could also be utilized for the data-driven clustering in high-dimensional feature spaces, e.g., for image classification and object recognition.

7. Conclusion

This study adds two novel methods for the quantitative assessment and the well-grounded comparison of urban patterns to the literature: (1) a non-parametric multi-scale evaluation of the heterogeneity of the built-up density and (2) a parametric hierarchical multi-scale clustering that models urban patterns by the help of Gaussian bell curves to form a Gaussian Mixture Model. Both approaches represent a sophisticated fusion of classical geodetic network adjustment, digital image processing with multi-scale directional filter banks, and information theory. These methods are applied to selected urban areas, i.e., 68 cities in Germany and 162 cities in the U.S. with (a) their functional extension and (b) a standardized region of 200 km by 200 km. At the end, 920 case studies are joint together in order to answer the research questions formulated in the beginning:

1. Yes, the multi-scale homogeneity reveals as robust measure to describe urban environments. The influence of the varying reference units in the interpretation is visible, but explainable, and does not impede the inter-continental comparison of the results. The scale of minimum homogeneity and therewith, maximum heterogeneity seems to be an essential variable in the description of spatial proximity. It denotes the spatial distance or the size of the patch that contains the most heterogeneous built-up density.

2. Yes, the decomposition into single Gaussian bell curves is reproducible, even if parts of the respective settlement agglomeration are missing in the actual sample. The method is able to clearly detect density clusters and join neighboring, but adjacent clusters to one thanks to its hierarchical multi-scale approach that applies step-wise coordinate refinement without any further time-consuming iteration step. The procedure is purely deterministic and thus, easily transferable to other geographical locations and to other, potentially higher-dimensional feature spaces as well.

3. Yes, there are distinct and replicable differences in the urban pat terns of the U.S. and Germany independent of the chosen reference unit. Some expectations were confirmed like the large extension of U. S. metropolitan areas or the small-scale urban structures in Germany. Other findings like the scale of the maximum heterogeneity – that also stands for a high quality of living – are completely new. We found a significantly higher heterogeneity in a much smaller environment in Germany. Furthermore, German metropolitan areas are characterized by numerous small, more or less equal-sized clusters whereas U.S. metropolitan areas contain one large and dominant center and only few small adjacent urban clusters.

This study provides the methodological basis for a nation-wide, full coverage comparison between Germany and the U.S. at first and a continental-wide comparison between Europe and North America at second. As the GUF-DenS – the only database needed for this study – ensures a global coverage, these two approaches enable even a world wide comparative study. Including additional data sources like infra structure or census data will refine the

results for dedicated regions where this kind of geo data is available. The fused evaluation of diverse data sets will probably lead to a new perception of spatial proximity in urban environments at local, national, continental, and global scale.

Appendix A

The 68 urban regions in Germany under study: Aachen, Aschaffenburg, Augsburg, Berlin, Bielefeld, Bocholt Stadt, Bonn, Braunschweig Salzgitter Wolfsburg, Bremen, Bremerhaven, Cologne, Constance, Darmstadt, Dresden, Düren Stadt, Düsseldorf, Erfurt, Flensburg, Frankfurt am Main, Freiburg im Breisgau, Giessen, Görlitz, Göttingen, Halle an der Saale, Hamburg, Hanover, Heidelberg, Heilbronn, Hildesheim, Ingolstadt, Iserlohn, Kaiserslautern, Karlsruhe, Kassel, Kiel, Koblenz, Leipzig, Lübeck, Magdeburg, Mainz, Mannheim Ludwigshafen, Mönchengladbach, Muenster, Munich, Neubrandenburg, Nuremberg, Offenburg, Oldenburg, Osnabrück, Paderborn, Pforzheim, Regensburg, Reutlingen, Rosenheim, Rostock, Ruhr, Saarbrücken, Schweinfurt, Schwerin, Siegen, Stuttgart, Trier, Ulm, Wetzlar, Wiesbaden, Wuppertal, Würzburg, Zwickau.

The 162 urban regions in the U.S. under study: Ada, Alachua, Albany, Albuquerque, Allen, Atlanta, Atlantic City, Austin, Bell, Benton AR, Benton WA, Berks, Boston, Boulder, Brazos, Brevard, Brown, Butte, Caddo, Cameron, Charleston, Charlotte, Chatham, Chicago, Cincinnati, Collier, Columbus, Cumberland ME, Cumberland NC, Cuyahoga, Dallas, Dane, Dauphin, Davidson, Denver, Detroit Greater, Douglas NE, Durham, East Baton Rouge, El Paso CO, El Paso TX, Erie NY, Erie PA, Escambia, Fayette, Forsyth, Fresno Greater, Genesee, Greene, Green ville, Guilford, Hamilton TN, Hampden, Hartford, Hidalgo, Houston, Indianapolis, Ingham, Jackson MO, Jacksonville, Jefferson AL, Jefferson KY, Jefferson TX, Kalamazoo, Kent, Kern, Knox, Lackawanna, Lafayette, Lancaster NE, Lancaster PA, Lane, Larimer, Las Vegas, Lee, Lehigh, Los Angeles Greater, Lubbock, Lucas, Luzerne, Madison, Mahoning, Marion FL, Marion OR, McLennan, Memphis, Merced, Miami Greater, Milwaukee, Minneapolis, Minnehaha, Mobile, Monterey, Montgomery AL, Montgomery OH, Muscogee, Nashville, New Haven, New Orleans, New York Greater, Newport News, Nueces, Oklahoma, Onondaga, Orange, Peoria, Philadelphia Greater, Phoenix, Pima, Pittsburgh, Polk, Portland, Potter, Providence, Pulaski, Richland, Richmond Greater, Roanoke, Rochester NY, Sacramento, Salt Lake, San Antonio, San Diego, San Francisco Greater, San Joaquin, Santa Barbara, Santa Cruz, Sarasota, Scott, Seattle, Sebastian, Sedgwick, Sonoma, Spokane, St Joseph, St Louis, St Lucie, Stanislaus, Stark, Summit, Tallahassee, Tampa Hills borough, Tampa Pinellas, Thurston, Tulare, Tulsa, Tuscaloosa, Utah, Vanderburgh, Ventura, Virginia Beach, Volusia Daytona Beach, Wake, Washington Greater, Washoe, Washtenaw, Webb, Weld, Winnebago IL, Worcester, Yakima, York.

Appendix B. Supplementary data

Supplementary data to this article can be found online at
https://doi.org/10.1016/j.compenvurbsys.2022.101920

Chapter VI

Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception

Environment & Planning B: Urban Analytics and City Science, 2021, vol. 48(2)

Michael Wurm, Jan Goebel, Gert G. Wagner, Matthias Weigand, Stefan Dech & Hannes Taubenböck

© The Author(s) 2019.

https://dx.doi.org/10.1177/2399808319869341

Published 20 Aug 2019

Abstract

The morphology of today's cities is the result of historic urban developments and on-going urban transformation resulting in complex urban spatial structures. While functionally as well as spatially, cities are structured into sub-units such as the city center, business districts, residential areas, or industrial and commercial zones, their precise localization in the geographic space is sometimes difficult. City centers in particular are difficult to grasp because they stand for many different urban functions. Generally, they are characterized as areas of high densities, such as employment, population, or commerce, but also as areas of high structural densities. In past studies, various approaches, data, and cut-off values have been presented to separate city centers from the surrounding transition zones in geographical information systems based on density values. To overcome the difficulty in defining the right density threshold, the current study presents an approach which integrates the subjective perception of citizens on the urban spatial structure and relates it to the floor area ratio to delineate city centers to infer a characteristic density threshold. In a large empirical study of 67 monocentric cities in Germany, we observe a decreasing floor area ratio gradient toward the urban fringes. The result of the study reveals a relative threshold for city centers at around 30% of the density of the "central place."

Keywords

City center; cognitive perception; urban spatial structure; density; floor area ratio; SOEP

1. Introduction

The spatial structure of cities has undergone fundamental transformations over the past centuries. Today's cities appear often as decentralized, polycentric (Garreau, 1991) urban structures; they are sometimes even seen as an intermediate step to a dispersed urban world (Batty et al., 2004). This is mostly a result of planning policies in the 20th century where urban functions have been segregated and decentralized shaping modern cities as complex urban systems connecting urban functions and urban forms throughout spatial and functional networks (Batty, 2008). While these described processes are mainly observed in North America and recently in Asia, European cities mostly still resemble what they looked like "*in the lifetime of St. Thomas Aquinas [. . .] built in the 13th century*" (Holmes, 2001). Medieval European cities were characterized as places with a mix of urban functions (e.g. housing, trade and commerce, consumption, employment) in very close spatial proximities resulting in the majority of jobs being concentrated in central manufacturing areas (Anas et al., 1998). Since city sizes were regulated by walking or horseback riding distances, spatial congruency between these functions could be enabled inside city walls which were constructed in the shortest, thus in a circular shape around a "*central place*" which served as nuclei of urban development, e.g. the central market (Schlüter, 1899). This has led to the development of monocentric urban spatial structures in European cities which could be preserved until the 19th century when the Industrial Revolution gave way to new usages of urban spaces and processes of de-concentration and suburbanization gained in prevalence (Siedentop, 2005). A key prerequisite for the latter was the decreasing importance of spatial proximity as a necessity for economy due to new transport technologies. Thus, the importance of the historical city centers in terms of trade and commerce has eventually declined over time (Besussi & Chin, 2003; Schiller, 2001; Thomas & Bromley, 2003), but they are still important for the provision of urban functions and soft economies today, favoring high density and mixed locations (Batty et al., 2004).

In fact, high density and high diversity are considered as the most important characteristics when defining city centers. Most commonly, functional characteristics are used to define and delineate city centers, such as high concentrations of commerce, economic activities, employment, high individual and public traffic (including pedestrians), as well as high diversity of activities and touristic attractions, respectively (Anas et al., 1998; McMillen, 2004; Murphy & Vance, 1954; Sert, 1952; Thurstain-Goodwin & Unwin, 2000). Besides urban functions, also morphological characteristics were used to delineate city centers, mostly in terms of structural density (Krehl et al., 2016; Murphy, 1971; Murphy & Vance, 1954; Taubenböck et al., 2013, 2016). As a result of this plurality of characteristics for defining city centers, a commonly agreed and consistent definition of city centers has not evolved (Ralphs & Wyatt, 2003), and moreover, no commonly accepted density threshold has been found.

This ambiguity in defining and delineating city centers is addressed in the present paper. We aim at taking up on the fuzziness for delineating city centers by analyzing cognitive perceptions of citizens on city centers and relate it to the floor area ratio (FAR) to empirically infer a density threshold value for separating city centers from the transition zone. In this way, we analyze people's cognitive perceptions to spatially define city centers, since citizens

recognize and realize the presence of a city center when they are standing in a city center even if they cannot define it. The use of subjective perception to create "mental maps" has been introduced by Lynch (1960) where he describes the concept of "legibility", making parts of the city recognizable to its dwellers and thus triggers to spatially organize the city in the sense of mental maps. They are a subjective interpretation of the environment as a function of individual social values and constraints (Rengert & Pelfrey, 1997). This concept is related to the direct perception theory (Gibson, 1979) where specific objects are related to specific functions or activities. For example, a park affords relaxing or a street affords walking (Jiang, 2012). In this way, a city creates also a mental map for citizens (Haken & Portugali, 2003).

Based on these concepts, we search empirically for a characteristic degree of FAR which is commonly accepted by citizens to separate city centers from their surrounding areas. Thus, this paper contributes to research on the delineation of city centers. The work provides new perspectives on delineating city centers following an empirical, quantitative analysis of 67 monocentric major cities in Germany. In an experimental setting, we calculate structural density represented as FAR based on 3D building models and relate it to cognitive perceptions of citizens on the city center. To demonstrate the proof of concept, we divide the statistical population of all cities into four groups based on population density since it provides a proxy value for the urban spatial structure. We argue feasibility of this proof of concept if the empirically derived distances are similar among the four groups.

The remainder of this article is structured as follows: we begin by reviewing how city centers are defined and localized in previous studies based on functional and morphologic features ("Delineation of city centers: Related work" section). The "Data and methods for delineating city centers" section describes the used data and the workflow for the calculation of urban morphologic features such as the FAR and the integration of survey data on cognitive perception of city centers. This is followed by a discussion on the proposed analysis in the "Discussion of empirical findings" section and the final section concludes the paper with an outlook.

2. Delineation of city centers: Related work

To delineate city centers, we need to have an understanding of how and based on which criteria they are separated from their surrounding areas. But these surrounding "zones of transition," as introduced in the classical models of urban land-use (Park & Burgess, 1925), remain also rather fuzzy in definition and shape. They are described as areas with lower densities than the city center, e.g. population density (Borruso & Porceddu, 2009) and also lower land prices as a function of the accessibility to the central business district (CBD) as introduced in the classical, monocentric urban bid rent model by Alonso (1964) and still applied in current studies (D'Acci, 2019). From the aspect of historical urban development, the transition zone represents the former suburbs of the historic city, mostly located outside former city fortifications. In some cities, these fortifications could be preserved until today, e.g. Neuf-Brisach (France) or Lucca (Italy), and still delimit a sharp border between the city center and its surroundings, mostly visible by significant differences in density (Fig. 1(a)). In other cases where the city walls have been torn down, we find today other structuring

elements in their place such as boulevards, e.g. the “Ringstraße” in Vienna (Austria), the “boulevard peripherique” in Paris (France) or parks, e.g. Cologne (Germany) which still keep the separating function of the former city walls alive, mostly still visible by changing morphological densities (Fig. 1(b)). Another example are cities with no more clearly visible separating element in their today’s urban spatial structure between their former city centers and their surroundings since their urban morphology has been transformed more severely over the past centuries (Fig. 1(c)), e.g. due to large-scale destruction and rebuilding, e.g. Rotterdam (The Netherlands). In these cases, the physical circumscription of the city center cannot be easily identified since the transition zone is an area of flux which is characterized by a continuous change in urban function or land-use as well as a decrease in density without any obvious threshold.

From a scientific point there exists neither a single, commonly agreed feature nor cut-off value to delineate the city center from its surroundings. Due to the fuzziness of city centers, the use of geographical information systems for their delineation has been criticized for a long time (Taylor, 1995), although more recent approaches have successfully addressed this issue (Batty et al., 2004). In related work, city center detection was performed based on a large spectrum of features, such as functional characteristics (Borruso & Porceddu, 2009; Lloyd, 2009; Thurstain-Goodwin & Unwin, 2000), travel flows (Ratti et al., 2006; Roth et al., 2011), location-based social networking (Sun et al. 2015), social media data (Chen et al., 2017; Hollenstein & Purves, 2010), urban land-use data (Zhu and Sun, 2017), points of interest, and road network data (Lei et al., 2015; Yu et al., 2015). Besides functional characteristics, also morphological features such as the density were used to delineate CBDs (Krehl et al., 2016; Murphy & Vance, 1954; Taubenböck et al., 2013). These related works approach city center delineations either in terms of a clear-cut spatial delineation based on a broad set of different kinds of data or approach the fuzziness of city centers understanding it as higher order geographical phenomenon. In this way, Burrough & Frank (1996) find city centers based on the cognitive perception of citizens such as mental maps, as introduced by Lynch (1960). The approach of using social studies participants’ individual perception for spatially structuring urban areas by the features ‘location’, ‘density’, and ‘function’ was successfully adapted by Hoffmeyer-Zlotnik (2000). Further studies related citizens’ perceptions to the appearance of specific urban functions of city centers (Alarasi et al., 2016; Le et al., 2005; Lüscher & Weibel, 2012), assessed the quality of a center (Rengert & Pelfrey, 1997), or delineated the area of city centers for a small number of cities (Le et al., 2005; Lüscher & Weibel, 2012; Montello et al., 2003). Despite the related work in this context, one remaining drawback is that the individual choice of criteria for describing a city center remains very subjective for each individual. Thus, it is very unlikely that two citizens agree on a similar set of criteria (Thurstain-Goodwin & Unwin, 2000) making the search for an optimal feature set a very difficult task. Therefore, instead of searching for a commonly accepted set of functional criteria to delineate city centers for all cities under investigation, we base our approach to citizens’ perceptions on their household locations to delineate city centers and relate it to the FAR as a very visible and objective criterion. Moreover, we aim at finding a generalizable, characteristic FAR-threshold for separating city centers from their surroundings.

Fig. 1. Examples for city centers: (a) preserved city walls: Lucca (Italy), (b) former walls replaced by structuring elements such as boulevards: Vienna (Austria), and (c) entirely transformed: Rotterdam (The Netherlands). Source: Google Earth.

3. Data and methods for delineating city centers

The proposed workflow for empirically inferring a characteristic FAR-threshold to delineate city centers is based on citizens' perceptions and decreasing FAR gradients in monocentric cities (Barr & Cohen, 2014).

To evaluate the proposed proof of concept, we base our study on a large sample of major cities in Germany. Selection of sample cities was based on several criteria: (1) a minimum population size of 100,000 inhabitants to ensure a significant size of the city center and the transition zone, (2) a monocentric urban structure following a radial FAR gradient, (3) a historic city center to avoid recently planned city foundations, and (4) their spatial development has not been too impaired by natural barriers such as mountains or water surfaces. All German major cities have been individually evaluated with regards to their historic development and their urban spatial structure based on (historic) maps and satellite images. Cities that have undergone significant territorial reforms (e.g. Berlin, Wuppertal, Remscheid), with decreasing population sizes (e.g. Zwickau, Wilhelmshaven), with a spatial development which is affected by natural barriers such as the sea (e.g. Bremerhaven, Kiel, Flensburg) or cities without a historical city center (e.g. Wolfsburg, Karlsruhe, Leverkusen, Ludwigshafen) have been removed resulting in a total of 67 major cities for the current study.

For the selected cities, 3D building models are used to calculate the FAR on the spatial level of city blocks. Further, citizens' perceptions from a socio-economic survey are used to spatially delineate the extent of city centers. In a consecutive step, the spatial extents of the perceived areas of city centers are spatially superimposed with the values to infer a characteristic FAR-threshold value at the edge between the city centers and the transition zone.

In the absence of objective validation data for the size of city centers, we test the proof of concept with the help of creating four groups from the 67 cities based on the following assumption: if the inferred FAR-threshold values are similar among these groups, an interrelation between FAR and cognitive perception on city centers can be argued. The four groups are based on the population density of the 67 cities since the population density integrates both, population size and area of a city. Using Jenks' natural breaks classification algorithm (Jenks & Caspall, 1971), the following four groups are created: (1) ≤850 inh./km^2, (2) 850–1600 inh./km^2, (3) 1600–2500 inh./km^2, and (4) >2500 inh./km^2. The proposed workflow is presented in detail below and an overview of the required steps is depicted in Fig. 2.

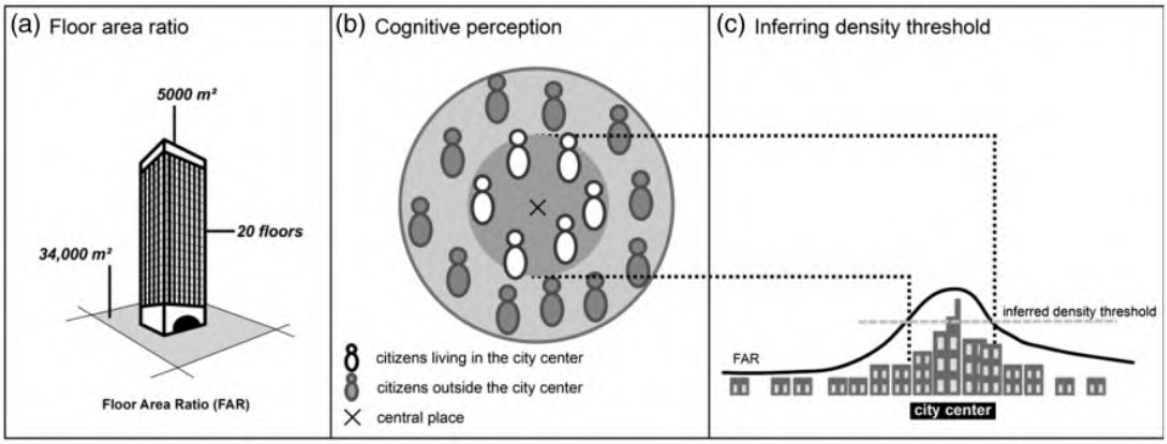


Fig. 2. Workflow for delineating city centers based on FAR and cognitive perception: (a) Calculation of FAR ("Inferring the FAR-threshold" section), (b) using citizens' perceptions to delineate city centers ("Cognitive perception of citizens" section), and (c) inferring the FAR-threshold ("Inferring the FARthreshold" section).

3.1. Calculation of FAR

Structural density represents a key feature for the characterization of urban morphology and is crucial for describing urban activities (e.g. Acioly & Davidson, 1996). It stands for proximity because with rising density of a city, distances are decreasing and citizens cover more distances walking in the city, by bike or public transport. Density serves as a trigger for synergies and concurrences, thus having a positive impact on prosperity and innovation. Despite its wide application in urban studies and planning, structural density still represents a fuzzy concept (e.g. Krehl et al., 2016; Taubenböck et al., 2016). For its calculation, usually building footprint areas or volumes are related to the area of the land lot or the city block.

An objective and comparable measure of structural density for analyzing the urban spatial structure represents the FAR (Barr & Cohen, 2014; Wurm et al., 2014). It is defined as the ratio of total usable floor space to the size of the plot, e.g. the FAR of a plot with a 5-story building constructed over the entire plot would equal to 5, while a 10-story building over half the plot size would also equal to 5 (Barr & Cohen, 2014). For the current study, FAR is calculated based on ministerial 3D building models at the Level-of-Detail-1 (LoD-1) and data on city block geometries. LoD-1 building models represent the shape of buildings as block model incorporating building footprint geometries (e.g. area) and a metric building height value; however, these data usually do not incorporate the number of floors. Thus, these data represent a simple 3D representation of buildings. For detailed information on the related LoD-1 building model, we refer to the official product catalogue for federal agencies on the website of the data provider.1 Since for calculation of FAR, the buildings' areas and the number of floors are required, we infer the number of floors based on empirically established linear regression models with a sample of 600,000 buildings including data on metric height and number of floors. The approach for inferring floor numbers based on LoD-1 building models is related to Wurm et al. (2011) and Krehl et al. (2016), and we refer to these studies for a detailed description of the workflow.

The model is validated in a twofold cross validation procedure, yielding an adjusted R^2 of 0.746. After application of the model to all 9.5 million individual buildings, FAR is

calculated for all 417,000 city blocks in all 67 cities (Fig. 3). Since the range of FAR values varies significantly among the cities, we use the normalized FAR for better comparability. Normalized FAR (FARN) values are calculated in relation to the FAR around the "central places," where usually the highest densities are found in monocentric cites. Thus, FARN is set to 1.0 for all cities at the "central place" and all other density values for the same city are related to this density. As a result, decreasing FARN gradients between the city center and the periphery are observed. The gradients vary also significantly between cities of varying sizes, e.g. in cities with higher population densities, higher FAR values are observed.

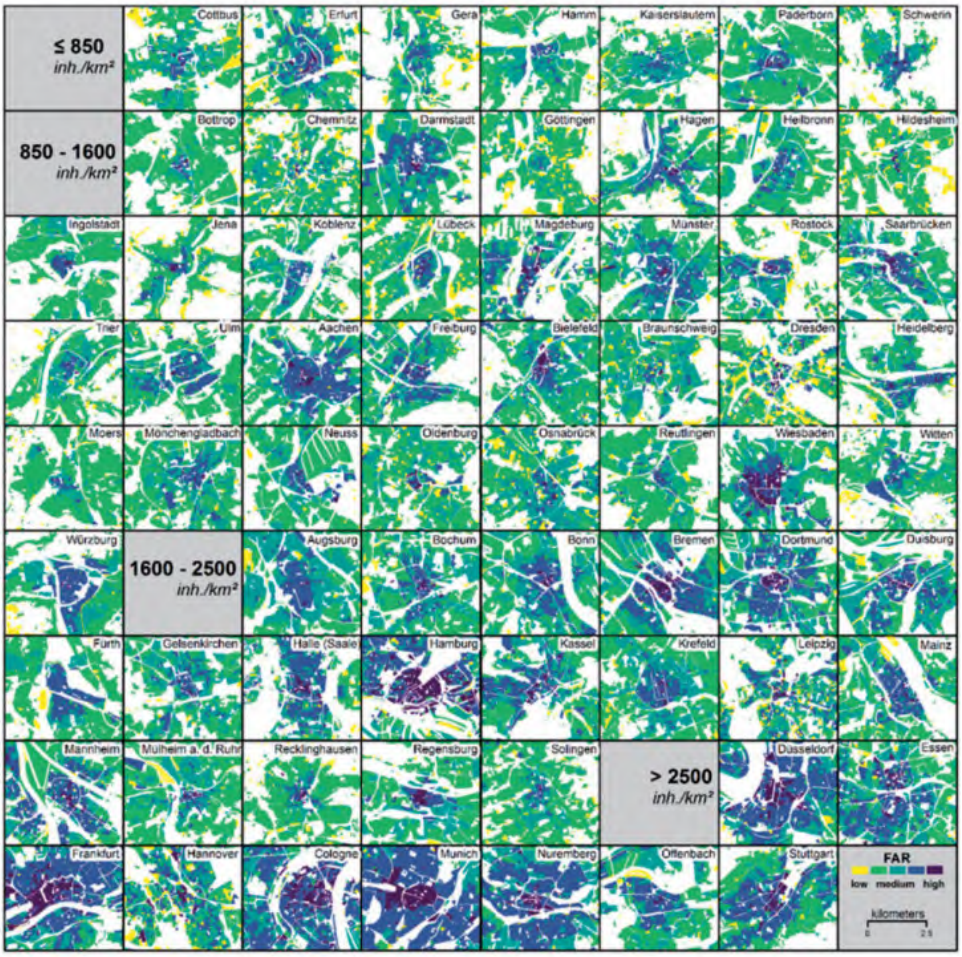


Fig. 3. FAR for all 67 cities of the study, presented in four groups of population density.

3.2. Cognitive perception of citizens

For all 67 cities in this study, we use cognitive perception of citizens to create a common understanding of the size of the city centers. These "mental maps" are based on national household survey data from the Socio-Economic Panel Study (SOEP) on more than 23,000 individuals in about 11,000 households in Germany (Goebel et al., 2018).

In SOEP, detailed socio-economic data about objective and subjective life conditions are surveyed on a yearly basis. One question asks participants to locate their house in the city as being located either "in the city center" or "outside the city center." In total, 2865 households claim to live in a city center. For each citizen's house located in the city center, the distance between the geographic location of the house and the "central place" is calculated using

anonymized geographically referenced household data (Goebel & Pauer, 2014). In the last step, the mean distance value for all the households is calculated. This distance value stands for the commonly perceived size of the spatial extent of the city center by the citizens. In the following analysis step in the "Inferring the FAR-threshold" section, this distance value is used to infer the FAR-threshold. Fig. 4 displays the households' distances to the "central places" for each population density group: with rising population density, also the distances to the central place rise for the households claiming to live in the city center. For inferring the FAR-density threshold ("Inferring the FAR-threshold" section), the mean distance values are used for the four groups: (1) 1.40 kilometers, (2) 1.94 kilometers, (3) 2.55 kilometers, and (5) 2.85 kilometers.

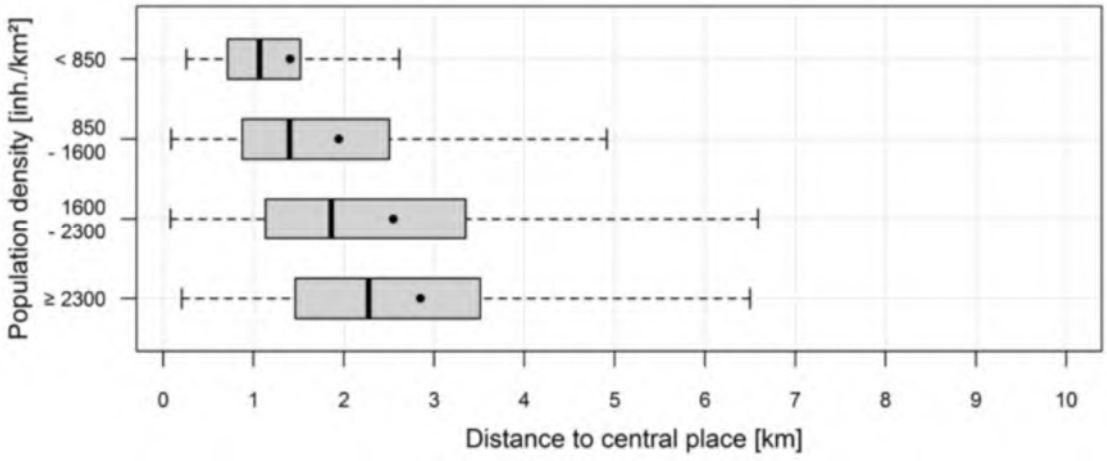


Fig. 4. Boxplots of distances between geographic locations of households and the "central place" grouped by the population density. The mean distance values are depicted as black dots inside the boxes.

3.3. Inferring the FAR-threshold

In this step, derived distances based on citizens' perceptions and normalized FAR-density FARN are mutually analyzed for inferring the threshold to delineate city centers from the surrounding transition zone. The threshold is inferred at the intersection of the distance value with the FARN modeled gradients for each of the four city groups. The modeled functions are characterized by a smoothing effect which can suppress local changes of the density measure, e.g. one tall building in the transition zone. Best fitting functions were sought as exponential functions. The FARN gradients are modeled for each of the four groups separately (Fig. 5). Goodness of fit for the modeled functions is measured as R2: (1) 0.985, (2) 0.992, (3) 0.985, and (4) 0.985.

The shape of the modeled functions varies between the four groups, e.g. FARN decreases at shorter distance from the "central place" toward the periphery for smaller cities than for larger cities (Fig. 5), e.g. FAR at a distance of 1.5 kilometers varies between 0.35 and 1.0 for city group 1 and group 4, respectively. Consequently, similar FAR values can be found at varying distances between the groups. We make use of this observation in our proof of concept in the following way: if the inferred FAR values are similar among the four city groups, we argue that structural density influences citizens' cognitive perception on city centers to a certain degree.

Inferred FAR values for the four groups are depicted in Fig. 5. The derived subjective distances to the central place vary between 1.4 kilometers (for the lowest population densities in the first group) and 2.85 kilometers (for the highest population densities in the last group). This accounts for an almost 250% difference in distance from the "central place" across the 67 sample cities. Regarding the inferred FARN values at the intersections, however, very similar values can be observed despite the large difference in distances: (1) <0.33, (2) 0.28, (3) 0.26, and (4) 0.30. Over all four groups, only very small differences in the FARN values are observed. Based on these inferred density values for the four groups, we observe FARN thresholds of approximately 0.3 for delineating city centers from the transition zone. As this threshold is tested for a large sample of 67 cities, it appears to be independent from varying city sizes or their spatial extent, respectively.

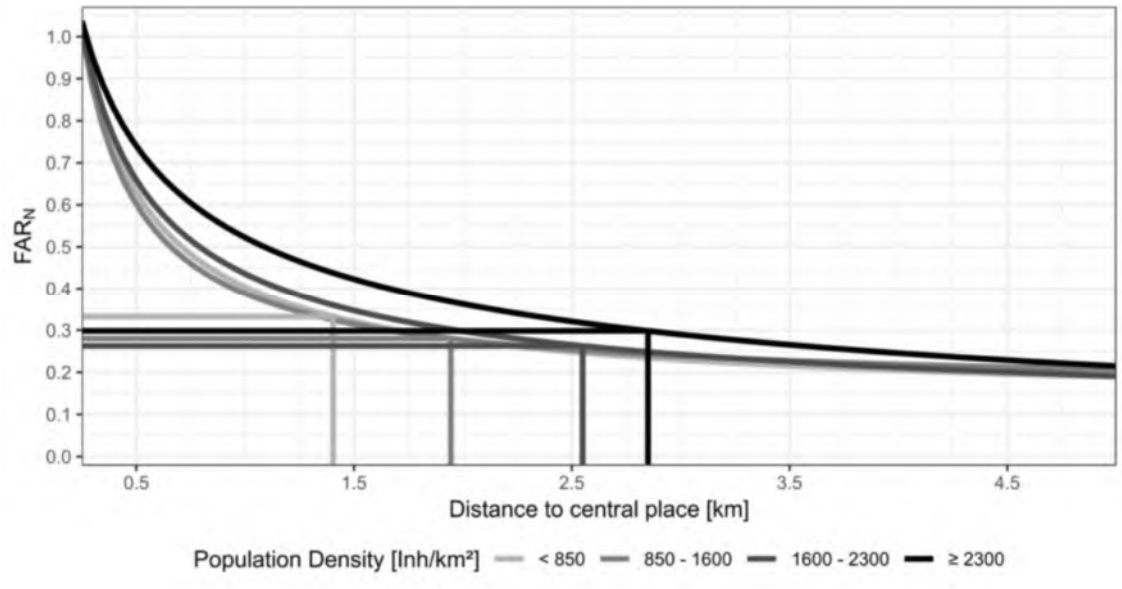


Fig. 5. Inferred FAR-thresholds for the four population density classes: gradients represent an exponential function of the decreasing FARN densities and the vertical bars represent the derived subjective distances as presented in Fig. 4. The horizontal bars indicate the inferred FAR value at the intersection point between the subjective distance and the FAR gradient. The difference in distance (x-axis) varies to a much larger degree than the difference in inferred FAR (y-axis).

4. Discussion of empirical findings

4.1. General remarks on the approach

What from an initial perspective seems to be a very trivial issue—pointing a city center on a map—becomes easily very complex when generalized descriptions and spatial circumscriptions of city centers are sought. While historical walled cities made a precise spatial demarcation of the city a feasible task, morphological transformations of the urban spatial structure and observed deconcentration processes of urban functions toward polycentric urban structures in European cities complicate this task to a high degree. The debate on defining and spatially delineating city centers takes up on the debate on the fuzziness of geographical places in urban areas which are more or less only qualitatively defined or are merely just spatial concepts at higher level. The underlying challenge of this ambiguity in definitions comes to light especially when entirely different aspects of these

geographical places are addressed, such as functional and morphological characteristics. While their precise definition and thus also their spatial delineation is subject to such a broad heterogeneity among scholars, their presence in urban areas cannot be neglected. For most of those geographical concepts, it holds true that they are "hard to define, but you know it when you see it" (Cervero, 2000). This perspective on the issue follows the empirical research on central places, which aims at exploring the organization of urban systems in an analytical way without the restrictions of formal theory (Burger & Meijers, 2012).

In the direction of these related studies, we also embed the presented study on cognitive perception and the urban spatial structure. This study presents an empirical approach to link cognitive perceptions of citizens and the structural density of cities. In this way, the concept of mental maps is integrated into the study design. Mental maps are very relevant because every citizen has his/her own individual perspective of the urban spatial structure and the location as well as the spatial circumscription of the city center. The individual perception of the location and extension of the city center is influenced by each citizen's unique subjective, personal history and understanding of the urban spatial structure (Rengert & Pelfrey, 1997). From the vast amount of morphological or functional attributes which characterize city centers, each citizen structures the city wittingly or unwittingly in his/her mind.

The presented approach allows for integrating a larger quantity of samples and cities than approaches using in-depth interviews. The latter mostly aim at precise spatial circumscriptions of city centers drawn by participants on a map, mostly as a case study for just one or a few cities (e.g. Lüscher & Weibel, 2012). These approaches may eventually lead to higher spatial precision and they can be analyzed in geographical information systems, but results are mostly restricted to a handful of case studies. Contrarily, our approach cannot predict a precise delineation for individual cities, but aims at associating urban structure and density on a more generalizable level. Therefore, it must be underlined that the inferred distances from cognitive perceptions cannot be transferred back to single cities for comparison with the actual situation since there is only one value for all cities from the same group of population density.

For a meaningful interpretation of this empirical study, general preconditions must be considered and therefore we highlight the following points:

- the approach derives a Euclidian distance for finding the extent of city centers. Since only one distance is used, a radial extension of the city center around the most "central place" is assumed together with a circular circumscription of the transition zone around it. Hence, the approach requires the assumption of a monocentric city model which can be more or less accepted for most German cities with similar historic city development (Schlüter, 1899). A linear transfer of the approach to cities with a completely different urban spatial structure, e.g. to polycentric urban structures, appears at this point under consideration of the used data not feasible.
- for all cities where a discrete delineation of the city center is not possible, a precise association of the respective functional and morphological urban areas into city center and transition zone is beyond feasibility since these measures do not follow a strict concept. Moreover, in today's cities the transition between the city center and the

transition zone is mostly continuous. Therefore, approaches for finding city centers use selected physical and/or functional, quantifiable characteristics of city centers. The perception of citizens integrates all objective and subjective characteristics of each individual's surroundings. Thus, the vast number of attributes influencing people's perception is difficult to fit into a model. To account for all these features, the current study relates to citizens' cognitive perceptions of the city center by the location of their houses.

4.2. Findings

In due consideration of the above mentioned aspects, the presented empirical study aims at relating the urban spatial structure to the cognitive perception of citizens. Cognitive perceptions of 2865 households living in a city center are analyzed with regard to their distances to the most "central place" of each city. Subsequently, the derived distance is used to delineate the city centers and intersected with the FAR gradient to infer a FAR-threshold for separating city centers from the transition zone. For comparability between cities, we use the normalized FARN which is represented as a decreasing density gradient from the most "central place" toward the periphery. The inferred FAR-threshold values are very similar for all cities among the four groups of city sizes: they range between 67 and 74% of the density of the most "central place" for all cities. In other words, we argue that structural density might play a role in the cognitive perception of citizens for separating city centers from the transition zone.

5. Conclusion and Outlook

The empirical experiments in this study indicate a relation between structural density and cognitive perception of city centers. The study brings data of great heterogeneity together for seeking new insights on the complex urban spatial structure. In detail, an interdisciplinary analysis of spatial and non-spatial data was conducted in the form of a combination of GIS-data and socio-economic survey data. With respect to the data, we performed the analysis for 67 cities with 9.5 million individual buildings to calculate the FAR. Furthermore, response from 11,000 households participating in the SOEP were used to analyze the city centers providing a very large database for the experimental setup of study.

For related studies in the future, to gain even more reliable and transferable results, we propose the use of a higher spatial density of household observations. Furthermore, with respect to future works, while this approach was empirically setup and tested for German cities with similar historic developments, a comparison to cities with different historic development would be of great interest, e.g. to North-American cities featuring CBDs or polycentric urban structures; however, the generation of a representative amount of households data on the individual perception might be the limiting factor. Further, since the current study focuses solely on the morphologic aspects of the urban spatial structure, we propose to extend the study by the integration of additional functional characteristics, e.g. employment density or commerce.

Chapter VII

Searching for the DNA of urbanization

A material perspective

Cities, 2024, vol. 151, 105079

Luis Inostroza & Hannes Taubenböck

© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

https://doi.org/10.1016/j.cities.2024.105079

Received 19 Oct 2023; Accepted 25 Apr 2024; Published 11 May 2024

Abstract

Broadly accepted categorical differentiations of urbanization understand cities as well-defined objects containing urban spaces in contrast to their hinterlands. However, urbanization's multidimensional complexity challenges these approaches in the context of increasing social issues marked by rapid urban expansion, uneven development, ways of life, inequality, commodification, etc., that require fresh scientific answers grounded in innovative empirical evidence. Here, we analyzed the population-based and the land-use/land cover-based categorical understandings of urbanization, looking at their origins and main shortcomings. Our analysis makes a generalized description of urbanization's spatial complexity, with an emphasis on the problematic spatial delimitation of urban boundaries; urbanization occurring in remote wild areas; and the missing third spatial dimension. We discuss these shortcomings based on recent scientific developments, providing reasons why the categorical approach needs to be changed and how. We propose a continuous indicator of urbanization which is based on the accumulation of anthropogenic materials, a physical, rather than a spatial or demographic characteristic. Our proposal allows the analysis of socio-ecological systems' spatial organization, pursuing comparative studies across geographies and times, informing globally generalizable patterns of urbanization processes and giving a material body to address claims for sustainable urban development.

Keywords

Urban development; GIS spatial analysis; sprawl; industrial ecology; technomass

1 Conceptual grounding

What is urbanization? Most often it has been understood as a process of population shift from rural to urban areas. This very simplified understanding, however, is complicated by conceptual imprecisions – in spatial, thematic and temporal terms. This unclear conceptual foundation upon which the urban analysis apparatus is built has been posing tremendous practical challenges to urban research for decades (Castells, 1977), without substantial improvements reflected in how urbanization has been conceptualized and measured (Bai et al., 2017).

A stronger conceptual grounding of urbanization remains a shared claim across different urban epistemologies, from archaeology (Ortman et al., 2020; Smith, 2015) to urban ecology (McPhearson et al., 2016; Pickett & Zhou, 2015), while searching for clarifying effects on the practical side of urban research (Brenner & Schmid, 2015). On the one hand, in disciplines such as urban studies, urban economics or urban planning, urbanisation is understood as an exclusive social phenomenon (Batty, 2013; Storper & Scott, 2015) without any ecology. However, in the context of increasing social issues marked by uneven development, inequality, ways of life, commodification, etc. the multidimensional complexity of urbanization challenges such a socially biased urban analytical apparatus. To tackle such challenges, we call for fresh scientific evidence to advance meaningful scientific answers. On the other hand, increasing global issues and scientific concerns about manifold ecological problems, including climate change, land consumption and degradation, pollution, loss of biodiversity and ecosystem services, etc., have been linked with urbanization (Haase et al., 2018; Seto et al., 2016). These negative ecological effects are evidencing that urbanization cannot be considered purely as a social process. We claim here that re-examining the theoretical grounding is in demand to provide new innovative conceptualizations enabling better integration of social and natural sciences. This integrative approach can be central to advance urban research and providing meaningful answers to the current environmental crisis. The pathways for addressing these environmental challenges require changes in the current way of conceptualizing the relationships between socio-cultural constructions – i.e. the concepts used to understand reality- and material processes – and the idiosyncratic phenomenologies possible to observe (Hornborg, 2001).

Current conceptualizations of urbanization have relied on broadly accepted categorical differentiations in which cities are conceived as well-defined objects containing urban spaces and in contrast -in terms of drivers and processes- to their hinterlands, the so-called rural areas. Conceptually speaking, these categories reflect the Western intellectual tradition's bias of isolating society from nature (Alberti et al., 2003; Lin et al., 2012; Young, 2009). Thus, empirical studies are biased towards spatiality. The attempts to find urban boundaries fail in this conceptual trap, which is resolved by using context-specific spatial assumptions that are limited to performing comparisons across different geographical settings or historical times. Such spatial units, whether generated administratively or data-driven, have their idiosyncratic justifications. They serve as the basic information vehicle for scientific research, policy development and planning processes. These spatial reference units are operationally so powerful that they shape the current understanding of urbanization. However, grasping the full picture requires looking beyond such specific 'urban' spatial

units because the impacts and dependencies arising from urbanization go far beyond the limits of such urban spatial units. Furthermore, current spatial empirical attempts do not unveil the true continuous spatiotemporal nature of urbanization (Inostroza et al., 2019) involving not only a social but also an ecological dimension that are both shaped in a very material sense. The DNA of urbanization, what makes urbanization possible as such, is what we discuss in this piece.

The remainder of this essay is as follows: Firstly, we refer to the population-based and land-use land cover-based understandings of urbanisation. The emphasis here is on how the categorical understanding arises and why it needs to be changed. We revise the complexity and challenges posed by urbanisation. These reflections, which are based on empirical findings, discuss mostly the shortcomings of using categorical variables to define urbanisation; spatial complexity, with an emphasis on spatial delimitation; urbanisation occurring in remote wild areas; and the missing third spatial dimension of urbanisation. In Section 3, we provide an overarching analysis of the morphogenetics of urbanisation, to illustrate the relevance of its material structure. This overview provides the basis for an analysis of urbanisation with a focus on the material dimension, which is discussed in Section 4. We conclude that rather than the spatiality or the demography, urbanisation requires a particular materiality, and it is here where the DNA of urbanisation is to be found.

2 The urbanization we know and measure

Over the last 50 years, the understanding and measurement of urbanization has mostly relied on population change. The second approach used to analyses and measure urbanization comes from spatially explicit sciences, such as GIS and remote sensing, in which the definitory aspects are land-use/land-cover change (LULCC).

2.1. Urbanization as population change

The demographic understanding of urbanization was first proposed by King Davis, an American sociologist who developed a systematization of global population statistics. Over this basis, the UN population statistics were developed (Brenner, 2013) using administrative borders that often do not match urban areas' physical extension. Furthermore, to differentiate what is 'urban' and what is 'rural', the demographic approach requires cutting population values that each country defines arbitrarily. Indeed, there is a large variability of the threshold values separating urban from rural settlements (Inostroza & Zepp, 2021; Taubenböck et al., 2019). For instance, in some Latin American countries, the population threshold for an area to be considered urban is 2500 inhabitants; in other countries, the threshold is 5000 inhabitants; and in India, the threshold is 100,000 inhabitants. Such a definition of what urban is leads to inconsistent, non-comparable statistics originating measurement problems (Inostroza et al., 2019; Taubenböck et al., 2022). Surprisingly, such statistics are today broadly cited in almost every article dealing with an urban issue.

In a strictly phenomenological sense, there is nothing special about these population values: our current understanding of urbanization remains without a clear analytical distinction. Rather than being an empirical fact, what is today called 'urban' is floating on thin air. Such

arbitrary demographic assumptions leave the most common sentence currently used in scientific articles – "today more than 56% of the population is urban" – meaningless, without analytical power to define urbanization. Criticism of the demographic approach to urbanization dates back to Castells, who, more than 50 years ago, called this approach "*the statistical empiricism in the delimitation of the concept of urban*" (Castells, 1977):10). The ground for new conceptualizations of urbanization needs to be fertilized with fresher innovative concepts.

2.2. Urbanization as a Land-use Land-cover change

Land-use Land-cover change (LULCC) is considered a key aspect of urbanization (Bai et al., 2012; Grimm et al., 2008). Thus, LULCC methods are commonly used in spatiotemporal assessments of urbanization (Grimm et al., 2008; Schneider & Woodcock, 2008). From this strict operational perspective, land can be classified as urban or nonurban by using a Boolean approach, which leads to a categorical description of urbanization that does not allow for direct quantitative comparisons between different geographies and time (Inostroza et al., 2019).

Homogeneous land cover patterns are obtained using remote sensing methods that delineate discrete land units that compile several land surfaces in a categorically coarse manner (Anderson et al., 1976). This is the approach used in broadly known data sets, like the Corine land cover (https://www.eea.europa.eu/publications/COR0-landcover) the Urban Atlas (https://land.copernicus.eu/local/urban-atlas), the Global Human Settlement layer (Pesaresi et al., 2013), the Global Urban footprint (Esch et al., 2012), or Local Climate Zone classifications (Bechtel et al., 2015; Zhu et al., 2022). These datasets use the same homogenization principle to combine artificial and natural surface types, classifying space into different categories that are mutually excluding. The problem lies in classifying space rather than measuring phenomenological changes continuously.

2.3. The shortcomings

Urbanization has been understood based on urban and rural spatial classes. Such categories describe a set of idiosyncratic features found in particular places that are considered mutually exclusive. These categories cannot describe the continuous spatial complexity of urbanization, hiding the multi-scale and multidimensional interrelations of urban cores, peripheries, hinterlands and more. Thus, these categories are not an adequate measure of urbanization with all its dependencies in time, space and processes (Inostroza et al., 2019). They remain as common categories and also important landmarks in political and planning discourses (Parr, 2007). In science, it is also a common strategy to use categories to provide a better understanding of complex physical phenomena. One can understand this as a permissible tool to make the complexity of the world tangible. For instance, scientific disciplines addressing temperature make categorical qualitative statements like hot or cold. However, to measure temperature, we use a continuous scale and unit—Celsius, Kelvin or Fahrenheit degrees—for which science provides frameworks to perform such measurement trivially. In the measurement of urbanization and its spatial complexity and impacts, a truly

continuous measurement based on a physical – i.e. material – property is still widely missing. In consequence, urbanization's DNA in a material sense is missing in current analysis.

A scientific measurement of urbanization does not need to find an urban space in opposition to a rural one, but rather to measure the urban intensity on a continuous scale. That is a scientific outcome that can serve as a basis for further analysis, where categorizations might fall out of what can be properly called scientific and be more appropriated in a policy context (statistical urban for instance).

The categorical understanding of urbanization is challenged by three overarching features: (1) urbanization does not follow strictly monocentric urban development patterns but rather a dispersed manner. With it the standard administrative and spatial definitions used to measure urbanization are blurred at several scales (Brezzi & Veneri, 2015; Kloosterman & Musterd, 2001); (2) urbanization is not constrained to urban areas and their hinterlands but rather spreads over all types of landscapes, including remote areas, where it is found in completely different spatial forms of anthropogenic alterations and modifications (Parry et al., 2014; Salgueiro Barrio & O'Shea, 2022); and, (3) urbanization has a fundamental third spatial dimension normally unacknowledged in standard population or LULC-based approaches due to the bi-dimensionality of the used indicators (Nichol et al., 2013).

2.3.1. Blurred urban limits

The limits of administrative urban areas do not reflect the real area covered by cities and therefore can mislead any space-based calculations like population density or per capita green areas. Commonly, administrative areas might include large agricultural and forest land and some other surrounding natural areas biasing comparative approaches between different cities. Indeed, the mismatch between the physical extension of cities and the respective delimitation of administrative units (Fig. 1) biases spatially based comparisons (Taubenböck et al., 2019). Figure 1 illustrates this mismatch where neither the administrative area of the metropolitan municipality nor of the province follows the settlement patterns. To avoid these shortcomings researchers should delineate the real extension of the urban area before performing spatial assessments (Kabisch & Haase, 2013).

Research studies dealing with the urban heat island (UHI) effect provide another good example: The UHI accounts for the difference in temperature between cities and their hinterlands (Oke et al., 2017). Here the causality of the UHI lies in the physicality, such as high building density, paved surfaces, building materials, etc., that do not match the peripheral urban spatial units. To avoid this methodological shortcoming, the continuous urban fabric must be strictly separated, before spatial indexes for analysing the UHI or green space patterns are estimated). The use of administrative units without the spatial delimitation of the physical extension of the urban continuum brings relevant bias to urban analysis, whether for urban expansion, green space, UHI, or many other aspects.

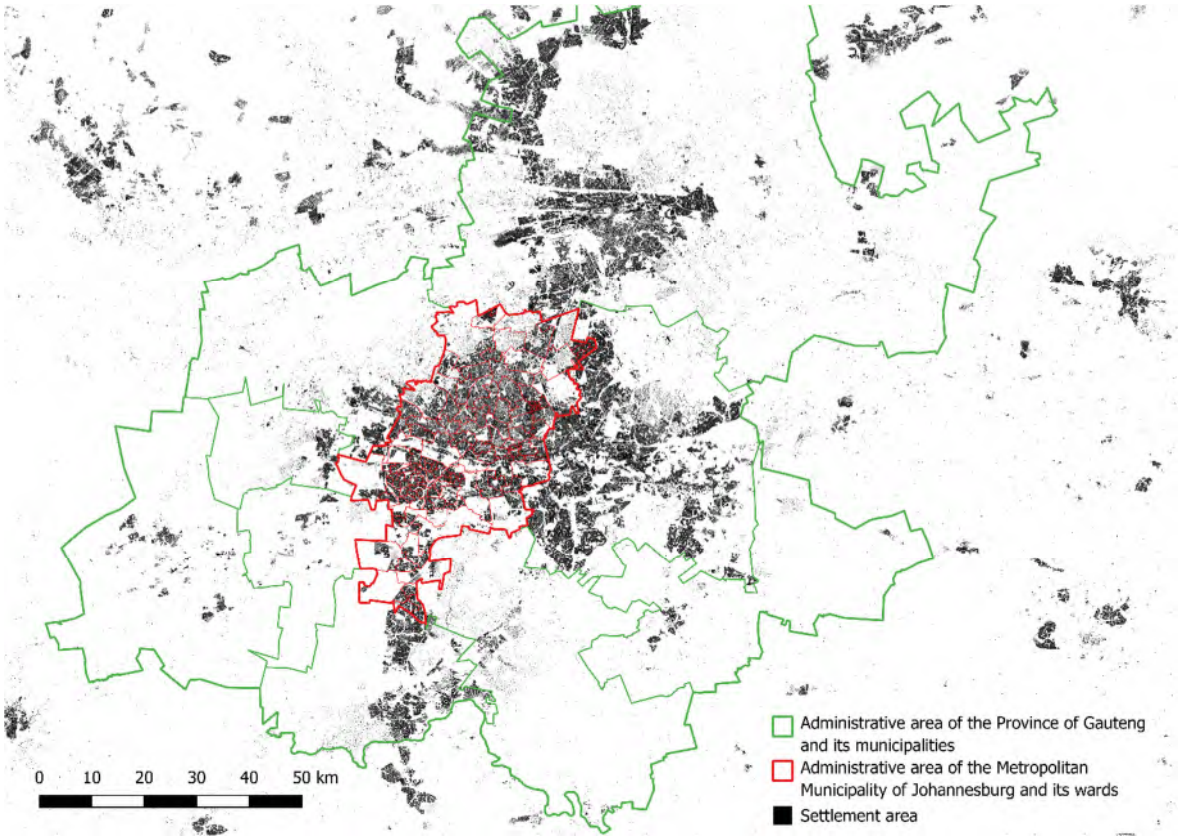


Fig. 1. The province of Gauteng and its municipalities. Source: Administrative units are taken from the GADM – Database of Global Administrative Areas data (2018), the settlement areas are taken from the Global Urban Footprint mapping product (Esch et al., 2012), own elaboration, unpublished.

2.3.2. The urbanization of the last of the wild

The planetary scope of urbanization makes the current urban-rural categorical differentiation highly insufficient for describing its spatial structure and the urban colonization of remote wild areas. This categorization introduces an incompleteness or even a bias to analytical results related to the common conceptualizations of a spatial urban category.

Qualitative and quantitative spatial analysis understands urbanization as a dichotomy between LULC classes, where economic activities not located in cities are not considered to be urban. For example, current land systems science does not classify mining camps, isolated industrial complexes, industrial farming, refugee camps, small settlements, roads, etc., as urban land use categories, as can be seen in analyses of low technomass amounts linked to urbanization degrees, as illustrated in Fig. 4. But these are places where the urban process is co-determined materially, economically, ecologically, but also socially and politically. Thus, the current understanding of urbanization is simplifying the multi-dimensional relations of the urban on local to planetary scales.

2.3.3. The missing third spatial dimension

Common urbanization assessments rely mostly on 2-D LULC indicators where patches of LULC are used to describe spatial heterogeneity (Forman & Godron, 1986; Forman, 2014; Zhou et al., 2017). While these assessments only use one or two spatial dimensions, urbanization as a process unfolds in three spatial dimensions and time (Wentz et al., 2018).

From a phenomenological perspective, the 3-D traces of urbanization are to be found in buildings but also in technological infrastructure including roads, train tracks, electrical lines, water and sewer pipes. These are on the contrary cartographically reflected in one dimension and thus difficult to include in spatial analysis vis a vis dominant 2-D entities (Inostroza et al., 2019).

Increasing attention has been given to building density in its 3-D representation which has produced interesting findings (Lin et al., 2012, 2020). However, to date, these studies mostly remain at a non-multi-temporal stage. To continue in this promising direction such at tempts should take a stronger compromise with real continuous measurement, abandon the replication of urban-rural dichotomies, and work towards 3-D multi-temporal studies. Nowadays, artificial intelligence can substantially contribute to 3-D and 4-D analysis and calculations.

3 The material spatiotemporal behavior of urbanization

In the previous section, we have analyzed the conceptual and operational shortcomings of current standard approaches to urbanization. What is at the core of urbanization? Where can we find its DNA, the structure determining its reproduction and perpetuation in time? Ac cording to (Scott & Storper, 2015), the core of the urban definition is "the urban land nexus". While this approach brings space as a fundamental component of the urban definition, it does not provide any clue for its measurement, remaining trapped in the spatiality that we have criticized in Section 2. Furthermore, space is a high-level abstraction that doesn't allow for the necessary inclusion of qualitative change that is fundamental to understanding urbanization and its impacts, and the underlying ecology that follows the laws of thermodynamics. We acknowledge that urbanization is materialized in space, and relates to transformations of ways of life, and social or economic structures, among other issues (Nassehi, 2002; Tonkiss, 2013) encompassing a complex, even confusing morphology. Figure 2 shows examples of urban complexity unfolding in manifold different urban forms in different places around the world. Such great morphological complexity makes it difficult to derive generalizable results or comparisons. For instance, in the case of Trento's surroundings such urban morphology might be archetypical for northern Italian landscapes but not necessarily generalizable to all European landscapes.

While built-up structures encompass paramount volumes of anthropogenic materials that are accumulated and remain for decades or centuries, we argue that the morphogenetics of urbanization produces a context-specific material structure, archetypical material assemblages composed of infrastructures and buildings, that precede and underpin other development forms. This materiality is a common characteristic of urbanization across historical times and geographies that has been measured even in archaeology (Ortman et al., 2020). All these different kinds of humanly processed matter are what we call 'technomass' (Hornborg, 2001; Inostroza, 2014). Entangled with all types of designed and remnant green space, they account for the very urban material structure, that follows a set of regularities (Lemoine-Rodríguez et al., 2020; Ortman et al., 2020). The following section looks at the material spatiotemporal behavior, i.e., how the material structure unfolds in space and time, in short (inner) and long (extended) distances.

Fig. 2. Urbanization possess a particular phenomenology. Examples from different cities. Source: own elaboration, unpublished.

3.1. Omnipresent urban expansion made of matter

Urban expansion is currently one of the most powerful forces modifying landscapes, spreading around existing urban cores, developing in a polycentric manner, or spreading out at larger distances (Angel et al., 2011; Taubenböck et al., 2012). Urban expansion is a spatial process that implies the increase of the physical size of cities—surface under a 2-D approach and material stocks under a 3–4-D approach—into surrounding land areas.

While urban landscapes feature regional or cultural specifics (Taubenböck et al., 2020), the arising expansion patterns are common and lie beyond social or cultural context and have been characterized as a global homogenization of urban form (Lemoine-Rodríguez et al., 2020). At the same time, urban expansion implies a permanent increase in the built-up material structures that make up the body of urbanization. Materials have been acknowledged as one of the six fundamental components of urban form (Wentz et al., 2018). Furthermore, the expansion of urban areas implies a direct higher, faster consumption of raw materials that are accumulated in old and new urbanized areas but also in natural landscapes (see next section).

It stands to reason that there is a direct link between urban expansion and their respective material accumulation rates and volumes. This has, however, not yet been acknowledged neither by urbanization studies nor by urban metabolism analyses that considers urban systems to be in a steady state concerning mass and energy (Baccini & Brunner, 2012). We argue that traditional assessments of urban expansion should be linked with urban

metabolism analysis to ascertain the dynamics of material accumulation at the local, regional, or even planetary scales and interrelations.

3.2. The extended urbanization found in the last of the wild

Buildings, roads, infrastructure, machines, etc., are the material characteristics of urbanization regardless of their location, even considering that they are spread across the whole Earth. Indeed, while different material intensities are to be found all over the world, many of them are not considered properly urban, like mining camps and isolated infrastructures (Brenner, 2013; Inostroza & Zepp, 2021). However, there is an extended type of urbanization that makes resource extraction possible in locations that are far away from urban cores and is occurring even in remote areas of the world. Urban areas are always connected to areas of resource extraction through particular economic activities (Inostroza, 2018; Inostroza & Zepp, 2021). Classifying the urban tissue as an ultimate and independent attribute of space neglects its entanglement with the non-urban ecosystems providing the material basis for urban life.

This broad scope of dispersed urban fragments belongs to a planetary urban fabric that connects even the remotest places to an urban network of material and energy flows mediated by monetary transactions (Inostroza & Zepp, 2021). Analyzing the traces of urbanization in remote areas using the background of industrial ecology will acknowledge where energy and materials consumed in urban areas come from. The extended urbanization is both, an urbanization pattern in itself and the necessary condition that allows the building, sustaining and reproduction of the urban space.

3.3. The production and reproduction of the urban material structure

Spatially, there are regularities followed by urban systems all over the world. In the process of urban expansion, homogenization of urban form has been observed (Lemoine-Rodríguez et al., 2020). However, looking at the materiality of urbanization patterns, it is possible to observe profound socio-ecological asymmetries emerging within urbanization, either in the materiality of the grey or the green (Fig. 2). Such asymmetries have been addressed in a large body of scientific literature concerned with environmental justice, inequality, informality and the like (Behrens & Robert-Nicoud, 2014; Castells-Quintana & Royuela, 2015; Kabisch et al., 2017). Urbanization is internally a process of spatial differentiation, what (Scott & Storper, 2015) have called the urban land nexus, that segregates different social classes with distinctive materiality. Both processes occur simultaneously producing a combi nation of materials in specific volumes, rates and intensities that ultimately give rise to the apparently unique but fundamentally similar urbanization patterns seen in different parts of the world and even historical times (Lemoine-Rodríguez et al., 2020; Ortman et al., 2020). For instance, technomass would allow the identification of trends and specificities of urbanization concerning disparities between regions due to differential material means resulting from natural capital availability, trade and institutional capacities, technological development and levels of social complexity, among others. Such an approach could bring light to pivotal periods like the last century, a period marked by planning analyses and regulations prevalent across the world.

Indicators linking these two contradictory aspects, the spatiality, and the materiality of urbanization, are operationally needed to unveil the regularities and homogenization of urban form (Lemoine-Rodríguez et al., 2020). At the same time, however, it is needed to consider the socio-ecological asymmetries that are produced. We argue that indicators integrating the spatiality as bounded to the materiality reveal the urbanization's DNA. Furthermore, materials have a lower degree of abstraction than space, making them more tangible to be integrated into different scientific approaches, as has been shown in archaeological research on proto-urbanization (Ortman et al., 2020).

4 The material body of urbanization

Urbanization produces a material accumulation that occurs in a highly structured manner contained in archetypical material assemblages. The aim of what we call '*technomass*' is to provide a meaningful scientific continuous measurement of such assemblages. At the same time, archetypical material assemblages possess a set of structural properties such as 1) *particular shares of building materials* like glass, steel, concrete, plastics, and the like that can be identified for different locations and dates. These ratios can be easily obtained based on standard cubage calculations from engineering sciences or using general assumptions (Haberl et al., 2021; Tanikawa et al., 2015; Tanikawa & Hashimoto, 2009); 2) the *energy profile* they imply, for cooling, heating and other forms of energy consumption and embodiment (Gros et al., 2014; Palme et al., 2018); 3) the *labour power* they contain (Ortman et al., 2020); 4) the *economic (monetary and non-monetary) value* they imply (Pettit et al., 2020); 5) a specific *material quality and complexity*; and 6) an *emerging urban morphology* that encompasses a rich phenom enological description of urban environments. We argue that these properties can establish meaningful connections with many scientific disciplines dealing with urbanization and sustainability issues and can be linked to urbanization impacts, like biodiversity loss, air pollution and greenhouse emissions and manifold outflows of wastes and emissions.

'*Technomass*' is an indicator that embodies such structural properties. It is conceptually and operationally different from other indicators that have been used in industrial ecology (Fig. 3), accounting for material stocks (Haberl et al., 2021; Kalt et al., 2021; Miatto et al., 2024). This analytical distinction between technomass and material stocks can be illustrated with the example of a building's demolition, which encompasses a phase change from a standing building (archetypical material assemblage) to a pile of demolished materials. Demolition entails a change in the particular way materials are assembled, while their amount remains essentially the same. However, a fundamental shift in the material properties takes place. For instance, a building as an archetypical material assemblage, embodies and communicates (matter that carries information) its own teleology, as a purposive designed object to fulfil a specific function, and it produces specific impacts accordingly. It possesses lower entropy and higher economic value, and is operationally expressed therefore using the material assemblage as a carrier of the unit of measurement, i.e. tons per hectare of technomass contained in buildings, roads, etc. (Inostroza, 2014). On the contrary, a material stock is a pile of materials without teleology (random), higher entropy and lower economic value. It is expressed in the weight or volume of single disaggregated materials, like plastic, concrete,

iron, copper, etc. Therefore, a material assemblage is, by definition, different from a simple material stock. Meaningful definitions matter while addressing the DNA of urbanization.

4.1. Urbanization is a continuous spatiotemporal process

Contrary to a dichotomic approach such as urban-rural, a continuous measurement accounting for the differential degrees and intensities of urbanization is the cornerstone to analyzing and measuring its impacts (Inostroza et al., 2019). For instance, the concept of a suburban area – without conceptual agreement within the scientific community as to which this particular space is exactly meant by this term and thus a comparatively abstract construct – it does as such not provide any clue of the type of urbanization patterns it contains. On the contrary, an urbanization degree in the order of 10,000 ton*hectare^{-1} allow several scientific disciplines to understand and quantify a broad set of social and ecological impacts, such as UHI, carbon emissions, impact in water cycle and the like. However, urban science at large is still missing appropriate indicators to measure urbanization on a continuous basis using continuous variables that reflect a measurable physical property of urbanization. Advancing such indicators is crucial in developing a new science of cities.

Fig. 3. Demolition in Brussels. Source: author's personal archive.

4.2. Detectable material and social structural asymmetries through technomass differentials

The urban space is produced through socio-ecological processes, in particular urban conditions that are produced and reproduced by specific actors assuming particular roles in producing socio-ecological configurations (Heynen et al., 2006). Two overarching types of socioecological asymmetries can be mentioned: (1) asymmetries between social classes,

with the encapsulation of those few appropriating and accumulating the benefits, such as better material conditions or access to green space; (2) asymmetries between urban tissues within the same city, where particular areas developing capsules of prosperity co-exist with areas falling into obsolescence and deterioration. Thus, it is replicating and deepening internal economic and socio-ecological asymmetries. Such asymmetries are usually higher within the same city than between different cities. In the distribution of benefits and impacts, asymmetries are more internal than external to urban areas. The materiality of urbanization is one way to unveil such socio-ecological asymmetries.

4.3. Urbanization is made of technomass

Urbanization is a spatiotemporal process producing and reproducing a socio-biophysical transformation of space and leaves phenomenological traces in the form of material structures. Urbanization unfolds in very specific material assemblages, purposive objects like buildings and infrastructures that are highly complex themselves at smaller scales. Such urbanization traces can be measured by weight and/or volume (ton, m^3) per unit of space (m^2, hectare, km^2) and time (per minute, per day, per year). Technomass is a quantification of the materials that have been already accumulated and can be directly measured in an aggregated manner. A differentiation into different materials using standard knowledge from engineering is also possible (Inostroza, 2014). Technomass quantification includes only materials that have been trans formed by human labor, being this a physical or chemical transformation that is always purposive, i.e. it does not include any excretions nor organic material that can be accounted as biomass or energy.

Technomass describes the continuous spatial structure of urbanization (Inostroza et al., 2019) based on a physical property and not in arbitrary spatial or demographic assumptions. Continuous measurement of urbanization can describe the broad scope of the spatiotemporal intensities of urbanization in a comparable manner, from distant geographies and times (Figs. 4 and 5). The indicator offers a solid operational alternative to other existing indicators that rely on land use change or similar two-dimensional representations. In this way, urbanization's materiality can be linked to manifold socio-ecological impacts. Conseuently, by definition and without further assumptions, this indicator provides a continuous measurement of urbanization based on a self-evident physical property that possesses a high comparative value. Technomass (Inostroza, 2014) allows a continuous, unambiguous and non-categorial measurement of urbanization that is not biased by socioeconomic and geographical diversity (Fig. 5) and can be applied in any type of landscape and city.

Fig. 4. Full spectrum of u.rbanization degrees found in southern Patagonia. Source: own elaboration, unpublished.

4.3.1. In landscapes

The biophysical conditions of landscapes have been largely analyzed (Wrbka et al., 2004). However, there is a disciplinary disconnection between studies that focus either on the natural or the built aspects that produce a dismissal of human-made infrastructures like buildings and infrastructures as components of landscapes (Inostroza et al., 2019). Even the smallest infrastructure in a remote area (Fig. 4) requires paramount amounts of technomass to be extracted, processed, trans ported and built, with manifold ecological effects that can be scientifically addressed by looking at this material dimension. However, technomass present in remote areas does not count as belonging to the urban space. Looking at urbanization as a material accumulation process allows for a continuous measurement that is also closer to reality and thus more meaningful than aggregated, dichotomous categories. This approach enables the systematic accountability of manifold manifesta tions of urbanization intensities, like differential infrastructures through landscapes, equipped

natural areas, farms, (Fig. 4) small and large cities (Fig. 5). Indeed, the construction materials' global demand has grown exponentially since the 1950s to be transformed into buildings and technological infrastructure that will remain stocked in landscapes for centuries (Krausmann et al., 2009; Steinberger et al., 2010; Wiedenhofer et al., 2015). It is necessary to use novel indicators (Fig. 4), to make them visible for landscape science and related disciplines.

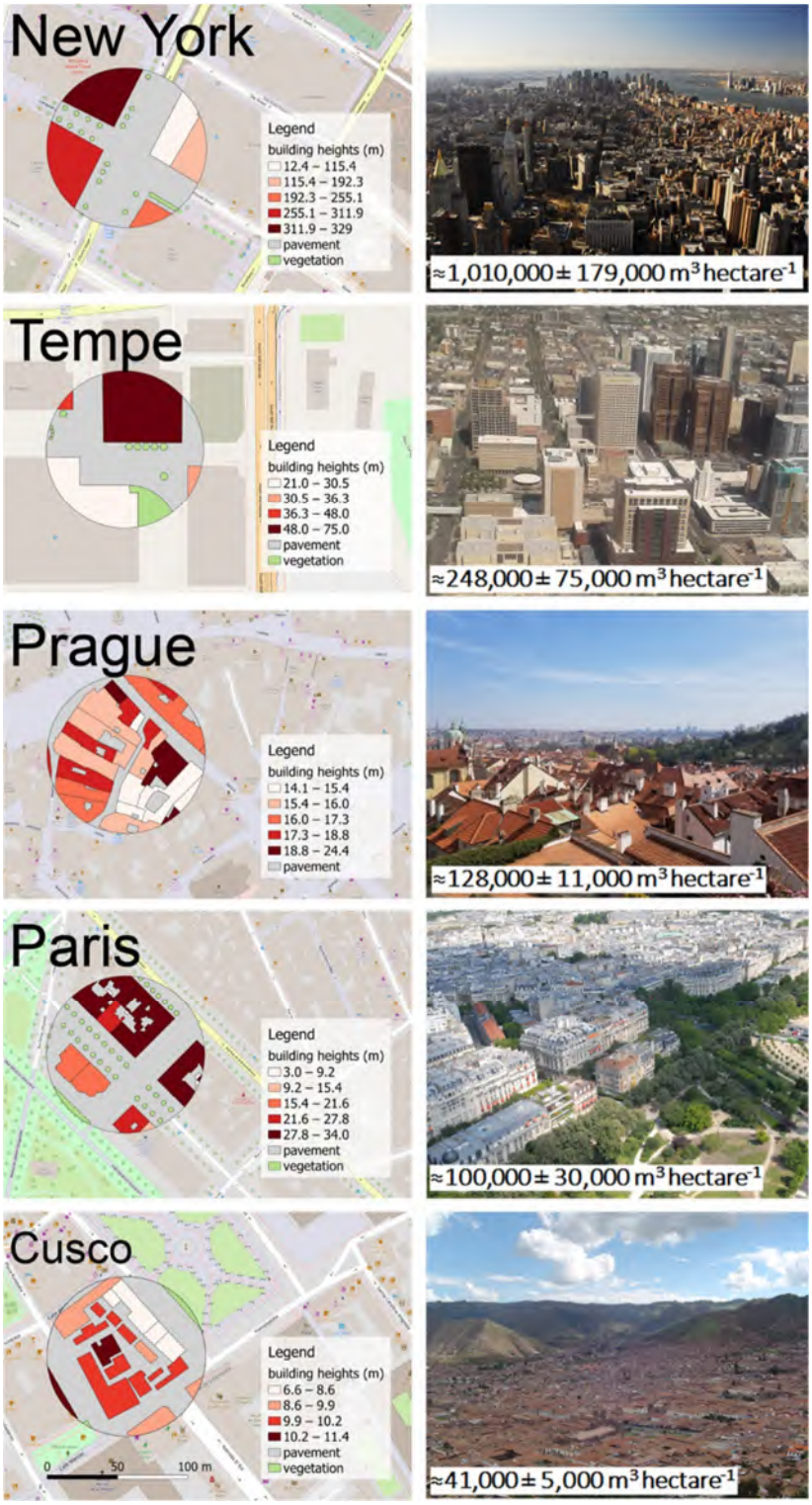


Fig. 5. Amounts of technomass found in 5 different cities around the world. Source: own elaboration over Google Maps ©, Google Earth Pro © and own datasets from previous research, unpublished.

Technomass and biomass in the form of built structures and natural elements are highly intertwined at several spatial scales. It produces a high material and social complexity of urbanization that can be dis closed by using technomass to unveil manifold material effects (Inostroza, 2018).

4.3.2. In urban cores

The high spatial, material, morphological and social complexity of urbanization occurs at very short distances in four dimensions. The capacity to capture the city in three physical dimensions has been greatly enhanced by recent developments in remote sensing (e.g. Frantz et al., 2020; Geiß et al., 2019; Li et al., 2020). However, the temporal aspect of 3D urban change remains mostly neglected due to data issues, and thus plays a subordinate role and must be given more attention (Wentz et al., 2018). The 'Technomass' makes it possible to measure different landscapes with their material assemblages in a single unit – from 1 to 10 tons per hectare in remote areas up to more than 1,000,000 tons per hectare in Central Business district urban cores (Fig. 5). This allows the DNA of the location to be decoded and measured consistently and uniformly.

4.4. Limitations

We have analyzed urbanization as a phenomenological material transformation of space, that allows humans to perform urban activities. DNA is a material structure, a molecule, that allows life its reproduction. We used the DNA as a metaphor to express that, anthropogenic materials, i.e. technomass, provide the necessary conditions for the repro duction of urban space and the continuity of all its functions. Therefore, the material conditions are determining the reproduction capacity of urban space.

We acknowledge that urban space is highly complex and multidimensional including manifold idiosyncratic manifestations that produce the great urban diversity that we observe in cities across the world. While technomass can provide a robust continuous description of the urbanization process, it cannot certainly encompass all angles of its complexity. Therefore, such material description might be complemented with other socio-cultural dimensions that concern other disciplines interested in urbanization.

5 Conclusion

Addressing the high complexity of the urbanization process in the XXI century requires conceptually and analytically different approaches than in the past. Urbanization has been understood mostly as a demographic and categorical phenomenon. In this essay, we have pro posed an understanding of urbanization as a continuous spatiotemporal process, without dichotomies and categorical concepts by using indicators that are based on a physical and continuous unit of measurement. We propose technomass, an indicator that measures the amounts of anthropogenic materials accumulated in space as a phenomenological continuous expression of the urbanization process. Our approach is both, data-driven and spatially explicit, not requiring any further assumptions and allowing a strong continuous representation of the complex het erogeneity and the observed spatial patterns at several scales. On this basis, the differential impacts of urbanization can be linked across the globe,

it can be spatialized and compared across distant places and time, looking for instance, to the impacts of UHI and technomass, or ecosystem services and technomass. Technomass is a process-based conceptualization where the fluxes of materials from source to sink can be analyzed and measured linking urbanization and sustainability. The strength of technomass measurement relies on its comparability, continuity, and independency of settlement morphologies. However, communications and networks are not always proportional to the technomass and are thus expressed in the form of flows rather than technomass which is a measure of stocks.

A robust operational framework to analyze, measure and ascertain urbanization can be constructed using the process of material accumulation over time—history in a broader sense—depending permanently on pre-existing ecological attributes—i.e., natural capital—located either in the hinterlands or even in faraway locations. Material accumulation is at the vortex of the urban organization, producing the very urban material structure allowing its reproduction. Rather than space, which is ubiquitous and highly abstract, it is anthropogenic matter what makes urbanization possible, observable and qualitatively idiosyncratic. It is an approach that unveils the underlying thermodynamics implicit in urbanization processes.

Technomass reveals urbanization's DNA, as unfolding in a continuous manner that follows regular, teleological and archetypical material assemblages that transcend geographies and time. The presence of technomass reveals urbanization's DNA indicating the presence of

humans, now, or in the past. Exactly like in the case of DNA, that indicates the current or past presence of an organism (Taberlet et al., 2018). This materiality is objective and can be directly measured and compared providing a solid ground for a theory of urbanization encompassing even the origins of cities as such. On the practical level, it can foster comparative international research on urban margins, new suburbs, that are no longer concentrated around cores but spread across increasingly vast urban regions shedding light on qualitative imbalances between the center and the periphery. At the same time, it can open new perspectives on spatial planning implications, the obsolescence of administrative boundaries, and the reinforcement of spatial planning for large areas or urban regions, whose dynamics and new resulting boundaries can be better addressed using technomass and material flow analyses. A continuous measurement of urbanization, i.e., a measurement based on a clear, physical measure, as we have proposed in this essay, makes evident the direct and indirect socioecological effects of urbanization, over the basis of robust, comparable and reproducible empirical evidence. Such an indicator can bring together different research communities in integrative approaches aimed at addressing the elusive sustainable urban development. This is an operational conceptualisation that places the process of material accumulation as the urbanization's DNA.

Part III

Urban dynamics

CHAPTER VIII	Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios	171
CHAPTER IX	The digital urban frontier: Disparities in social media activity between consolidated and newly urbanized areas in Africa	201
CHAPTER X	Border regions across the globe: Analyzing border typologies, economic and political disparities, and development dynamics	227
CHAPTER XI	NO_2 air pollution trends and settlement growth in Megacities	253

Chapter VIII

Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios

Cities, 2024, vol. 151, 105117

Hannes Taubenböck, Johannes Mast, Christian Geiß, Michael Wurm, Thomas Esch & Karen C. Seto

© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

https://doi.org/10.1016/j.cities.2024.105117

Received 19 Jan 2024; Accepted 11 May 2024; Published 4 June 2024

Abstract

Cities are expanding all over the world. In this study we aim to understand how the pace and scale of urban growth dynamics vary across geographies. This study contributes to the literature in four key ways: First, the study is comprehensive including all 1567 cities worldwide that had >300,000 inhabitants in 2015. Second, comparison of settlement dynamics of cities is performed in relation to continents, to geographic regions, to national levels and to city sizes. Third, a data-driven approach to delineate city boundaries allows for consistent comparisons across geographies. Fourth, the measured past urban trends are projected into the future and linked to climate scenarios used by the Panel on Climate Change (IPCC) to estimate where spatial focal areas of urbanization and climate-change meet. The study is based on remote sensing-derived geodata such as the World Settlement Footprint Evolution dataset for the period 1985 to 2015 and morphological urban areas. The key findings are: At continental scale, highest relative urban growth rates are measured in Asia. Among the IPCC regions, highest relative urban growth rates are in East Asia. In contrast, lowest growth rates are in Oceania and Europe. As expected, across most regions, there is a general decline in relative growth rates over the study period. One surprising finding is that absolute settlement expansion rates are not declining over time, but are rather relatively constant. It is also interesting to see that across and within countries growth rates of urban settlements vary widely in Asia, whereas in Europe there are only marginal differences. And finally, a look into the future reveals that certain regions can expect particularly high exposures. In South Asia and West Africa, for example, very high urban growth rates and a very high increase in predicted maximum rainfall are forecast. The greatest increase in extreme temperatures is predicted for Europe and North America. Settlement growth has already slowed there, so adaptation measures are taking center stage.

Keywords

Urbanization; urban growth; remote sensing; comparative urban research; future exposures

1 Conceptual grounding

Urbanization is one of the megatrends of global change. It is, as Zhang and Seto (2011) argue, simultaneously a demographic, economic, and land-use change phenomenon. The pull effects of cities, predominantly driven by economic developments (Glaeser, 2010), societal changes affecting the entire lifestyle (Sennett, 2018) and the ecological roles of cities both as causers but also as problem solvers (Bettencourt, 2007) testify to the crucial role of global urbanization processes.

The often-used slogan '*the future will be urban*' makes first of all clear that cities in general have key roles to meet development and climate goals. Urban areas are now home to >56 % of the world's population (United Nations, 2018) and they account for >80 % of the gross do mestic product (World Bank, 2022). However, urban areas also account for 66 % of global energy consumption (World Bank, 2022), and, according to the latest IPCC report, they contribute to two-thirds of global greenhouse gas (GHG) emissions (IPCC, 2022). The contribution of each city to these numbers depends, among many aspects, on the size of urban settlements as well as on their land use patterns. These spatial patterns of urban land, including the two- and three-dimensional form and structure, have significant impact on GHG emissions (IPCC, 2014), land consumption (Hu, Wang, Taubenböck & Zhu, 2021), land use efficiency (Schiavina et al., 2022), the urban heat island effect (Massaro et al., 2023), mobility patterns (Kang, Ma, Tong, & Liu, 2012), quality of life (Sapena et al., 2021), among many other issues. Thus, the spatial patterns of urban settlement growth influence a city's social, economic, and ecological conditions. However, the temporal and spatial dynamics (e.g. Angel, Parent, Civco, Blei, & Potere, 2011; Melchiori et al., 2018; United Nations, Department of Economic and Social Affairs and Population Division, 2018), the evolving size dimensions (e.g. Florida et al., 2008; Taubenböck et al., 2014), the geographic locations (e.g. Haase, Güneralp, Dahiya, Bai, & Elmquist, 2018; Zhou, Li, Asrar, Smith, & Imhoff, 2018) and the structural forms (e.g. Bechtel et al., 2015; Seto, Sanchez-Rodriguez, & Fragkias, 2010) of global urban settlement growth manifest differently in different parts of the world (e.g. Debray, Kraff, Zhu, & Taubenböck, 2023; Güneralp, Reba, Hales, Wentz, & Seto, 2020).

Numerous studies point to a rapid pace of urban expansion. The most commonly used data to measure urban expansion are population figures (Biłozor & Cie´slak, 2021; Liu, Huang, Xu, et al., 2020; United Nations et al., 2018). However, as many studies point out, urban population growth and urban land growth convey very different aspects of urbanization. A recent study of over 300 cities finds that urban population growth and urban land expansion are correlated, but this relationship varies for countries at different stages of development (Mahtta, Fragkias, Güneralp, et al., 2022).

Many studies document the dynamic growth of settlement areas (e.g. Angel et al., 2011; He et al., 2019; Liu et al., 2020) and its regional differences (e.g. Güneralp et al., 2020; Huang et al., 2021; Liu et al., 2020; Melchiori et al., 2018; among others). In terms of structural patterns, studies also reveal differences across the world (e.g. Angel, Lamson-Hall, Blei, Shingade, & Kumar, 2021; Bechtel et al., 2015; Taubenböck, Debray, Qui, Schmitt, & Zhu, 2020; Zhu et al., 2022; among others): in a meta-analysis, for example, Seto, Fragkias, Güner alp, and Reilly (2011) report that India, China and Africa have been the foci of global urban

land expansion between 1970 and 2000 and that land expansion rates are higher than or equal to urban population growth rates. Furthermore, Gong et al. (2020) measure China, the United States, India, Russia, and Brazil as the top five countries of impervious surface expansion. Zhang and Seto (2011) measure consistently higher rates of urban growth in India than in China for the time period 1992 to 2000 and vice versa from 2000 and 2008. Moreover, urban growth rates in these countries exceeded the rates of the United States and Japan. Huang and Xu (2022) derive, in general, lower growth rates in higher-income countries. Taubenböck et al. (2012) document highest growth rates in Asian megacities, followed by African and South American ones. However, Sun, Chen, Li, et al. (2020) note that there are still only few systematic and comprehensive analyses of these trends, especially for all cities and not just the very large ones.

In these studies on physical expansion and structural patterns of cities, remote sensing data is the core data source. The question of how urban areas should be delineated to provide a consistent spatial reference for comparative studies is the subject of considerable debate (see e. g. the works of Liu, He, Zhou, et al., 2014; Angel, Lamson-Hall, Galarza, & Blei, 2018; Melchiori et al., 2018; Dijkstra et al., 2021; Taubenböck et al., 2022; Inostroza & Taubenböck, 2024). Overall, the literature on urbanization is not consistent in their spatial references and comparisons within and between studies are therefore susceptible to statistical bias (e.g. French, Giacinti, Robinson, et al., 2022; Raciti, Hutyra, Rao, & Finzi, 2012). Geographic studies based on a data-driven, consistent delineation of cities across the world thus still remain few. The main reference units of measurement were administrative boundaries or standardized extents and to a lesser degree locally specific, data-driven delineations of spatial units. The former, however, do not allow consistent comparisons between the cities as they have grown historically and/or politically and are therefore not a spatially comparable spatial unit (Taubenböck et al., 2022). Commuter catchment areas, as one data-driven example, allow for consistent reference units (e.g. Heider et al., 2022); however, these remain challenging on global scale due to limited consistency and coverage of the data (Herfort, Lautenbach, Porto de Albuquerque, et al., 2023; Masucci, Arcaute, Hatna, Stanilov, & Batty, 2015). Recently, manifold methodologies using remotely sensed data have been developed for data-driven and thus consistent delineations of cities (e.g. Corbane et al., 2019; Lemoine-Rodríguez, Inostroza, & Zepp, 2020; Montero, Tannier, & Thomas, 2021; Taubenböck et al., 2022) allowing for conceptually coherent comparative urban geographic analysis on a global scale (Dijkstra et al., 2021; Moreno-Monroy, Schiavina, & Veneri, 2021; Taubenböck et al., 2019).

This study aims to contribute to existing studies on global urbanization in four ways: First, referring to the above, comparative urban research at city level has often been conducted using inconsistent spatial units (cf., e.g., Deuskar, 2015; Masucci et al., 2015). Here, we rely on a data-driven and spatially consistent reference unit delineating cities (Taubenböck et al., 2019), that is entirely independent of administrative units. Second, for a spatially differentiated assessment of settlement dynamics of cities across the world, we refer to various areas: to continents, to geographic regions as used by the IPCC, as well as to the national level. Furthermore, we relate settlement dynamics to city sizes and we relate the dynamics of capital cities to the urban growth rates of their respective countries. Herewith

we contribute a multi-scale geographical differentiation of urbanization dynamics. Third, many studies are limited to a sample of urban areas (cf., e.g., Seto et al., 2011; Kuffer, Pfeffer, & Sliuzas, 2016; Zhu et al., 2019). Our set of study sites consists of all cities worldwide that had >300,000 inhabitants in 2015. Fourth, we project the past urban trends into the future and link them to climate scenarios from the IPCC. In doing so, we aim to spatially and quantitatively estimate where which urban growth rates will coincide with which temperature and precipitation changes. Studies modelling future urban projections are emerging (Acuto et al., 2018; Chen, Li, Liu, et al., 2020), but they remain underrepresented in literature. In our study we aim to add to the literature by localizing where spatial focal points of urbanization meet projected focal areas of climate-change.

By addressing these four issues, we aim to reduce knowledge gaps by a spatially coherent analysis of global urbanization trends from the past. And, by projecting these urban growth trends to the future, we aim to sharpen knowledge with respect to climate change scenarios.

2 Data

The data sets for analyzing urban growth are (1) the *World Settlement Footprint Evolution (WSFEvo)*, (2) the *morphological urban areas* as reference units and (3) *geographical reference areas*. To localize areas where increased urban growth will encounter particularly intense climate change, the study is based on (4) *IPCC projections*. And, in order to be able to generate a statement about how many more people can be expected to live in these areas, (5) *UN population forecasts* are used.

The *WSFEvo* dataset maps the extent of human settlements on a yearly basis from 1985 to 2015. It is derived from Landsat im agery at a 30 m spatial resolution (Marconcini, Gorelick, Metz-Marconcini, & Esch, 2018). Using the settlement area in 2015 as a baseline, an iterative classification based on the temporal signatures of several spectral indices is retroactively applied to determine the built-up areas for each pixel and each year between 2015 and 1985. This method is suited to capture the growth or persistence of settlements; however, it needs to be noted that this approach is not capturing their shrinking. The availability of Landsat data varies across the world and over time, leading to a varying confidence of the mapped settlement extent. Accordingly, the WSF Input Data Consistency Score (IDC) layer is used to exclude data points where only few scenes were available at a particular year for the respective sample city. Figure 1 illustrates the expanding settlement patterns as mapped by the WSFEvo by the examples of the metropolitan regions of the Pearl River Delta (PRD M.R.) in China, Tokyo in Japan, and Jakarta-Bogor in Indonesia. These three examples impressively show how dynamically some regions have grown (PRD M.R.) and how little horizontal urban expansion has taken place in other urban landscapes (Tokyo).

(1) The analysis is based on a spatial reference unit that delineates cities across the globe in a data-driven way. Based on the Global Urban Footprint (GUF) and the GUF Density (GUF-DenS) data sets (Esch et al., 2018), city boundaries have been delimited for the year 2015 (for details on data and methods, we refer to Taubenböck et al., 2019). Therefore, the density of the built landscape along an urban-rural gradient is used to delineate ‘the urban’ from ‘the rural’ hinterland. Of course, this urban-rural transition is

fluid and there is no consensus on where 'the urban' ends and 'the rural' begins. One benefit of using a mathematical approach is that it creates spatial units in a consistent way that make cross-city comparisons more credible. These units are called '*morphological urban areas*' (MUAs), as illustrated in Fig. 1. In this approach, if two cities intersect in space, they are combined into one metropolitan region within the MUA. Hereafter, these cities will be referred to as Metropolitan Regions (M.R.).

This study is based on 1567 cities. This selection relates to the freely available city boundaries from the study by Taubenböck et al. (2019). Originally, all 1692 cities worldwide that had >300,000 inhabitants in 2015 were included in the analysis (United Nations, & Department of Economic and Social Affairs, 2015). Due to the aforementioned combi nation of coalesced cities into metropolitan regions, the quantity is reduced to just these 1567 cities (Fig. 1a).

(2) For a spatially differentiated analysis of settlement dynamics of cities across the world, continents, geographic regions, as well as nations are used. As continents are by no means homogeneous spatial units, the spatial reference analysis is complemented with geographical regions (RC10) as categorized by the IPCC Working Group II in its Fifth Assessment Report (Krey et al., 2014) (cf. Fig. 1b). Of course, these geographic regions are also not homogeneous spatial units; however, they subdivide the land masses according to similar features in geographical, political, economic, legal, religious, cultural, linguistic and artistic characteristics and therefore suggest that similar processes and dynamics exist. Thus, they offer one analytical structure for evaluating and comparing settlement growth rates of cities. We acknowledge that any definition of geographical regions over simplifies places, as these are partly fluid, intertwined and/or overlapping. However, our basic research question whether certain continents or geographical regions actually have different urban growth trajectories, can be examined in this manner.

(3) For data on future climate stress, projected changes in annual maximum temperature and annual maximum daily precipitation are used. These data are provided by the Coupled Model Intercomparison Project (Eyring et al., 2016) in support of the sixth IPCC assessment report. The data are available via the IPCC-WG1 Atlas (Iturbide et al., 2020; Gutierrez et al., 2021).

(4) In order to be able to make statements about the number of expected additional exposed city dwellers, the population projections of the United Nations are used. Here, it is relied on the World Urbanization Prospects (United Nations et al., 2018).

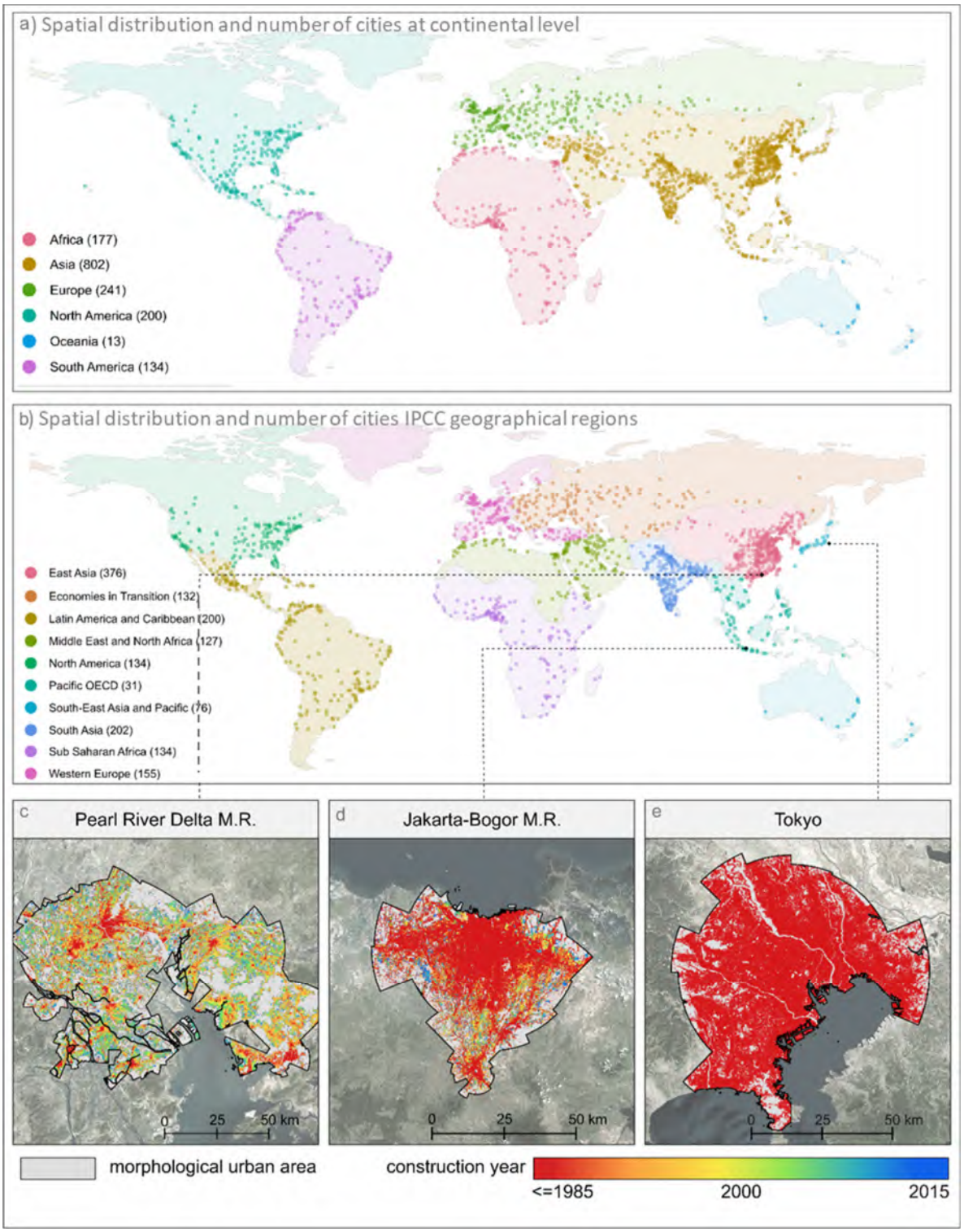


Fig. 1. Top: The set of the 1,567 largest cities across the globe and their assignment to (a) continents (in brackets the number of cities) as well as to (b) the IPCC geographical regions. Bottom: Settlement extent and expansion based on the WSFEvo from 1985 to 2015 and the morphological urban areas (MUAs) as reference units as of 2015. The data sets are exemplified by the three most populous metropolitan areas of the world: The metropolitan regions of the Pearl River Delta in China (c), of Jakarta-Bogor in Indonesia (d) and of Tokyo in Japan (e).

3 Methods

3.1. Indicators for measuring urban growth trajectories

We seek to quantify the settlement growth patterns of cities using a variety of indicators, each selected to highlight a particular aspect of growth histories.

For the measurement of urban growth trajectories, the settlement extent per year as a fraction of its settlement extent in 2015 within each MUA (Growth to Target, GTT) is calculated, i.e. the 2015 settlement extent is set to 100 % per city. As we rely on annual settlement extents, the relative share of the 2015 settlement extent in the previous years is derived in a backwards analysis (see formula (1)). By using the extent at the end of the time period as the baseline, growth histories of cities can be compared irrespective of their absolute sizes.

(1) $G_{TT}(y) = \frac{X_y}{X_{2015}}$

where X_y is the settlement extent in year $y \in N \cap [1986, 2015]$.

Changes in settlement growth rates over time can be highlighted by calculating the growth in each year in relation to the current extent at that point in time. The relative growth rate for a year (G_{YR}) is calculated as the increase in settlement area compared to the previous year. To increase the stability towards noise, a rolling mean of a three-year window is applied (formula 2):

(2) $G_{YR}(y) = \frac{1}{3} \times \left(\frac{X_{y-1} - X_{y-2}}{X_{y-2}} + \frac{X_y - X_{y-1}}{X_{y-1}} + \frac{X_{y+1} - X_y}{X_y} \right)$

Likewise, the absolute yearly settlement growth rate in a three-year window (G_{YA}) can be calculated as follows in formula 3:

(3) $G_{YA}(y) = \frac{X_{y+1} - X_{y-2}}{3}$

At the beginning and ending of our monitoring period, a two-year window is used. Note that no growth rates can be calculated for the year 1985, as it includes all built areas for that year, and not just the growth of that year.

In addition to the past urban growth trajectories, the measured urban growth trends from 1985-2015 per city are projected into the future. To predict the urban area added between 2016 and 2045, the trend is modelled in a linear manner relying on the past yearly absolute growth rates at the level of geographic regions (formula 4). Using the model, the absolute growth for the years 2016-2045 and sum the predictions into a single value per region G_{EX} is predicted (formula 5).

(4) $\hat{G}^r_{YA_{year}} = \beta_0 + \beta_1 * year + \varepsilon$

(5) $G^r_{EX} = \sum_{year=2016}^{2045} \hat{G}^r_{YA_{year}}$

where $G^r_{YA\,year}$ is the sum of absolute yearly growth G_{YA} for all cities in the region r, and G^r_{EX} is the predicted added area in the region r between 2016 and 2045.

As alternatives, linear extrapolation of urban areas at city level was tested. A comparison of the approaches is presented in appendices a.11 and a.12. These projected settlement sizes are complemented with projections for the cities' populations derived from the UN World Urbanisation Prospects (United Nations, 2018). However, these projections are only available until 2035.

Furthermore, the analysis is based on IPCC scenarios for different climate reference regions. These regions are intended to represent consistent climatic regimes and physiographic settings (Iturbide et al., 2020). As indicators of future climate stress, projected changes in annual maximum temperature (TXx) and annual maximum daily precipitation (Rx1day) are used, both for the scenario of 2°C global warming compared to the 1850-1900 baseline. TXx and Rx1day respectively serve as indicators for the intensity of extreme temperature events and precipitation events (Seneviratne et al., 2021).

3.2. Analysis strategy

Whether a city is growing at a fast or slow pace is by definition relative. One rate of growth may be fast relative to the global average but slow relative to trends in the region. To give expression to this relativity spatially in a global, regional and national comparison and temporally over the monitoring period, a three-track strategy to analyze urban settlement growth is pursued: First, *growth* for the period from 1985 to 2015 is analyzed. Second, a more detailed look at *changes in growth rates within the study period* is given. Third, these dynamics are categorized with respect to *continents*, *geographical regions*, *national levels* and *city sizes*.

In spatial terms, the six *continents* are Africa, Asia, Europe, North America, Oceania and South America. The ten *geographical regions* correspond to the region categorization 10 (RC10) by the IPCC Working Group III in its Fifth Assessment Report (Krey et al., 2014). They are North America, Western Europe, Pacific OECD (Japan, Australia, New Zealand), Economies in Transition (Eastern Europe and part of former Soviet Union), Latin America and Caribbean, Sub-Saharan Africa, Middle East and North Africa, East Asia, South Asia, and South-East Asia and Pacific. Beyond this, the indicators are also grouped at the world level. In this way, a global reference for settlement development is created that allows to relate developments on the level of continents or geographical regions. To identify characteristic patterns for each group, the growth indicators are aggregated using the median as well as the mean and the standard deviation within each group (σ).

Furthermore, cities' growth rates are compared within their respective countries (*national level*). This further spatial refinement of the geographical region highlights extremes which at coarser scales would be lost in aggregation. Thereby, we quantify on the one hand the variability of growth rates of cities within countries. This is to show how urbanization dynamics vary within and between countries, whether they are homogeneously and equally distributed in a country, or whether a country has both, fast-growth centers and low-growth or even cities with stagnating growth. In addition, particular attention is given to the capital city of the respective country. We specifically ask whether capital cities as political centers have a special position in relation to other cities in the respective country in terms of growth

rates. Therefore, the city whose MUA contains the capital of that country is determined and it is put in relation to other cities.

In terms of city sizes, settlement growth is analyzed in relation to *groups of cities of different population size* and in relation to *city area extent*. Here, we relate to three groups of cities of different population size: small cities with less than one million inhabitants in 2015, large cities with populations between 1 and 10 million, and megacities with >10 million inhabitants, as the UN (2018) define megacities quantitatively as conurbation having >10 million inhabitants. With respect to the *city area extent*, the significance of the relationship between settlement growth and settlement extent of cities is determined. Therefore, the relative growth rate of a city in a year GYR is modelled as a function of its extent (measured in km^2) in the previous year. As the WSFEvo timeseries can be seen as a type of panel data where observations from the same city are not independent, we apply a linear mixed model with random effects (Bates, Mächler, Bolker, & Walker, 2015) to account for the un explained city to city variation in intercept and slope. We calculate approximate *p*-values via Satterthwaite's degrees of freedom method (Kuznetsova, Brockhoff, & Christensen, 2017) and use a threshold of 0.05 to determine the significance of the linear relationship. As indicators for the model fit, conditional and marginal coefficients of determination are calculated according to Nakagawa, Johnson, and Schielzeth (2017). The conditional determination coefficient (R^2_C) rep resents the degree of variance explained by the entire model, including city-specific effects, while the marginal determination coefficient (R^2_M) represents the variance explained by the city extent only. The analysis at large scale can identify broad trends at group-level.

In a final analysis, we relate projected future settlement dynamics to projected climate change scenarios. Here the aim is to localize regions where both, increased climate change is predicted and high urban growth rates are recorded. We try to underpin this with figures in order to estimate the extent of exposure in 2045. Beyond this, we also select population figures in order to assess the dimension of exposure. This is done because the assessment via this variable is more accessible than via area growth.

4 Results

We find that the settlement areas of the 1567 cities in our study have on average doubled in size over 30 years from 1985 to 2015. Combined, this doubling of the settlement area totals >144,000 km^2 – an area roughly five times as large as Belgium, about the area of New York state, or half the area of Italy. This underscores the massive physical dimension of urban expansion.

The results also show a wide range of settlement expansion dynamics. While some cities have remained relatively constant in settlement area, others have expanded by >1000 %. These differences in urban growth dynamics have a certain systematic occurring between continents, the IPCC geographical regions, national levels and city sizes. In the following, we elaborate on these findings in more detail.

4.1. Urban growth trajectories 1985–2015

4.1.1. Asia outpaced other continents, East Asia outpaced other geographic regions

One unexpected finding is that 30 years ago, Asian cities were on average smaller in settlement area than the global median (and mean). However, during this 30-year period, these Asian cities have on average nearly tripled their settlement area (in 1985 they had on median only 39% (mean 41 %, σ 19 %) of their area in 2015). With 802 cities in Asia, the region also has by far the greatest number of cities with >300,000 in habitants (51.2 % of the 1567 largest cities across the globe).

In contrast, we measure on median the lowest settlement expansion in Oceania (from 71 % (mean 70 %, σ 16 %)) followed by Europe where cities already had 70 % (mean 69 %, σ 12 %) of their 2015 settlement area in 1985. The given variability of settlement expansion dynamics across cities is also lower in Oceania (it needs to be considered that the number of observations is with only 13 cities very low here). Also, in Europe (n = 241) and, with some caveats, North America (n = 200) the variability of settlement expansion dynamics across cities is measured lower. Finally, we illustrate the drastic differences at the continental level by the examples of Asia and Europe: in a ranking of the highest relative settlement growth rates of the 1043 cities among our set of cities of the two continents, the first 208 ranks are taken by Asian cities. The fastest growing city in Europe (Tirana, Albania) would only be ranked at 209.

Of course, continents are not homogeneous spatial units and, as shown by the standard deviations and illustrated in Fig. 2a, the variability of growth rates is enormous. Thus, we now detail our analysis with respect to the IPCC geographical regions. In Asia, where we measured the highest average settlement growth rates at continental scale, we identified highest settlement expansion rates specifically in East Asia. On median, the cities there have even more than tripled from 28 % (mean 30 %, σ 15 %) of their settlement area to reach their 2015 extent. And, South Asia is the only other IPCC region where cities have had higher expansion rates during the monitoring period (from 46 % (mean 46 %, σ 15 %)) than the world average (on median from 53 % (mean 51 %, σ 21 %)). In contrast, we measure the lowest growth rates in Western Europe and the Pacific OECD: These regions had already reached on median 73 % (mean 70 %, σ 15 %) and 72 % (mean 72 %, σ 12 %) of their respective 2015 settlement area in 1985 (Fig. 2b).

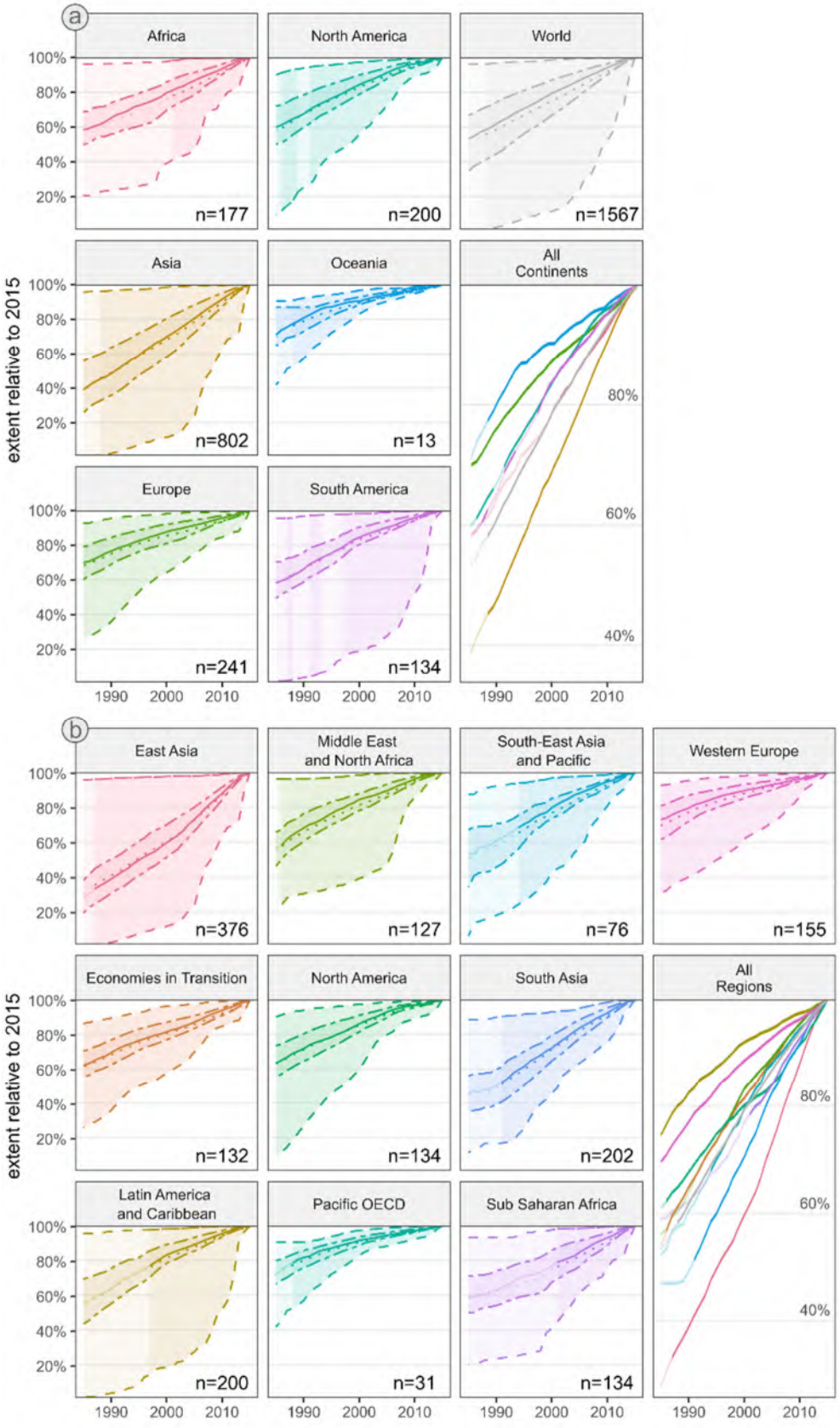


Fig. 2. City growth trajectories towards their size in 2015. The dotted lines correspond to the mean, the dashed, dot-dashed, and solid lines correspond to the minimum, 25%, 50%, 75% quantiles, and maximum of the GTT: a) per continent and b) per IPCC geographical region, in reference to the world. Periods of low WSF data consistency (cf. Appendix a1) are shown in lighter colours.

4.1.2. Relations to city size categories

We find that the most dynamic growth rates occur in the smallest cities. Specifically, 1064 small cities with fewer than one million in habitants have the most dynamic growth trajectories. Their settlement extent in 1985 is on median 53 % (mean 50 %, σ 21 %) of 2015. The 468 large cities between 1 and 10 million have settlement extents in 1985 of 55 % (mean 54 %, σ 19 %) and the 35 megacities have the lowest rate starting from 66 % (mean 61 %, σ 25 %) (Fig. 3). Of the 100 cities with the highest settlement increase relative to their starting extent, 76 are small cities, 21 are large cities and there are only 3 megacities.

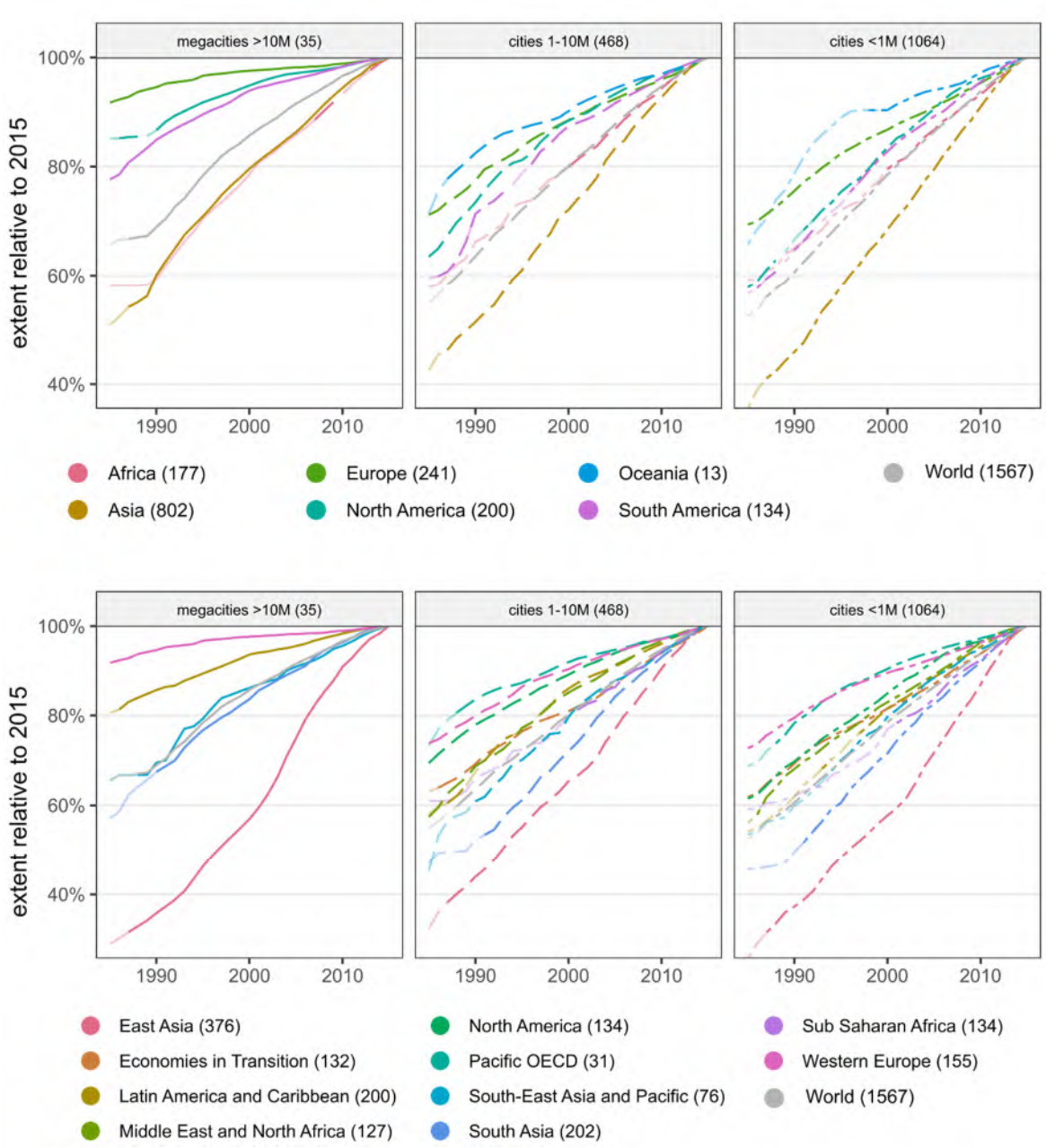


Fig. 3. City growth trajectories in relation to city size.

It is interesting to note that these trends in relation to the size groups of cities hold up on continental level. Again, the Asian cities have on median the most dynamic growth trajectories within their specific city size group (small: 36 % (mean 39 %, σ 19 %), large: 43

% (mean 44 %, σ 19 %), megacities: 51 % (mean 48 %, σ 22 %)). Among the small cities, growth trajectories in Asia and particular in East Asia are by far the highest (Fig. 3). These are representatives of cities that were previously either towns or even villages or only founded after 1985. A small portion of cities is measured at only 1–3 % of its 2015 extent. These became large cities with >300,000 inhabitants within a few decades. Of the 100 cities with the highest settlement increase relative to their starting extent, 93 are measured in Asia, whereas 81 of them in are China. The small cities on all other continents have on median larger settlement extents in 1985 than the global median. This shows, above all, that on these continents all cities of 300,000 inhabitants or more already were comparatively large 30 years ago.

It is also interesting to note that among the megacities, the lowest growth rates from 1985 to 2015 are measured in Europe (92 % (mean 88 %, σ 7 %)), North (85 % (mean 85 %, σ 5 %)) and South (78 % (mean 79 %, σ 6 %)) America, i.e. in these parts of the world the megacities have been large cities for a long time. Of the 100 cities with the largest extents in 1985 compared to 2015, 41 are in Europe, 21 in North America and 4 in Oceania. This means the majority of large, mature cities with low growth trajectories are in the western world. In contrast, among the megacities in 2015, we observe by far the most dynamic growth trajectories for Chinese megacities. The largest urban agglomeration, the Pearl River M.R., grew by 816 % followed by the Hangzhou M.R. with 710 %, Suzhou M.R. with 640 % and the Shanghai M.R. at fourth rank with 390 %.

Although small cities have the most dynamic growth trajectories, the large cities with populations between 1 and 10 million still absorb the largest percentage of settlement area growth during the study period. With 49.6 % of settlement area growth they carry almost half of the entire added settlement area of the 1567 cities. The smaller cities, however, still carry 31.5 %. Nevertheless, it is noteworthy that the only 35 megacities still contributed 18.1 % of the settlement growth. This ranking applies to all continents, except for Europe. There, the small cities shoulder with 55 % the bulk of the settlement growth compared to 42 % by the large cities and only 3 % of the megacities. For the IPCC geographic regions, this ranking basically also applies. Exceptions are Economies in Transition where the bulk of the settlement growth (59 %) is shouldered by small cities or East Asia, the Pacific OECD, and Southeast Asia and Pacific, where the megacities absorb with 30 %, 24%, and 33 % more settlement growth than small cities with 26 %, 22 %, and 23 % (see Table 1 for all details).

Table 1. Share [in %] of settlement expansion in absolute terms from 1985 to 2015.

Continent	Megacities > 10 M	Cities 1–10 M	Cities < 1 M	IPCC geographic region	Megacities > 10 M	Cities 1–10 M	Cities < 1 M
Africa	8.5	58.2	33.4	East Asia	30.4	43.4	26.2
Asia	27.1	45.2	27.8	Economies in Transition	1.9	39.4	58.7
Europe	2.7	42.1	55.3	Latin America and Caribbean	10.1	52.5	37.4
North America	5.5	62.0	32.6	Middle East and North Africa	10.2	56.7	33.1
South America	13.1	49.6	37.3	North America	5.5	63.2	31.3
Oceania	–	82.3	17.7	Pacific OECD	24.2	53.8	22.1
				South Asia	17.3	50.8	31.9
				South-East Asia and Pacific	33.4	43.4	23.3
				Sub Saharan Africa	6.4	60.4	33.2
World	18.1	49.6	31.5	Western Europe	7.2	46.1	46.7

4.2. Development of annual settlement growth rates over time

4.2.1. A comparison at continental level, IPCC geographic regions and national levels

We also analyzed settlement dynamics at the continental level, across IPCC geographical regions and national levels by using *relative* (G_{YR}) and *absolute* (G_{AA}) *annual growth rates*.

For the *relative settlement growth dynamics,* it is particularly interesting to see that decreasing trends over time on all continents are measured except for Africa. Annual growth rates (G_{YR}) in Asia were over 4.1 % (mean 4.5 %, σ 4.3 %) at the beginning of the monitoring period from 1985 onwards. Since then, a steady decline in these dynamics to 1.3 % on median (mean 1.5 %, σ 1.3 %) at the end of the monitoring period in 2015 is observed. A steady decline of relative annual settlement growth rates and a leveling off at around 1 % per year is also measured for Europe, Oceania, North and South America. Although rates started at over 2 % (respective medians 2.01 %, 2.23 %, 2.53 %, and 2.43 %; means 2.12 %, 4.00 %, 2.38 %, and 3.03 %; σ 1.13 %, 5.14 %, 2.08 %, 2.71 %), a decreasing trend to around 0.5–1 % in 2015 is measured (respective medians 0.95 %, 0.46 %, 0.77 %, 0.56 %; means 1.26 %, 0.54 %, 0.90 %, 0.70 %; σ 1.0 %, 0.29 %, 0.60 %, 0.58 %; Fig. 4a). As mentioned above, a different development is detected only for Africa: At the beginning of the monitoring period, annual settlement growth rates were around 0.54 % (mean 1.9 %, σ 2.6 %) and rising to 1.9 % (mean 2.8 %, σ 3.1 %) in 2000. However, this low annual growth rate is based on poor data consistency from 1985 to 2000 and therefore we must view this statement with caution. From 2000 onward, however, the data situation is good. Growth rates since then have – what seems to be a general trend – also steadily decreased from 1.9 % (mean 2.8 %, σ 3.1 %) in 2000 to 1.0 % (mean 1.3 %, σ 1.3 %) in 2014.

These observations can be further differentiated across IPCC geographical regions: Specifically, cities in East Asia have grown most dynamically. Growth rates there were at 7.23 % (mean 8.9 %, σ 6.8 %) annually at the beginning of the monitoring period. Since then, we observe a decline in growth dynamics to around 1.6 % (mean 2.9 %, σ 1.5 %) at the end of 2015. For South Asia, South-East Asia as well as in Sub Saharan Africa, low growth rates of 0–1 % in the earlier period are measured. Again, we need to point to the low input data consistency until the year 2000 that makes these values unreliable and their low growth rates unlikely. However, beginning in 2000, we observe significantly higher settlement growth rates of 3.24 (mean 4.7 %, σ 2.1 %), 2.23 (mean 2.45 %, σ 1.5 %), and 2.53 % (mean 3.27 %, σ 3.4 %) which steadily decrease to 0.86 % (mean 0.12 %, σ 0.1 %) – 1.14 % (mean 1.45 %, σ 1.4 %) (Fig. 4b). In principle, however, the trend towards decreasing relative growth rates remains consistent across regions.

With respect to the city size groups, we observe also steadily decreasing relative growth rates since 1995. More or less parallel descending curves are measured, with the annual growth rates of smaller cities running at a higher level than those of large or megacities (Fig. 4c).

If one takes a look at *absolute settlement growth* G_{YA}, a different picture emerges: Absolute settlement expansion in Asia is comparatively low and more or less constant annually between 100 (mean 250 ha) and 150 ha (mean 400 ha) over the monitoring period. In contrast, the absolute settlement expansion is measured with 200 (mean 200 ha) to 400 ha

(mean 600 ha) for Oceania and with around 250 ha (mean 500 ha) for North America on average by far highest from 1985 to 1995. Over the course of time, they steadily decrease there to around 100 ha (mean 200 ha) per year (Fig. 4d). The high relative growth rates in Asia are therefore derived from smaller starting sizes of the cities. The high ab solute growth rates in Oceania and North America result in low relative growth rates due to the large city sizes. With the poor data availability and consistency in Africa until the year 2000, the low growth rates in the earlier period contain some uncertainty. However, after this time period data availability and consistency is high and we find the cities on this continent feature similar high annual growth rates as the other continents.

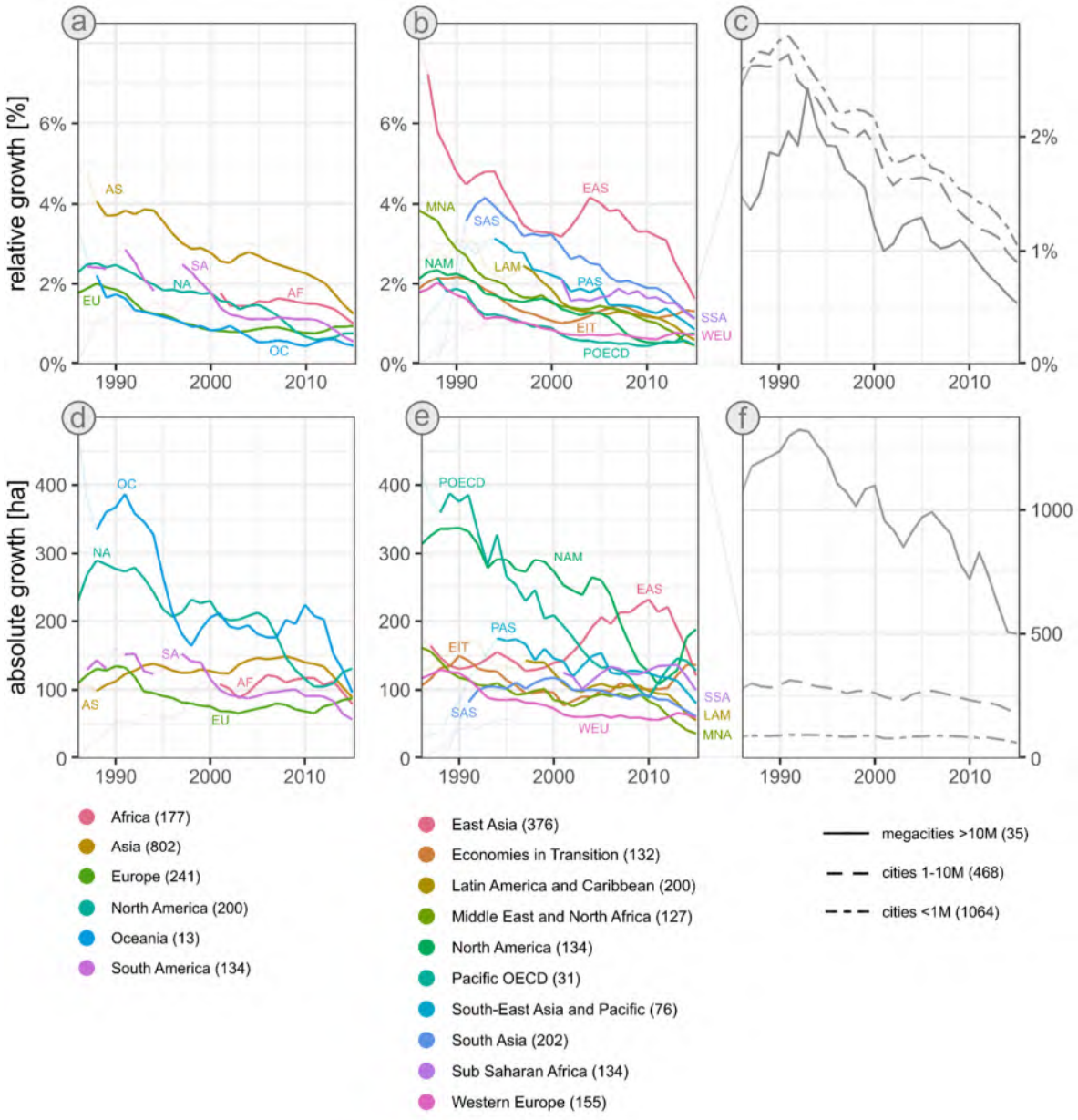


Fig. 4. Three-year rolling average of relative growth rate G_{YR}, median of a) continents, of b) IPCC geographical regions, and of c) size categories; three-year rolling average of absolute growth rate G_{YA}, median of d) continents, e) IPCC geographical regions, and f) size categories. Periods of low data consistency are shown in lighter colors.

If one details to the IPCC geographical regions, three groups can be categorized (Fig. 4e): The Pacific OECD and NAM are an own group with extraordinary high average annual settlement expansion rates. There >300 ha (mean 600 ha) in 1990 are measured; however, with a steady decline until 2010 to around 150 ha (mean 300 ha), after which the data suggests rising rates for North America. East Asia is another group experiencing a doubling of absolute settlement growth per year in the 2000s. The last group are the cities of the remaining IPCC geographic regions which have a settlement growth of around 100 ha (mean 200–250 ha) per year settlement growth with a slight decreasing trend over time. With respect to city sizes, we observe absolute growth rates remaining reasonably stable for small (ca. 80 ha (mean 120 ha) annual settlement expansion) and large cities (ca. 250 ha (mean 400 ha) annual settlement expansion). Megacities, in contrast, show a decline from the early 1990s onwards from 1300 ha (mean 2500 ha) per year to 500 ha (mean 1800 ha) per year (Fig. 4f).

In the last step, we refine the spatial levels of continents and IPCC geographical regions even further. We now look at the mean annual settlement growth rates G_{YR} of cities and their distributions at the national level. Of course, Asia also features the countries with the highest urban growth rates, while there are countries in Europe or North America that have, as expected, lower growth rates. And yet, it is interesting to see that certain countries from other continents represent outliers: Mexico, for example, is the only non-Asian country in the top ten of urban settlement dynamics. The Asian countries Japan and Indonesia, in contrast, are in the lower ranks and comparable to European countries (Fig. 5). This shows that in order to understand urban growth in its regional differences, different spatial target units are in demand.

Many Asian countries, such as Malaysia, Pakistan, China, South Korea, Iraq, Türkiye, Iran or India but also Mexico or Brazil in the Americas, show large variabilities of growth rates of cities. The countries are characterized by centers of highest settlement dynamics versus cities with very low development trends. This reveals an inequality in the settlement growth rates of cities within the countries. In contrast, countries such as Germany, the United Kingdom, Japan or France generally have very low urban growth rates and the variability is also very low. This shows developments that are more evenly distributed across the cities.

When we examine capital cities, it is interesting to observe that they tend to exhibit comparatively low growth rates compared to other cities within their country (cf. Appendix a.3). As most capitals rank among the largest cities in their country, this is in line with the observation that growth decreases with city size (see Section 4.1.2). Notable exceptions are Almaty, Rome, Brasília, Cairo and Abuja, the capital cities of Kazakhstan, Italy, Brazil, Egypt and Nigeria, which rank above the median of growth rates for their countries.

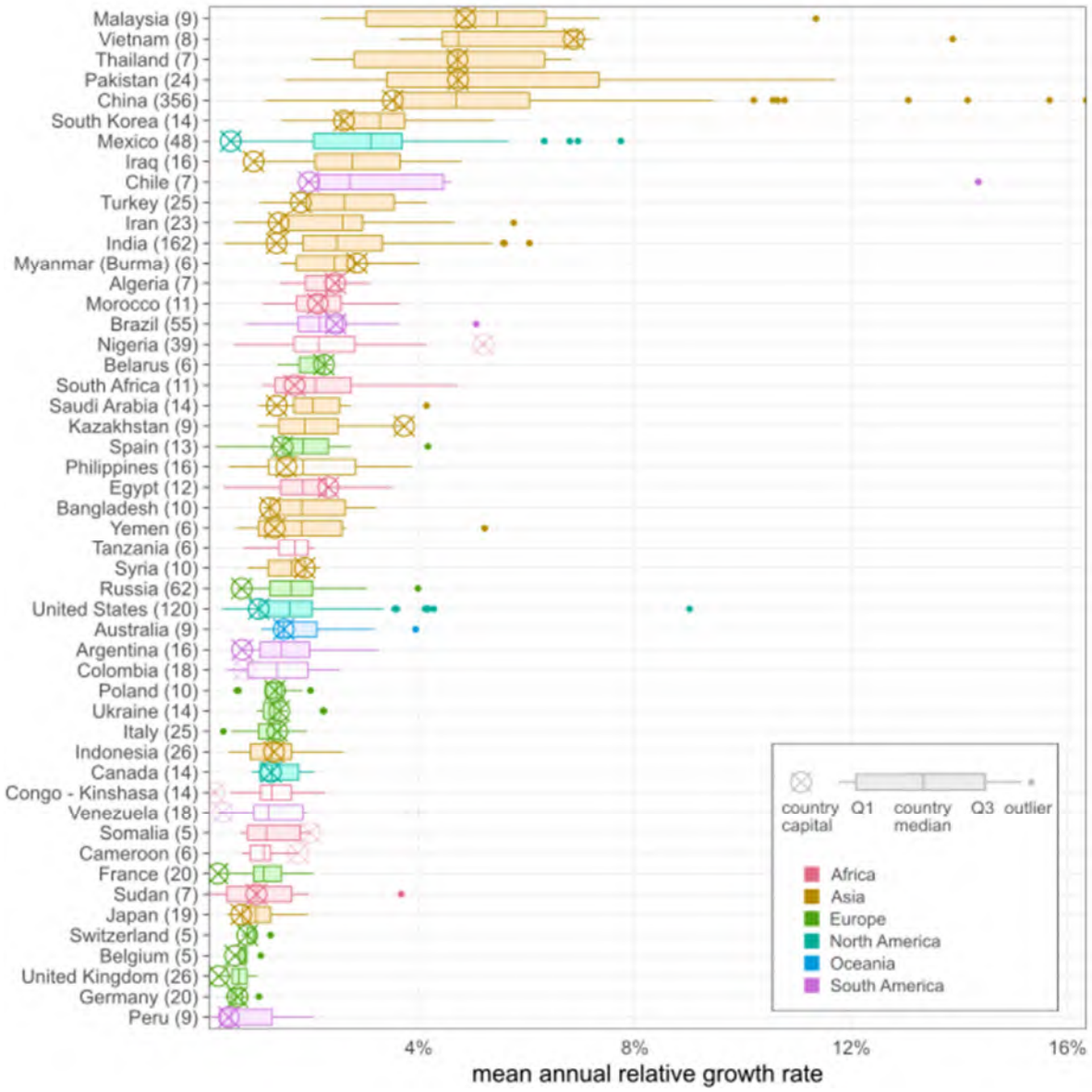


Fig. 5. Mean annual relative growth rate (G_{YR}) of cities at national level, with capitals highlighted, for all countries with at least five cities in our dataset. The number in brackets refers to the country's number of cities in our dataset. Objects of low data consistency are shown in lighter colors.

4.2.2. Relations to city size

A systematic statistical evaluation of relative annual settlement growth dynamics and city sizes reveals two main findings: first, it be comes clear that relative annual settlement growth dynamics decrease as the city size, i.e. the settlement area of cities, increases (Fig. 6). This trend is sometimes more pronounced and sometimes less so, but it is a given in principle (see Appendix a.2 for examples of linear regressions). In a linear mixed regression model, the relationships are highly significant for all continents (Table 2). And the same is true for all IPCC geographical regions (Table 3). For the subgroups of capital cities and all other cities, respectively, the correlations remain equally statistically significant. It is interesting to note that the correlation is weaker for smaller cities than for larger cities (Table 4), and weaker for capitals (Table 5).

Second, it becomes clear that cities on different continents or IPCC geographical regions in fact have on median very different settlement extents. This shows that spatial patterns of cities vary greatly by region. This also makes it clear that regionally different approaches are needed with regard to sustainable development. On continental level, cities in Oceania, North America and Europe show on median the largest settlement areas. It is interesting to see how relative settlement growth rates decline when cities are getting larger in extent. On the level of IPCC geographical regions, the Pacific OECD has by far the largest settlement extents, followed closely by North America. The smallest settlement extents are found in South Asia, at least in the >300,000 population category of cities under study here.

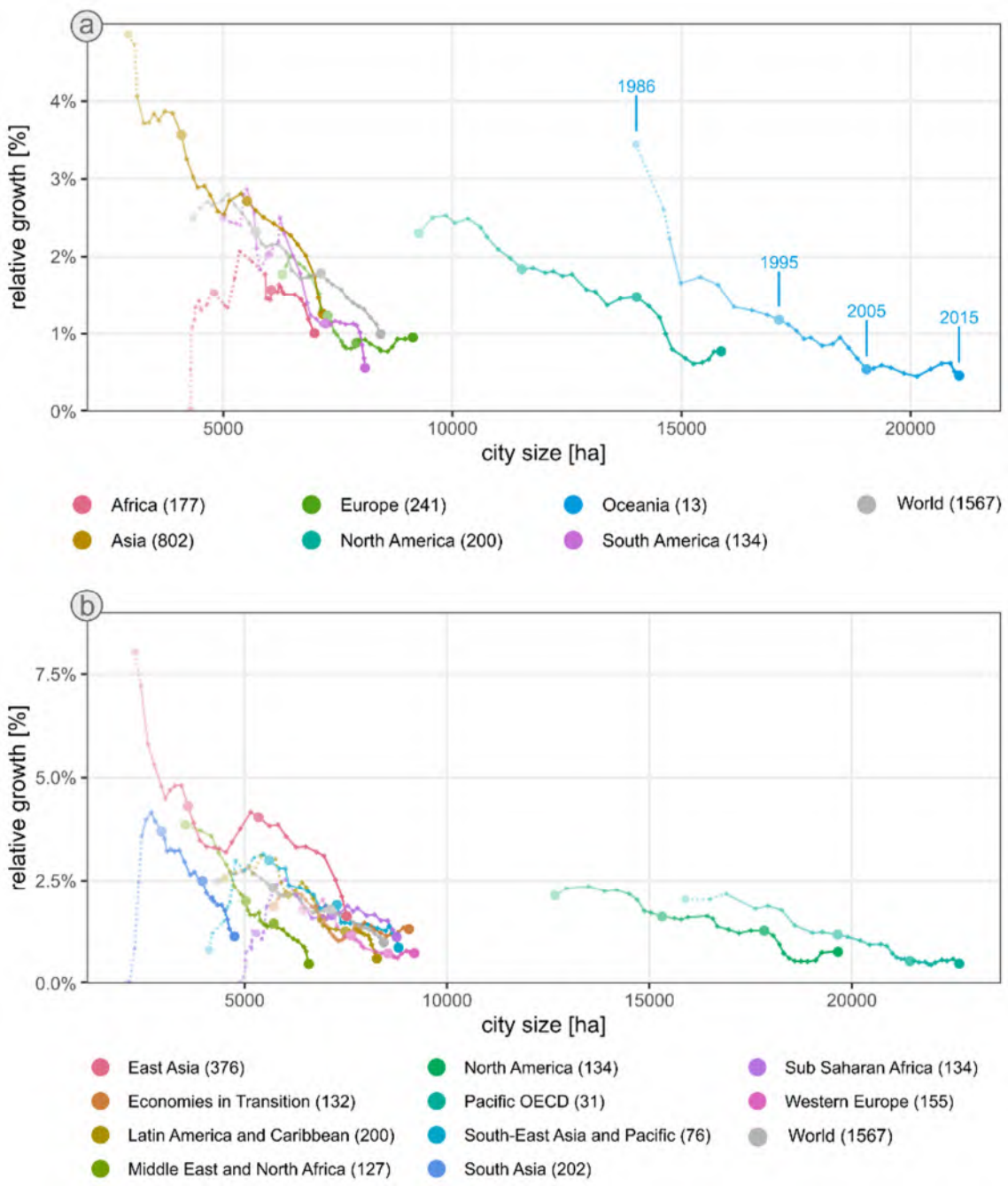


Fig. 6. Relative growth (G_{YR}) in relation to city sizes: a) Three-year rolling average, median of continents and b) of IPCC geographical regions. Periods of low data consistency are shown in dotted lines.

Table 2. Coefficients of linear regression models of relative growth rates versus settlement extent per city, for each continent. Significance encodes the p-value as: * p<0.05; *** p<0.001. R^2_C and R^2_M refer to the conditional and marginal determination coefficient respectively.

Continent	City count	Coefficient	p-value	Significance	R^2_C	R^2_M
Africa	176	−0.06	1,71E-11	***	0,9957	0,2183
Asia	802	−0.11	1,5E-67	***	0,9990	0,3119
Europe	241	−0.05	3,83E-35	***	0,9971	0,4147
North America	199	−0.06	9,61E-13	***	0,9998	0,1936
Oceania	13	−0.03	0,047592	*	0,9998	0,1953
South America	133	−0.10	1,61E-11	***	0,9991	0,2560
World	1564	−0.09	1,9E-115	***	0,9991	0,2718

Table 3. Coefficients of linear regression models of relative growth rate (G_{YR}) versus settlement extent per city, for each cultural sphere. Significance encodes the p-value as: * p<0.05; *** p<0.001. R²C and R²M refer to the conditional and marginal determination coefficient respectively.

IPCC geographical region	City count	Coefficient	p-value	Significance	R^2_C	R^2_M
East Asia	376	−0.10	1E-27	***	0,9983	0,2756
Economies in Transition	132	−0.04	7E-06	***	0,9968	0,0888
Latin America and Caribbean	199	−0.11	6E-16	***	0,9991	0,2508
Middle East and North Africa	127	−0.11	2E-22	***	0,9976	0,4814
North America	133	−0.03	9E-20	***	0,9995	0,4092
Pacific OECD	31	−0.05	4E-05	***	0,9999	0,3766
South Asia	202	−0.16	2E-21	***	0,9967	0,3219
South-East Asia and Pacific	76	−0.08	2E-10	***	0,9992	0,4341
Sub Saharan Africa	133	−0.04	2E-06	***	0,9957	0,1445
Western Europe	155	−0.07	2E-27	***	0,9993	0,4475
World	1564	−0.09	2E-115	***	0,9991	0,2718

Table 4. Coefficients of linear regression models of relative growth rate (G_{YR}) versus settlement extent per city, for each city size category. Significance encodes the p-value as: * p<0.05; *** p<0.001. R²C and R²M refer to the conditional and marginal determination coefficient respectively

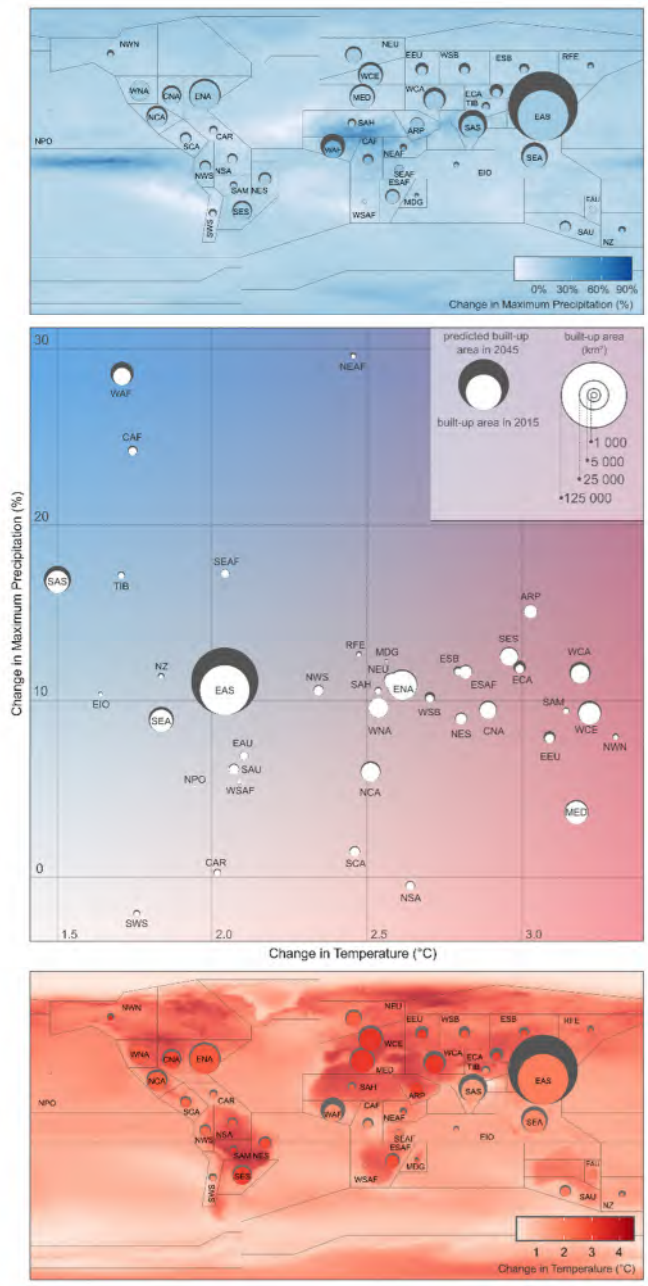
Capital	City count	Coefficient	p-value	Significance	R^2_C	R^2_M
Megacity >10 M	35	−0.01	1,9E-08	***	0,9919	0,4348
Cities 1–10 M	468	−0.03	4,17E-62	***	0,9948	0,3668
Cities <1 M	1061	−0.12	1,55E-87	***	0,9889	0,2154

Table 5. Coefficients of linear regression models of relative growth rate (G_{YR}) versus settlement extent per city, grouped for all capitals and all other cities. Significance encodes the p-value as: * p<0.05; *** p<0.001. R²C and R²M refer to the conditional and marginal determination coefficient respectively.

Capital	City count	Coefficient	p-value	Significance	R^2_C	R^2_M
Non-Capital	1421	−0.10	9,9E-103	***	0,9991	0,2574
Capital	143	−0.04	2,01E-06	***	0,9997	0,1075

4.3. Future urban growth and exposure to IPCC climate change projections

Finally, we venture an outlook on the future for these 1567 largest cities of the world: We do this by putting the IPCC climate change scenarios until the year 2045 in relation to the projected settlement growth rates. We discuss the scenarios of 1) maximum precipitation change (y-axis) and 2) maximum temperature change (x-axis) in relation to the projected settlement growth of the cities in the respective IPCC climate reference regions (Fig. 7).



		population [M]		built-up area [km²]		RMSE [km²]
		2015	2035	2015	$G_{EX\ 2045}$	1985 2015
East Asia	EAS	617	839	71788	56983	264
Western Africa	WAF	75	154	8843	5936	26
South Asia	SAS	276	446	14621	5521	39
South East Asia	SEA	120	175	14554	2856	40
West Central Asia	WCA	95	144	9830	2480	24
East Central Asia	ECA	15	23	2057	1799	10
Western & Central Europe	WCE	101	108	13830	1567	29
Eastern Europe	EEU	15	16	2443	1136	10
West Siberia	WSB	12	14	1691	1005	5
Eastern North America	ENA	125	148	24298	901	42
North Central America	NCA	70	90	10038	855	27
Mediterranean	MED	158	206	14878	770	32
Central North America	CNA	31	40	7996	636	20
East Siberia	ESB	15	19	1328	557	5
Central Africa	CAF	34	74	1922	406	3
Northern Europe	NEU	47	54	6191	362	10
North Eastern Africa	NEAF	7	16	528	309	4
S.E. South America	SES	90	107	8937	301	19
New Zealand	NZ	2	3	567	243	2
N.E. South America	NES	37	5	3333	223	8
N.W. North America	NWN	3	4	635	221	3
S.W. South America	SWS	10	12	974	189	7
Russian Far East	RFE	4	5	473	189	2
Tibetan Plateau	TIB	10	17	1067	159	3
South Central America	SCA	24	33	2612	159	7
Sahara	SAH	13	23	1079	137	5
East Southern Africa	ESAF	33	55	4619	127	13
Caribbean	CAR	12	15	1192	120	5
N.W. South America	NWS	44	57	2632	105	8
North South America	NSA	17	21	2016	93	4
West North America	WNA	42	51	11474	71	20
South Australia	SAU	8	11	2623	66	4
Arabian Peninsula	ARP	36	57	5026	38	14
South Eastern Africa	SEAF	23	56	1923	15	8
Equatorial Indian Ocean	EIO	1	1	417	13	7
South American Monsoon	SAM	6	9	947	11	3
Eastern Australia	EAU	8	11	2056	13	4
Madagascar	MDG	3	7	192	7	2
North Pacific Ocean	NPO	1	1	178	0	1
West Southern Africa	WSAF	6	9	601	0	1

Fig. 7. Left: Current urban extent and future urban expansion G_{EX} under climate stress for the cities in the respective IPCC climate reference regions. Growth until 2045 was projected based on growth trends from 1985 to 2015. Projected changes in annual maximum daily precipitation RX1day(y-axis) and projected changes in annual maximum temperature TXx(x-axis) at 2°C of global warming compared to the 1850–1900 baseline.

1) In African regions close to the equator (Western Africa (WAF), Central Africa (CAF) and North Eastern Africa (NEAF)), the projected maximum precipitation increase is 25 % or more. In these cities, therefore, increased likelihood of flooding or heavy rainfall events is to be expected in the future. Of particular significance is that the exposure there will increase, as we also expect very high urban growth rates in terms of settlement and population for the WAF and CAF regions. In the following, we quantify the increased exposure in relation to the population, as this is easier to classify. This means that 85 cities among our set of cities with >116 million inhabitants as of 2015 will expose >242

million inhabitants by 2035 to these extreme precipitation changes. These regions are followed by South Asia (SAS), South Eastern Africa (SEAF) and the Tibetan Plateau (TIB) with about 16–18 % projected maximum precipitation change. Especially for SAS, we also expect very high urban growth rates. Thus, another 195 cities among our set of cities with >309 million inhabitants in 2015 will expose >519 million in habitants by 2035 to these precipitation changes. For most other regions between 5 and 15 % maximum precipitation changes are projected. Only for very few regions, such as the Caribbean (CAR), Southern Central America (SCA) and Northern-South America (NSA), no significant precipitation change is projected.

2) With respect to temperature increases, the cities in North-Western North America (NWN), Western and Central Europe (WCE), Eastern Europe (EEU), West Central Asia (WCA), the Arabian Peninsula (ARP) and the South-American Monsoon (SAM) are projected with >3 degree Celsius. In terms of heat stress these cities need to be at the forefront of developing adaptation measures. The exposure is particularly high in WCE and EEU, where growth of built-up area is expected to continue. In these regions 246 cities with >256 million inhabitants as of 2015 will expose 80 million more to these extreme temperature changes by the year 2035, i.e. 336 million inhabitants. Here it is particularly noteworthy that especially the highly developed Western world, and Western and Central Europe in particular (in WCE it is 108 cities with projected >107 million inhabitants by 2035) will be increasingly exposed. This risk will increase even more for cities in countries with a comparatively old population (NWN, WCE, EEU). Nevertheless, 2–3 degree Celsius of temperature increase can be expected for the majority of cities around the globe. Adaptation measures for our cities on our planet must become a central element to face these changes. Few regions are expected with lower temperature increases, especially in the SAS, TIB, WAF or CAF regions. However, it must be clearly stated that these regions are still confronted with expected temperature rises between 1.5 and 2 degree Celsius.

5 Discussion

5.1. Three decades of urban growth: the perspective of settlement expansion

This study reveals that Asia is the continent with the highest urban settlement growth rates over the 30 years of monitoring from 1985 to 2015. In contrast, the cities in Oceania, Europe and North America are measured as least dynamic. These findings are generally in line with other studies (e.g. Angel et al., 2011; Güneralp et al., 2020; Taubenböck et al., 2012), but now we confirm them at a higher degree of consistency enabled by the data-driven and globally comparable delineation of the analyzed cities. The continents, however, are not homogeneous in terms of city growth: this study in addition reveals that Asia features with East Asia globally the most dynamic geographic region and at the same time with Japan or North Korea countries that are among the least dynamic areas at national level. This shows that the aggregation of trends to global or continental levels masks the diversity of growth patterns, and highlights the need for region-specific urban development patterns. With reference to city sizes, the study reveals that the highest growth rates are achieved by the

smaller cities under one million. This trend within the world's 1567 largest cities found in our study is consistent with the findings of Tian, Tsendbazar, van Leeuwen, and Fensholt (2022) which showed that the same applies to villages. They found settlement growth rates in peripheral rural areas far exceeded those in urban centers However, we also found that the large cities between 1 and 10 million inhabitants contribute the largest share of the overall settlement growth among our set of cities across the globe. The results show that the focus for sustainable urban development must move away from the very large and megacities as most development can be expected in today's smaller cities.

Over the monitoring period a clear trend of declining annual settlement growth rates over time is characteristic on all continents, but less pronounced for Africa. In Africa, we saw increasing dynamics, starting from a very low level. Overall, we basically see urbanization continuing in these large cities larger than 300,000 people and leveling off at about 1 % per year. This general trend of declining growth rates over time suggests urbanization at reduced rates. However, the relative growth only provides one view of things, as we found the absolute annual gains in settlement areas in our set of cities remains largely constant (or only slightly decreasing). Furthermore, despite constant absolute growth, we also clearly see declining relative growth rates with increasing city sizes. Thus, we observe increasing city sizes are not accompanied by proportionally increasing growth rates (G_{YR}). This inverse relationship con firms that observations made for the relationships between cities' population size and growth (cf. Egidi, Salvati, & Vinci, 2020) also exists in the built domain. It would be interesting to investigate whether the point of peak growth is governed by a universal rule, or primarily dependent on site-specific characteristics and global trends. As always, there are exceptions to this rule: 40 cities of our set of cities feature increasing growth rates with increasing city sizes (see Appendix a.4). These 40 cities were among the smaller ones in our set, with, on median, 0.232 million inhabitants in 1985 and 0.561 million inhabitants in 2015 (most of them belonged to the group of medium (32) or large (8) cities).

And, we reveal that capital cities predominantly feature lower settlement growth rates than the median of the respective country. Whether this is due to a causal relationship with their status as seat of government must be left open due to the plethora of other factors which might influence growth rates. For instance, most capitals are usually cities of large size, which – as shown in this study – is likely to have a diminishing effect on relative growth rates. Nevertheless, it will be interesting to analyze in future studies which other factors, such as political or economic significance, primacy, location, etc. determine different growth rates of cities.

The study also shows that the highest share of settlement growth is in today's large cities of 1 to 10 million people. With the trend towards decreasing growth rates with increasing size and the further urbanization to be expected, it can be assumed that the majority of the settlement growth in the next decades will come from the small cities in our set of cities or in cities that have reached similar sizes by now. A large part of this growth is expected to occur in Africa and Asia, which will continue to be the dominant centers of urbanization.

5.2. Three decades of urban growth: the relation of settlement expansion to population growth

The focus on this study was on purpose on physical expansion of settlements at city level. Of course, urbanization also relates to population numbers. Thus, we spotlight here the increase in settlement area in relation to the increase in population: According to the World Urbanization Prospects (United Nations, & Department of Economic and Social Affairs, 2015), the combined population with our set of cities numbered 2.246 billion in 2015, an increase of 103 % compared to the 1.104 billion in 1985. Over the same time frame, the settlement area within the cities' MUAs increased by 85 % from 14.1 million hectares to 26.1 million hectares and was thus lower.

Interestingly and in contrast to these results, Liu et al. (2020) as well as Güneralp et al. (2020) measured the rate of urban expansion at global level notably faster than that of population growth, indicating that land consumption per person is increasing. In our study, however, we identified for the world's 1567 largest cities a different trend: there, population growth is faster than settlement growth. From this, it can be deduced that the larger cities on our planet have a more efficient population-settlement ratio.

However, we need to be aware that substantial variations exist across the globe as, for example, we measure different city sizes for different continents and geographical regions. Hu et al. (2021) showed how diverse land consumption is with respect to the structural configuration of cities across the globe. Using the extremes of their study, Mumbai in India is measured with 46,406 inhabitants per km^2 and Melbourne in Australia with 4069 inhabitants per km^2, allows to project theoretically what that would mean for our set of all cities across the globe larger than 300,000 inhabitants: In the three decades from 1985 to 2015 these cities expanded their settlement area by 120,178 km^2 and increased their population by 1.142 billion. Globally land consumption was thus more efficient than in cities like Melbourne, because if the actual expanded settlement areas would have a population density as in Melbourne only 0.489 billion people could live there. Vice versa, however, if the added settlement area was as densely populated as Mumbai, 5.577 billion people would call these cities their home. This theoretical thought experiment reveals how decisive the structural configuration of cities is.

5.3. Focus areas of the future – where high urban growth will meet strong climate change

The combination of the IPCC climate change scenarios with our forecast of urban growth allows to identify regions where developments will generate particularly high exposures. Nevertheless, it should be noted at this point that exposure and climate change will have an impact almost everywhere, resulting in increased risks. However, focus areas with the expected highest increases in precipitation and very high settlement growth projections are found in the WAF, CAF and SAS regions. In terms of expected highest temperature increases, the regions of NWN, WCE, EEU, WCA, ARP and SAM will have highest exposures.

Climate change will have a particular impact in cities as the high concentration of people, assets and the economy there makes societies vulnerable (Güneral, Güneralp and Liu, 2015). Our city- and region-specific analysis clarifies, in a spatially differentiated manner, which adaptation measures are important for which cities. For example, measures for temperature regulation such as greening, albedo management with white-painted buildings and cool roofs or fresh air corridors (e.g. Leichtle et al., 2023; Massaro et al., 2023) are predominately important for cities in Western and Central Europe. The vulnerability there will increase even more as these regions have comparatively old populations which are much more susceptible to heat stress (Chapman, Watson, Salazar, Thatcher and McAlpine, 2017). For cities in Sub Saharan Africa and South Asia, flood protection measures, such as sustainable urban drainage systems (Fletcher et al., 2014) or the sponge city concept (Nguyen, Ngo, Guo, & Wang, 2020), among others, must also be increasingly considered in planning processes (e.g. Kakinuma et al., 2020).

5.4. Data-technical and conceptual issues

We argue that the results in this study are insightful and different from other studies as the analysis has been based on consistent spatial reference units, delineating cities uniformly using a data-driven method. Thus, the reference units of morphological urban areas are a comparable unit, which does not produce statistical distortions, such as those inherent in administrative extents. Nevertheless, it is worthwhile to critically question other issues of this approach:

First of all, our input data, i.e. the WSFEvo, is neither 100 % accurate nor has consistent accuracies across space. Although the WSF provides high accuracies in general for urban areas, it has been shown that ac curacies vary across regions. In general, Marconcini et al. (2020) reported accuracies of 83–89 % (Kappa 0.55–0.78) for the WSF 2015 dataset. We deliberately made the analyses over the complete monitoring period of 30 years and explicitly noted existing uncertainties in the data (Appendix a.1 and marked in the Figs. 2, 3 and 5). The extent to which these inaccuracies and inconsistencies affect our results, however, must be left open. If we look in particular at the rising relative settlement growth rates in Africa at the first half of the monitoring period – this measured trend deviates from the general global developments –, we are aware that this result is based on comparatively uncertain data (Appendix a.1). This results from low numbers of suitable satellite images. Against this background, we suggest to interpret the annual or three-year growth rates with caution. Nevertheless, to assess whether these trends are generally valid, we compared the settlement growth rates with another, independent data set. We compared relative growth rates of the WSFEvo to an alternative, independently available product at global scale, the Global Human Settlement Layer Built-Up Grid (GHSL), which provides information at 250 m resolution for the years 1975, 1990, 2000, and 2015 (Pesaresi et al., 2015). The GHSL is suitable for validation because it is a more aggregated data set in space and time and is thus less likely affected by periods of low availability of satellite images. Comparing both products' relative growth rates for the periods of 1990–2000, where some WSFEvo data is poor, and 2000–2015, where WSFEvo data is almost always good, we find that the relation between the GHSL and WSFEvo growth rates is similar across both periods. We conclude

that for aggregations over longer time periods such as from 1985 to 2000 the general settlement development trends which we measured are coherent even in regions where yearly data is poor.

Second, and beyond classification errors, we must be aware that the WSF classifies the built environment of cities; i.e., other aspects of urbanization such as demographic growth rates might be decoupled, e.g. in ghost cities (e.g. Shi et al., 2020). It is also worth noting that with the WSEvo, we cannot address the cases where settlement shrinking occurs; nevertheless, this is a considerably minor phenomenon compared to the high growth rates of global urbanization almost negligible.

Third, our database only allows to analyze the two-dimensional expansion of cities. The expansion into the third dimension, which has a non-negligible effect in many urban areas, must therefore be left out at this point (Wentz et al., 2018). Recent studies show the capability to derive large-scale height information from remote sensing data that studies on 3D growth appear possible on a large scale in the foreseeable future (e.g. Frantz et al., 2021; Geiß et al., 2019; Li, Koks, Taubenböck & van Vliet, 2020).

Fourth, the applied MUAs provide a consistent reference unit that takes the respective city sizes into account. However, due to data availability, we refer in our entire study on the city extents of the reference year 2015. A systematic calculation of the MUAs for each year remains a task to be done. It would be interesting to see to what extent dynamic rather than static MUAs affect the results. Further studies with other data-driven city delineations (e.g. Corbane et al., 2019) and other input data (e.g. Pesaresi et al., 2015) are in demand to confirm the results here.

Fifth, we have conceptually selected our set of cities covering all cities that had >300,000 inhabitants in 2015. This is due to the avail ability of MUAs. The extent to which the measured trends continue for smaller cities is thus left aside in our study. We must bear in mind that this approach is not a cohort of cities for the monitoring period, but rather allows a retrospective analysis of the growth history of those cities which have come to a significant size in our present day.

Sixth, there are endless possibilities to divide cities into groups for analysis. It is clear that any kind of classification of cities is fraught with conceptional assumptions and subjectivity. However, we believe that by continents, IPCC geographical regions, national levels as well as by the city size groups a differentiated spatial picture on urbanization dynamics across the globe can be drawn. Alternative divisions of cultural spheres such as for example done by Huntington (1997), by structural types of cities (Taubenböck et al., 2020), or other city size groups to name a few in the domain of this paper are in demand to complete this analysis. And beyond, studies such as from Sun et al. (2020), who related urbanization rates to different economic levels, are suggested for further research.

Seventh and last, the prognosis from the IPCC regarding climate changes, as well as the temporal extrapolation of urban growth rates are, of course, subject to a high degree of uncertainty. With regard to urban growth projections, we have chosen three different methodological approaches. We admit that these are very simplistic. Although growth rates differ across the scenarios, the areas that are expected to grow the most remain similar across

approaches (see appendix a6-a11). These un certainties that develop from the combination of the IPCC projections with our urban growth projections were not explicitly addressed in this work. However, since the projection is only intended to offer a general assessment, we argue that it allows the identification of the focus regions that will have to deal with the most drastic changes in urban growth and climate change. We acknowledge that more complex urban growth models focusing e.g. on density (e.g., Li, Zhou, Eom, Yu, & Asrar, 2019) or considering socioeconomic data may improve the accuracy of the projections. Also, the intra-urban dimension of that growth through spatial growth models such as SLEUTH (Clarke, Hoppen, & Gaydos, 1997) would add to knowledge.

Settlement developments in cities show very different dynamics around the world. These differences are meaningful as they impact demographic, economic, social, ecologic, among other issues of human kind's coexistence. For sustainable development, we need to know where to look at and where to act. In this study, we systematically documented and analyzed these dynamics for the 1567 most populous cities on our globe in relation to continents, IPCC geographical regions, national levels and city sizes. Thus, this study is not biased by the city selection or the focus on specific geographic areas. And, as we use on a data-driven delineation of cities, the comparisons across cities are coherent.

6 Conclusion

Urbanization is one of the drivers of global change. Humankind – as Saunders (2010) described it – is in the midst of the greatest migration it has ever undertaken: from the countryside to the cities. In 2015 –the reference year for this study– 7.38 billion people lived on our planet and 54 % of them were urban dwellers (UN, 2015). Among these 54 %, 46.1 % lived within the morphological urban areas of cities with at least 300,000 inhabitants. This makes it clear that not only does the majority of humanity live in cities, but one in four (1.84 billion) lives in these large cities. So, it is not just urban life in general that is increasingly defining our coexistence, but quite explicitly metropolitan life.

More and more cities of this dimension are emerging. And these cities have on average roughly doubled their settlement area over the 30-year period from 1985 to 2015. These dynamics – as we were able to show – are highly variable worldwide: The highest relative settlement growth rates have occurred in Asia and especially in East Asia, the lowest in Oceania and Europe. And, we were able to show that with increasing size, cities grow slower. At the same time, however, shifts can also be observed from the analysis – African cities are remaining dynamic in terms of growth. With this study, we were able to empirically compare cities based on a uniform delineation of urban areas. Thus, we are convinced that the measured trends are statistically not biased and close to the truth. If we project these growth trends into the future and relate them to the IPCC's climate change scenarios, then it becomes clear where the exposures and risks for city dwellers will increase most: The fastgrowing cities in Sub-Saharan Africa will be confronted with highest maximum precipitation increase. In contrast, cities in Europe, while less dynamic in growth rates, will be most exposed to temperature increases globally.

At best, this empirical study is a continuation as well as a starting point to further elaborate urbanization trends in comparable studies: In doing so, it will be interesting to see whether alternative yet consistent spatial delineations of cities yield stable trends, whether alternative or expected better data from remote sensing in the future further confirm these trends, and whether the trends measured here continue over the next several decades, among other questions. Urbanization will continue; using better data and methods to drive analysis remains a key challenge; but the ultimate challenge remains to turn this knowledge of dynamics into sustainable solutions of city making.

Appendix

Supplementary data to this article can be found online at
https://doi.org/10.1016/j.cities.2024.105117

Chapter IX

The digital urban frontier: Disparities in social media activity between consolidated and newly urbanized areas in Africa

Applied Geography, 2025, vol. 181, 103687

Johannes Mast, Marta Sapena, Henri Debray, Justice Nana Inkoom, Richard Lemoine-Rodríguez, Christian Geiß & Hannes Taubenböck

© 2025 German Aerospace Center (DLR). Published by Elsevier Ltd. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.1016/j.apgeog.2025.103687

Received 26 Dec 2024; Accepted 2 June 2025; Published 11 June 2025

Abstract

Africa's cities can be seen as the sites of the world's current urban frontier, where potentials for urban growth and digital transformation are high. This study compares older and newer urban areas in Africa to investigate whether there is a digital divide in social media activity between them. Across 135 sites in Africa's fastest-expanding cities, we identify age of built-up areas using remote sensing data and relate it to social media activity, measured by the density of geotagged tweets. We quantify disparities in tweet density and use permutation tests, descriptive statistics, and linear regressions to relate them to geographic regions, degrees of urban plannedness, and spatial structure of settlements. Findings show clear evidence of disparities. In most sites, more recently built areas featured significantly lower tweet densities than older ones, while the inverse case was not observed. These disparities were more frequent in Western Africa than in Northern and Southern Africa and decreased with increased plannedness and compactness of the settlement structure. We conclude that in many African cities, the digital urban frontier is lagging behind the physical one, creating a hitherto undocumented form of digital divide that should be considered in plans for the transformation of Africa's cities.

Keywords

Social media; remote sensing; Africa; intensity of plannedness; city; urban frontier; urbanization

1. Introduction

In the 21st century city, urban space is both physical and digital. Terms such as "smart city", "digital twin", and "digital citizenship" illustrate that urban communities are increasingly complemented by information-rich overlays and digitally represented (Robertson & Feick, 2016) on social networks, web-mapping sites and content-sharing platforms. The spread of information & communication technology and geospatial software applications make the digital space increasingly important for how individuals and organizations perceive the city and, by extension, make decisions within it (Törnberg & Uitermark, 2022). In this way, the power to shape the cultural, social, and economic reality of urban areas is partly dependent on the visibility they have in the digital sphere. But this visibility is not evenly distributed, neither among citizens nor among the spaces in which they live (Boy & Uitermark, 2017; Zhu & Lerman, 2016). The competition for attention has led to the development of entire industries and technologies around an "urban attention economy" (Törnberg, 2023). Online visibility is a form of power, and it is in limited supply: not every place can be the center of attention. But which places are?

The spatial distribution of online visibility on social media can be found in explicit form on web platforms that offer a geolocation functionality, such as Twitter (now X) or Instagram. Compared to other social media services like Facebook or WhatsApp, these platforms are notable because they allow users to explicitly link their contribution to specific geolocations, creating geolocated social media (GSM) data. The result is a layer of GSM activity that covers cities like a perceptible "digital skin" (Rabari & Storper, 2015) – a skin that exists in both virtual and geographic space and has the potential to connect the two worlds. As Kelley (2011) describes it, GSM are positioned to fill the gaps between the material landscape and the socio-cultural facts of how we perceive, experience, and interact with this environment. For this reason, they are frequently used sources of information for individuals and indicators for places' positions in the urban attention economy. Importantly, online visibility can be approximated via GSM activity (the number of GSM content pieces attached to a certain location) and reception (number of responses, typically in the form of likes, shares, or replies).

The distribution of GSM activity is highly heterogeneous across and within cities (Robertson & Feick, 2016). Understanding variation across urban space can support urban studies in several ways: Firstly, it helps identifying types of intraurban areas for which GSM data is promising for research and planning. Secondly, understanding what drives differences in GSM activity can reveal digital disparities and underlying socio-cultural inequalities (Lemoine-Rodríguez, Mast, et al., 2024). Previous works identified inequalities in social media use along demographic lines such as gender, ethnicity, and age (Hargittai, 2020; Malik et al., 2015; Wentrup et al., 2016). In geographic space, disparities exist between countries (Huang & Carley, 2019), within countries (Sanderson et al., 2024), and even within cities, between more and less prosperous parts of neighborhoods (Taubenböck et al., 2018) and from centers to peripheries (Jiang et al., 2016). However, one aspect of the digital urban divide was not researched yet, despite its increasing relevance considering the rapid growth rates of many cities: The digital disparity between older and newer parts of settlements.

New settlements are no negligible part of the urban landscape: As humanity is undertaking its largest migration ever – from rural to urban areas (Saunders, 2010) – cities are growing at unprecedented rates due to a combination of natural population increase, migration, and the transition of formerly rural to urban areas (UN DESA 2019). The way in which cities accommodate the newly arriving population is mostly through the creation of new built-up areas (henceforth NBA), usually close to existing older built-up areas (henceforth OBA). Although population densities are also increasing in OBAs, NBAs absorb the majority of urban population growth. Analyses of satellite imagery over the past decades have shown that the urban extent of many cities has increased at astonishing speed and in a wide variety of patterns (Taubenböck et al. 2020, 2024, 2025). This diversity is the result of differentiated paths of historical processes that directly impacted urban developments, such as demographic growth, colonization (Şen, 2024), economic growth, and changes in planning culture (Taylor, 2013). Even within the extent of a same city, differences in urban morphology testifies to these historical processes. Trying to summarize them in a general and reproduceable way, Debray et al. (2023) showed that the connection between historical processes and intra-urban morphology can be captured through the concept of intensity of plannedness (henceforth IoP). The concept of IoP summarizes the combined historical processes influencing intra-urban layout on a gradient of how forcefully it has been planned. This gradient is conceptualized in five categories ranging from very spontaneous urban developments resulting from the decisions of many individuals to extremely planned development resulting from the manicured planning of a single institution or estate company.

While the underlying processes are manifold, they affect settlements' form and density, which are related to socio-economic factors and the quality of life of their inhabitants (Debray et al., 2023; Sapena et al., 2020).

For all these reasons, NBAs across the world, even when sharing a same development date, can be very heterogeneous in terms of morphologies. Therefore, NBAs cannot be associated with a specific and unique morphological type of urban fabric: they stand on their own as a spatiotemporal frame in the dynamic process of urbanization. NBAs are characterized by two factors: 1) They are new i.e., they consist of recently constructed buildings; 2) They are places where many (or all) new residents have only recently settled. While we cannot assume anything about the distance people moved to relocate to these newly built housing areas, this still makes them places of comparatively high population dynamism, and thus an interesting unit of analysis for migration research. And it is the first characteristic, their recency, that allows their large-scale analysis through remote sensing time series. The temporal granularity of satellite data supports fine-grained analyses of settlement age which have revealed the massive extent to which cities have grown in recent decades (Lemoine-Rodríguez et al., 2020; Liu et al., 2020; Taubenböck et al., 2024). While the distribution of GSM in the urban space has been analyzed in several studies (e.g., Taubenböck et al., 2018; Jiang et al., 2016; Cai et al., 2017; Lv et al., 2021; Robertson & Feick, 2016; Yin and Guangqing, 2021; Lang et al., 2022; for reviews of the use of Twitter data in urban geography, see Zhu et al., 2022; Smith et al., 2025), to our knowledge the link between GSM activity and settlement age remains unexplored.

This is especially relevant on the African continent, which has, together with Asia, the highest urban growth rates (Kamana et al., 2024; Taubenböck et al., 2024) and between 2020 and 2050 is predicted to double its population from 704 million to 1.4 billion people by 2050 and increase its total urban footprint from 175,000 km^2 to 450,000 km^2 (OECD, 2025). Because of this, Africa can be considered the place where we find the world's urban frontier: A region where new territory is transformed into urban space at unparalleled pace and at the same time, much urban potential yet to be realized (Goodfellow, 2022). Compared to elsewhere, much of Africa's urban expansion is unregulated (Güneralp et al., 2017), with substantial heterogeneity in urban morphology (Taubenböck et al., 2020). At the same time, a high degree of urban primacy indicates that it is traditional urban centers that still dominate most African countries, with disproportionately less resources and attention given to other urban areas (Güneralp et al., 2017), suggesting a divide that may also extend into the digital sphere. In other words, the digital urban frontier may lag behind the physical urban frontier. Whether this disparity exists, and whether regional characteristics or the plannedness and structure of settlements influence it, has not been investigated so far.

In this study, we address this research gap. Our hypothesis is that the settlement age affects how GSM is used, influenced by a possibly increased affinity of new residents to discuss their new environment, NBAs' novelty providing initial advantages in terms of economic attention, or, potentially, disadvantages resulting from delayed development of cultural, social or economic places and telecommunication infrastructure. To validate our hypotheses, we address the following research questions:

How does the density of GSM activity differ between new and old areas of African settlements? If imbalances exist, do they differ in frequency and magnitude (2) between regions of Africa, (3) between spontaneous and planned settlements, and (4) based on the spatial structure of NBAs?

By identifying if –and where– a digital divide can be measured in Africa's cities, we aim to supply the knowledge base on urban development with a digital perspective to enable effective urban strategies (Kamana et al., 2024) that adequately incorporate all parts of urban agglomerations in a digital transformation (African Union Commission, 2021; OECD, 2025).

2. Data acquisition and processing

2.1. Study period and study sites

We conducted this study for a ten-year period between January 2010 and December 2019. We selected this timespan due to the availability of a consistent set of geolocated social media posts from the platform Twitter where the geotagging feature became first widely used in 2009 and substantially changed in 2019, reducing availability of fine-grained geotags (Kruspe et al., 2021, pp. 212–221). To identify the age of all built-up areas that already existed before this study period or were built-up during the study period, we used remote sensing data to map the built-up year not just for the study period, but as far back in time as data allows (see Section 2.3). It was not possible to analyze the entire African continent due

to data limitations. While remote sensing data is available, access to social media data was less abundant. Consequently, we conducted a site-selection process to identify 135 study sites of 5 km radius (Fig. 1) where the most urban area was added between 1985 and 2019, according to the World Settlement Footprint dataset (Marconcini et al., 2020). Focused on capturing any areas where strong growth occurred, the sites are not intended to reflect any particular urban boundaries and their centroids often do not correspond to a city center. Additionally, we ensured the representation of different geographic and cultural contexts (Fig. 1) by stratifying sites across United Nations (UN) georegions (United Nations Statistical Office, 1982), selecting at least 20 sites for each georegion. A description of the site-selection process and the site outlines can be accessed in Supplement A.

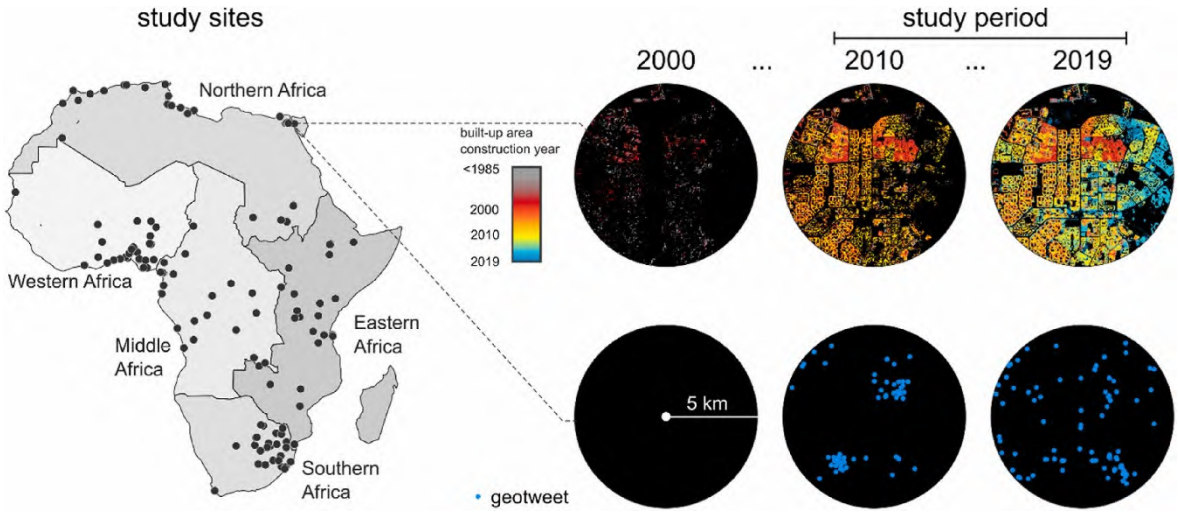


Fig. 1. Overview of the 135 study sites (left) and the primary data sets: geotweets (lower right) and built-up areas with construction year (upper right), mapped by the World Settlement Footprint Evolution (WSF_{Evo}) and the Do-it-yourself built-up (DIY-BU) mapping tool.

2.2. Built-up expansion

For each site, we used remote sensing data to map the yearly expansion of built-up area. Due to the absence of frequent long-term data on built-up expansion in Africa, we extended the World Settlement Footprint Evolution (WSF_{Evo}) dataset by Marconcini et al. (2021) which provides settlement extents from 1985 to 2015, using the Do-it-yourself built-up (DIY-BU) mapping tool (Sapena et al., 2024) for the years 2015–2019. Unlike WSF_{Evo}, which is derived from Landsat imagery at 30 m resolution and global scale, DIY-BU uses Sentinel imagery to produce local urban expansion maps at 10 m resolution. Compared to global multi-temporal products, its locally fine-tuned approach using building footprint data has been shown to provide more accurate results (for more details, see Sapena et al., 2024). We resampled the WSF_{Evo} to match the spatial resolution of the DIY-BU maps. Through spatial overlap, we assigned the earliest date from WSF_{Evo} to the built-up pixels in DIY-BU maps, resulting in urban expansion maps covering 1985 to 2019 at 10 m resolution (Fig. 2).

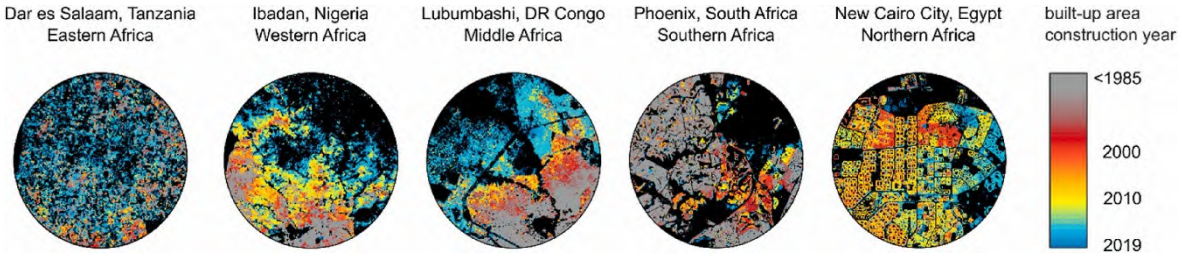


Fig. 2. Examples of urban expansion in study sites.

2.3. Twitter data

As a source for GSM data, we used geolocated tweets (geotweets) from the microblog platform Twitter, a favored data source in many fields because of its coverage and widespread use (Karami et al., 2020). Twitter users may attach geolocation in several ways, primarily through place-based tags of various precision levels (point-of-interest, neighborhood, city, administrative, and country), or through precise coordinates obtained from the device or third-party applications (Kruspe et al., 2021, pp. 212–221). We were granted access to Twitter's API via Twitter's Academic Research Track (Twitter Inc, 2022), which allowed for the query of 10 million tweets per month via V2 of the Twitter API. We used the R package academictwitteR (Barrie & Ho, 2021) in December 2022 and January 2023 to download all geotweets available for the study period (1. January 2010–1. January 2020) and bounding boxes of the study sites. From this set, we kept those that we could precisely locate within the circular boundary of a site via a point-of-interest-level place tag or precise coordinates. If a geotweet contained both, we geolocated the tweet using precise coordinates.

The availability of sufficiently precise geotweets varied over time, likely due to changes in Twitter's API, applications, and policies, or user preferences. Of 14,391,264 geotweets initially returned by the query, we discarded those with insufficient spatial precision (coarser than point-of-interest-level) and excluded accounts producing many geotweets with clear evidence of automation (unusually high user activity, speed, and concurrency of tweets, compare Petutschnig et al., 2020; Lemoine-Rodríguez, Mast, et al., 2024; Mast et al., 2024). This resulted in a final total of 6,997,346 geotweets for the 135 sites. No metadata was recorded other than account ID and the calendar year in which each geotweet was posted. We integrated the geotweet and remote sensing datasets by annotating each geotweet with the construction year of the built-up pixel at its location. Geotweets posted from non-built-up pixels were excluded from the analysis.

3. Analysis strategy

After the datasets were thusly integrated, the disparity in geotweet activity between OBAs and NBAs could be addressed analytically. The simple distinction between OBA and NBA permitted limited nuance, thus, we instead treated built-up age as a continuum. Along this continuum, the disparity in geotweet density was quantified and evaluated with significance tests at the site-level to answer research question 1 (Fig. 3). Subsequently, research questions 2–4 were tested by relating the significance and magnitude of the disparity to geographic

regions, site-level planning, and spatial metrics of settlement structure, as explained in sections 3.4–3.6.

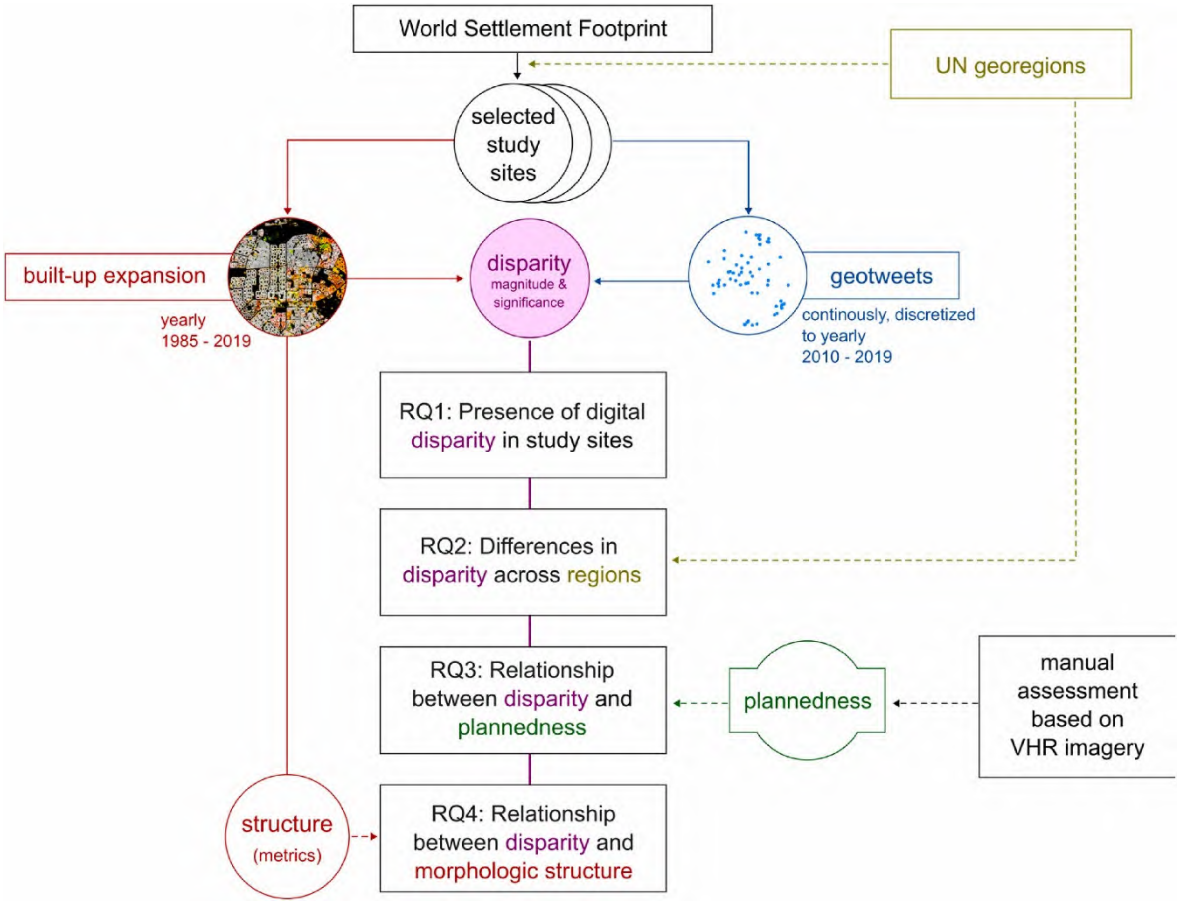


Fig. 3. Schematic workflow of the study.

3.1. Geotweet density

We sought to attain a general understanding of the relationship between built-up age and geotweet activity in our study sites. Therefore, we combined information on geotweets and on the spatial dynamic of settlement expansion. We operationalized this for the geotweets by weighting them by their impacts. We structured each AOI into a spatial grid of 100-m-wide square cells and computed the weight for each geotweet as:

$$W_{tweet} = \frac{1}{n_{per\ user}} + n_{retweets} + n_{replies}$$

with $n_{per\ user}$ being the amount of geotweets posted by a given account within the same month in the same grid cell, $n_{retweets}$ being the number of times the geotweet was re-tweeted and $n_{replies}$ the amount of replies the tweet generated. In this way, we reduced the local impact of highly active users and shifted the focus from the geotweets' production to their visibility, assuming that a retweet, quote, or reply of a geotweet increases the digital representation of an area just like a new geotweet would.

From the geotweet data, we associated the weight of the geotweet, its geolocation and its publication date. With this, from the built-up expansion data, we retrieved the date of

construction of the built-up environment of the tweets' geolocation to compute the age of the built-up environment at the publication date of each tweet. After aggregating this information by publication year of the geotweets and by age of the built-up environment, we computed the weighted density of geotweets as:

$$density_X^{age} = \frac{Ntw *_X^{age}}{area_X^{age}}$$

with $Ntw *_X^{age}$ being the count of tweets adjusted by their weights for the year X (between 1985 and 2019) for the category of built-up age of their geolocation, denoted age. Accordingly, $area_X^{age}$ is the surface occupied by the built-up environment of the category of built-up age age as of the year X of the geotweet publication. In other words, for each site, we computed the distribution of the weighted geotweet counts across built-up age of the settlement in the period 1985–2019 (see Fig. 4).

3.2. Quantifying disparity

This procedure yielded a sequence of geotweet density values for each site and year, which can be seen as a distribution of density by age of construction (Fig. 4f).

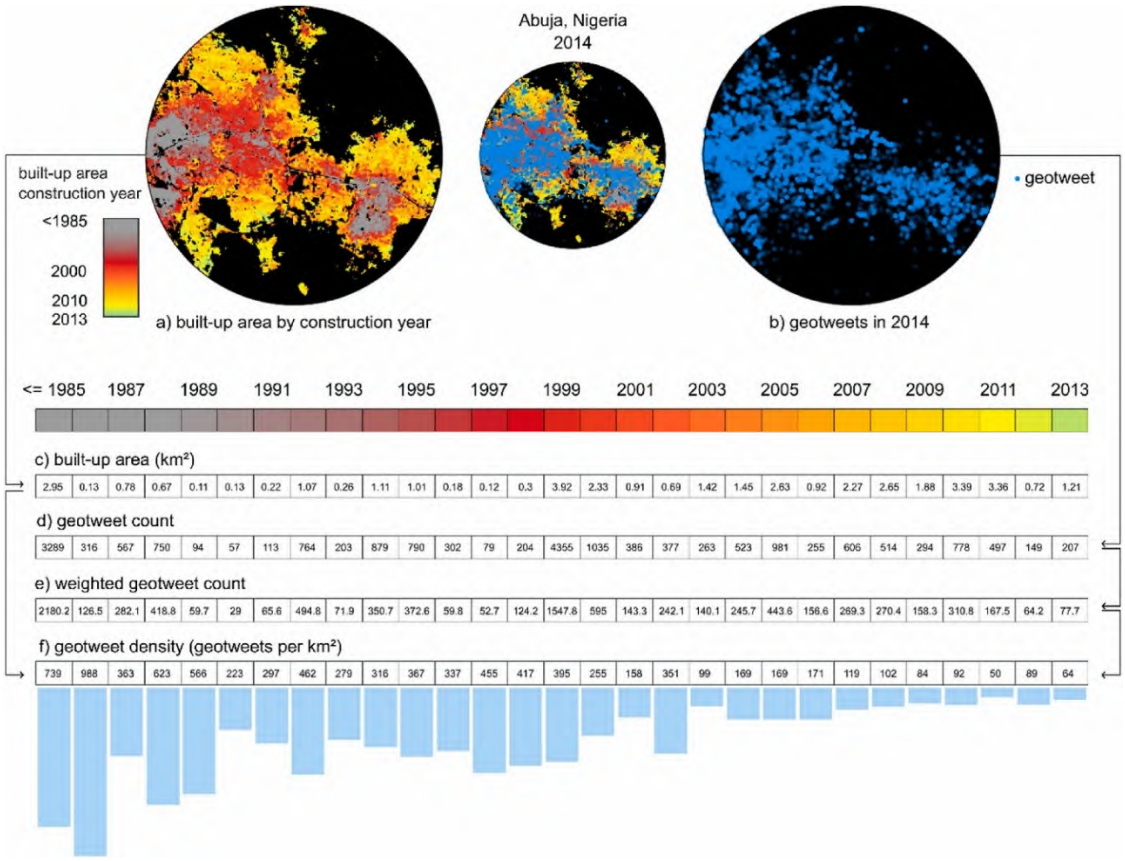


Fig. 4. Derivation of density distribution for the year 2014 in a single site (Abuja, Nigeria). Items (a) and (b) display the two input data sets used to compute the built-up area (c) and the geotweet count (d & e) per construction year, which are joined to compute the geotweet density across construction years, where the geotweet density is higher in older areas (left) than newer areas (right).

Our first research question concerned the balance of this distribution, which we hypothesized to be skewed towards the new or old areas. In order to compare this age-related digital disparity across study sites, we designed a new measure. This measure is similar to skewness metrics and represents the degree to which the measured density disparity approaches the theoretically maximum possible disparity. We refer henceforth to this measure as the digital density skew (DDS) and describe its computation in detail in Supplement B where we also compare it with alternative measures. The DDS is bounded between −1 and 1, respectively corresponding to the cases where all tweets posted for a given year are located only in the newest (DDS = −1) or the oldest parts (DDS = 1) of a settlement (in the year the tweets were posted). Accordingly, a value of 0.5 indicates that the disparity is half of its theoretical maximum. The value of zero indicates balanced geotweet density between older built-up areas and newer built-up areas of the site (relatively to the other values in the observed sequence). While this interpretation of new and old built-up as relative to the observed sequence is not necessarily appropriate in the historical context of every site, it is consistent and comparable. Figure 5 illustrates how the DDS responds to various distributions.

After calculating the DDS separately for each of the 10 calendar years of the study period, we calculated the overall DDS for each study site as the mean DDS over all years, excluding years that did not contain at least 100 geotweets produced by at least 10 different accounts. Further, we excluded study sites that did not contain tweets by at least 100 different accounts overall.

Additionally, we tested the robustness of the DDS towards GPS-inaccuracies, low numbers of geotweets, low numbers of users, and the impact of highly active accounts (Supplement C).

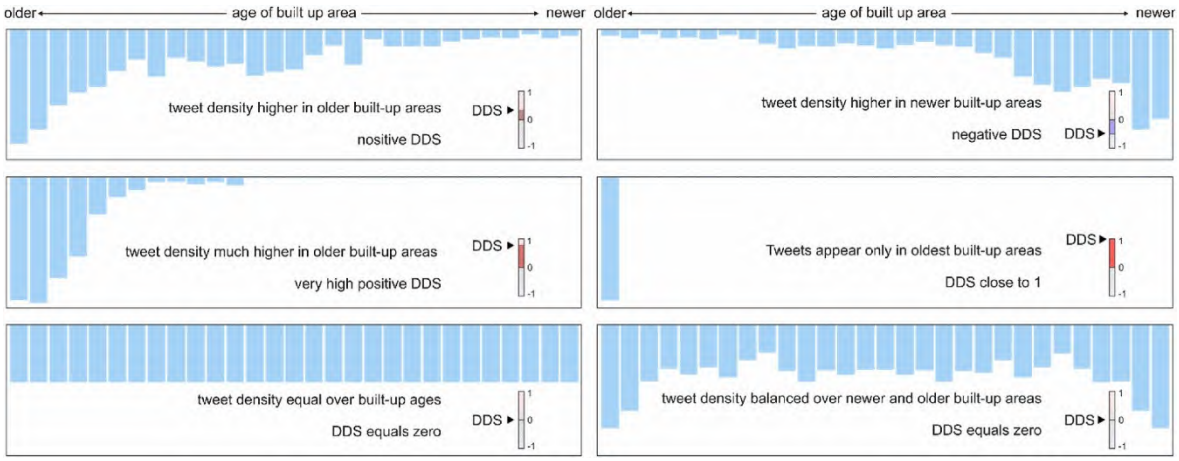


Fig. 5. Several fictive tweet density distributions (compare Fig. 4 above) and their quantification by the disparity measure digital density skew (DDS).

3.3. Testing disparity at site level

The magnitude of the DDS indicates the degree to which there is an observed association between age and geotweet activity within a site, but it does not indicate its statistical significance.

To assess the significance, we compared the DDS against a null distribution. Because the weighting approach and topologic effects might affect new areas differently than older ones, the probability distribution of the DDS is not necessarily symmetric and centered on zero. Thus, to create a null distribution that accounts for these effects, we used a permutation test (Supplement D). We used the R package spatstat to model the spatial distribution of geotweets for each site as a clustered point process with a Cauchy distribution (Baddeley & Turner, 2005). Unlike a completely random spatial sampling, modeling the geotweet distribution as a clustered point process reflected empirical observations about the distribution of tweets (Steiger et al., 2015).

From the model, we simulated 1000 spatial point patterns for each site, each having a spatial overall clusteredness similar to the original distribution, but with different locations of clusters. Reassigning the geotweets to these new coordinate-pairs, we repeated the process of extracting the built age, the geotweet density, and the DDS, resulting in 1000 reference DDS values which approximate a null-distribution of DDS values for each site. On this basis, we assessed the significance of the DDS via a right-tailed p-value (fraction of permutations whose DDS value was at least as high as the observed DDS) and a left-tailed p-value (fraction of permutations whose mean DDS value was at least as low or lower than the observed mean DDS), applying a significance threshold of 95 %.

A positive DDS indicated that geotweet activity was denser in older parts of settlements and a significant right-tailed test for the DDS indicated that this is higher than could be caused by chance. Conversely, a negative DDS along with a significant left-tailed test indicated that the activity was significantly denser in newer settlements.

3.4. Comparing across African regions

We tested whether digital disparity varies significantly across geographic regions. We followed the UN geoscheme to group sites into five geographic regions: Northern Africa, Eastern Africa, Middle Africa, Western Africa, and Southern Africa (United Nations Statistical Office, 1982; Fig. 1). These regions are one of many possible geographical stratifications, and like any, cannot accurately represent all cultural, historical, or ethnic divisions. Acknowledging this limitation, we still considered these regions the best suiting available systematic division scheme to the best of our knowledge, corresponding to United Nations standards, and being widely used for statistical purposes.

We applied both parametric and non-parametric ANOVAs of the DDS values over the five regions. The non-parametric ANOVA is a robust alternative that does not, like the parametric version, assume normality (assessed with a Shapiro-Wilk test (Royston, 1982; Razali and Yah 2011) and heterogeneity of variances across groups (assessed with a Levene, 1960). We report p-values and ω^2 and η^2 as effect sizes for the parametric and non-parametric ANOVA, respectively. We also applied post-hoc two-sample t-tests to identify between which regions differences can be found.

We further compared prevalence (number of sites) and frequency (proportion of sites) of significant disparity for each region.

3.5. Comparing across intensity of plannedness

According to the intensity of plannedness (IoP) ontology (Debray et al., 2023), we assessed the IoP of each site using experts' knowledge supported by visual inspection of up-to-date aerial imagery, OpenStreetMap data and Google Street View imagery.

The assessment was firstly performed individually by the authors and, in a subsequent step, differences between the authors' assessments were discussed until an agreement was reached. Each site was assigned one of the five IoP degrees ranging from completely spontaneous to completely planned at block level and at street-level separately with respect to the morphologic and structural appearance in the image data. Figure 6 shows examples for each of the five IoP degrees.

To test RQ3, similarly as for RQ2, we performed both nonparametric and a parametric ANOVA of the DDS by the five IoP degrees.

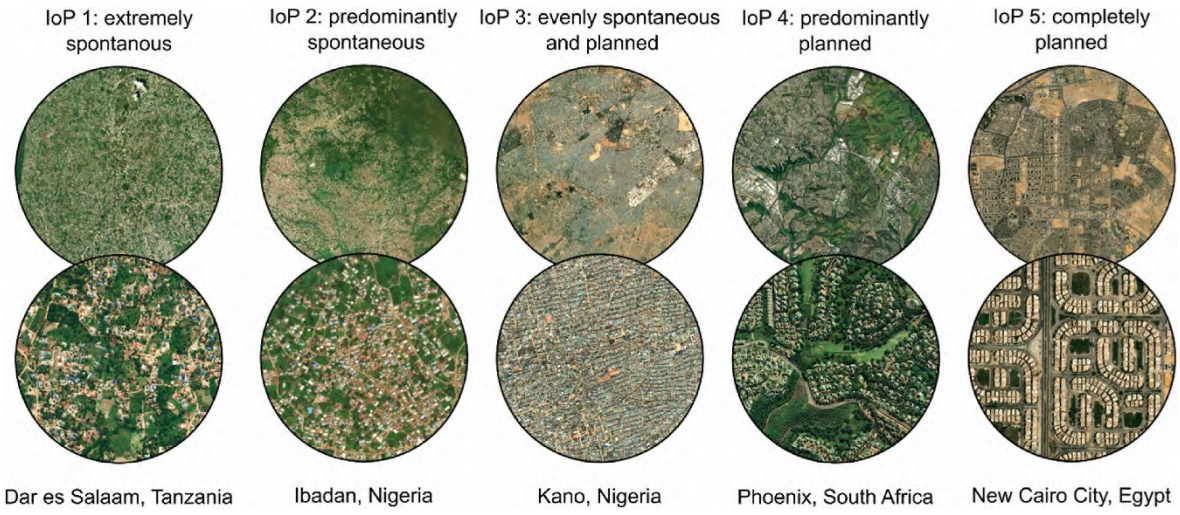


Fig. 6. Example sites with varying intensity of plannedness (IoP) (similar for both street-level and block level). Data: Microsoft (Bing Maps).

3.6. Modeling the relationship between disparity and the spatial structure of settlements

Finally, we assess whether digital disparity varied significantly according to the spatial structure of the built-up area in each site. This spatial arrangement can indicate socioeconomic processes that may affect digital disparities but are challenging to analyze directly due to a lack of data. Thus, to explore whether any links exist between spatial structure and digital disparities, we used spatial metrics to quantify the spatial structure of built-up areas in the sites. We calculated the metrics for the built-up area in 2019. The selected metrics have been shown to successfully quantify urban morphology (Sapena et al., 2020). These metrics focus on the size, shape, distribution, and connectivity of built-up areas. Metrics such as mean patch size (MPS), urban density (UD), and object density (OD) describe the extent and concentration of built-up patches, while the shape index (SI) and urban compactness (UC) capture the complexity and cohesion. Measures like the Euclidean nearest-neighbor distance (ENND) and area-weighted standard distance (AWSD) describe

the spatial arrangement and centrality of built-up patches. Additionally, the porosity index (PI), effective mesh size (EMS), and dispersion index (DI) quantify fragmentation, openness, and distribution patterns, offering a comprehensive view of how built-up areas are organized within the sites (see Supplement E for a detailed description of the metrics).

Additionally, we calculated the spatial partial information (SPI) component of entropy, as described by Altieri and Cocchi (2022) in the analysis of urban fragmentation patterns. Compared to the spatial metrics suggested by Sapena et al. (2022), the SPI can be calculated at arbitrary scales, parametrized by distance ranges, and it is inherently scaled between 0 and 1, allowing us to compare built-up compactness at various scales. A high SPI, especially at short distance ranges, corresponding to fine scales, hints at compact urban structure. Using the R package SpatEntropy (Altieri et al., 2021), we calculated the SPI at four scales based on the distance-ranges: less than 100 m; 100–500 m; 500–2000 m; 2000 m and more.

To test for a relationship between the DDS and the spatial metrics, we fitted a series of linear regression models between DDS and each metric. We report the significance of the relationships and the coefficients of determination (R^2) of these models. We fitted three different models. First, we used all sites regardless of the significance of the DDS to model the one-to-one relationship between the DDS and the spatial metrics and assess their relationship. Second, we repeated this only for the sites with significant DDS and explored the intensity of the relationship. And third, as a single metric can only represent a single aspect of urban form, we fitted a multiple regression model to uncorrelated metrics using a forward stepwise regression approach to receive a simpler model with only the significant metrics.

4. Results

4.1. Disparity at site level

We quantified the digital disparity using the DDS measure to test whether geotweet density varies by built-up age (RQ1). Overall, the mean DDS was measured at 0.201 (standard deviation = 0.16), significantly higher than the mean of any of the 1000 permutations (right-tailed p-value = 0.000), making it very unlikely to result from spatial randomness. Similarly, the mean DDS for significant sites was measured with 0.308, i.e., 30.8 % of the theoretical maximum disparity. Both perspectives support the hypothesis of an overall trend towards disparity between OBA and NBA. With the mean DDS being significantly higher than the distribution of mean DDS under the null, the results provide unambiguous evidence of a widespread, but not uniform, disparity (see examples in Fig. 7).

At the level of individual sites, 63 of 135 showed a significant right-tailed test and a positive DDS, indicating higher digital density in OBA. For 55 sites, the result was not significant, and 17 sites were excluded because twitter data was insufficient for most years. Few sites were measured with a negative DDS, and their left-tailed tests were not significant, such as *Zango, Angola* (−0.272), *Lumpundu, Congo* (−0.217), or *Al-Fashir, Sudan* (−0.195), suggesting that there were no cases where geotweet activity was higher in newer areas. Results for all sites are listed in Supplement F.

The sensitivity analyses indicate that the reported results of the significance tests are robust towards realistic magnitudes of GPS inaccuracy and the impact of highly active accounts (see Supplement C for details).

To complement these overall statistics, we provide detailed description of sites which exemplify different cases. Firstly, an example of a site where twitter activity is far denser in OBAs is *Nador, Morocco* (DDS = 0.367, p = 0.01, B2 in Fig. 7). This mediterranean city is surrounded by a lagoon to the east and mountains on the other sides. The majority of Nador's built-up area originated before the start of our study period and forms a consistently settled core area. Recent expansion occurred predominantly in A) the western suburbs in the form of scattered rows of connected houses, or in B) a highly planned area on the south-eastern shore, where many hotels and restaurants suggest a touristic focus. While this new planned area contains several hundred of the city's geotweets, the aforementioned suburbs contain very few. The vast majority of the 17 K tweets were distributed throughout the core, with geotweet density increasing along the main roads and declining towards the edges.

Ginti, Nigeria (B1 in Fig. 7), is the site with the largest settlement expansion in our study. A moderate disparity was measured here (DDS = 0.257, p = 0). Located in Ikorodu, north-east of Lagos, the site displays a continuous settlement expansion that extends outwards from the main road and is only roughly bounded by green spaces and wetlands. In contrast to Nador, almost all expansion occurred in recent decades. The roughly 62 K geotweets were distributed throughout all the built-up area, but most densely in the central area near the roads, which are spaces of social and commercial activity. Notable was a single concentration of 17 K geotweets in one location close to the center of the site, but from different accounts and with different content. While the concentration of tweets along the core areas resembles Nador, many of the core areas in Ginti are more recent, leading to a weaker measurement of disparity.

In *Sheikh Zayed City, Egypt* (B3 in Fig. 7), west of Cairo, large neighborhoods were developed in recent decades in a comparatively brief time and highly planned manner. The 110 K geotweets were distributed fairly evenly across the built-up area, with notable concentrations in malls (the two largest built around 2010) and a high-tech business district in the north (built around 2000). As these sites were built-up throughout the study period, no significant disparity was measured for this site (DDS = −0.093, p = 0.85).

Bani Walid, Libya (B4 in Fig. 7) is an oasis town divided by a valley into a northern and a southern half. Despite heavy fighting in the aftermath of the Libyan civil war in 2012 (Gumuchian, 2012), the town has expanded in south-western, south-eastern, and northern direction. Although the built-up area has doubled during the study period, almost all of the 2.5 K geotweets are located in the OBAs. There, they were highly concentrated, and in some cases, dozens of tweets were attached to the same precise coordinate. This phenomenon was occasionally observed in other sites, but particularly impactful here due to the comparative sparsity of tweets. Coupled with the low number of users, this led to a high variance of the permuted disparity. Consequently, the measured disparity, although moderately high, was not significant (DDS = 0.281, p = 0.22).

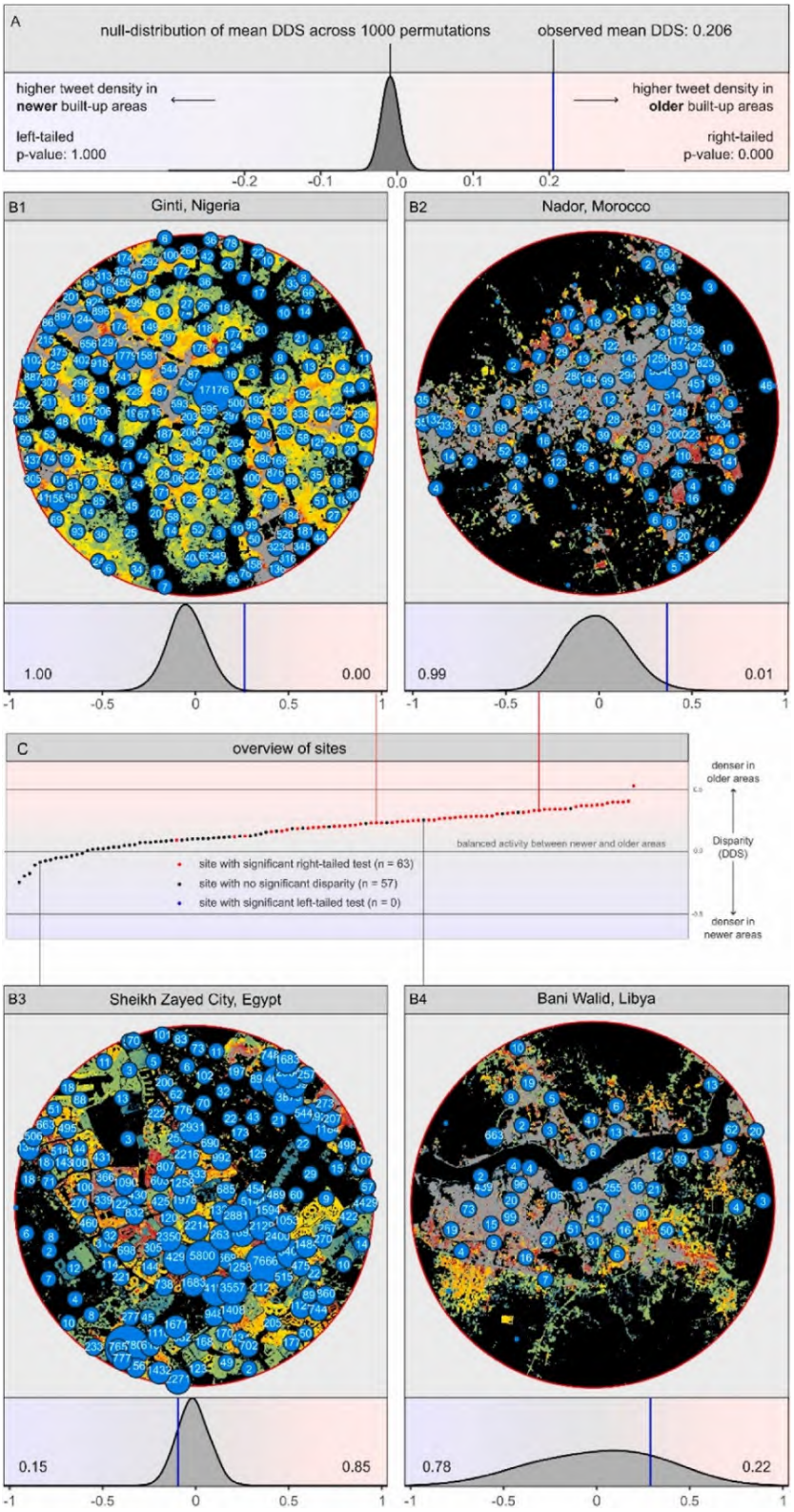


Fig. 7. Results of the significance test. Top (A): Mean DDS across all sites (blue line) compared to simulated null distribution (gray). B1-B4: Four example sites with geotweet distribution. Tweets are clustered for visualisation, with the number of tweets in a cluster indicated as white numbers. The curve below each map displays the simulated reference distribution (gray) relative to the measured disparity (blue line), with values to the left indicating higher density in newer built-up areas and values to the right indicating higher density in older built-up areas. Numbers above the curve indicate left-tailed, and right-tailed p-value respectively. Middle (C): Distribution of DDS values across study sites. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

4.2. Disparity across African regions

Both parametric and non-parametric ANOVA tests for geographical differences in digital disparity were significant (with p-values of 0.002 and 0.001, respectively), at a moderately high effect size ($\omega^2 = 0.108$ and $\eta^2 = 0.1295$, respectively). There were no extreme outliers, and the distributions of sites' DDS were approximately normal (Shapiro-Wilk test p-value>0.05) within each of the five regions, but evidence of heterogeneity of variances was found across regions (Levene-test p-value = 0.012). Altogether, this provides compelling evidence of differences between some, but not all, regions.

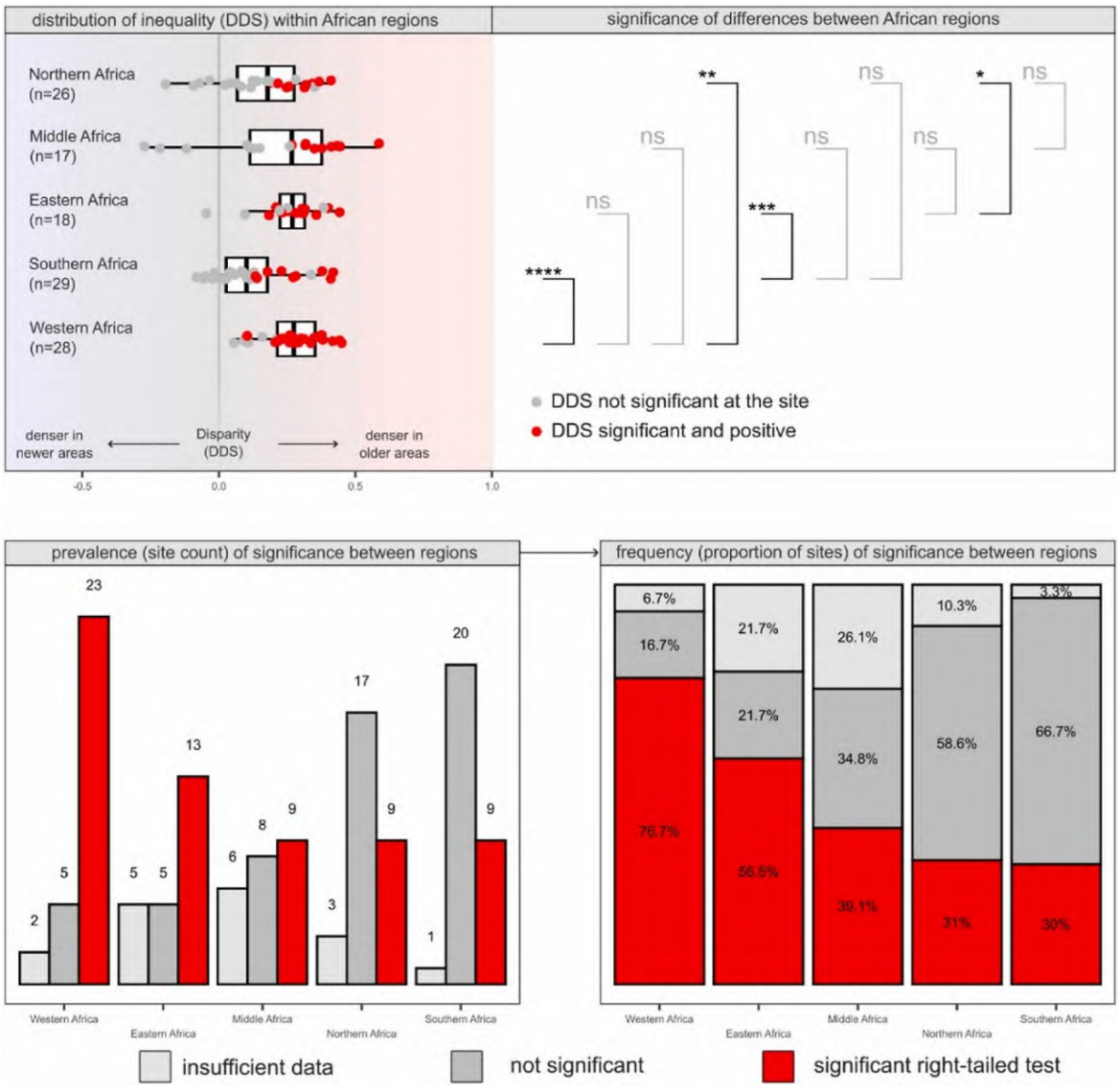


Fig. 8. Comparison of digital disparity (DDS) by UN regions. Top: Comparison of intensity via parametric ANOVA. DDS between regions was assessed with a *t*-test and are encoded as ns: no significant difference; ∗p < 0.05; ∗∗p < 0.01; ∗∗∗p < 0.001; ∗∗∗∗p < 0.0001. Bottom: Prevalence (number of sites) and frequency (proportion of sites) of significant tests within each group.

Examining the frequencies illustrated in Fig. 8, we found that the frequency of disparity varies by region. In Western Africa, the DDS was significant for most sites (23 of 29). In Northern and Southern Africa, the majority of sites did not display significant DDS. In

addition, data availability varied across regions. In Eastern and Middle Africa, more than a fifth of sites provided insufficient data, meaning that in no year, more than 100 geotweets from at least 10 accounts were available.

4.3. Disparity across degrees of IoP

Regarding the digital disparity between different levels of spontaneous and planned settlements, significant differences in DDS were found between degrees of IoP, both at street-level and at block-level (Fig. 9).

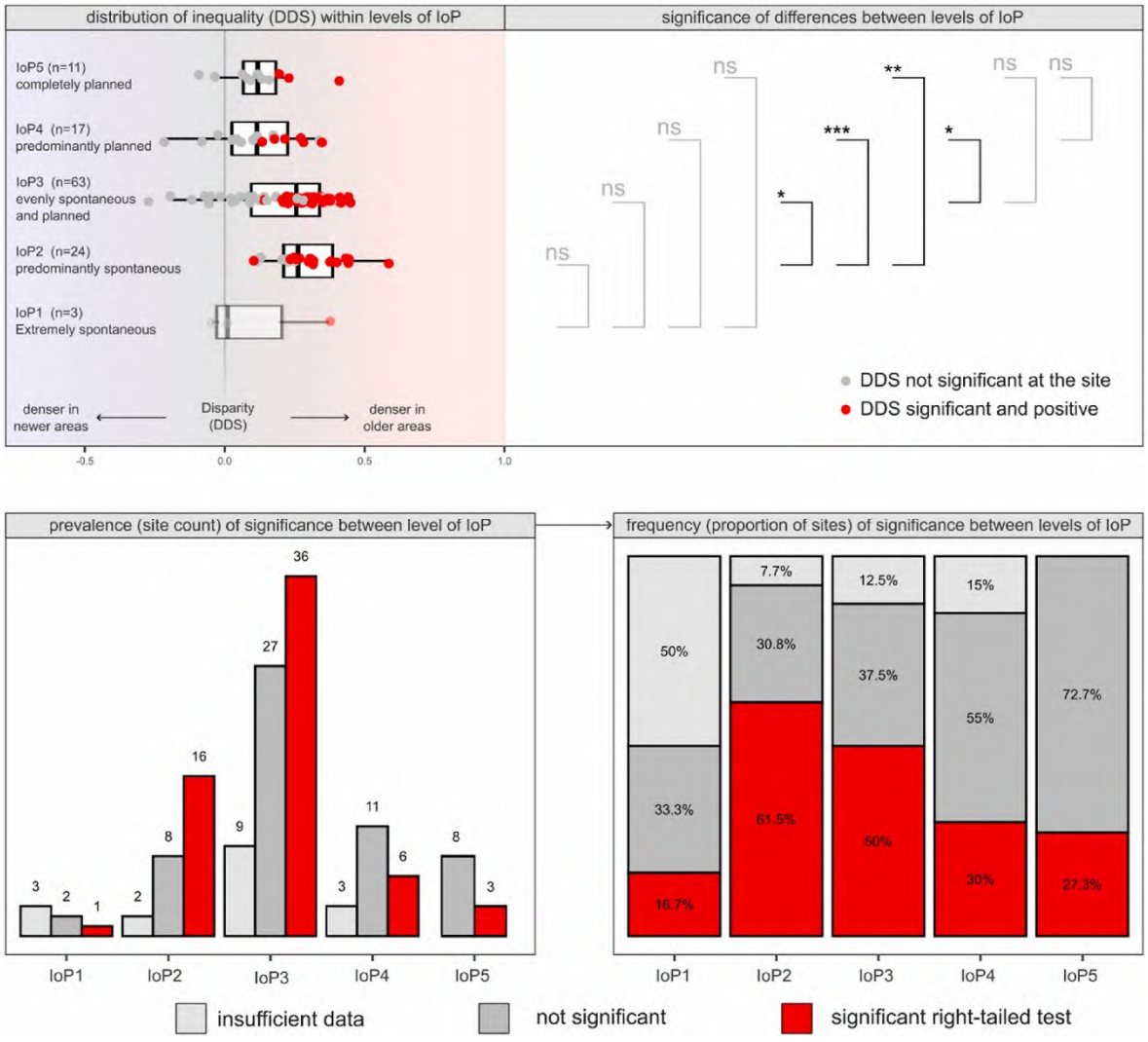


Fig. 9. Comparison of disparity by degrees of intensity of plannedness (IoP) at street-level. Top: Comparison of intensity via the parametric ANOVA test. Differences in disparity (DDS) between degrees of IoP was assessed with a *t*-test and are encoded as: ns = no significant difference; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$. Bottom: Prevalence (number of sites) and frequency (fraction of sites) of significant tests within each group.

Firstly, at street-level, the ANOVA test for differences across degrees of IoP was significant ($p = 0.003$) with a moderate effect size ($\omega^2 = 0.099$). There were no extreme outliers, and no evidence of unequal variances (Levene-test p-value = 0.593). The distributions of sites' DDS-metrics were approximately normal except in IoP 1 (completely spontaneous)–a small group with only 6 sites of which half have sufficient data. The nonparametric ANOVA

confirmed the differences between groups (Kruskal test p-value = 0.004, $\eta^2 = 0.1016$). The post-hoc two-sample t-test found the strongest evidence of differences between IoP 4 and IoP 2 (p = 0.0006), as well as between IoP 5 and IoP 2 (p = 0.0027).

The frequency of disparity varied across street-level IoP degrees. While for the predominantly spontaneous sites (IoP 2), 16 out of 24 sites (~67 %) showed significant disparity, only 9 out of 28 (~32 %) of the more planned sites with IoP 4–5 did.

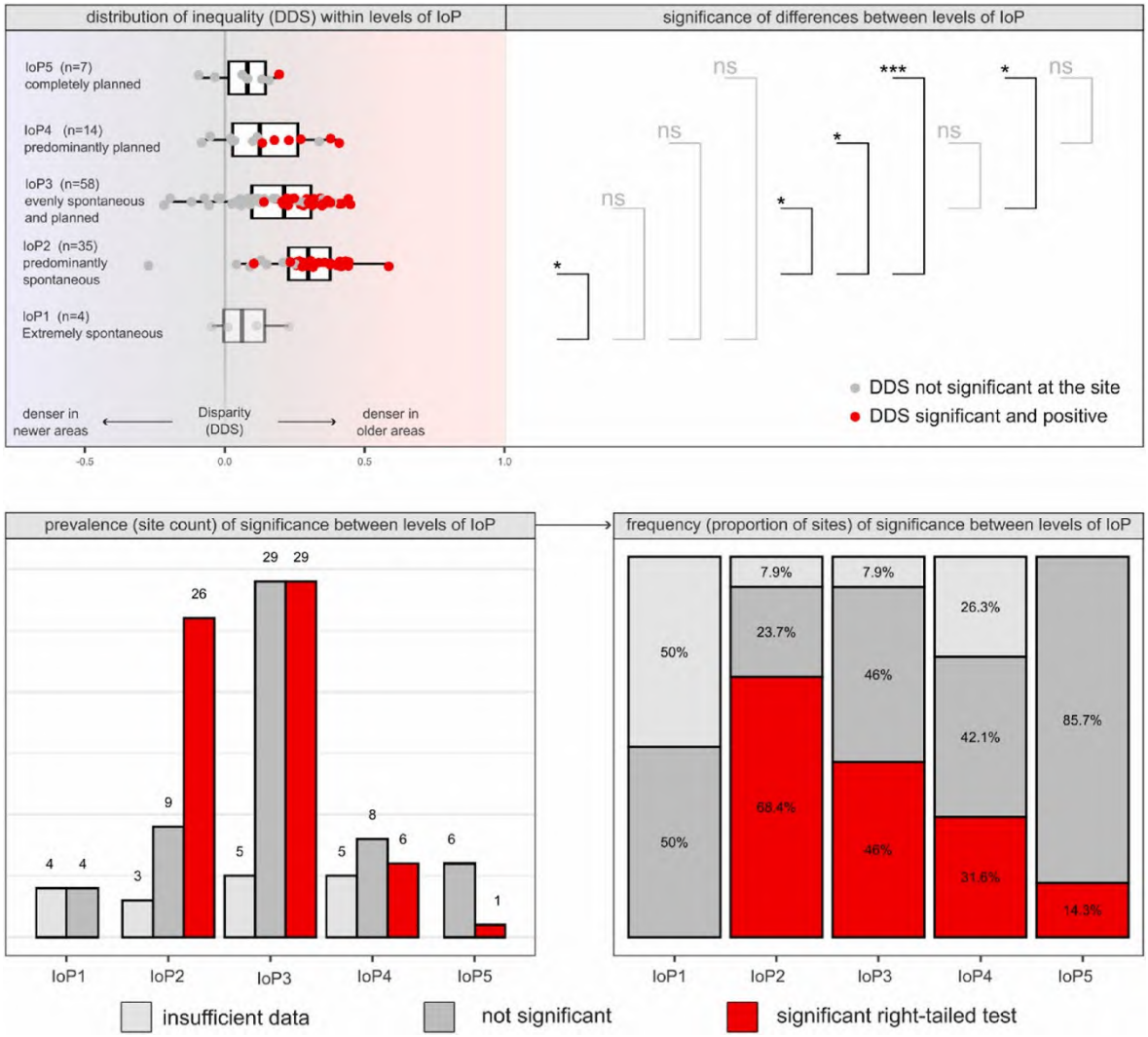


Fig. 10. Comparison of disparity by degrees of intensity of plannedness (IoP) at block-level. Top: Comparison of intensity via ANOVA. Differences in disparity (DDS) between degrees of IoP was assessed with a *t*-test and are encoded as: ns: no significant difference; ∗p < 0.05; ∗∗p < 0.01; ∗∗∗p < 0.001; ∗∗∗∗p < 0.0001. Bottom: Prevalence (number of sites) and frequency (fraction of sites) of significant tests within each group.

At block-level, results were similar. The ANOVA of DDS by degrees of IoP was significant (p = 0.002) with a medium effect size ($\omega^2 = 0.109$). There was no evidence of unequal variances (Levene-test p-value = 0.607), but IoP 2 had one extreme outlier (Zango, Angola, in IoP 2) and evidence of non-normality. Hence, we refer to the nonparametric ANOVA which found similar differences between groups (Kruskal test p-value = 0.001, η^2 was moderate at 0.1354). In the post-hoc two-sample t-test, the most significant differences were found between IoP 5 and IoP 2 (p = 0.00096), although differences also existed between

IoPs 4 and 2 (p = 0.01670), IoPs 3 and 2 (p = 0.01145), IoPs 3 and 5 (p = 0.02260), and IoPs 2 and 1 (p = 0.03601), although the latter should be interpreted with caution because of the small size of group 1, of which only 4 sites had sufficient data.

Just like at street level, the frequency of disparity varied across block levels (Fig. 10). While for the predominantly spontaneous sites (IoP 2) almost three quarters showed significant disparity, only 1 out of the 7 highly planned sites did. Altogether, the tests strongly indicate that digital disparity differs across IoP degrees.

4.4. Disparity's relationship with the spatial structure of settlements

Out of the fourteen models between metrics and DDS, eleven were significant at the 95 % confidence level (Fig. 11), indicating that the digital disparity is linked to aspects of the built-up spatial structure. However, the significant models had limited explanatory power (R^2 = 0.09–0.20), suggesting that other factors also influence DDS. Focusing on sites with significant disparity, only 9 out of the 14 models were significant (R^2 = 0.04–0.22).

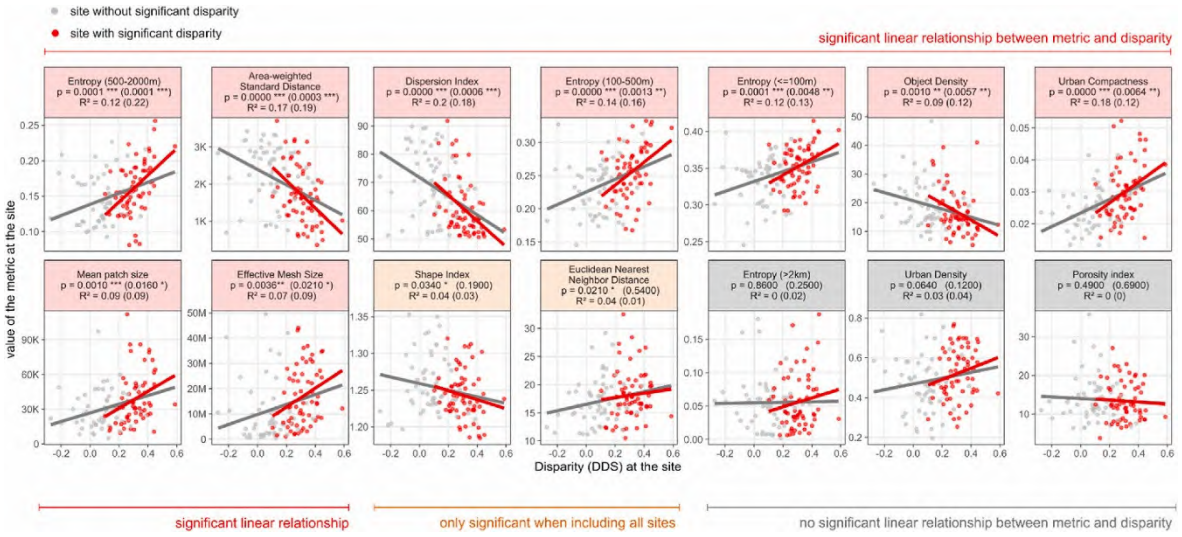


Fig. 11. Relationship between Digital Density Skew (DDS) and spatial metrics of settlement structure, reported by the coefficient of determination (R^2) and p-value (p) of their linear regression. Significance is encoded as × p < 0.05; **p < 0.01; ***p < 0.001. Values in parentheses refer to models that only include significant DDS as observations.

Notably, the spatial metrics Urban Density, Porosity Index, and SPI (>2 km) showed no significant relationship with DDS at the 95 % level. This indicates that built-up density and connectivity are not directly related to the digital density disparity. In contrast, DDS demonstrated a strong correlation with Area-Weighted Standard Distance (R^2 = 0.17, p < 0.01), Dispersion Index (R^2 = 0.20, p < 0.01), Urban Compactness (R^2 = 0.18, p < 0.01), Object Density (R = 0.09, p < 0.01), Mean Patch Size (R^2 = 0.09, p < 0.01), Effective Mesh Size (R^2 = 0.07, p < 0.01), and the SPI at <100 m (R^2 = 0.12, p < 0.01), 100–500 m (R^2 = 0.14, p < 0.01), 500–2000 m (R^2 = 0.12, p < 0.01). This suggests that a higher degree of built-up centrality, compactness, and larger and thus less contiguous patches are associated with higher values of DDS, while a higher degree of disaggregation, dispersion, and more numerous smaller built-up areas are associated with lower DDS.

The model resulting from the stepwise multiple linear regression procedure (Fig. 12 and Table 1) showed that the metrics DI, Euclidean Nearest Neighbor-Distance, and Urban Compactness together explained around a quarter of the variance in disparity ($p < 0.001$, multiple $R^2 = 0.2911$, adjusted $R^2 = 0.2724$). When only including significant sites, the only significant metric was Area-Weighted Standard Distance. The model performed better than single metric models, suggesting that a combination of complementary metrics has a higher association with DDS.

This set of metrics suggests that more compact settlements (high Urban Compactness and SPI, low Dispersion Index) with greater distance between built-up clusters (high Euclidean Nearest Neighbor Distance) and less open spaces (low Urban Density) are associated with greater disparity in digital density. Overall, an urban form which is linked to high digital disparity is characterized by a large and dense compact urban core, with several isolated settlements in the immediate periphery, but far from each other.

Table 1. Coefficients of the multiple linear regression between DDS and spatial metrics. Significance encodes the p-value as: ∗p < 0.05; ∗∗p < 0.01; ∗∗∗p < 0.001. Multiple $R^2 = 0.2911$, Adjusted $R^2 = 0.2724$.

term	coefficient	std. error	t value	p-value (significance)
Intercept	*0.296*	*0.113*	*2.626*	*0.0098* **
Dispersion Index	−0.005	0.001	−4.077	0.0001 ***
Euclidean Nearest Neighbor Distance	0.007	0.003	2.218	0.0290 *
Urban Compactness	4.000	1.845	2.168	0.0320 *

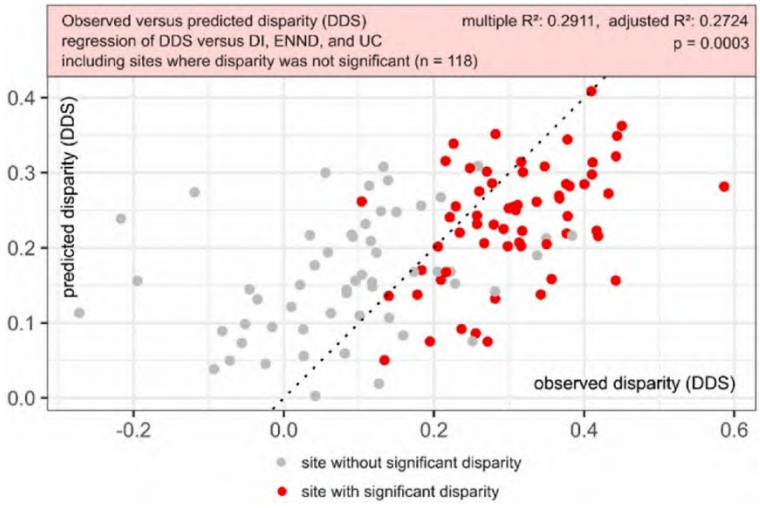


Fig. 12. Observed versus predicted DDS using a multiple linear regression.

5. Discussion

This analysis of geotweet distribution in old and new built-up areas across Africa revealed a hitherto undocumented form of multi-faceted urban digital divide. First, in many of the fastest expanding sites across Africa, we found that digital representation increases with built-up age, at least insofar as measured by geotweet density. Second, this disparity varies

in intensity across geographic regions, being strongest in West Africa and weakest in Northern and Southern Africa. Third, this urban digital divide is related to plannedness of the urban structure, with a higher intensity of plannedness linking to lower disparity. And fourth, it is related to the two-dimensional structure of the settlement, particularly built-up fragmentation.

The *first observation*, that a significant disparity exists in many study sites, is the most robust and has two main implications: From a data perspective, it indicates that geotweets, and perhaps GSM in general, are a scarcer data source in new neighborhoods compared to old ones. This does not discourage the use of GSM data altogether, as for 63 of our 135 sites, we find more than 10 K geolocated tweets in the areas that were built-up since 1985. Clearly, data scarcity is not a general rule. Therefore, our conclusion is not that NBAs are digital deserts (Taubenböck et al., 2018), but A) that NBAs may be underrepresented in urban studies relying on GSM; B) that data availability should not be taken for granted when designing studies based on GSM; and C) that biases resulting from unequal data coverage should be considered.

Secondly, our findings can be seen as indication of a digital divide between new and older neighborhoods of African cities. In other words, new neighborhoods are less represented online than older neighborhoods in the same city. There is no single reason for this, as many factors can play a role, from socioeconomic to technical. Based on inspection of our study sites, we assume major influences by density of activity, population, and infrastructure. Unfortunately, there is no sufficiently fine-grained data capturing population density and infrastructure development at Africa-wide scale. Thus, we cannot assess these factors in our study. Even if data were available, the large range of potential factors and the diversity of urban spaces in Africa would require a larger sample than 135 sites to analyze. Our results, however, provide some initial indications which can guide future research: Disparity differs between regions, being greater and more frequent in the study sites of Western Africa than in other parts of Africa, and least frequent in Northern and Southern Africa. Possibly, the higher penetration of internet and communication technologies (Nchake & Shuaibu, 2022) and the slower pace of urbanization in Northern and Southern Africa (UN DESA 2019) have a mitigating effect on the disparity. There is likely some link to the intensity of plannedness, as the disparity is lower and less frequent in more intensely planned sites, which are more common in Northern and Southern Africa. While our approach is insufficient to determine causalities in the relationship, it is possible that the uniformity in design simplifies the provision of infrastructural resources across the settlement. Further, we find that settlement structure is likely to play a role in the magnitude of the digital divide, although the models' explanatory power is limited. Our interpretation is that there are other (confounding) factors linked to plannedness and settlement structure which influence the magnitude of digital disparity. Almost certainly, the availability of infrastructure and the density of population and cultural or financial capital play a role. Given sufficient data, statistical approaches could disentangle the causal relationships between the physical form of settlements, their built-up age, their socioeconomic structure, and their visibility in the digital sphere. As our study shows, time-series of the settlements' physical form and digital activity already exist at an Africa-wide scale and intra-urban resolution. Unfortunately, it is unlikely that

socioeconomic data can be gathered consistently at this scale. Instead, interview-based surveys remain relevant for the study of digital divides. Existing surveys could be expanded to inquire about barriers to social media use and preferences if they do not already do so. Additionally, the inclusion of data on urban functions, land cover, and land use is a promising next step. Such data have been used to assess the impact of local urban functions on digital urban vibrancy (e.g., Lang et al., 2022). The combination of remotely-sensed settlement age with land cover and land use classifications has revealed that the land cover of recently urbanized area differs substantially from older urban areas and between geographic regions (Taubenböck et al., 2025). It will be worth investigating whether the divide between the digital and physical frontier on the African continent is caused by such differences in urban function and land use between newer and older urban areas.

While our big-data approach allowed a comparative study at continent-spanning scale, we acknowledge its limitations. The urban expansion maps are not perfectly accurate and do not distinguish built-up ages earlier than 1985. Data quality varies between regions, potentially affecting cross-region comparisons of DDS. Hence, we implemented a permutation-based test which provides robust estimates of significance. The precision of GPS devices used for the precise geotagging is imperfect, with most studies finding horizontal errors of 5–20 m, but such minor spatial errors lead only to minor errors in the determination of the construction year, and a minor shift in DDS (see Supplement C). More concerning are cases where the number of geotweets and contributors for a site are low. This exacerbates the well-documented demographic bias of Twitter, where a vocal minority contributes most of the tweets (Lemoine-Rodríguez, Biewer, & Taubenböck, 2024; Wojcik & Hughes, 2019). We mitigate the impact of highly active accounts through a weighting procedure and confirm that their behavior is not substantially different from less active accounts for the purposes of this study (see Supplement D). Nevertheless, we acknowledge that geotweet density does not represent the entire digital skin of the urban areas but must be interpreted as a proxy for the online visibility on a particular platform. The choice of Twitter is sensible for our study because it is public and offers geotag functionality. However, in terms of market share, Facebook was found to be the leading social media platform in Africa in 2022, and there are substantial regional variations in the use of platforms and penetration of web technologies in general (Bhanye et al., 2023). Consequently, researchers should choose the platform(s) to fit the geographic setting and goals of the study based on data access. If precise coordinates are not required, geoparsing approaches (Hu et al., 2024) can serve as an alternative to the use of geotags, and provide alternative perspectives. Additionally, the access modes and capabilities of Twitter's API were substantially changed following the rebranding of the platform as “X”. As a consequence, the geotweet datasets used in our study can now only be reproduced imperfectly and at substantial cost (Davidson et al., 2023), highlighting the need to diversify web data sources in geographic research.

And just as one platform does not represent social media use as a whole, our selected study sites are not representative of all settlements. Rather, they exemplify the class of fast-growing settlements in Africa where the digital divide we investigated is most acute. Consequently, they are not a representative sample of all growing settlements, and it is possible that patterns might be different in areas of low to medium growth. We also

acknowledge that assigning a single IoP or compactness value to each study site is a strong simplification of reality. Each site covers considerable heterogeneity in almost 80 km^2 area and their circular boundary cannot be adapted to morphological subdivisions. Nevertheless, we consider it to be a suitable choice for this study because it is objective and can be derived for any location. Finally, our data is limited in that geotags primarily capture the attention given to a place by local people. While nothing prevents users from commenting on more distant places, the geotagging functionality of Twitter, for most of the study period, only allowed users to tag content with their current location. Consequently, geotagged data includes less content written about places from elsewhere, and thus provides only an incomplete picture of online representation. Geoparsing approaches (Hu et al., 2024) could alleviate this issue by enabling the inclusion of text based on mentions of places. However, most of these approaches rely on existing datasets, such as gazetteers, knowledge-graphs, or training text corpora, which may be biased themselves (Graham and De Sabbata, 2020), making them unsuitable for this study.

Our findings are plausible in the African context, and we do not expect them to generalize to the global level because geographic regions differ substantially in their urban history, planning, political systems, and telecommunication infrastructure.

In summary, despite limitations, the converging indications found in the study point strongly towards a first and clear indication of a type of digital urban divide at continental level: A divide between new and consolidated parts of African settlements.

6. Conclusion

By relating maps of urban expansion to GSM posts from Twitter, we find clear evidence of a hitherto undocumented type of digital divide–one that occurs between older and newer parts of the rapidly growing African settlements. Going forward, the perspective provided by GSM posts should be complemented by other information sources on local digitality, and further research with additional socioeconomic data is required to understand the underlying causalities of this digital divide. As a starting point, this study finds that settlements' plannedness and structure are linked to the magnitude of this disparity, and that it should be considered a trend—rather than an inescapable fact—of urban development.

By documenting this disparity, we show that the "urban frontier", which is expanding into the hinterland through the built landscape, often does not run parallel to the "social media frontier", which is lagging behind. It appears that it takes time for these new urban spaces to be filled with places, people, and activities that are representable on the web, and the population with the means and affinity to represent them. Thus, the study reveals a further layer—in the digital sphere—in urbanization processes and contributes to the inclusion of digital perspectives into the broader African Union led 2020–2030 digital transformation strategy, plans and programs, and actions for the transformation and sustainability of Africa's urban spaces.

Funding

This study has been conducted as part of the European Union's Caroline Herschel Framework Partnership Agreement on Copernicus User Uptake (FPCUP) under Grant Agreement No. FPA 275/G/GRO/COPE/17/10042, Action 2019-2-11 “Downstream service/application development for monitoring of environmental indicators”, the project MIGRAWARE, Grant No. 01LG 2082C, funded by the German Federal Ministry of Education and Research (BMBF) as part of the programme WASCAL WRAP 2.0, and the research and teaching initiative “A New Focus in English Linguistics: Geolingual Studies”, Grant No. 98662, funded by Volkswagen Foundation as part of the programme “Momentum—Funding for Recently Tenured Professors”.

Appendix A. Supplementary data

Supplementary data to this article can be found online at
https://doi. org/10.1016/j.apgeog.2025.103687

Chapter X

Border regions across the globe

Analyzing border typologies, economic and political disparities, and development dynamics

Applied Geography, 2023, vol. 151

Hannes Taubenböck, Christoph Otto, Fabian Gülzau & Steffen Mau

© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.apgeog.2022.102866

Received 18 May 2022; Accepted 27 Dec 2022; Published 3 Jan 2023

Abstract

The cosmopolitan dream of a borderless world has little to do with reality. Today's borders bear witness to regulatory intervention in the circulation of goods, information, capital and people. These interventions, naturally, have an impact at border regions. For analyzing these impacts, we map, quantify and relate border typologies, development dynamics near borders, and economic and political indicators of neighboring nation-states. We do so on global scale for all current 315 land borders. We rely on data from a mix of border dossiers, in-depth literature review, censuses and multi-temporal mapping products from satellite imagery. Our analysis strategy is two-fold: First, in a descriptive analysis, we map the various border typologies. And, we also compute development dynamics over a 15-year period from 2000 to 2015. Since there are few consistent, appropriately spatially resolved, and globally available datasets, we measure development by the proxies 'settlements' and 'population' instead of the usual economic characteristics. We use an ensemble of metrics that show not only the developments in the border region but also the dynamics in the border region relative to the respective nation-state. By means of a global ranking, we show the variability of development dynamics at borders across the globe. Second, we relate these dynamics to the different border typologies, and to economic and political differences of neighboring nation-states. We find the following trends: higher political or economic differences of neighboring nation-states relate to stronger border fortification, greater economic or political disparities relate to stronger population or settlement accumulation at the poorer or less free side of the border, and stronger fortification hinders settlement and population development to a certain degree. These empirically measured trends, however, are only partially statistically significant and not as strong or unambiguous as assumed. In a critical discussion, we reflect on the capabilities and limitations of such an empirical global approach.

Keywords

Border typologies; settlements; population; politics; economy; remote sensing; spatial metrics; spatial concept

1. Introduction

From space, astronauts see our beautiful blue planet in its entirety. They observe the distribution of land as a borderless picture. In contrast, the most common cartographic representation of our world divides land into distinct territories of nation-states. These borders demarcate jurisdictional territorial units, but they also testify to political, economic, social, cultural, religious differences in space (Popescu, 2012). This striking juxtaposition naturally leads to the questions of how and to what extent developments are influenced by these borders and, vice versa, how border typologies are defined by these developments, and if so, what the effects are near these borders.

In this paper, we approach these questions in a global approach and embed them in the sense of an existing dialectic related to borders be tween nation-states: On the one hand, the 21st century hallmarks of globalization (Beck, 1997), cosmopolitanism and networked individualism (Wellman, 2001), networked people, communities and cities (Sassen, 2002), mobilities (Urry, 1999) and flows (Castells, 2000) reveal an increasingly interconnected, and perhaps in this sense increasingly uninterrupted, borderless world beyond territorial units of seemingly less importance. Some have predicted that we are moving into the direction of a '*borderless world*' (Ohmae, 1990) and that borders have effectively lost their function, not at least due to the fact that the global flow of capital, finance, goods, services, people, ideas, and cultures was a common argument for economic growth (e.g. Friedman, 2006; Moore, 2003). With the internationally celebrated fall of the Berlin Wall and the Schengen Agreement abolishing stationary border controls at the EU's internal borders, all of this seemed to herald an era of open borders, i.e. the abolition, dissolution and crossing of borders (Kamwela & van Bergeijk, 2020). On the other hand, exactly this expectation of falling borders or even a borderless world has not materialized. A larger strand of border research has argued that though globalization indeed impacted on cross-border flows, it did neither abandon borders nor make all cross-border transactions possible. Moreover, most scholars agree that borders are still powerful institutions of demarcation and separation, in some sense, borders have even gained importance through heightened cross-borders flows and the interest of states in exercising control (e.g. Mau, 2021; Newman, 2006; Paasi, 2019). States have reacted to increased human mobility, economic interdependence and perceived external challenges by heavily investing in border infra structure and border technology (e.g. Dijstelbloem, 2021; Simmons, 2019). A world without borders, as Haselsberger (2014) argues, is a utopia: Territory cannot be managed if it is not clearly delimited. It is claimed that a '*border orientation*' has become a salient feature in international politics and that states heavily invest in infrastructure to "secure" their borders and to control movement across borders, even into fences and walls (Simmons & Kenwick, 2021; Vallet, 2016). And, according to Stiglitz (2017), the positive effects of econometric expectations in globalization have also been overestimated: Trickle-down economics assumed to ensure that everyone benefits when GDP in creases were overestimated, and the costs, including negative distributional effects, were underestimated. It resulted in the opposite of '*borderless*': more than 50 new walls between nation-states have been built in the last 30 years (Bissonnette & Vallet, 2020; Vernon & Zimmermann, 2019). These borders are there to separate the 'inside' from the 'outside', between 'us' and 'them' (Brown, 2010). That means in addition to the

dissolution of borders, a simultaneous trend towards closure, border selection and control in various border typologies exists (Mau, 2021).

In this paper, we investigate what this dialectic indirectly implies: Borders have in their differences, different implications – on many levels such as political, economic, demographic, jurisdictional, to name just a few. Since we conduct our investigation on a global level, we obviously cannot consider all levels. Here, we focus the empirical investigation to the following: we assume that border typologies, economic and political situations between neighboring nation-states, and the development of settlements and population in borders regions relate with measurable effect to each other.

Such dependencies and relations have been at the focus in manifold research studies in domains such as border rules (e.g. Simmons, 2019), border security and control (e.g. Simmons & Kenwick, 2021), customs and trade (e.g. Carter & Poast, 2020) or migration and immigration (e.g. Mau et al., 2015). Kamwela & van Bergeijk (2020), e.g., investigated economic effects of physical border walls. For nation-states separated by a wall, on average 46 to 73 percent less trade has been measured than it would ceteris paribus be the case if the border wall did not exist. Other studies, e.g. on migration, measure, contrary to expectations, that fortified borders even lead to an increase of immigration flows (Pecoud & de Guchteneire, 2006; Schon & Leblang, 2021), but it is also apparent that '*hard*' borders increase cross-border selectivity (Mau, 2021; Korte, 2021). Beyond this, there are studies investigating environmental effects of borders. For example, Trouwborst et al. (2016) or Linnell et al. (2016) document how borders cut up connected natural environments and affect habitats negatively. The cited studies investigated the impact of heavily fortified borders. A gradation of border types across this dichotomy between heavily fortified and open border, however, has not yet been considered.

The cited studies are focused on selected individual or few border regions. We are not aware of any studies on a global scale on effects related explicitly to the vicinities of borders. This is primarily due to inconsistent or missing comprehensive data sets on economics, migra tion or other relevant topics at the spatial scale of border regions across the world. The capability to spatially investigate the developments in the vicinity of borders regardless of administrative or other large reference units is challenging. Earth observation (EO) from space is currently the only data source allowing for a consistent, global analysis at these spatial scales. But still, a global analysis in this field of research is unknown. However, since economic (e.g. Chen & Norhaus, 2019) or demographic (Sapena et al., 2022) information can only be derived indirectly from EO-data, the focus in existing studies is on proxies. For example, Bennett & Faxon (2021) used nighttime light emissions as proxy for development along the fortified border of Myanmar revealing significantly lower light emissions than in the nation-state as a whole or in the border regions of neighboring nation-states. Kolosov et al. (2018) used remotely sensed information on road networks, land use and density of settlements as geographical attributes to visualize that the difference in urban development is greater at Russian borders that were also part of the Soviet Union than at post-Soviet borders. Röder et al. (2015) documented urban development on the Namibian-Angolan border based on multi-temporal land cover classifications. By and large, these studies remain limited to the specific situation at transitions between two nation-states at regional level. At

a continental level, Taubenböck et al. (2017) mapped the interconnected settlement areas of Europe across open borders using EO-data. They showed how the proxy '*settlement density*' reflects economically interconnected spaces, such as the so called 'Blue Banana' crossing many European borders. Stephenne et al. (2009) developed a border permeability model for the EU-25 and Central African land borders based on accessibility, hiding opportunities and probability of arrest by border police using a variation of EO-data, statistical databases and surveys.

However, to the best of our knowledge, assessing developments at border regions based on EO-data and derivatives such as settlements or population in combination with other data sets such as a differentiated categorization of border typologies as well as economic of political data sets has not been systematically investigated at global scale.

The remainder of the work is structured as follows: Section 2 briefly presents the conceptualization of this study. Section 3 introduces the empirical set-up, i.e. the data, the spatial concepts and methods. In section 4 the results on border typologies, rankings of settlement and population dynamics at border regions across the world are presented. Furthermore, relationships of various border typologies, economic and political differences of neighboring nation-states to settlement and population dynamics are illustrated. This is followed by a discussion in section 5, where we try to evaluate which influence determinants have on border developments and we discuss the limitations of our data, conceptual assumptions, and the related main results. Section 6 concludes with a perspective.

2. Conceptualization of this study

In this study, we investigate border regions at global scale. Unlike existing studies, our analysis is performed on all 315 current international land borders of nation-states and takes account of the two-sidedness of borders. i.e. their dyadic nature.

As basis for the investigation, we, on the one hand, record and map border typologies for all land borders. On the other hand, we compile a set of economic, political, physical and demographic indicators that are globally available and which describe the nation-states or the particular border regions: the gross development product (GDP) per capita and types of political regimes for the year 2017; the settlements and the population development dynamics for a time period of 15 years.

In the analysis, we approach border regions first through a descriptive study of border typologies and development dynamics (1), and second, we investigate relationships between the economic and political indicators, border typologies, and the settlement and demographic dynamics (2).

(1) We basically document, map and quantify the *variability of border typologies* for the current time step (cross-sectionally), as well as *settlement and population development dynamics* over 15 years in a descriptive analysis.

Conceptually, we equate higher growth rates over time with stronger development. Development is typically measured by economic indicators (e.g. Shashoua et al., 2015) or in border regions often by quantities of border crossings (e.g. Hodges, 2007) or movement of

goods (e.g. Carter & Poast, 2020). How ever, economic analyses are on the one hand mostly based on data at administrative spatial units, i.e. data are aggregated to heterogeneous spatial reference units and therefore they are not always appropriate for comparison due to differences in size and form (e.g. Openshaw, 1984). And on the other hand, these data are often only inconsistently or not available everywhere for a global study. As studies have shown that there are positive declining correlations between settlement area growth/urbanization and GDP per capita (e.g. World development indicators, 2005), we understand in this paper the variables *settlements* and *population* as proxies for development. We do so, as these data sets have been derived in comparatively high spatial resolution from remotely sensed and census data at global level. That means, these two variables allow us to spatially define the vicinity to border areas independently from any administrative unit and thus it is possible to create a uniform, consistent and thus com parable spatial reference unit near borders.

(2) For the investigation of relationships between the selected indicators, we follow two guiding hypotheses:

a) Borders separating the territories of nation-states from each other come with different typologies. We hypothesize that the way in which borders between nation-states are physically built and crossings managed – among other historical and geopolitical factors – are a direct result of political or economic disparities between neighboring nation-states. We formulate as guiding hypotheses: The higher the political or economic differences between neighboring nation-states, the stronger the border fortification, while not excluding that fortifications in place may strengthen existing differences. To put this homology assumption differently: The smaller the disparities are in economic and political terms, the less likely walled borders and strong fences should be in place.

b) Developments in settlement and population vary in the vicinity of borders and we assume the greater the economic and political disparities between neighboring states, the greater are differences in growth rates. And, we specifically assume that in the case of neighboring nation-states separated by heavily fortified borders and with extreme differences in political and economic systems, a strong pull effect leads to accumulation effects on the poorer or more unfree side of the border. This relates to the fact, that fortified borders are in most cases built only by one state and are associated with gaps in wealth and political disparities so that the function is to deter unwanted cross-border transactions and movements. Accumulation on the poorer and more unfree side would indicate a retention or magnet effect of strong border barriers.

3. Empirical set-up

3.1. Data basis

3.1.1. Typologies of borders

Borders feature a wider variety of forms than the often-used dichotomic classification of ‘open’ and ‘closed’. They may be partially physically non-existent, unguarded and crossed completely without control, but they may also be fortified and allow passage only under

certain conditions for certain groups. In this study, we rely on a categorical classification of five different border types.

In general, we understand border types according to their physical material design and the associated possibility of crossing the border. We apply a typology which is based on various data: 1) on *border dossiers*, which were compiled for all nation-states and their borders between April 2018 and October 2019. The border dossiers contain information on the physical-material layout of borders collected through an in-depth literature review, i.e. digital repositories (e.g., Scopus) and online searches; 2) *satellite imagery* of georeferenced border crossings were queried for a visual check; 3) *existing studies* have been reviewed to compare our coding (cf. Avdan, 2019; Carter & Poast, 2017; Gülzau & Mau, 2021; Hassner & Wittenberg, 2015; Jellissen & Gottheil, 2013; Jones, 2012; Simmons & Kenwick, 2021).

The typology is conceptualized in five different types of borders:

1) *Frontier borders*: borders with no infrastructure or only rudimentary infrastructure. These no-man's-land borders are often located at border crossings that are difficult for state officials to reach because they run in remote regions (e.g., jungles or deserts). The border is left to its own, but this does not preclude the development of lively informal economies.

2) *Landmark borders*: borders characterized by little state control architecture. This includes borders that have been dismantled through bilateral or multilateral treaties, for example, to facilitate the movement of goods and people (Nita, 2017).

3) *Checkpoint borders*: these are characterized by border posts installed at specific border crossings. In satellite imagery, we often find roads that become multi-lane, eventually ending in a control point where documents are checked and customs are collected.

4) *Barrier borders:* these are installed by states that want to ensure that such controls are not undermined. Barriers are used to direct border traffic toward regular border crossings. However, physical barriers such as fences and ditches that characterize such barrier borders do not extend along the entire border route, but are rather located at neuralgic border crossings.

5) *Fortified borders:* these are borders where states have erected fences or walls along the entire borderline to further discourage uncontrolled border crossings. In contrast to barrier borders, which are characterized by the installation of obstacles at neuralgic points, states with fortified borders attempt to reinforce the border line as complete as possible.

Figure 1 illustrates these five different border types.

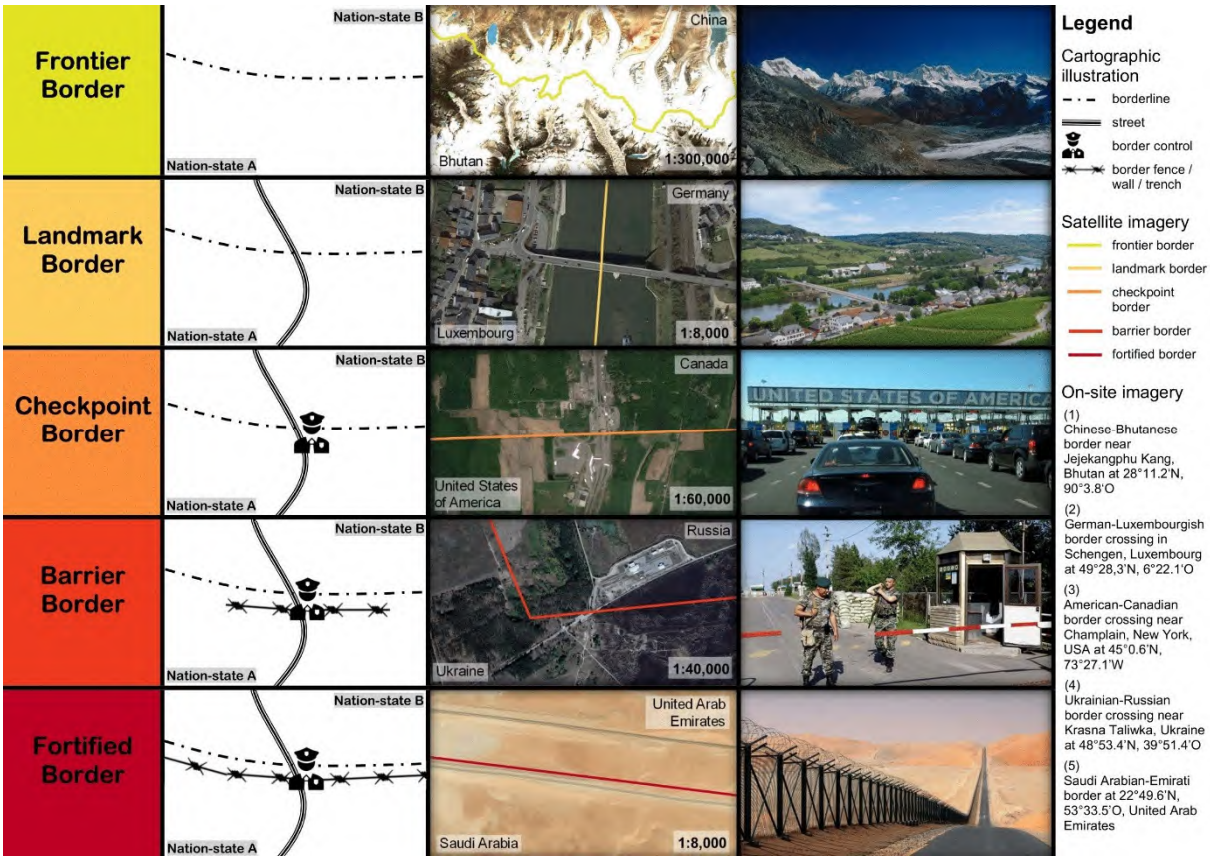


Fig. 1. Border typologies and cartographic as well as photographic illustrations: (1) Frontier border: © Steve Razzetti, http://www.razzetti.com/; (2) Landmark border: © LFT_OliverRaatz; (3) Checkpoint border: © Bobby Hidy; (4) Barrier border: © Kyiv Post; (5) Fortified border: © John A. Kelley, Soil Scientist, Environment Agency-Abu Dhabi, UAE.

3.1.2. Settlements and population

We rely on global data sets derived from remote sensing and census data. We apply the product suite of the Global Human Settlement Layer (GHSL) for measurement of two parameters: the settlements and the population. We measure changes over the time period 2000 until 2014 for settlements and until 2015 for population. We rely on the GHS_BUILT_LDSMT_GLOBE_R2018A (Corbane et al., 2018) and the GHS_POP_MT_GLOBE_R2019A (Schiavina et al., 2019) datasets, both covering the whole world with a spatial resolution of 250 m (in the following the time period 2000 until 2015 generally stands for the time period of both development proxies). The former provides the percent age of settlement area per pixel. It is derived by fully-automated symbolic machine learning from Landsat satellite data (Corbane et al., 2019). The settlement layer is used to disaggregate population census data to a population grid showing the absolute amount of people per pixel (Freire et al., 2016). Figure 2 illustrates the used data sets by the example of the western part of the Mexican-US-American border.

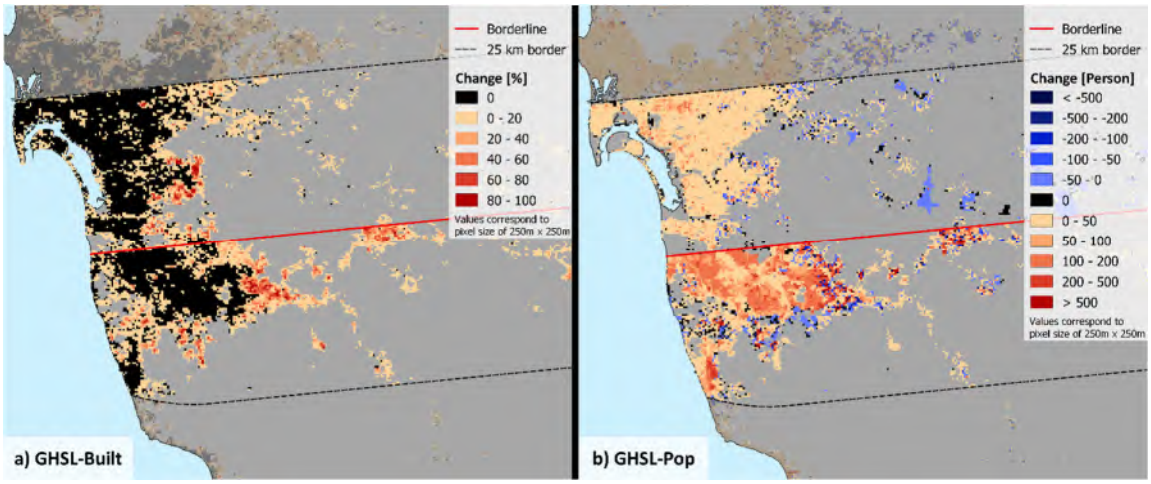


Fig. 2. Development for a 15-year period from 2000 to 2015 at the border vicinity (25km) using proxies: a) the settlements, and b) the population. Data are based on the GHSL-product suite. The example illustrates parts of the Mexican-US-American border with San Diego (USA) and Tijuana (Mexico) along the coastline. In this example it is remarkable how pronounced the area growth and especially the population increase on the Mexican side in the south is in comparison to the USA in the north.

3.1.3. Economic and polity data

As stated in our hypothesis, it is assumed that economic and political differences between neighboring nation-states affect which border typologies are implemented and how dynamically development occurs. In our study, we understand the parameters *GDP* per capita and the *classification of political regimes* as our proxy variables in this regard.

We rely on the *GDP* per capita calculated in current US-Dollars for the year 2017 (The World Bank, 2021). With this variable, we approximate the degree of economic disparity between neighboring states. And, we rely on a *classification of political regimes* for the nation-states using the situations in the year 2017. Here, we apply the political conditions to a scale from - 10 to +10: The values - 10 to - 6 are a gradation for autocratic regimes from maximum unfree and restrictive; - 5 to 5 is a gradation for anocratic regimes and +6 to +10 is a gradation to a maximum free democracy. The classification relies on Marshall & Elzinga-Marschall (2017) and is referred to below as polity indicator. These data are not available in grid resolution. Here, we refer for both variables to the reference unit of the respective nation-state.

3.1.4. Border lines and related land areas

For the spatial definition of all borderlines worldwide, we base our analysis on the *Large-Scale International Boundaries 10.0* (Office of the Geographer and Global Issues at the U.S. Department of State, 2021) data set. This data set features all international borders and special lines on land with an estimated positional accuracy of 100 m. They are the respective starting point for defined border regions (cf. section 3.2.).

We consider the total number of land borders worldwide, which amounts to 315, or 630 in total, respectively (based on the Direct Contiguity Dataset (Version 3.2) of the Correlates of War Project). Each borderline occurs twice in the dataset, as border infrastructures need not

continue across the border. The crossing from one nation-state to another might have a fortified form with complete checks. The return way might only be a checkpoint border with punctual controls (e.g. Mexican-American border). Accordingly, our border lines dataset has a dyadic format.

For the definition of the land areas, we use the OpenStreetMap dataset to spatially define the extent of the nation-states (OSM, 2021).

3.2. Spatial reference units

Our input data on border typologies, settlement and population dynamics, and economic and political situations have different spatial references: Border typologies refer to a border line, settlement and population dynamics have a grid resolution and can therefore be related to freely configurable reference units such as border regions or entire nation-states, and the economic and political data refer to the respective nation-states. While the reference units of border lines and entire nation-states are self-explanatory, the reference unit related to the spatial vicinity of borders needs to be defined.

In general, we have the border line in its territorial course. This line, however, has no spatial extension. As it is nowadays the case that the border as a linear control post has been replaced by an ensemble of control sites, technologies, and infrastructures that can enable, channel, or prevent mobility or exchange, the border of the 21st century moves away from the border line and spatially reaches far into the adjacent space or even beyond to distant areas where mobility is controlled (e.g. Shachar, 2020; Mau, 2021). This makes control monitoring spatially porous and spatially elusive. It is very difficult to grasp this spatially, especially since mobility control often takes place in distant regions. Against this background, we limit our spatial reference unit to the vicinity of border regions and thus we follow the basic idea that a border has an impact on the geographical area directly surrounding it. We do so, because for our global approach, precise spatial coverage of control points away from the border line is not data feasible.

By these considerations, we analyze the developments in the *spatial vicinity of the respective border lines*. Of course, it is a priori not clear which distance is meaningful to possibly measure border effects. Rather, it can be assumed that there can be no uniform distance value for this. In order to create global uniformity and thus consistency in the methodological approach, we have defined a border region of 25 km, which is considered to spatially approximate the catchment area of the borders. As there are two border typologies for each border line (see chapter 3.1.4.), the border region of a nation state is assigned the typology of the border that defines immigration. However, the analysis exclusively on the border region of 25 km does not go far enough, since dynamics are very different across the globe. The extent to which the border inhibits or strengthens developments must therefore also be assessed relative to the development dynamics in the specific national context. Against this background, we rely on an additional spatial unit where we relate the development in the *spatial vicinity of the border lines* to the development of the *respective entire nation-states*.

3.3. Concept for analyses

Our concept of analyses is two-fold: first, we apply a *descriptive analysis* on border typologies as well as on settlement and population dynamics across the globe. And second, we test the *relations* of our variables to each other along our guiding hypotheses, as stated in chapter 2, to reveal global trends.

3.3.1. Descriptive analysis: the distribution of border typologies and ranking of border dynamics

We map and quantify, on the one hand, the border typologies based on the introduced classification scheme across the globe. We relate the typologies to the quantitative occurrence at global and continental level and present this in cartographic representations as well as in pie charts. On the other hand, we quantify which nations across the world have the highest dynamics in border regions. To do this, we develop a ranking of the highest dynamics by the variables 'settlement' and 'population'.

However, a single metric seems to be unsuitable for this purpose. For example, relative settlement or population growth in the vicinity of the border with virtually no existing settlement area in the year 2000 could be classified with very high dynamics, although the actual absolute growth is very little. And, at the same time, absolute growth naturally includes a bias due to the unequal lengths of borders.

To address these challenges, we develop an index to compute the dynamics based on an ensemble of relative metrics. We calculate the relative growth of settlements and population for our period of analysis for two different spatial units:

Firstly, the relative growth is calculated at the *spatial vicinity of the respective border lines*, i.e. at the border regions of 25 km distance to the border line (cf. equations (1) and (2)).

(1) $$Settlement\ growth_{border\ region} = \frac{settlement\ area\ border\ region_{2014}}{settlement\ area\ border\ region_{2000}}$$

(2) $$Population\ growth_{border\ region} = \frac{population\ border\ region_{2015}}{population\ border\ region\ _{2000}}$$

In general, the results can be assigned to three classes: A growth rate lower than 1 means decrease of settlement areas or population within the border region over the monitoring period, a growth rate equal or close to 1 stands for no development, and a growth rate higher than 1 signifies increase of a development proxy.

Secondly, we apply two further metrics which put the growth rates in the vicinity of the border (border region of 25 km distance to the border line) in relation to the growth rates of the respective *entire nation-state* (cf. equations (3) and (4)). Growth rates are highly variable around the world and this is also true in areas close to borders. By calculating these in relation to the entire particular nation-states, we try to assess the dynamics in the national or supra-regional context. With this, we quantify whether we see different development tendencies near the border than for the entire respective nation-state:

(3) $$Settlement\ growth\ ratio_{border\ region\ vs.\ nation-state} = \frac{Settlements\ growth_{border\ region}}{\frac{settlement\ area\ nation\ state_{2014}}{settlement\ area\ nation\ state_{2000}}}$$

$$(4) \qquad Population\ growth\ ratio_{border\ region\ vs.\ nation-state} = \frac{Population\ growth_{border\ region}}{\frac{population\ nation\ state_{2015}}{population\ nation\ state_{2000}}}$$

Again, the results can be assigned to three classes: A growth rate lower than 1 means lower settlement or population growth rates within the border region than in the entire nation-state, a growth rate of 1 stands for equal development dynamics, and a growth rate higher than 1 signifies higher development rates.

For the four formulas, we interpret a growth rate in the range from $\frac{1}{1.05}$ to 1.05 as no development or equal development dynamics in the border region as in the entire nation-state, respectively.

For the calculations of all these metrics, we apply two exceptions: Firstly, we remove border regions from the analysis that have a growth rate of 0 or infinity (4 border strips for the proxy settlement and 3 border strips for the proxy population are removed). And secondly, we specify border regions where less than 0.1% of the area in 2000 (2015, respectively) is settlement. If the settlement area shares are below 0.1% of the total area and growth is so low that the uncertainty of the input data exceeds the measured growth, we set development to 1. We do so as existing errors in the GHSL would distort growth rates with such low area shares (95 border strips for the proxy settlement and 45 border strips for the proxy population are set to a growth rate of 1).

From these four metrics, we generate two rankings, each based on a different composition to estimate the highest dynamics near the border. The first ranking is composed of the mean of the two metrics related to the vicinity of the border (eqs. (1) and (2)), the second one is composed of the mean of the two metrics for the ratio of the developments in the vicinity of the border to the entire nation-state (eqs. (3) and (4)).

3.3.2. Testing relations between variables

The dynamics in the vicinity of borders – as diverse as they are on our planet – are related to a multitude of factors which are often even very locally specific such as border typologies, demographic, economic, social, political, environmental or topographic situations, among many others. In this second methodological part, we relate the presented global data (cf. sec. 3.1.) of border typologies, political and economic status quo of the nation-states, and settlement and population dynamics to each other following the guiding hypothesis formulated in section 2. It is therefore clear that our data sets do not cover all the various influencing factors in a comprehensive manner and that locally specific

circumstances cannot be depicted. Rather, we search for these specific relationships at a more aggregated, general level.

For the analyses, we present the variables in boxplots via the respective reference variable. We evaluate the extent to which the medians in our boxplots show truly significant differences by means of p-values to the rank-sums based Kruskal-Wallis H tests. We apply this test as we do not have normal distributions in the samples. Where applicable, we evaluate pair-wise significance in variance differences through post-hoc Tukey-Kramer-Nemenyi all-pairs test with Tukey-Dist approximation.

When linking the data sets, it must be considered that the different variables are available at different spatial reference units (section 3.2.) and that the variables are accordingly spatially linked to each other: For the relation of border typologies and economic and political differences among nation-states, naturally, the spatial reference units are the entire nation-states. For the relation of economic and political differences and settlement and population dynamics, two metrics referring to different spatial reference units are applied: Relative to the border region and relative to the ratio of border region to the entire nation. In this specific case, we calculate the relationships as follows: The relative settlement (or population) growth in the border region (or the relative growth in the border region vs. the entire nation) to the relative growth in the border region of the neighboring nation-state (or the relative growth in the border region vs. the entire nation for the neighboring nation-state). For the relation of economic and political differences and settlement and population dynamics, we also apply these two metrics referring to the two spatial reference units. For the analyses, we present the variables in boxplots via the respective reference variable.

4. Results

4.1. Descriptive results: distribution of the border typologies and the dynamics in border regions across the world

Border typologies vary across the globe – from frontier to fortified. From the cartographic representation (Fig. 3) and the quantitative shares of border typologies at global and continental levels, we explicitly point out the following aspects: On *global scale*, checkpoint borders (59.4%) are the dominating border typology, while the other four border types more or less share about 10% among all borders. Every fifth border on our planet, thus, belongs to the most fortified border categories (barrier, fortified) (20.6%). On *continental level* there are spatial differentiations: Landmark borders only exist in Europe and were created by the entry into force of the Schengen Agreement in Europe (Felbermayr et al., 2017). In contrast, frontier borders do not exist in Europe and North America at all. The most fortified border categories (barrier, fortified) occur predominantly in Asia (71 out of 130); however, they exist on any continent.

Near these borders, just over 610 million people lived in 2015. This means that 8.2% of the world's population lived within 25 km or less of a border.

In general, we measure border regions not as dynamic in terms of settlements compared to the global development. Settlement areas in border regions increased by 16.1% between 2000 and 2015, while globally they increased by 20.3%. It is interesting to note, however, that for population it is vice versa. The population has grown by 24.5% near the border, while globally it has increased only by 20.0%. Thus, at global scale, we measure a balanced growth in settlement and population. In border regions, however, we observe that population growth is higher than settlement area development. A trend that points to population accumulation effects.

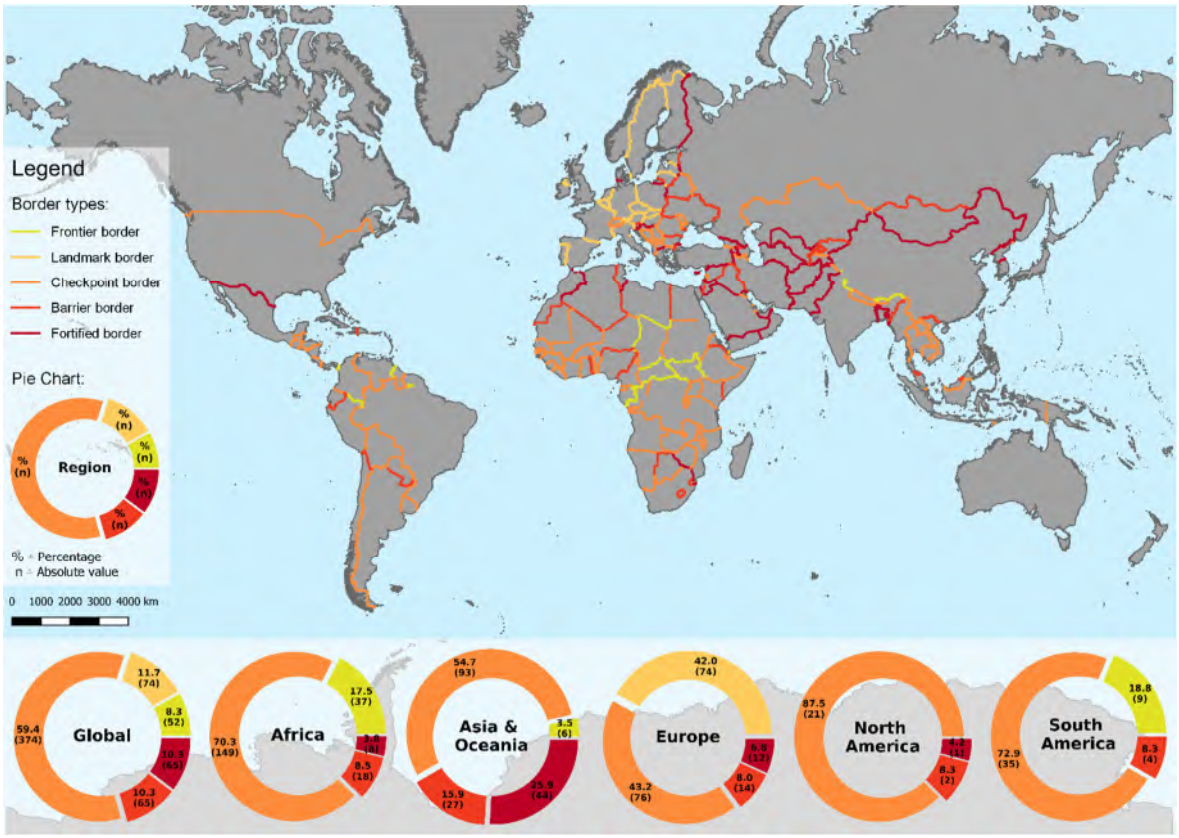


Fig. 3. Border typologies and their distribution across the globe in the year 2017 (the map shows the color of border type of higher fortification per border line).

Settlement and population densities and trends over time in the vicinity of borders, however, are highly variable in relation to border typologies or to regions or continents (Fig. 4). Although, many people live close to borders, globally, border regions are of comparatively low-density, at just over 0.5% settlement density. Here, the landmark borders have by far the highest settlement density at 3.24% (in 2015), while checkpoint, barrier and fortified borders have much lower settlement shares at about 0.5%. The same is true for population density. Landmark borders have 104 people per km2 in 2015, almost twice as many as the global average. The other three types of borders have 58, 40 and 84 inhabitants per km^2, respectively. It is interesting to see that frontier borders are in fact no man's land. In the border regions, we see a settlement density of only 0.01% for this type and only 7 people per km^2 in 2015.

With respect to continents, it is interesting to see how Europe and North America feature higher settlement densities than the global average in the border vicinities. Due to a large amount of frontier borders in Africa, Asia and South America, the average densities are below the global averages. For population, we see similar trends with Europe above the global average. Over the 15 years of monitoring, Asia features high growth rates reaching population densities above the global average.

As shown, border typologies and development dynamics for the 15- year study period are not equally distributed around the world. We now illustrate these differences per border line of two neighboring states. For this purpose, we use rankings according to the growth

dynamics of *settlements* and *population* in the *vicinity of the borders* and as *ratio in the vicinity of borders to the growth rates of the respective entire nation-states*.

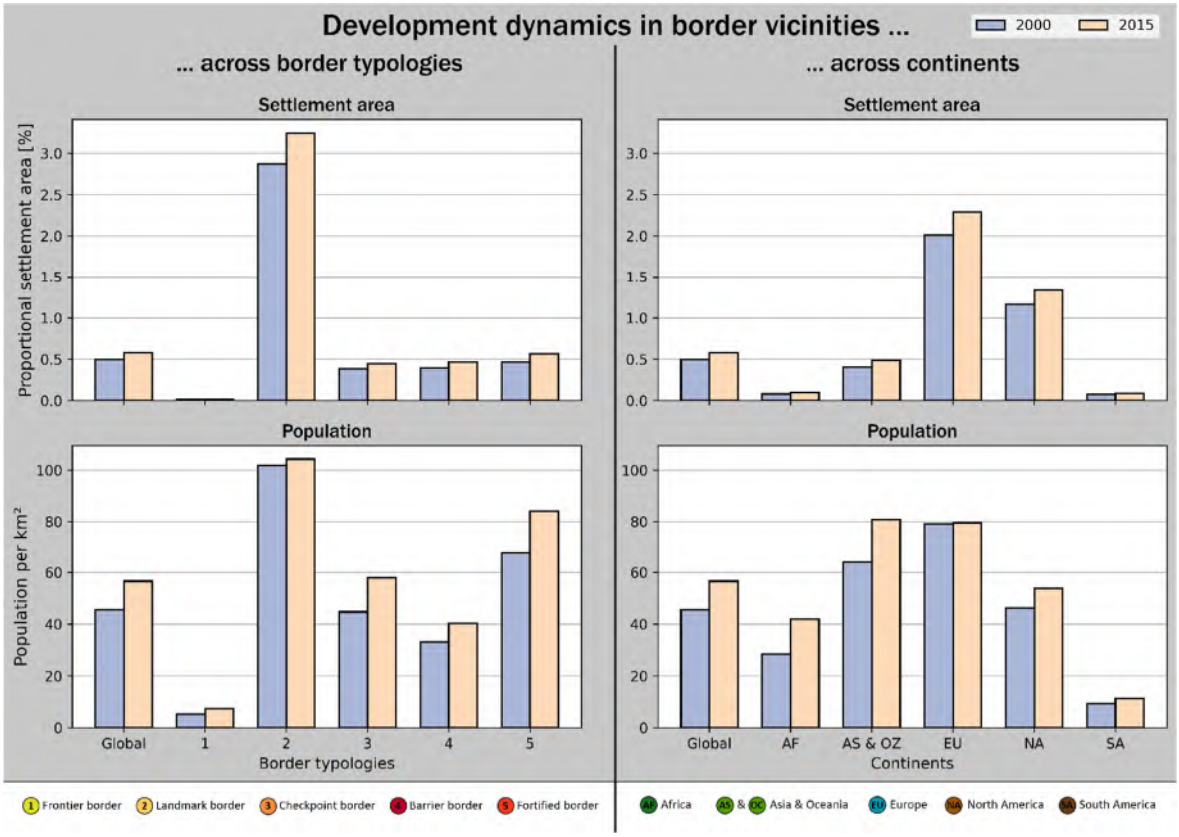


Fig. 4. Development dynamics for settlements and population from 2000 to 2015 in the border vicinities with respect to border typologies and continents and in global comparison.

In the *vicinity of the borders* (Fig. 5a), we measure highest *settlement growth rates* for the borders of Bhutan-China (settlement growth rate in the border region of first named nation-state: 15.35-fold), Bhutan-India (5.02) and Norway-Finland (4.35). 59 borders (i.e. 9.5%) of the 624 border regions across the world are measured with higher settlement growth rates than 1.50-fold. Of these borders of highest dynamics, 89.8% are located either in Africa (29) or in Asia (24). The remotely sensed input data on settlement development results in virtually no single border with a setback. This may reveal a correct trend that built-up landscapes are inert and mostly even remain so, even with population loss. But it also shows that the algorithms and data probably cannot detect these low levels of deconstruction activity, if they exist. The border of Surinam to Brazil is an example of no development at all.

Further, it is interesting to see that high settlement growth rates do not necessarily mean high population growth rates, as the top 3 in this ranking are other border regions: For *population*, we measure highest growths rates for the borders of Angola – Republic of Congo (population growth rate in the border region of first named nation-state: 6.55-fold), Venezuela-Brazil (5.26) and Chad–Niger (3.91). Expanding the list, the Oman-Yemen (3.71), Chad-Libya (3.53), Libya-Chad (3.53), Libya-Niger (3.53) or Kenya-Somalia (3.46) borders follow – all regions with generally high population growth and particularly high numbers of refugees. For population, 114 borders (i.e. 18.27%) are measured with higher growth rates than 1.50-

fold. Dominating shares among them are with 57% in Africa, 28.1% in Asia, and 10.5% in South America. In contrast to the settlement development, where in principle there is no decline measured, we are measuring a decline in terms of population for 129 border regions, i.e. for 20.7% of all of them. Among them, 70.5% are located in Europe. The strongest population declines were recorded in the Iran-Turkey border region with 0.45 followed by Surinam – Brazil (0.65) and Albania – Kosovo (0.69).

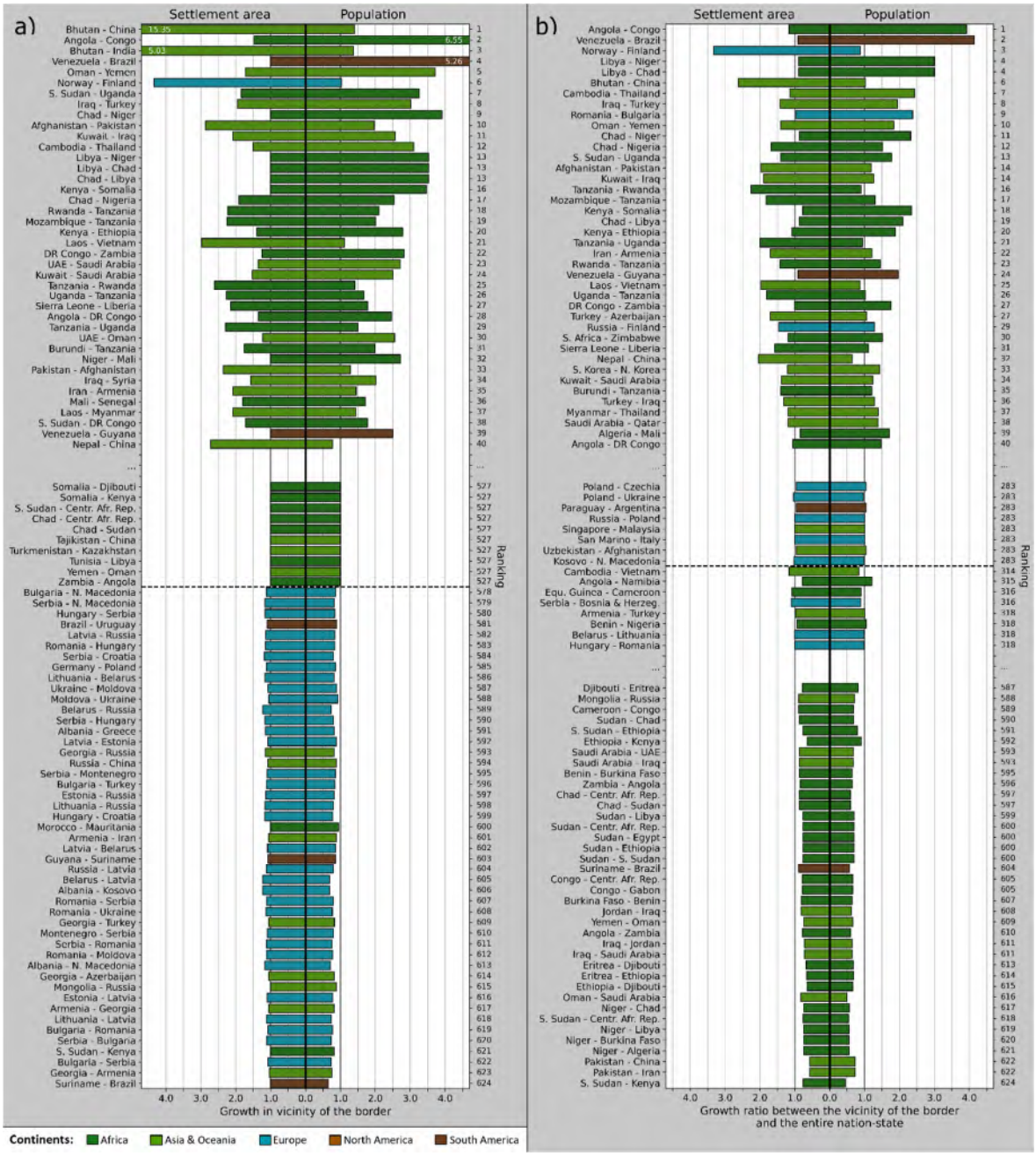


Fig. 5. Ranking based on development dynamics: a) relative growth rates of settlements and population in the border regions; b) relative growth rates of settlements and population in the border regions in relation to the relative growth rates of the entire nation-state. Above the dashed line we measure a positive joint growth of population and settlements, below a negative rate.

In the joint ranking of both variables in the *vicinity of the borders*, we measure for 578, i.e. for 92.63% of all border regions increases. The Bhutan-China border region (settlement

growth rate in the border region of first named nation-state: 15.35-fold; population growth rate: 1.41-fold) is at rank 1, and the lowest rank has the Suriname-Brazil border region (1.00; 0.65). It is again striking that 28 of the 30 border regions with the highest growth rates in this respect (top 5%) are solely located in Africa (17) and Asia (11), while 20 of the 30 border regions with the lowest growth rates in this respect (lowest 5%) are located in Europe. The measured growth rates in the vicinity of the borders basically reflect uneven growth rates on our planet.

As a second statistic on border development, we look at *growth rates in the vicinity of borders and consider them in relation to the growth rates for the respective entire nation-states* (Fig. 5b). Here, the *growth rates of settlements* on the border Norway to Finland are the highest (3.33-fold), followed by Bhutan - China (2.62), and Tanzania - Rwanda (2.26). The lowest settlement growth rates relative to the nation-state are measured at 0.56-fold for the border from Pakistan to Iran, Pakistan to China (0.57), and Ethiopia to South Sudan (0.65). In total, there are 301 border regions (i.e. in 48.2% of all border regions) where growth is higher near the border than in the nation-state (30.9% of these are in Africa, 28.2% in Asia and 26.6% in Europe). This result at the global level is a first indication that our measures of settlement development do not reflect a general difference between growth rates in border regions and nation-states. Nor does their distribution across continents suggest any obvious trend for border regions of different political, economic, or cultural circumstances to be an explanatory factor here.

In terms of *population* trends, we see that 359 border regions (i.e. 57.5% of all borders) are experiencing lower population growth rates in border vicinity than in the entire nation-state between 2000 and 2015. Even if a somewhat stronger trend toward lower population development near the border can be identified here, it is still not very pronounced. The Iran-Turkey border is the frontrunner in this respect: population development is only 38% as strong in border vicinity as in the entire nation-state. It is followed by South Sudan - Kenya (0.45) and Oman – Saudia Arabia (0.5). For 11.7% of the sample, growth near the border is at least one-fifth lower than in the nation-state as a whole. 54.8% among them are in Africa.

For the combination of both variables (settlement and population) for the *ratio in the vicinity of borders to the growth rates of the respective entire nation-states*, we measure for 314, i.e. for 50.3% of all border regions increases. Highest growth rates are measured in the Angola - Re public of the Congo border region (settlement growth rate in the border region of the first named nation-state: 1.16-fold; population growth rate: 3.93-fold), and the lowest in the South Sudan - Kenya border region (0.76; 0.45). In this ranking, it can be emphasized that 25 of the 30 border regions with the highest growth rates in this respect (top 5%) are again solely located in Africa (16) and Asia (9). However, 29 of the 30 lowest growth rates in this respect (lowest 5%) are also exclusively located in Africa (22) and Asia (7). This indicates that predominantly African (and by extension also Asian) nation-states developed compared to each other quite differently between 2000 and 2015.

Overall, this ranking shows that the variability of settlement and population development at border vicinities is highly variable across the globe and even highly variable on regional levels. While we identify certain region-specific tendencies (e.g. lowest growth rates and partly even population decline in border vicinities in Europe, or highest growth rates

predominantly in Asia and Africa), regional specifics (economy, demography, topography, conflicts, etc.) are too diverse and do not reveal obvious, simple geographical relations. On this basis, no clear-cut pattern can be discerned, though one should not jump to conclusions and infer that border infrastructures do not matter for settlement and population issues. A straight forward relationship might be thwarted by a larger number of context variables.

4.2. Relations of border typologies, economic and polity indicators, settlement and population dynamics for neighboring nation-states

4.2.1. Relations of border typologies with GDP per capita and polity indicators

With regard to our first guiding hypothesis (a), i.e. that we assume the higher political or economic differences between neighboring nation-states are, the stronger the border fortification is, we see this thesis basically confirmed in the global empirical analysis (Fig. 6).

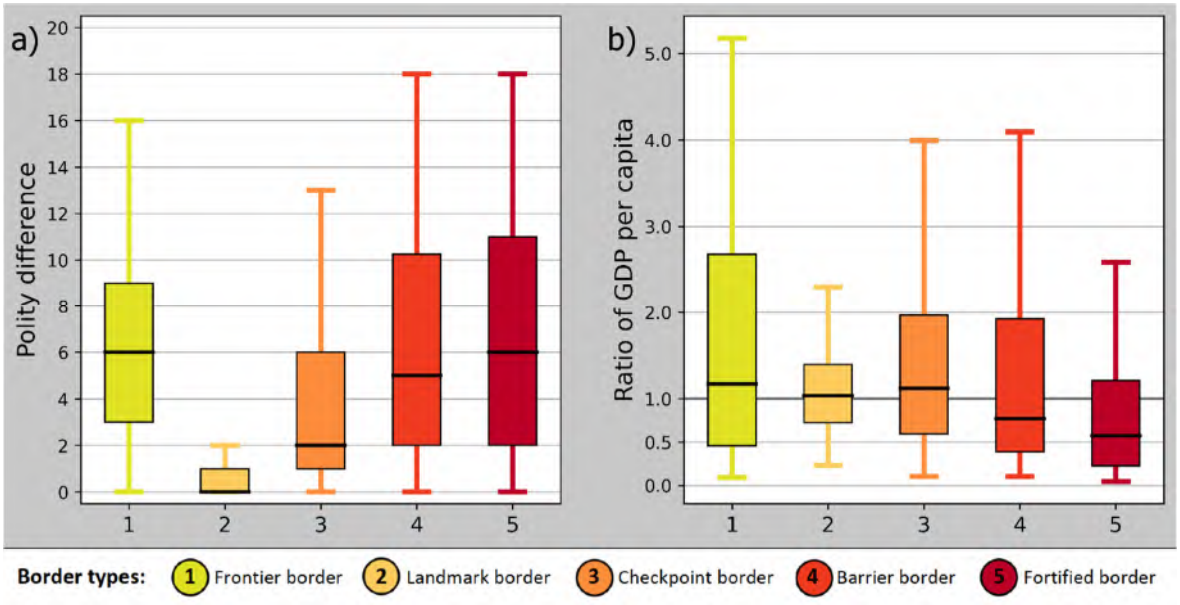


Fig. 6. a) Relation between polity difference and border typologies of neighboring nation-states, Krukals-Wallis test p-values < 0.001 ***; b) Relation of the ratio of GDP per capita between neighboring nation-states and border typologies, Krukals-Wallis test p-values < 0.001 ***.

For both variables, relations are found: For *political difference* among neighboring nation-states, we find the greater the differences in the political systems of the respective two bordering nation-states, the more fortified the border. While the measured variance per border typology is, except for the landmark borders, high, we still see a clear trend with rising medians. The medians increase from 0 for the landmark borders to 2 for the checkpoint borders ($p < 0.001$) to 5 ($p < 0.001$) and 6 ($p < 0.001$) for barrier and fortified borders and these are based on the Krukals-Wallis test statistically significant. The frontier borders, however, seem to represent a special case: Although the borders are not physically fortified or strongly controlled, their inaccessible terrain then takes over this control function of nation-states featuring on median (6) great political differences. (Appendix, Table 1). For *economic ratio* among neighboring nation-states, we find the smaller the GDP per capita of the nation-state facing the border in relation to the GDP at the neighboring nation-state is, the more

fortified a border is. These results indicate a trend that richer nation-states shut themselves off from poorer neigh boring nation-states. Again, the variances among our sample across the world per border typology are high and capture the influence of various factors not considered in the study. This means that the relationship is neither unambiguous nor linear. However, a recognizable trend of lower medians with greater negative disparity in GDP per capita is given; however, the p-values of the Kruskal-Wallis H test show only significance for pair-wise comparison to fortified borders (see Appendix, Table 2). The medians decrease from 1.17 for the frontier borders, to 1.04 for landmark and 1.12 for the checkpoint borders, respectively to 0.77 and 0.57 for barrier and fortified borders.

4.2.2. Relation of settlement and population dynamics with GDP per capita and polity indicators and with border typologies for neighboring nation-states

For the first part of our second guiding hypothesis (b), we assume that the greater the political or economic disparities of neighboring nation-states, the greater are differences in settlement and population growth rates in border regions. We see this thesis also basically confirmed in the global empirical analysis (Fig. 7).

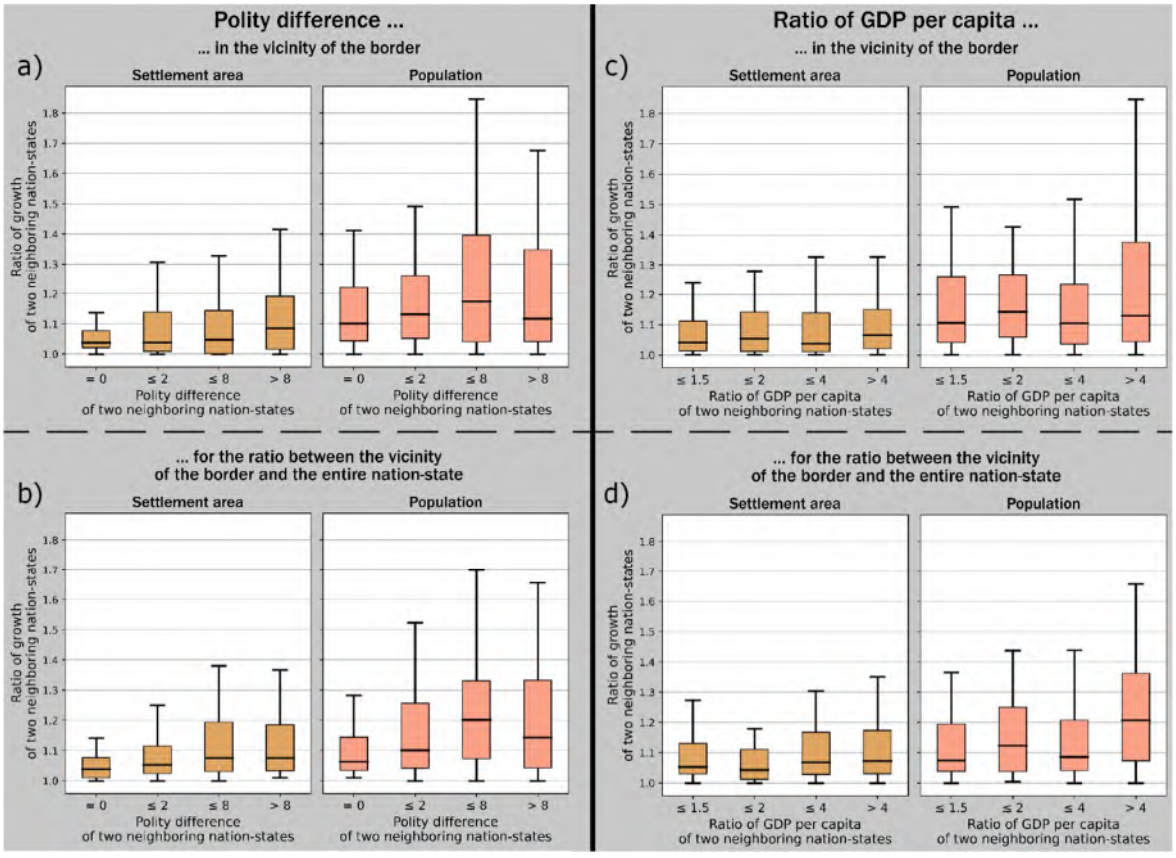


Fig. 7. Ratio of settlement and population dynamics of two neighboring nation-states: a) in the vicinity of the border in relation to the polity difference of two neighboring nation-states, Krukals-Wallis test p-values: left 0.356, right 0.415; b) for the ratio between the vicinity of the border and the entire nation-state in relation to the polity difference of two neighboring nation-states, Krukals-Wallis test p-values: left 0.002 **, right 0.001 **; c) in the vicinity of the border in relation to the ratio of GDP per capita of two neighboring nation-states, Krukals-Wallis test p-values: left 0.719, right 0.217; d) for the ratio between the vicinity of the border and the entire nation-state in relation to the ratio of GDP per capita of two neighboring nation-states, Krukals-Wallis test p-values: left 0.162, right 0.004 **.

With regard to *polity differences*, we observe the higher political differences between neighboring nation-states, the higher on medians the growth rate disparities of settlements and population. While this confirms the hypothesis, the statistical relationship is not significant (Appendix, Table 3 and Table 4) for the spatial unit of the vicinity of the border (Fig. 7a). In relation to the whole nation-state, this trend is also observed. Here, the Krukals-Wallis test even shows statistical significance (Fig. 7b; Appendix, Table 5 and 6). P-values for pair-wise comparison of settlement and population growth rates between 'states without political disparities' to 'neighboring states with political dis parities' in the range between 2 and 8 ($p = 0.009$, settlement growth ratio; $p = 0.001$, population growth ratio) and by more than 8 ($p = 0.004$) return statistical significance. With regard to the *ratio of GDP* per capita, we generally observe that with higher disparities among neigh boring nation-states, we have predominantly higher growth rate dis parities on medians among neighboring border regions. This confirms the hypothesis, but the relationships are not significant for the ratios of growth rates in the vicinity of the border (Fig. 7c; Appendix: Table 7 and 8). In relation to the whole nation-state, this trend is also observed and partly even statistically significant. For example, we see the highest medians when GDP disparities are larger than 4-fold. The Kruskal-Wallis H p-values for pair-wise comparison between states with low economic disparities (≤1,5) to neighboring states versus states with high economic disparities (>4) ($p = 0.003$; population growth ratio) reveal significance (Fig. 7d; Appendix: Table 9 and 10).

It can be concluded that large disparities between nation-states in economics or political regimes also implicate higher growth rate dis parities than in nation-states that are more similar in these domains. However, it should also be noted that the global variability is again high, trends are given but only in parts with statistical significance and thus, one-to-one conclusions cannot be drawn due to the wide variety of local, regional or national specifics.

For the second part of our second guiding hypothesis (b), we specifically assume that in the case of neighboring nation-states separated by barrier or fortified borders and with extreme differences in political and economic systems, a strong pull effect leads to accumulation effects on the poorer or more unfree side of the border. The assumption here is, that hard borders structure space and population primarily on the economically weaker and less democratic side.

Therefore, we specifically analyze all border regions across the globe where democratic (polity indicator >5) and undemocratic (polity indicator <5) states are separated by barrier or fortified borders (in total 20 border regions). The results show that the magnet effect of democratic states is very pronounced. The median growth rates in the border regions on the autocratic side are throughout all variables well above those on the democratic side (Fig. 8), although not statistically significant ($0.381 < p < 0.849$).

The same is measured when we analyze border regions where the ratio of the GDP per capita between the two neighboring nation-states is larger than 5.59. This value accounts for the sum of the mean of all ratios of the GDP per capita between two neighboring nation states that are greater than 1.0 and its standard deviation. This scenario applies for 23 border regions in total. Here, we see for all variables on medians a higher accumulation of settlements and population on the poorer side of the border between 2000 and 2015 (Fig. 8), although again not significant ($0.232 < p < 0.752$). Thus, the observed trend goes in the

direction of the hypothesis: when there are strong political or economic differences among neighboring states separated by barrier or fortified borders, an accumulation effect occurs on the poorer or less free side of the border. However, the results are statistically not significant and therefore we cannot fully confirm the hypothesis.

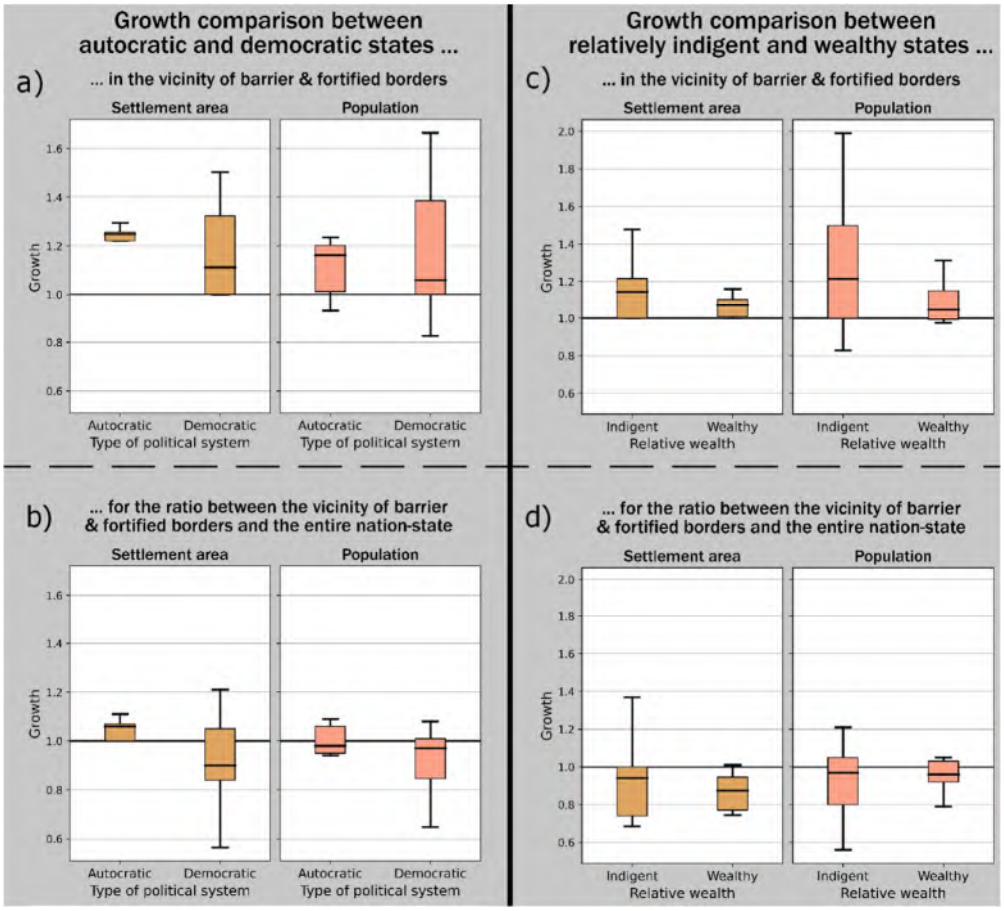


Fig. 8. Comparison of settlement and population development; a) in the vicinity of barrier and fortified borders between autocratic and democratic nation-states, Krukals-Wallis test p-values: left 0.616, right 0.849; b) for the ratio between the vicinity of barrier and fortified borders and the entire nation-state between autocratic and democratic nation-states, Krukals-Wallis test p-values: left 0.446, right 0.381; c) in the vicinity of barrier and fortified borders between indigent and wealthy nation-states, Krukals-Wallis test p-values: left 0.348, right 0.232; d) for the ratio between the vicinity of the border and the entire nation-state between indigent and wealthy nation-states, Krukals-Wallis test p-values: left 0.752, right 0.752.

5. Discussion

From a g*eographical* point of view, we were interested in the relation between border infrastructures and population and settlement development dynamics. Given that borders separate territories as well as populations and manage and modulate mobility, we departed from the assumption that they should matter. In this study, we documented the variability of border typologies across the globe and, we measured that in border regions settlement development is lower and population growth, however, is stronger than in a global comparison. However, we have generally not found a uniform pattern with regard to development paths of settlements and population in border regions, i.e. we have quantified very different development dynamics across the world.

Settlement areas have grown by 16.1% and the population by as much as 24.5% in the vicinity of the borders. Basically, we measured an increase in settlements in the border regions almost everywhere. When relating settlement growth rates in border regions to the entire nation-states, we found higher growth rates for about half of our sample and lower growth rates for the other half. Thus, initially and on this general level, no positive or negative influence on development can be assigned to border regions. For population, we registered a decline in 20% of all cases in the border regions, and we measured lower growth rates for 57.5% of all border regions than in the respective nation-state. An indication that borders hinder population development in their vicinity. However, these results reveal that simple conclusions about the causes of less and more development fall short.

When relating our variables to each other, we observed statistically significant trends that the greater political or economic disparities be tween neighboring nation-states are, the stronger the border fortification is. We also saw that the greater the economic or political disparities of neighboring nation-states are, the stronger population or settlement accumulation disparities in border regions are on medians, but mostly not with statistically significant relations. And, it becomes apparent that strong fortification in neighboring states of high discrepancies in polity or economy trigger settlement and population accumulation on the more unfree and poorer side of the border implying that the effect is much stronger on the side "against which" border walls and fences have been erected. This could be interpreted as a ‚*magnet effect*'.

However, we need to be aware that all of this is overlaid by so many factors that we can only show relationships and not causalities. The large variances in the boxplots capture the influence of various factors not considered in the study. Although certain trends can be interpreted from the figures, we must also state that settlement and population developments near the borders cannot simply be attributed to the border typologies or political and economic variables. The interpreted trends were only in parts statistically significant, not recognizable in correlation diagrams and there was little correlation measured. In other words, this trend in medians does not allow for an easy conclusion because there the high variances within the respective groups feature countless examples where these indicated trends are vice versa.

And yet there seems to be a statement here about the relations of the used variables. The fact that the results not ambiguously support assumptions shows that global studies remain in their aggregated format limited in their conclusions. This is because too many macro and micro indicators in the political (e.g. visa policies), demographic (e.g. migration), economic (e.g. trade), social (e.g. family ties), environmental (e.g. climate change, topography) or other spheres influence development lines, which can only be statistically represented here as variances. The incompleteness in our data set, the limited number of borders and the large number of covariates to be considered hinder the production of causalities. However, these mentioned data sets will be difficult to achieve at the global level with sufficient quality and spatial resolution. But, we believe, in spite of these shortcomings, it is necessary to first generate the empirical basis and descriptive statistics in order to be able to basically assess or even model the developments at the macro level.

Of course, the question arises whether the selected variables are really suitable as proxies for development. For a global study with this spatial resolution, we currently have, however, no alternative as the applied database is unique on a global scale.

Beyond the geographical aspects of this study, we are aware that results might be influenced by *uncertainties in the input data sets*. In general, our data sets contain slight temporal offsets. The borders relate to the reference year 2018, GDP and data on the political regime are from 2017, settlement growth rates relate to the time period of 2000–2015. Beyond this, our input data sets naturally contain errors: The accuracy of the GHSL, assessed by over 40 million building foot prints spread over different continents and settlement density characteristics, is reported with a balanced accuracy of 0.86 with slight regional differences (Corbane et al., 2019). However, large variations in accuracy due to location, surrounding land-cover, structural compositions, among others have been revealed (Klotz et al., 2016; Mück et al., 2017). For the population data, the quality assessment also has limits, due to the lack of worldwide independent but compatible reference data. However, a correlation analysis with 18 European nation-states using official GEOSTAT 2011 resident population data yielded a correlation coefficient of 0.83 (Freire et al., 2016). And, of course, for the economic and political data sets as well as for the border typologies errors or misclassifications may exist which cannot be quantified. What does this mean now? We believe that the database is solid especially for a global approach, but we cannot quantify and evaluate the impact of mis classifications. As example, we expect most misclassifications for settlements in peripheral areas with low densities, i.e. also in areas where some of our border strips are located. And still, as we measure to a certain extent that our assumptions are quite empirical in the data, we assume the general trends identified are plausible.

From a *methodological* point of view, this study can be considered as an entry point with expansion and development potential in many respects. The conceptual and methodological approaches applied in this study derive from the available data or are based on logical reasoning. However, these approaches can, of course, be adapted or extended to confirm or refute the identified trends: 1) The *variables*: other satellite data such as night light emissions have already shown the capability to also proxy economic development (e.g. Chen & Norhaus, 2019; Bachtrögler-Unger et al., 2021). Or data sets of higher spatial resolution such as from Landsat or Sentinel would allow to improve accuracies and spatial precision (e.g. Taubenböck et al., 2012). Thus, these data sets could extend thematic dimensions and spatial precision. Beyond, as we have seen, the resulting large variances in the boxplots capture the in fluence of various factors not considered in the study. This means, for example, extending the feature sets by data such as on the particular country's policy of purposefully supporting or impeding population migration would allow to model local effects needed to better fathom the complex nesting of this wide variety of factors. However, the demand for global availability of these data sets in consistent manner is a prerequisite, not always easy to meet. 2) the *spatial metrics*: the four metrics used to measure development can be expanded almost arbitrarily and other statistical approaches could be tested for comparison; 3) the *spatial concept* can be systematized, since it is a priori unclear how far the in fluence of borders on the surrounding area goes (e.g. Brakman et al., 2012), i.e. tests of different spatial units near

the border would allow to systematize the derived trends beyond the here used border regions of 25 km distance to the border line. Beyond this aspect, nowadays borders consist more and more of an ensemble of control sites, technologies, and infrastructures that sometimes replace the linear control post. The border line thus virtually reaches far into the adjacent space or even beyond to distant areas where mobility is controlled and thus calls for other spatial concepts and data types. In addition, the influence of adjacency, sub-adjacency and other spatial positional relationships among countries can be considered in future studies; 4) the *time period* of analysis can be extended, i.e. due to the availability of data, we referred in this study only to the type of border fortification in 2017. However, this ignores the fact that the degree of permeability or closure of borders change over time (Brakman et al., 2012; Sohn & Licheron, 2018). Thus, a multi-temporal classification of border typologies is in demand. Or, satellite data such as from night-time lights or Landsat would allow longer time periods of monitoring development processes. These data, conceptual, and methodological ideas are one sampling of ways among many more possibilities to put global studies on a broader empirical footing in the future.

Global studies can certainly capture trends, as this study has proven. But, as this study also shows, the revealed trends are only partly of the expected signs, partly their signal is weak due to discussed data-related, conceptual or methodological issues. This calls for further research: This study should therefore be understood as a first empirical approach to flank local studies on border regions with general global trends. And, together with the ideas just discussed to expand this study, we want to lay the scientific foundation to analyze impacts more systematically.

6. Conclusion and outlook

This study attempted to describe and quantify border typologies, to document and quantify development with the proxies ‘settlement’ and ‘population’ for a 15-year period from 2000 to 2015 in the vicinity of borders and, to relate all of these variables to economic and political indicators. The analysis was carried out for all 315 land borders across the globe. All of this was based on a completely new empirical data set where heterogenous data from literature surveys, censuses and remote sensing are synergistically analyzed. Such a global study does not replace in-depths and case-study research, but allows discerning general patterns and associations. Our umbrella question was: How do border infrastructures relate to settlement and population growth? Do they matter?

In general, we found highly different development dynamics in border regions across the globe. From it, we state that the border effects related to our variables do show trends: higher political or economic differences relate to stronger border fortification, greater economic or political disparities of neighboring nation-states relate to stronger population or settlement accumulation disparities in border regions, and strong fortification separating nation-states of high economic or polity differences trigger settlement and population accumulation on the poorer or more unfree side of the border. However, these general trends are measured on medians, they only are partially statistically significant, and thus these trends are not as strong or unambiguous as assumed. Variations across the globe are large and one-to-one conclusion cannot be drawn. And still, these trends are an empirical statement towards the

effects of economic and political situations, border typologies and settlement and population development in border vicinities.

Of course, this study is not an all-encompassing test of the assumptions as the influencing factors are manifold and local specifics occur. Rather, we see this study as a first approach to globally systematize these relations and developments. This study and many further studies are needed not least to rationalize simplistic arguments with empirical in formation in a highly emotional political and societal debate.

Acknowledgements

The establishment of the border infrastructure data was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 290045248 – SFB 1265. Furthermore, we thank Matthias Weigand for his support.

Appendix

Supplementary data to this article can be found online at
https://doi.org/10.1016/j.apgeog.2022.102866

Chapter XI

NO_2 air pollution trends and settlement growth in megacities

IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing, 2024, 17

Thilo Erbertseder, Hannes Taubenböck, Thomas Esch, Lorenza Gilardi, Heiko Paeth & Stefan Dech

© 2024 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

https://doi.org/10.1109/JSTARS.2024.3419573

Received 12 Mar 2024; Accepted 18 June 2024; Published 26 June 2024

Abstract

Megacities exert substantial demands on the world's energy resources, thereby producing large shares of atmospheric emissions and air pollution. Despite several studies on nitrogen dioxide (NO_2) changes in megacities, a systematic analysis in relation to settlement growth is still pending. This study examines the trends of tropospheric NO_2 pollution in 38 megacities regarding settlement growth by analyzing Earth Observation data. Multiannual records of tropospheric NO_2 from Global Ozone Monitoring Experiment (GOME), Scanning Imaging Absorption Spectrometer for Atmospheric Cartography, GOME-2A, and GOME-2B are evaluated regarding the World Settlement Footprint evolution from 1996 to 2015. Compared to previous studies, this work strictly uses remote sensing data and the spatial concept of functional urban areas. Therefore, uncertainties due to inconsistent spatial references, incomparable administrative units, as well as heterogeneously reported local data are widely excluded to enable an appropriate comparison of megacities across the globe. On average, we observe an increase in tropospheric NO_2 burden of 5.06 ± 0.83%/year and a relative settlement growth of 2.87%/year for the time period considered. We find a wide range of trends: relative NO_2 pollution ranges from an increase of 44.0 ± 5.9%/year to a decrease of -3.2 ± 0.2%/year, settlement growth rates vary from 0.3%/year to 9.5%/year. Despite this variety, the results exhibit a pronounced relation to the income group following the world's economies classification of the World Bank. Complemented by a cluster analysis and ranking of the megacities, the study provides indicators to develop a better understanding of the drivers and impacts of urbanization and air pollution.

Keywords

Air pollution; income class; megacities; nitrogen dioxide (NO_2); time series analysis; urban growth; urbanization

1. Introduction

According to the World Health Organization (WHO), air pollution is now considered the largest single environmental health risk (WHO, 2021). Humans are exposed daily to a variety of environmental stressors, such as nitrogen dioxide (NO_2), ozone (O_3), and particulate matter (PM), to name a few. These can lead to a wide spectrum of adverse health effects ranging from cardiovascular, respiratory, and metabolic diseases to susceptibility to infectious diseases, such as influenza or COVID-19 (Manisalidis et al., 2020; Gilardi et al., 2023; Rittweger et al., 2022; Weismann et al., 2023). A recent study showed that an average of 2.9 years of life expectancy is lost worldwide due to (outdoor) air pollution (Lelieveld et al., 2020). Around 99% of the global population is exposed to air pollutant levels that exceed the WHO guideline limits, with low- and middle-income countries suffering from the highest exposures (WHO, 2021).

Although air pollution is a challenge of global importance, exposure to air pollutants is predominantly an urban threat (Manisalidis et al., 2020). Urban areas across the globe account for less than 2% of the land surface only; however, they have always been spatial focal points of air pollution and resource needs as they consume two-thirds of the global energy with a majority of fossil fuels (World Bank, 2022).

The rise in the world's urban population over the past few decades has led to an increase in the number of megacities, generally defined as cities with a population of greater than 10 million (United Nations, 2018). As such, megacities have become global centers of population growth, settlement growth, production, energy demand, and as a consequence emission of greenhouse gases, air pollutants, and their precursor substances (Molina & Molina, 2004). As these cities continue to grow (Taubenböck et al., 2012), emissions are a concern, not only for local air quality but also for regional and continental scales and even for global climate, due to interactions with processes linked to climate change, such as radiative forcing (Huszar et al., 2021; He et al., 2024). Evaluating past air pollutant trends in urban areas gives insight into the effectiveness of previous and current policies to curb emissions and control these levels, and possibly supports measures for future policies (Baklanov, Molina & Gauss, 2016).

This study focuses on the air pollutant NO_2 as it can be attributed mainly to anthropogenic combustion processes associated with transport, industry, energy production, and heating (Molina & Molina, 2004). Due to its short lifetime in the lowermost troposphere during daytime (2–8 h), it predominantly remains at the emission sources and its amount is sensitive to emission changes (Duncan et al., 2013; Lange, Richter & Burrows, 2022). NO_2 is a regulated air pollutant and directives demand routine monitoring by ground-based stations in ambient air. However, the global availability of these measurements is heterogenous, the spatial representativity of point measurements for large urban landscapes is limited, and such data is often scarce in the Global South. Since 1995, space-based observation of the tropospheric composition has enabled operational daily monitoring of NO_2 levels on a global scale (Burrows et al., 1999). Contrary to in-situ measurements, space-based observations are gathered homogeneously with the same instrument and are sampled on a consistent and spatially integrating grid (Erbertseder, Taubenböck & Meyer-Arnek, 2015). However, these differ from in-situ measurements at ground level as they yield indirect measurements of

vertically integrated quantities, such as tropospheric NO_2 column densities (Richter & Burrows, 2002; Valks et al., 2011).

Due to the short lifetime of NO_2 and its predominantly human origin from combustion processes, as outlined above, tropospheric NO_2 constitutes an invaluable tracer to examine the anthropogenic footprint in the troposphere as well as human activity in different geographic perspectives and scales (Erbertseder, Taubenböck & Meyer-Arnek, 2015; Müller, Erbertseder & Taubenböck, 2022; Erbertseder, Baier & Bergemann, 2015). Since natural contributions to the tropospheric NO_2 burden are small and are confined to lightning (Voigt et al., 2022) and microbial processes in soils (Lu et al., 2021), it has been used to delineate urban pollution islands and pollution hotspots (Müller, Erbertseder & Taubenböck, 2022; Voigt et al., 2022; Samad et al. 2023) and to reveal weekly cycles reflecting anthropogenic activity patterns in urban environments (Erbertseder, Baier & Bergemann, 2015; Beirle et al., 2003; Stavrakou et al., 2020). Furthermore, tropospheric NO_2 column densities have been extensively used to infer NOx emissions, see, e.g., (Lange, Richter & Burrows, 2022; Silvern et al. 2019; Goldberg et al. 2021; Ding et al., 2022).

Several satellite-based studies analyzed NO_2 changes and trends in selected regions or megacities by examining multiannual records from a single sensor or a multisensor approach. In the first study, Richter et al. (2005) performed a linear regression of tropospheric NO_2 annual means from the Global Ozone Monitoring Experiment (GOME) (Burrows et al., 1999) for several regions in the world. A detailed analysis on NO_2 trends over China combining data from GOME and the SCanning Imaging Absorption spectroMeter for Atmospheric ChartographY (SCIAMACHY) (Bovensmann, Burrows & Buchwitz, 1999) was published by van der A et al. (2008). They used the time series analysis approach of Weatherhead et al. (1998) in a simplified way to detect trends in monthly mean data and to examine their significance. For a selection of megacities and urban agglomerations, NO_2 trends were derived from single-sensor data records using the same time series analysis method by Schneider and van der A (2012) and Schneider et al. (2015), while Konovalov et al. (2010) applied a nonlinear trend fitting and Geddes et al. (2016) used least-squares linear regression. Finally, the combination of data records from different sensors enabled the generation of longer times series and the detection of robust trends (Hilboll, Richter & Burrows, 2013) and even trend reversals (Georgoulias et al., 2019).

However, the derived trends in all these studies have so far neither been related to consistent spatial units like functional (Moreno-Monroy, Schiavina & Veneri, 2021) or morphological urban areas (Taubenböck et al., 2019), nor sufficiently to socioeconomic status or demographic parameters. This limits the interpretation of the findings and the reliability of a global comparison of cities. In previous studies, the derived trends are often related to a city or megacity, but no transparent spatial delineation is defined (Hilboll, Richter & Burrows, 2013; Georgoulias et al., 2019). Whether the observed mean or trend refers solely to the city center or to a grid cell covering the urban area and its surroundings becomes increasingly important as cities grow in size. It is inevitable to define a spatial unit for geographic comparisons, in particular, for such large agglomerations, such as megacities or megaregions (Erbertseder, Taubenböck & Meyer-Arnek, 2015; Taubenböck et al., 2014). In addition, interpreting trends alone might result in ambiguous conclusions. On the one hand,

a decreasing trend in NO_2 could result from a decline in population or traffic, reduced production or wealth. On the other hand, it could result from a transition to a green economy or energy and the implementation of new measures mitigating emissions like the introduction of new filtering techniques or electromobility, to name a few. Beyond that, population numbers or the gross domestic product (GDP) generally refer to administrative units leading to inconsistent spatial delineations of cities introducing bias in comparisons across cities (Taubenböck et al., 2019). In addition, they strongly vary in quality depending on the country and result in heterogenous datasets for comparisons of cities across the globe. As such, these numbers, concepts, and findings are prone to large uncertainties.

Therefore, a global objective and consistent measure from Earth Observation (EO) data are required. The World Settlement Footprint (WSF) (Marconcini et al., 2015) is derived from satellite data and exhibits an objective and multitemporal estimate of settlement growth that is considered a proxy for urbanization and economic development in this work. EO has proven to be an effective instrument, in particular, with respect to large-scale mapping of human settlements, such as megacities at continental and global scales (Taubenböck et al., 2012; Esch et al., 2017). The WSF evolution (Marconcini et al., 2021) that is available yearly from 1985 to 2015 provides an essential component to develop a better understanding of the drivers and impacts of urbanization and air pollution (Esch et al., 2022). A first attempt to link the settlement footprint to NO_2 trends was made by Erbertseder et al. (2023).

Against this background, this study centers on a systematic analysis of 38 megacities across the globe in relation to NO_2 pollution trends and settlement growth from 1996 to 2015. The study purely relies on satellite-based EO data. To overcome the problem of missing and heterogeneously defined spatial units, these data are rigorously analyzed with respect to functional urban areas (FUA) that provide a delineation of the cities based on principles of urban economics (Moreno-Monroy, Schiavina & Veneri, 2021). Furthermore, the results are ranked, classified, and interpreted regarding the income groups according to the world's economies classification of the World Bank (2023). The study aims at drawing overarching conclusions concerning the full spectrum of megacities, encompassing aspects, such as anthropogenic pollution and settlement growth, serving as a proxy for both urbanization and economic development. In addition, it will facilitate an analysis of their transformative stages, including transitions toward environmental sustainability (green transition), and their position within the global economic landscape.

The rest of this article is organized as follows. Section II describes the study area and the datasets used, followed by Section III outlining the methodology. Sections IV and V present the results and discussion including strengths, limitations, and perspectives of the work. Finally, Section VI concludes this article.

2. Study area and data

2.1. Functional Urban Areas

Instead of conventional, but heterogeneously defined administrative city units or arbitrarily chosen grid cells or geographic domains, we apply the concept of FUAs. This approach

enables us to consistently delineate megacity areas and megaregions against their rural surroundings and to define their spatial extent and urban pollution island. An FUA is defined by aspects of urban economics and consists of a city and its commuting zone, where the latter represents the area of influence of the city in terms of labor market flows. The Organization for Economic Co-operation and Development defines an FUA as the union of the administrative units in which at least 15% of the population is commuting to the main urban center of the area. To improve international comparability of urban-related statistics and indicators with an explicit spatial dimension, this method was developed further to estimate the number and extent of FUAs worldwide based on objective characteristics (Moreno-Monroy, Schiavina & Veneri, 2021). The improved method does neither rely on information on local unit boundaries nor data on commuting flows, which are usually unavailable in most countries. The FUAs are estimated through an automated classification procedure based on uniformly measurable parameters, such as travel time from urban centers, area of urban centers, local population, and country GDP per capita. These classification results are available as the Global Human Settlement Functional Urban Area (GHS-FUA) layer. To be consistent with the other datasets in this study, we apply the GHS_FUA_UCDB2015_GLOBE_R2019A dataset for the reference year 2015 (Schiavina et al., 2020). These FUAs, defined as polygon for each megacity, are used consistently throughout the study for any spatial aggregation.

2.2. Megacities

Megacities are defined by the United Nations (UN) as a conurbation or metropolitan area with a population exceeding 10 million inhabitants (United Nations, 2018). This study comprises 38 megacities across the globe referring to the baseline year 2015 (see Table 1). By means of the World Urbanization Prospects, the Population Division of the UN provides time series of the population of the largest 30 cities or cities with over 300k inhabitants and their population (United Nations, 2019). However, as can be inferred from Table 1, the megacities examined in this study primarily follow the ranking of the largest cities with respect to the population of the FUAs. A previous study has substantiated that a global ranking of the largest cities is a challenging task (Taubenböck et al., 2019). This is evident by the differing population numbers given by the UN and GHS-FUA depending on the administrative areas under consideration. The list of megacities in this article includes the 30 cities with the largest population in 2015 regarding the FUAs of the Global Human Settlement Layer (Schiavina et al., 2020) and the 30 largest cities in the world as of 2015 according to the UN (United Nations, 2018). As such, the list can be considered as a representative selection of megacities and a valid combination of these two datasets. However, only the population numbers of the FUAs are considered in any further analysis here.

Table 1. List of the 38 megacities examined in the study, sorted in descending order by the population of the Functional Urban Areas in the year 2015 (United Nations, 2019; Schiavina et al., 2020).

Rank	Megacity	Country	Latitude [°N]	Longitude [°E]	Population FUA 2015 [Mio.]	Population UN 2015 [Mio.]	Area FUA 2015 [km²]
1	Guangzhou	China	23.1255	113.2574	45.64	11.69	12 567
2	Jakarta	Indonesia	-6.2118	106.8416	39.81	10.17	9601
3	Tokyo	Japan	35.6895	139.6917	36.47	37.26	11 971
4	Delhi	India	28.6667	77.2167	30.08	25.87	5569
5	Shanghai	China	31.2222	121.4581	26.92	23.48	5878
6	Dhaka	Bangladesh	23.7104	90.4074	26.74	17.60	5760
7	Manila	Philippines	14.6042	120.9822	24.96	12.86	4873
8	Kolkata	India	22.5335	88.3560	24.64	14.42	5237
9	Seoul	South Korea	37.5683	126.9778	24.29	9.90	7053
10	Cairo	Egypt	30.0392	31.2394	23.49	18.82	4348
11	Mumbai	India	19.0740	72.8808	22.28	19.32	2367
12	São Paulo	Brazil	-23.5475	-46.6361	21.68	20.88	6077
13	Mexico City	Mexico	19.4273	-99.1419	21.41	21.34	4831
14	Beijing	China	39.9075	116.3972	21.28	18.42	5536
15	New York	USA	40.7170	-74.0037	19.52	18.65	17 489
16	Osaka	Japan	34.6758	135.5538	17.62	19.30	10 090
17	Moscow	Russia	55.7550	37.6218	16.38	12.05	8459
18	Bangkok	Thailand	13.7220	100.5252	16.29	9.40	5161
19	Los Angeles	USA	34.0317	-118.2417	15.66	12.35	10 407
20	Buenos Aires	Argentina	-34.6051	-58.4004	14.98	14.71	5376
21	Istanbul	Türkiye	41.0138	28.9497	14.79	14.13	2540
22	Karachi	Pakistan	24.9056	67.0822	13.42	14.29	1348
23	Tehran	Iran	35.6944	51.4215	13.41	8.55	3332
24	Ho Chi Minh City	Vietnam	10.7500	106.6667	12.83	7.35	3036
25	Jieyang	China	23.5526	116.3721	12.66	0.91	5922
26	London	U.K.	51.5085	-0.1257	12.60	8.66	6605
27	Lagos	Nigeria	6.4531	3.3958	12.34	12.24	2284
28	Bengaluru	India	12.9719	77.5937	11.93	10.14	2054
29	Lahore	Pakistan	31.5497	74.3436	11.79	10.37	3154
30	Chengdu	China	30.6667	104.0667	11.75	8.36	4034
31	Paris	France	48.8534	2.3488	11.24	10.73	6778
32	Rio de Janeiro	Brazil	-22.9028	-43.2075	10.81	12.94	3598
33	Lima	Peru	-12.0432	-77.0282	9.67	9.81	2121
34	Chicago	USA	41.8500	-87.6501	8.80	8.77	13 185
35	Wuhan	China	30.5833	114.2667	8.53	7.90	26 821
36	Tianjin	China	39.1088	117.1886	8.32	12.52	2413
37	Chongqing	China	29.5628	106.5528	6.04	13.37	2267
38	Kinshasa	Democratic Republic of the Congo	-4.3276	15.3136	5.76	11.60	591

2.3. NO_2 pollution

Global observations of the tropospheric NO_2 burden are derived from GOME (1996–2003) (Burrows et al. 1999), SCIAMACHY (2002– 2012) [32], and GOME-2A and GOME-2B (2007–2015; 2013–2015) (Munro et al. 2016). These instruments are nadir-scanning spectrometers that orbit in a polar Sun-synchronous pattern with a fixed equator crossing time. The spectrometers are passive and sense backscattered solar radiation from the Earth's

atmosphere and surface as well as extraterrestrial solar irradiance ranging from the ultraviolet to the visible and near-infrared portion of the spectrum at a high spectral resolution of less than 0.5 nm. The retrieval algorithm for NO_2 is based on the Differential Optical Absorption Spectroscopy (DOAS) technique exploiting the wavelength range from 405 to 465 nm. This method is utilized to derive the total slant column of NO_2 from the radiance spectra measured by the instrument. By applying the Beer–Lambert law, the DOAS fitting procedure entails a least-squares inversion aimed at isolating absorption signatures attributable to trace gases from the influence of broadband background processes, such as extinction arising from scattering on molecules and aerosols. In the next step, the total slant column density of NO_2 is converted into a vertical column density. After separating the stratospheric NO_2 contribution from the total vertical column, the resulting tropospheric residue can finally be converted into the tropospheric NO_2 vertical column density by applying the ratio of stratospheric and tropospheric air mass factors. The NO_2 content is finally given as the number of molecules per area (molecules /cm^2), mol per area (mol/m^2), or area-related mass (μg/m^2).

To combine the tropospheric NO_2 observations given the different sensor and orbit characteristics, a methodology was developed in Hilboll, Richter & Burrows (2013). The approach was applied Georgoulias et al., (2019) to gain a self-consistent dataset of monthly mean values on a regular grid of 0.25° × 0.25° spatial resolution. This study is based on the dataset of Georgoulias et al., (2019) as provided by www.temis.nl. To be consistent with the other data sources of the study, the time period considered covers 1996–2015.

2.4. Settlement footprint

To track the rapid urbanization observed in the recent past, yearly high-resolution settlement footprint data are a prerequisite for any further analysis. The settlement growth of the megacities under consideration is, therefore, derived from the WSF (Marconcini et al., 2015). It comprises a binary mask outlining the extent of human settlements worldwide. In particular, the WSF evolution dataset is used to outline the global human settlement extent at a 30 m spatial resolution on a yearly basis from 1985 to 2015 (Marconcini et al., 2021).

The classification was performed on the Google Earth Engine platform using the complete record of Landsat imagery. For each year, all available Landsat-5 and -7 scenes that were acquired over a given area of interest were gathered. After performing cloud and cloud-shadow masking, spectral indices were derived. Among others, these include the normalized difference built-up index, normalized difference vegetation index, and modified normalized difference water index. In the next step, key statistics were calculated, including temporal mean, standard deviation, minimum, and maximum. Temporal features prove generally robust if computed over at least seven clear cloud-/cloud-shadow-free observations. Accordingly, if this constraint was not satisfied for a given pixel in the target year, the time frame was enlarged backward until the condition was met. The WSF data are available via the DLR EOC Geoservice [Available: https://download.geoservice.dlr.de/WSF_EVO].

2.5. Income group

The World Bank Group assigns the world's economies to four income groups—low, lower-middle, upper-middle, and high (World Bank, 2023; 2022). The classifications are updated each year on 1 July, based on the gross national income (GNI) per capita of the previous calendar year. GNI measures are expressed in U.S. dollars using conversion factors derived according to the Atlas method, which in its current form was introduced in 1989. The World Bank's income classification aims to reflect a country's level of development, drawing on Atlas GNI per capita as a broadly available indicator of economic capacity.

3. Study area and data

3.1. Conceptualization

The study aims at analyzing, ranking, and classifying megacities with respect to NO_2 air pollution and settlement growth and to examine their statistical relationship. The conceptualization and methodology are illustrated in Fig. 1 and can be summarized as follows.

The first step to achieve these objectives is the exploitation of globally consistent satellite-based long-term records of tropospheric NO_2 (see Section 2.3.) and the settlement footprint (see Section 2.4.) from 1996 to 2015.

Second, the FUAs (see Section 2.1.) for the 38 megacities (see Section 2.2.) and the epoch of 2015 were derived from the GHSL_FUA and represented by polygons in shapefiles. Consistently, additional urban parameters, such as population or area size of the FUAs, were extracted.

Third, the time series of tropospheric NO_2 was extracted for each megacity's FUA. Therefore, all the monthly means were oversampled on a $0.01 \times 0.01°$ grid without any interpolation. For each megacity and month, an FUA-weighted statistic was calculated, including mean, minimum, maximum, standard deviation, number of pixels, and number of missing values as a consequence of cloud cover in the corresponding month.

Fourth, the settlement footprint, i.e., WSF evolution data, was extracted for each megacity within the corresponding FUA and time period on a yearly basis (see Section 3.3.).

Fifth, the analysis of the resulting time series and the separation of the linear trend component were performed by time series decomposition, as described in Section 3.2. The absolute and relative settlement growth rates were derived based on the statistical moments of the settlement area increments.

Sixth, the megacities were systematically analyzed and ranked according to the relative tropospheric NO_2 trends (%/year) and the relative settlement growth (%/year). The megacities were further examined regarding the per capita pollution and the income group following the world's economies classification of the World Bank (see Section 2.5.).

Seventh, by means of regression analysis, the hypothesis was tested whether there is a statistical relationship between relative NO_2 trends and relative settlement growth in megacities (both in %/year).

Finally, the spectrum of megacities was analyzed by k-means clustering to yield a classification of the megacities with respect to NO_2 air pollution and settlement growth (see Section 3.4.). Can we identify groups of megacities that evolve similarly? Which megacities exhibit strong settlement growth while reducing the overall NO_2 burden by implementing measures of green transition?

The method proposed in this study will enable us to derive general conclusions for the entire range of megacities, regarding urbanization and settlement growth, their transformative stages toward environmental sustainability, and their positioning in the world economies.

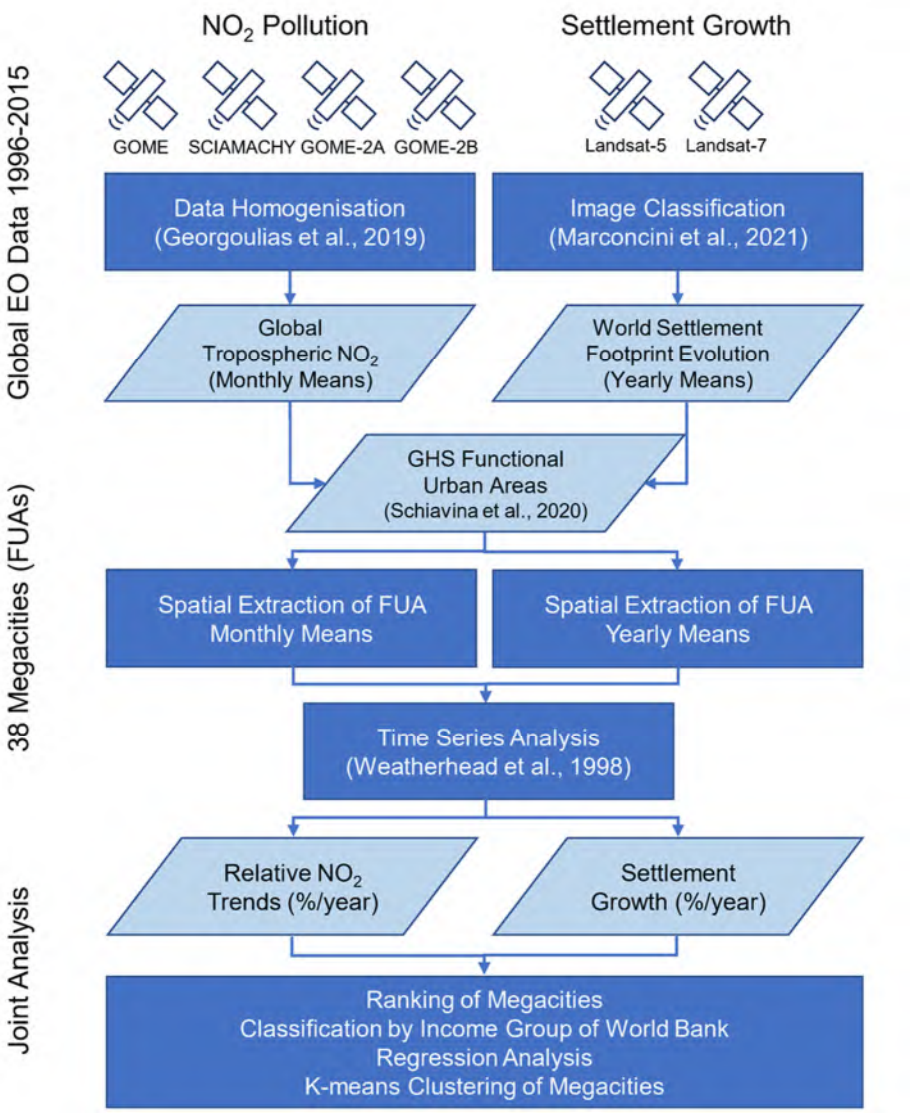


Fig. 1. Conceptualization of the methodology.

3.2. Time Series Analysis of NO_2 Data

Based on the satellite long-term records of tropospheric NO_2 (see Section 2.3.), a time series analysis with a multicomponent model (Weatherhead et al., 1998) was applied to identify linear trends of NO_2 and their uncertainties in megacities. This approach has been widely used to model geophysical quantities and, in particular, monthly means of observed

tropospheric NO_2 vertical column densities Y as a function of time t (van der A et al. 2008; Hilboll, Richter & Burrows, 2013; Georgoulias et al., 2019; Wendleder et al., 2020):

$$Y_t = \mu + S_t + \frac{1}{12}\omega + N_t \quad (1)$$

where μ is a constant term, S_t is the seasonal component, ω represents the linear trend per year, and N_t is the residual variability. The seasonal component is modeled as a sinusoidal and is assumed to be stationary, i.e., invariant in time

$$S_t = \sum_{J=1}^{4} \left[\beta_{1,j} \sin\left(\frac{2\pi jt}{12}\right) + \beta_{2,j} \cos\left(\frac{2\pi jt}{12}\right)\right] \quad (2)$$

where eight superimposed harmonics are fitted by the coefficients $\beta_{1,1}$–$\beta_{2,4}$. For the unexplained part of the time series, the noise N_t is considered to be autoregressive of the order of 1, i.e.,

$$N_t = \varphi N_{t-1} + \varepsilon_t \quad (3)$$

where ε_t are independent random variables and ϕ is the first order autocorrelation of the noise among successive monthly means, given as follows:

$$\varphi = Corr(N_t; N_{t-1}) \quad (4)$$

Like the seasonal component, the noise process is assumed to be stationary with $-1<\phi<1$. Following Weatherhead et al. (1998), the uncertainty of the linear trend component σ_ω can be derived as follows:

$$\sigma_\omega = \left[\frac{\sigma_r}{n^{3/2}}\sqrt{\frac{1+\varphi}{1-\varphi}}\right] \quad (5)$$

with σ_r being the standard deviation of the detrended residuals, n is the number of months considered, and ϕ is the first-order autocorrelation of the residuals. The significance of the linear trend ω can then be estimated at a 95% confidence level if

$$|\omega/\sigma_\omega| > t_\omega \quad (6)$$

where σ_ω is the uncertainty of the trend and t_ω is the value of Student t-distribution for a significance level of $\alpha = 0.05$ and the degrees of freedom defined by the number of months of the times series n (Tiao et al., 1990; Santer et al., 2000). The resulting relative trend and its uncertainty are given as percent per year (%/year).

3.3. Analysis of Settlement Footprint Data

Time series of the settlement area per megacity were calculated from WSF evolution data for the period 1996–2015 (see Section 2.4.) to match with the observation period of the NO_2 satellite data starting in 1996. In this study, we understand the terms "settlement footprint" or "urbanized areas" as the land directly occupied by a particular physical man-made structure. Thus, this definition represents a "settlement mask", defined by buildings, streets, and impervious surfaces (Taubenböck et al., 2012). The geographical extent of each megacity was delineated using the concept of the FUA. Consequently, the yearly growth in settlement area was calculated in km^2 in absolute terms and the relative growth rate in %/year.

3.4. Classification of Megacities According to Growth Dynamics and NO_2 Trends

K-means clustering is an unsupervised machine learning technique that allows the identification of clusters, i.e., similar groups of data points within the data cloud. It is the most common partitioning method for multivariate data (Loyd, 1982).

We apply this method to gain a classification of megacities according to growth dynamics and NO_2 trends. All inputs are purely satellite-based parameters, namely, the relative tropospheric NO_2 trend (%/year), the mean NO_2 burden (10^{15}molecules/cm^2), and the relative settlement growth (%/year). These data were normalized. In order to find the optimal number of clusters, the silhouette and elbow methods were used (Thorndike, 1953). Based on these methods, and a systematic test of various cluster numbers, five clusters were determined to be most appropriate for this study.

4. Results

As a first illustration of the results of this study, Fig. 2 shows examples of the WSF evolution depicting the extent and yearly growth of human settlements for the megacities Chengdu, China (left column, top), Bangkok, Thailand (middle), and Lima, Peru (bottom) within their FUAs. The panels on the right display the corresponding time series of monthly mean tropospheric NO_2 from 1996 to 2015 for the related FUAs. The red line illustrates the linear trend component as a result of the time series decomposition. The related relative trends (%/year) are annotated in the panels and are all significant at the 95% confidence level for the three cities. These three examples demonstrate a quite different overall development. Chengdu shows a very dynamic settlement growth and also a very strong increase in tropospheric NO_2 of 22.05 ± 2.50%/year. Bangkok visually also reveals a relatively dynamic settlement growth, but just a weak increase in NO_2 of 1.83 ± 0.49%/year can be observed. Contrarily, Lima exhibits almost no settlement growth within this period, but a strong increase in tropospheric NO_2 pollution of 10.71 ± 1.47% year.

This diversity highlighted here asks for a more systematic analysis and presentation of the results for the 38 megacities considered. Therefore, Fig. 3(a) illustrates the full spectrum of the relative NO_2 pollution trends (%/year) sorted by magnitude. In general, we find that the trends in the examined period from 1996 to 2015 vary—from strong increases to moderate declines.

When we compare the relative NO_2 trends, the values range from an increase in Chongqing of 44.05 ± 5.86%/year to a decrease in Los Angeles of −3.19 ± 0.24%/year. The mean NO_2 trend for all the 38 megacities is 5.06 ± 0.83%/year and the median is 3.16 ± 0.51%/year.

To put the trends into a socioeconomic context, the income groups are overlaid following the world's economies classification of the World Bank (2023). It is striking to see that the strongest increases in tropospheric NO_2 pollution occur predominantly in megacities of upper middle-income countries (UMICs) (Chongqing, Tianjin, Chengdu, and Wuhan) followed by cities in lower middle-income countries (LMICs) (Dhaka, Lahore, and Delhi). Megacities in high-income countries (HICs), however, solely appear at the very right of the bar chart. They all show declining trends, i.e., improving NO_2 air quality (Los Angeles,

Chicago, Paris, Tokyo, New York, London, and Seoul). The only exception is Quezon City (Manila, Philippines), which needs further clarification.

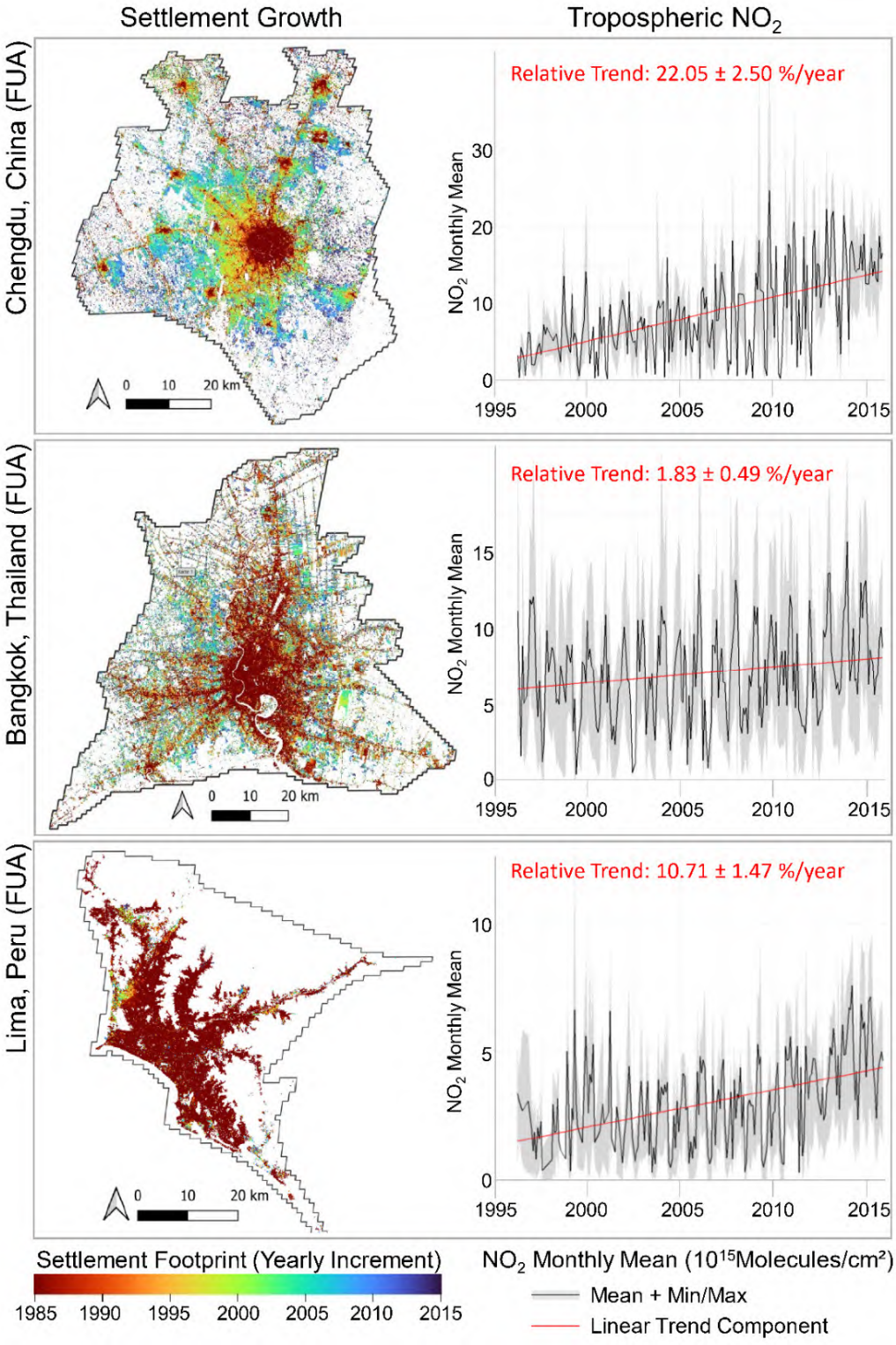


Fig. 2. Examples of the WSF evolution (left column) depicting the extent and yearly growth of human settlements from 1985 (dark red) to 2015 (dark blue) for the megacities Chengdu, China (top), Bangkok, Thailand (middle), and Lima, Peru (bottom). In the corresponding panels on the right, time series of tropospheric NO_2 from 1996 to 2015 for the related FUAs are shown. The red line illustrates the linear trend component. The relative trends are annotated in the panels.

When evaluating NO_2 pollution on a per capita basis (expressed as 10^{15} Molecules cm^{-2} 10^6 $Inhabitants^{-1}$), a population weighted metric for environmental quality, and subsequently arranging the values in descending order (see Fig. 3(b)), it becomes evident that the most elevated levels are prevalent in UMICs as well. Megacities in HICs follow in the hierarchy (Chicago, Los Angeles, and Seoul). Notably, all megacities from LMICs are found at the lowermost part of the spectrum with the lowest per capita NO_2 pollution (Jakarta, Kolkata, Quezon City, Dhaka, Lagos, Ho Chi Minh City, Mumbai, Bengaluru, Delhi, Karachi, and Cairo) with the exception of Lahore, Pakistan, which deviates from this pattern and exhibits higher per capita pollution levels.

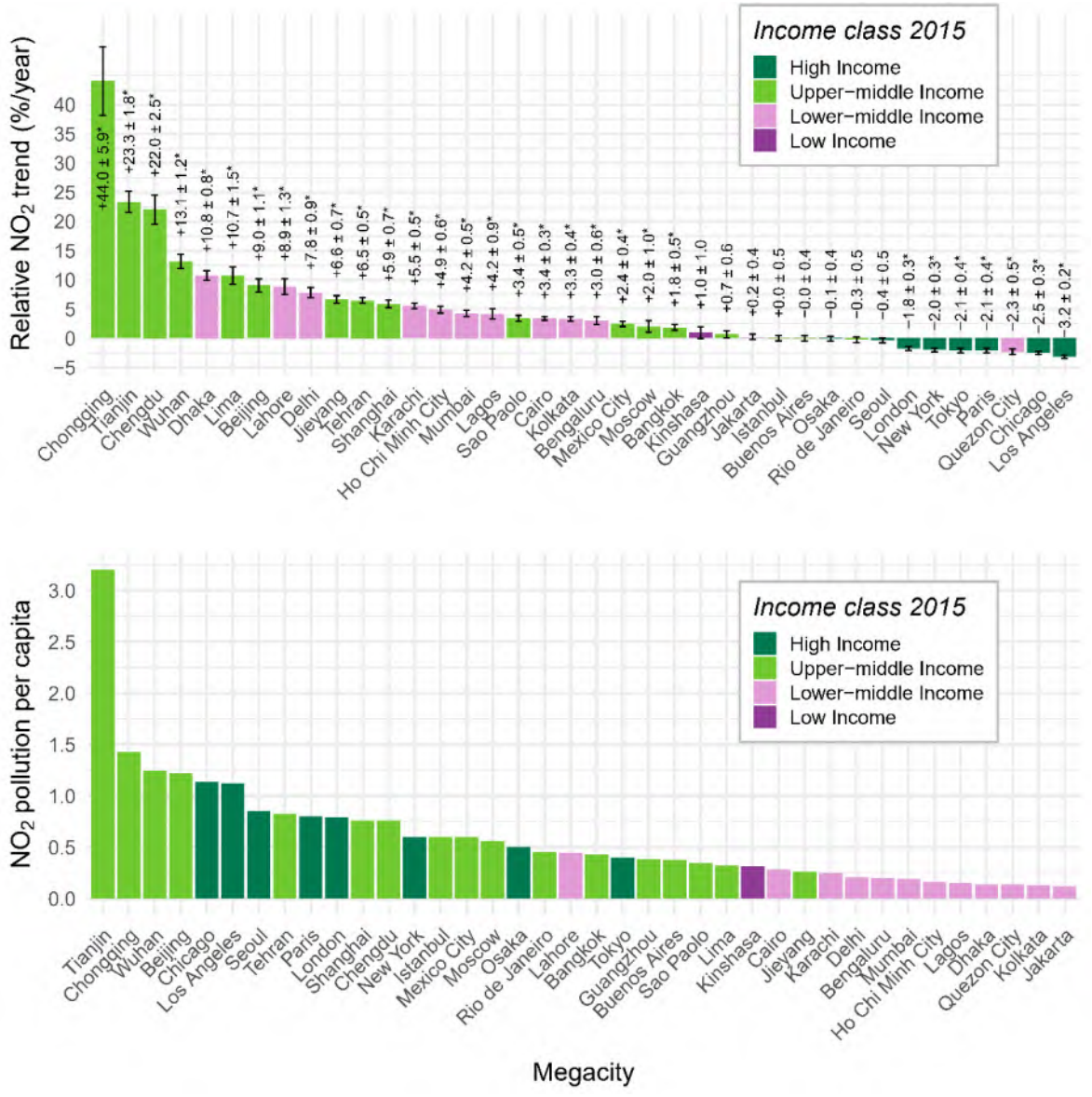


Fig. 3. (a) Relative tropospheric NO_2 trend (%/year) for each megacity FUA and the period 1996–2015, and (b) NO_2 pollution per capita for each megacity FUA and the same period expressed in (10^{15} molecules $cm^{-2}10^6$ $Inhabitants^{-1}$). Cities are sorted in descending order. Income groups are indicated according to the world's economies classification of the World Bank [52]. Asterisks in (a) indicate the significance of the trend at the 95% level.

As we reveal that the per capita NO_2 pollution levels in LMICs are the lowest, this should be interpreted in the context of their increasing pollution trends, suggesting potential

deterioration of air quality over time. Contrarily, although megacities from HICs show decreasing trends indicating improving air quality, the per capita pollution is still among the highest.

In general, we can identify two opposing developments: if the NO_2 trends of LMICs and HICs persist, with an increase and a decrease, respectively, the two groups will swap their positions in the hierarchy in the per-capita pollution (see bar chart Fig. 3(b)). Finally, megacities from UMICs, displaying the highest positive relative NO_2 trends, also exhibit high per capita pollution levels. This dichotomy highlights the intricate relationship between economic development and environmental quality, underscoring the complex nature of their interactions.

In comparison to the trends of NO_2 pollution, settlement growth rates are only positive throughout the 38 megacities under analysis. Nevertheless, we also observe a high variability—from close to stagnation to extreme expansion. Fig. 4 presents the yearly growth of settlement area (expressed in km^2) for a selection of the megacities and the time period of 1996 to 2015.

For this period of 20 years in the megaregion Guangzhou (Pearl River Delta), China, the highest growth in built-up area is measured. There, an expansion by 3511 km^2 (approximately the area of the city/state of Berlin, Germany) is mapped, followed by Shanghai (2423 km^2) and Beijing (1790 km^2). In contrast, Lima and Karachi show only a small increase in built-up area of 46 km^2 and 134 km^2, respectively. Over the years, most megacities exhibit a straight linear increase of built-up areas. However, we also observe a pronounced flattening of the slope for Los Angeles in 2008, Chicago in 2006, and Wuhan in 2011, indicating a slowdown of the quantitative urbanization, i.e., expansion of built-up area that could be attributed to economic developments like the economic crisis in 2008 in the case of Los Angeles. On the contrary, Bangkok indicates a sharp increase in 2013 as well as Tokyo in 2014.

Based on the times series of the absolute settlement growth, the relative growth of the built-up area (%/year) was calculated for the time period from 1996 to 2015. Accordingly, in Fig. 5, the relative growth of settlement area (%/year) is presented for each megacity sorted by size. Here, Chengdu (see also Fig. 2), Chongqing, and Shanghai indicate the highest growth rates of 9.54%/year, 9.03%/year, and 7.70%/year, respectively. The lowest relative growth rates can be identified for Kinshasa (0.27%/year), Lima (0.38%/year), and Los Angeles (0.56%/year). The average relative settlement growth rate for all megacities is 2.87%/year and the median 1.85%/year.

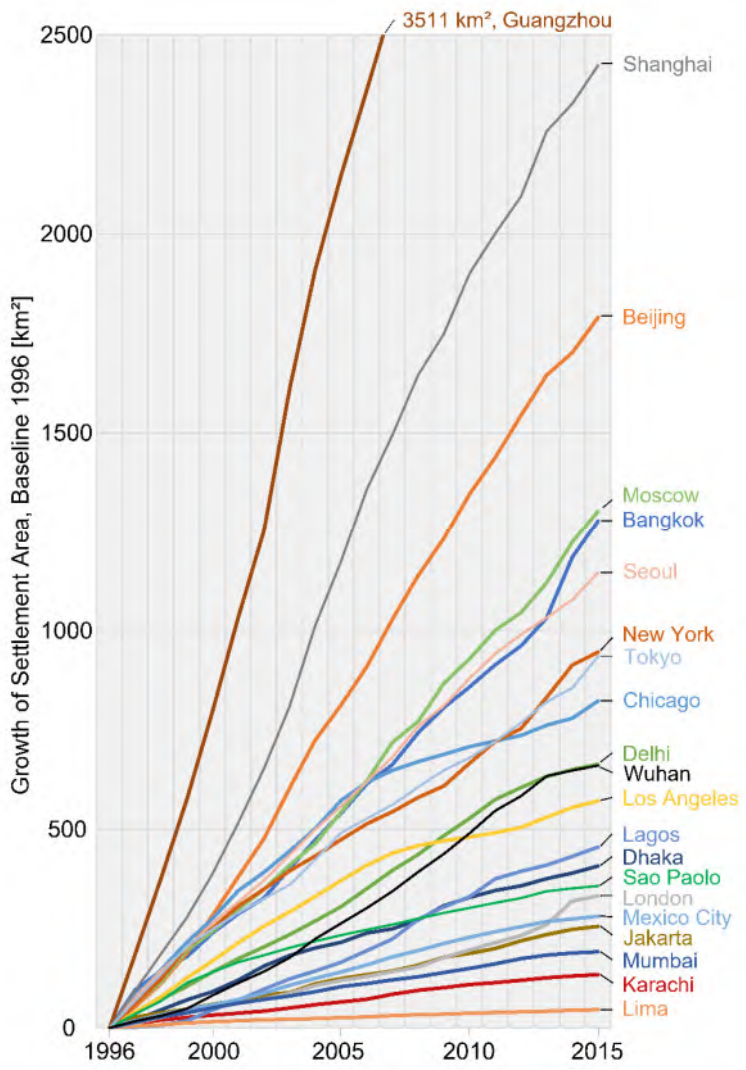


Fig. 4. Growth of settlement area in km2 for a selection of megacities from 1996 to 2015 within their FUA. Note that 1996 is used as the baseline year.

Consistently with the analysis of the NO_2 trends, we put these findings in relation to the income groups following the world's economies classification of the World Bank. It can be observed that among the top ten megacities with the highest relative increase in built-up area, eight cities belong to UMICs. Actually, seven out of these ten megacities are located in China (Chengdu, Chongqing, Shanghai, Wuhan, Tianjin, Guangzhou, and Beijing). In contrast, the megacities with the lowest growth rates are attributed to the HICs. Further, it should be noted that megacities of LMICs predominantly appear in the center of the spectrum except for Bengaluru, India, and Ho Chi Minh City, Vietnam.

Megacities from the HICs are found among the smaller growth rates, with the exception of Seoul, South Korea (HIC), on rank 12 with relative growth rates of 3.36%/year, which stands for a very dynamic expansion for a megacity of an HIC. At the lowest end of the spectrum, we find Kinshasa (LIC) that is limited in growth by the Congo River and Lima (UMIC) that is strongly confined in growth by the ocean on the one hand and the Andes mountain range on the other hand (see also lower panel of Fig. 2). From Fig. 5, we can infer a dynamic cycle pertaining to urbanization and economic development levels. Megacities in LMICs display

a moderate relative settlement growth, which, on average, intensifies during the transition to the UMIC class. In the mature stages of HIC development, the relative settlement growth regresses to the lowest values, indicative of a state of saturation.

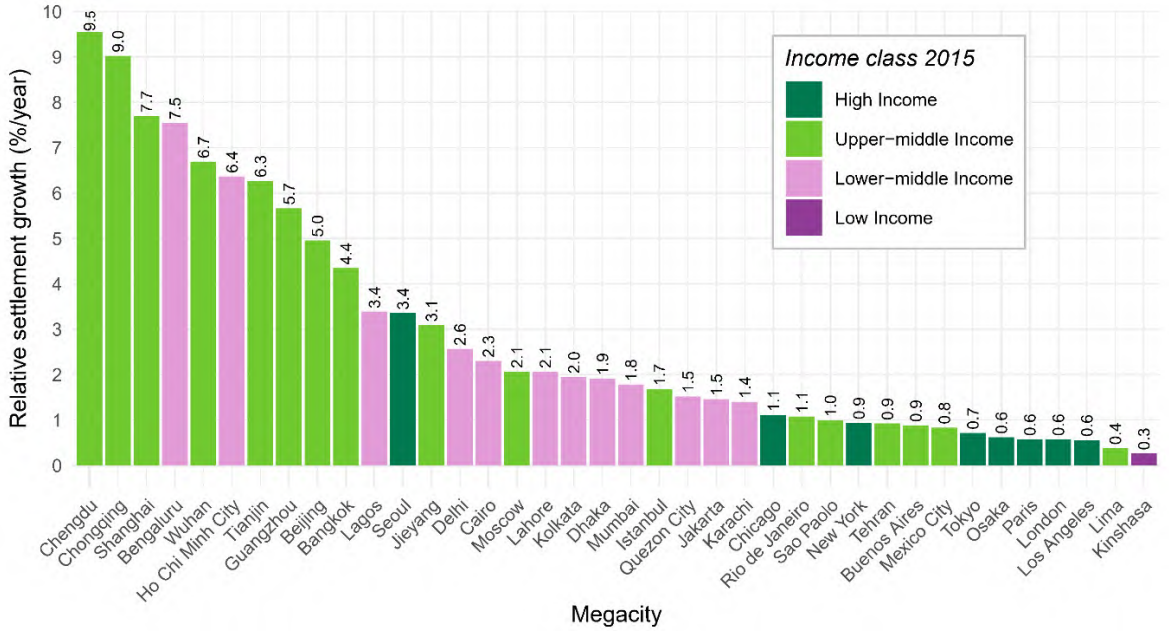


Fig. 5. Relative growth of the settlement area (%/year) for each of the 38 megacities and the time period 1996 to 2015 with respect to the FUAs. The income groups are overlaid following the world's economies classification of the World Bank (2022).

In Fig. 6, the results of the study are summarized by depicting the relative growth of settlement area (%/year) versus the relative NO_2 trend (%/year) for the 38 megacities under investigation. The error bars indicate the uncertainty of the NO_2 linear trend (5). The colors of the dots correspond to the income group classification of the World Bank, the size of the dots scales with the size of the FUA (km^2) according to GHSL_FUA. At a glance, we can see that megacities in HICs are concentrated in the lower left part exhibiting low relative settlement growth and predominantly negative relative tropospheric NO_2 trends (improving air quality). On the contrary on the upper right part of the plot, we find megacities of UMICs characterized by pronounced relative settlement growth and strong positive relative NO_2 trends. By means of the identity line (red dashed line), we can initially interpret the statistical relationship of the megacities concerning settlement and pollution development. While Bengaluru, Guangzhou, Bangkok, Quezon City, and all cities of HICs display a behavior with dynamic settlement growth and small increases or decreases in NO_2 pollution, Chongqing, Chengdu, Tianjin, Wuhan, Beijing, Dhaka, Lima, Tehran Lahore, and Karachi reveal a pronounced relation with disproportionately strong increases in NO_2 levels. Most strikingly Lima depicts a strong increase in air pollution while the settlement growth is the second weakest among the megacities. This can be explained by a strong internal concentration and agglomeration due to as said before limited expansion possibilities due to the coast and the Andes mountain range (see Fig. 2). Concerning FUA size, we observe that the megacities with the largest sizes can be found below the identity line. Only Moscow, Lagos, and Kinshasa are characterized by a proportional behavior (considering their levels of uncertainty).

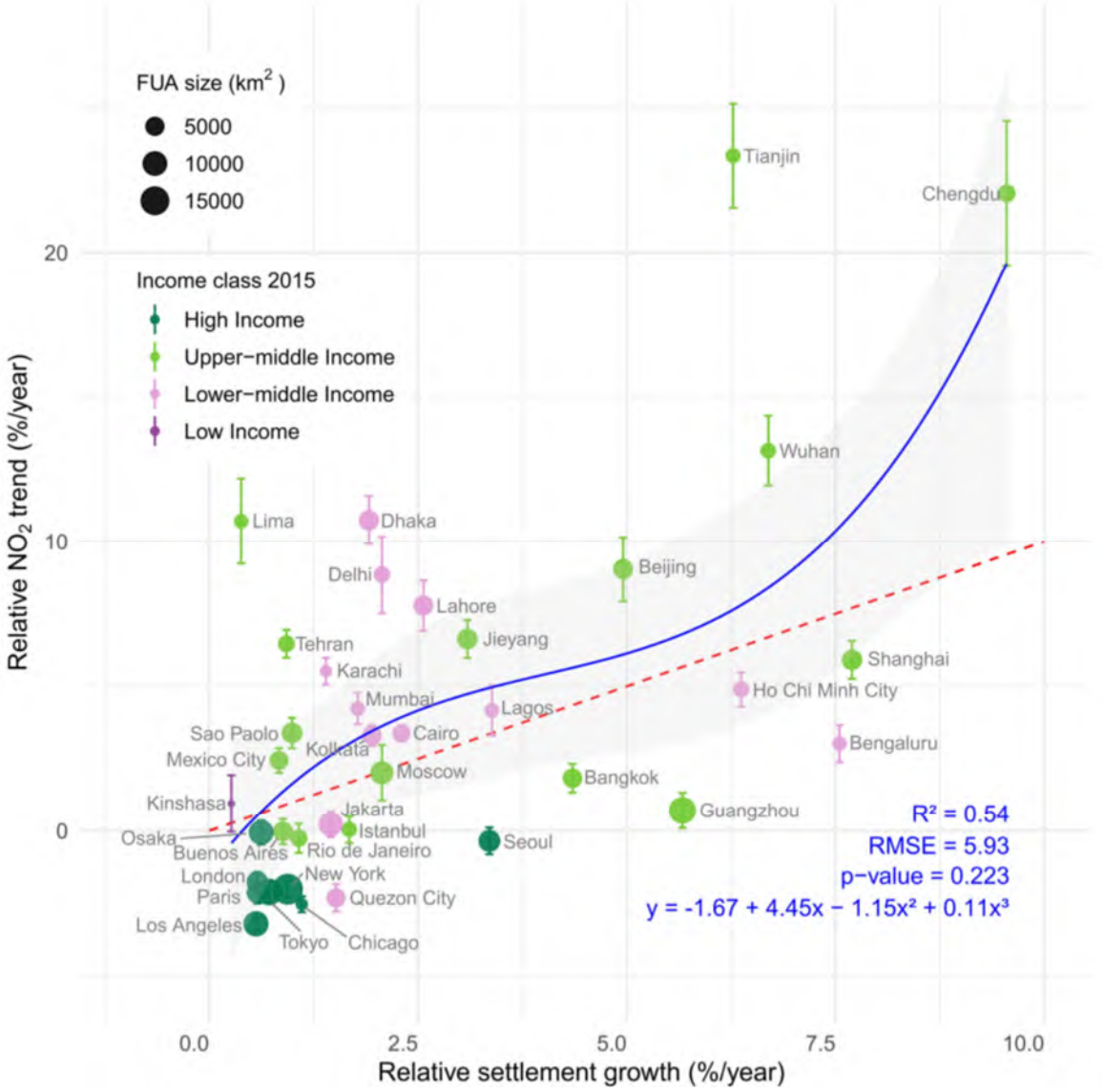


Fig. 6. Relative growth of settlement area versus relative tropospheric NO_2 trend for 38 megacities. The colors of the dots correspond to the income groups following the world's economies classification of the World Bank (2022), and the size of the dots scales with the size of the FUA (km^2) according to GHSL_FUA [50]. The error bars indicate the uncertainty of the linear NO_2 trend. The dashed red line indicates the identity line, the blue line shows the regression, with its parameters, R^2 and p-value, in the lower right. Note that Chongqing at 9.5/44.0 is omitted in the plot, but not in the statistics.

To further examine the statistical relationship between the two quantities, we tested different regression models, such as linear, quadratic, and cubic polynomial models, and evaluated their performance by the statistical parameters R-squared (R^2), root-mean-squared error (RMSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC). The linear regression is outperformed by both the quadratic regression and the cubic polynomial regression with R2 = 0.439 < 0.505 < 0.544, respectively. In addition, the RMSE is the lowest with the cubic fit (RMSE = 0.544), followed by the quadratic fit (RMSE = 6.177) and the linear regression (RMSE = 6.576). Regarding AIC and BIC, lower values indicate a better tradeoff between goodness of fit and model complexity. The AIC is the smallest for the cubic polynomial fit. For the BIC, the values are very close and as such an interpretation is difficult: with 260.77, the quadratic regression exhibits the lowest value compared to

261.30 for the cubic regression and 261.90 for the linear regression. Following the evaluation of the different models, we have finally chosen the cubic polynomial regression to best explain the statistical relationship between relative NO_2 trends and relative settlement growth. We find that it better captures the nonlinear relationship between the variables.

We test the null hypothesis that there is no statistical relationship between settlement growth and NO_2 trends. With a p-value = 0.223, we may accept the null hypothesis and suggest that there is no evidence in favor of a relationship between the two variables. Since statistical significance should be interpreted in the context of the specific analysis, we need to point out that the regression is strongly driven by a few cities with very high increases in settlement growth and NO_2 pollution from 1996 to 2015, namely, Chongqing, Chengdu, and Tianjin. Although they are no "outliers" in a statistical sense, their impact on the regression analysis asks for a careful interpretation of the findings.

Despite the observed variety among the megacities indicating different stages of, e.g., socioeconomic development or green transition, the megacities of the Global North are—as outlined before—clustered in the lower left characterized by decreasing pollution levels and small relative settlement growth rates (Los Angeles, New York, Chicago, London, and Paris). To derive a more diversified spectrum of megacity types, the grouping of megacities is further examined by k-means clustering using only the EO-derived parameters, including relative growth of settlement area (%/year), relative NO_2 trend (%/year), and mean tropospheric NO_2 column density (10^{15} molecules/cm^2). The unsupervised classification distinguishes five classes (Fig. 7). Class 1 (red) comprises four large Asian megacities from HICs (Seoul) and UMICs (Beijing, Guangzhou, and Shanghai) with high tropospheric NO_2 pollution and dynamic settlement growth. Class 2 (ocher) encompasses seven cities of MICs with moderate to high relative NO_2 trends in the range of 6.5 ± 0.5%/year to 13.1 ± 1.2%/year (Wuhan, Dhaka, Lima, Lahore, Delhi, Jieyang, and Tehran). Class 3 (green) corresponds to the abovementioned megacities from HICs and the Global North—except for Seoul—clustered in the lower left with improving NO_2 air quality and small relative settlement growth (Los Angeles, Chicago, Paris, Tokyo, New York, London, and Osaka). However, we also find in class 3 megacities from UMICs, namely, Mexico City, Moscow, and Istanbul. Class 4 (light blue) embraces the largest group of megacities. They are characterized by low to moderate relative NO_2 trends and low to moderate NO_2 mean values. Kinshasa from an LIC is located here. Class 5 (magenta) consists of three cities with the strongest relative NO_2 increase and settlement growth that started in 1996 with low values and small urban area sizes. Finally, Fig. 7 allows for a straightforward and comprehensive classification of the development trajectories of the megacities under investigation. This classification extends to discerning whether these megacities are progressing towards improved environmental conditions, such as the green transition of the economy, or experiencing deterioration, as well as whether they are in a state of stagnation, concentration, or expansion concerning their settlement areas.

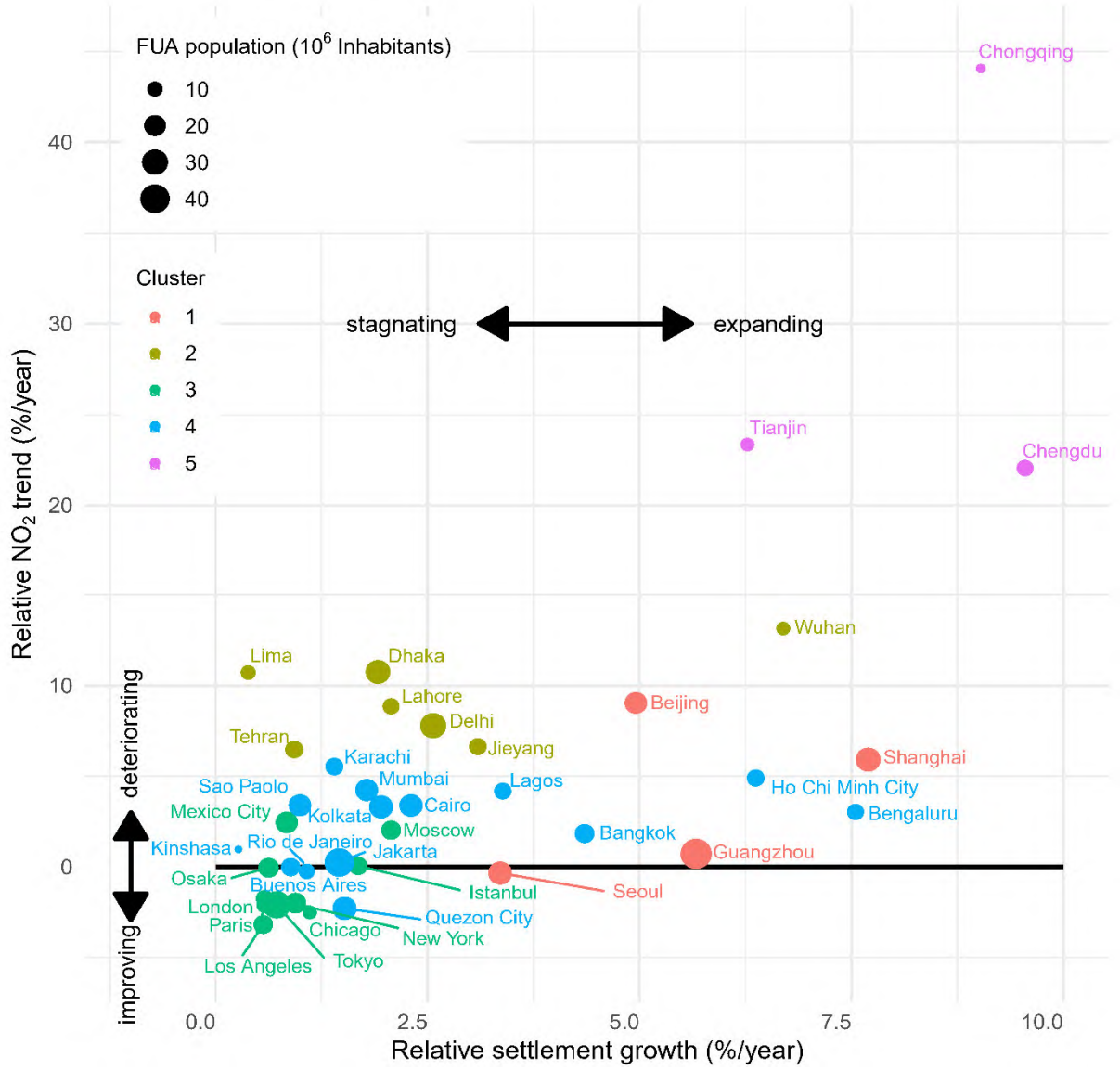


Fig. 7. Classification of megacities by k-means clustering. Five distinct groups were identified using the parameters relative growth of settlement area, relative NO_2 trend, and mean tropospheric NO_2 column density. The size of the dots scales with the FUA population (106 inhabitants).

5. Discussion

In this article, aspects of NO_2 air pollution trends as well as settlement growth were systematically examined and compared for and across 38 megacities worldwide. Therefore, multiannual records of satellite-based tropospheric NO_2 observations and the evolution of the settlement footprint were exploited covering a period of 20 years. A special feature in this study is the use of the spatial concept of FUAs enabling consistent comparisons without distortions due to simplified, nontransparent, or heterogeneous reference units. Compared to most previous studies, the results advance beyond a pure presentation of NO_2 trends by systematically incorporating variables on urbanization, socio-economy, and demography in the analysis.

5.1. Comparison of results with previous studies

A critical discussion of the results obtained against the body of literature, i.e., published trends and settlement growth rates, proves difficult due to missing or different spatial reference units and differing observation periods. To overcome parts of this challenge, we suggest the application of FUAs in this study. Due to the weak representativity of ground-based measurements for the large spatial extent of megacities, a comparison against our approach that integrates over the FUAs proves difficult, too. Furthermore, a satellite-based comparison of NO_2 trends and settlement growth does not yet exist to the best of the authors' knowledge. However, we can discuss in following the tendencies and identify agreements and differences. The direct reference regarding relative NO_2 trends is the work of Georgoulias et al. (2019) that investigated several megacities and urban agglomerations. However, the time period differs (1996 to 2017) and no spatial definition for the derived trends is provided. Overall, the tendencies of the relative NO_2 trends agree. However, the highest positive trends observed in Chongqing (44%/year), Tianjin (23.3%/year), Chengdu (22.0%/year), and Wuhan (13.1%/year) exhibit a smaller amplitude in Georgoulias et al. (2019) with 28.1%/year, 14.9%/year, 11.7%/year, and 7.0%/year, respectively. This could be explained by the extreme dynamic settlement growth and expansion that has not yet been considered in Georgoulias et al. (2019) or the extended period, including 2017, since a beginning decrease of NO_2 was identified by the authors around 2012 due to the implementation of environmental measures and policies. Notably, a large deviation was found for Dhaka, where we report a relative NO_2 trend of 10.8%/year compared to 16.6% in the previous study. Due to the elongated urban shape of Dhaka, the different spatial references could be an explanation. The negative relative NO_2 trends can be confirmed for all megacities of HICs, with the exception of the Japanese megacities of Tokyo and Osaka, where we derived negative trends (improving air quality) compared to small positive trends in Georgoulias et al. (2019). Interestingly, the exceptional behavior of Quezon City (Manila) for a megacity of LMICs with significant negative NO_2 trends of -2.3%/year is confirmed by the previous study (-2.2%/year). The regressing NO_2 levels in Tokyo and Los Angeles have also been identified by Bichler, Schönebeck & Bittner (2023) and linked to economic effects.

As for the NO_2 trends, it is also difficult to directly compare settlement growth rates among different studies. Again, the spatial reference unit and the time period considered play a dominant role and may distort the numbers. With respect to relative settlement growth, the authors in Taubenböck et al., (2012) and Taubenböck (2015) identified, for the periods 1990 to 2000 and 2000 to 2010, Dhaka as the megacity with the highest growth rate followed by Delhi, Jakarta, and Mumbai (where we only consider cities that occur in both studies). In the current study, Dhaka can be found in the midfield with 1.9%/year increase, while Delhi exhibits a stronger growth (2.6%/year) and Mumbai and Jakarta a weaker one with 1.8%/year and 1.5%/year, respectively. In our study, Shanghai (7.7%/year) and Beijing (5.0%/year) report the highest settlement growth among the megacities considered, compared to significantly smaller rates in [59]. However, both studies agree on the smallest growth rates for Tokyo, Buenos Aires, Rio de Janeiro, Paris, London, and Los Angeles. Notably, regarding the African megacities, Taubenböck (2015) reported a similar growth

dynamic for Cairo and Kinshasa, while they strongly differ in this study with Kinshasa revealing the smallest growth of all megacities (0.3%/year) and Cairo 2.3%/year, which is close to the average of 2.87%/year.

5.2. Strengths and limitations of NO_2 observations

The study proves the potential of today's satellite-based observations for studying the variability of the tropospheric composition. After all, these are retrieved from space-borne instruments that have not been specifically designed for the observation of tropospheric air pollutants in the beginning, but stratospheric O_3 (Burrows et al., 1999; Munro et al., 2016). However, even for the GOMEinstrument (Burrows et al., 1999) that started its operation in 1995 with its coarse native spatial resolution of 320 × 40 km^2, anthropogenic signals in tropospheric NO_2 column densities, such as the urban weekly cycle, could be identified (Beirle et al., 2003). This substantiates the sensitivity to the lower boundary layer. Apart from the coarse spatial resolution for urban environments, further limitations of the observations used in this study comprise 1) the low temporal resolution of mainly one (morning) overpass per day due to a sun-synchronous near-polar low Earth orbit, 2) the dependence on solar-backscattered radiation, and 3) the cloud coverage that often persists in particular during pollution episodes, e.g., as a consequence of inversion conditions in the lower boundary layer. Of course, this means that daily variations or local intraurban hotspots cannot be monitored.

However, these satellite-based observations enable the quantification of global tropospheric NO_2 pollution and its trends as was substantiated already by other authors, e.g., see Richter et al., (2005), Schneider, Lahoz & van der A (2015) and Hiboll, Richter & Burrows (2013). By design, one operating space-borne instrument can be considered as a global transfer standard delivering consistent and homogeneous measurements worldwide. It must be borne in mind, however, that tropospheric NO_2 column densities are observed, i.e., retrieved and not the NO_2 mass concentration near the ground (in ambient air) as in-situ sensors do. With these data, however, all regions and cities can be systematically examined in an area-wide, consistent way, appropriate for multitemporal and across-space comparisons (Georgoulias et al., 2019). Such comprehensive, spatially integrating, multitemporal datasets are an important data asset to complement ground-based measurements, as their limited spatial representativeness and often scarce availability lead to uncertainties in the interpretation.

5.3. Uncertainty of NO_2 observations

An important aspect of the trend analysis is the accuracy of the tropospheric NO_2 dataset. Following Boersma, Eskes and Brinksma (2004), the Level 2 processing chain enables to provide a realistic error budget for tropospheric NO_2 vertical column densities. The average uncertainty due to the tropospheric air mass factor over Europe is quantified by 18%–26% per pixel (Boersma et al., 2018). Other studies distinguish between uncertainties under polluted and unpolluted conditions [19], where the estimated uncertainty in the GOME-2 tropospheric NO_2 column for polluted conditions ranges from 40%–80% for a single observation. However, a detailed error analysis for individual retrievals exhibits a strong variation of these estimates (Boersma, Eskes and Brinksma, 2004). The total error budget

comprises contributions of the geophysical parameters cloud fraction, cloud pressure, surface albedo, surface pressure, the a priori NO_2 profile shape, and NOx emissions as well as aerosol-related errors. Following Valks et al. (2011), the most important uncertainties associated with the computation of the tropospheric air mass factor are cloud fraction, surface albedo, and the a priori NO_2 profile. Over polluted regions, air-mass factor errors contribute substantially to the overall error. The overall error of the tropospheric NO_2 columns is driven by error propagation of the error terms listed before. Since a monthly mean dataset is built from daily observations (with sample size n) (Georgoulias et al., 2019) and the standard error decreases by $1/n$, the overall error can be strongly reduced. For a campaign in Europe, monthly mean tropospheric columns are in very good agreement, with differences generally within 0.5×10^{15} molecules cm^{-2} (Valks et al., 2011). However, with respect to the systematic error, a comparison of tropospheric NO_2 column densities to global ground-based measurements shows, on average, a negative bias of typically -23% to -37% in clean to slightly polluted conditions, but reaching values as high as -51% over highly polluted areas (Verhoelst et al. 2021). A part of the biases can be explained by differences in the representativity of the stations and different area averaging and measurement times. It is also known that errors in the tropospheric columns result from shortcomings in the horizontally coarse a priori profile representation in the underlying chemical transport model used in the retrieval (Müller, Erbertseder & Taubenböck, 2022). Therefore, the tropospheric NO_2 retrieval could be significantly improved by using a priori profiles from a CTM with much higher spatial resolution and a more realistic cloud treatment (Liu et al. 2021).

In spite of the occurrence of NO_2 transport between urban areas and the surrounding regions, the relatively short atmospheric lifetime of NO_2, ranging from a few hours to a day, ensures that its concentrations predominantly reflect the local strength of NO_2 emissions (Lange, Richter & Burrows, 2022; Müller, Erbertseder & Taubenböck, 2022). Despite the limitations, satellite data of the last decades have already been used to analyze the background pollution and the trend development around the world (Richter et al., 2005; Schneider, Lahoz & van der A, 2015; Hiboll, Richter & Burrows, 2013).

5.4. Strengths and limitations of the WSF

The WSF has proven to be an effective geo dataset for large-scale mapping of human settlements at 30×30 m^2 spatial resolution on a global scale (Taubenböck et al., 2012; Marconcini et al., 2015; Esch et al., 2017). With the WSF evolution available on a yearly basis, this study quantifies and analyses the horizontal expansion of megacities. We believe that this is a valid and consistent systematic approach to study this aspect of urbanization and megacity evolution. However, although the mapping product has achieved high accuracies, spatially uneven distributed classification errors have been documented by Marconcini et al., (2015 & 2021).

In addition, it is known that in certain development cycles, a vertical growth in cities cannot be neglected. In particular, this occurs with the development and strengthening of (central) business districts or large-scale housing developments with high-rise or mid-rise buildings for living. However, multitemporal settlement data, including the evolution in the third dimension with sufficient accuracy, are still absent. Latest developments, however, show

that these mapping products will become more and more available (Esch et al., 2022; Frantz et al., 2021; Geiß et al., 2019; Li et al., 2020).

The WSF is a binary classification of 'settlement' and 'non-settlement'. It is clear that urban areas and especially large urban regions, such as megacities, exhibit intraurban structural and functional variations (Taubenböck et al., 2020). This is also associated with different sources of NO_2 emissions. Thematically, higher resolution classifications of urban structural configurations of urban landscapes are emerging (Zhu et al., 2022), but are not available multi-temporally for the monitoring period. This intra-urban differentiation, however, may provide new insights as a research topic in the future.

5.5. Multi-city studies and spatially consistent concepts

Many studies concentrate on a single or a limited number of selected cities, thereby raising concerns regarding the generalizability of conclusions and the applicability of findings to other urban or regional contexts, e.g., Bichler, Schönebeck & Bittner (2023). The need for multi-city studies was recently raised (Voogt et al., 2023). The current study contributes to serving this demand since it provides a systematic analysis of all megacities worldwide and empowers to derive general conclusions.

The demand for multi-city studies also raises the need for appropriate and geographically consistent reference units (Taubenböck et al., 2019). However, geographically inconsistent spatial city extents, such as administrative areas, remain the key reference units in most studies. Therefore, it is often difficult to perform a solid comparison without distorted statistics, to draw general conclusions, or to compare results across scientific contributions. The inconsistencies of population numbers are obvious in Table 1 when comparing numbers from the UN (United Nations, 2019) versus the GHSL-FUA (Schiavina et al., 2020) resulting from different statistical approaches and underlying spatial concepts. Contrary to previous studies on tropospheric NO_2 trends, this work advances on the approach by a systematic analysis of megacities worldwide considering pollution in combination with settlement growth and the spatial concept of FUAs providing comparable geographic units. To the best of the authors' knowledge, this consistent analysis is currently not available. Furthermore, a thorough investigation of this 20-year period by means of a consistent spatial concept lays also down the fundament for follow-up studies. On the one hand, the analysis that was carried out for the FUAs of the epoch 2015 could be expanded toward their multitemporal development. On the other hand, different data-driven approaches for consistently delineating megacities, such as the degree of urbanization (Dijkstra et al., 2021), morphological urban areas (Taubenböck et al., 2019), or spatially continuous ways of measuring the urban and the nonurban (Taubenböck et al., 2022), could be applied, and the stability of the results tested.

5.6. Dependence of results on time period considered

One might argue that the time period considered in the study is outdated and that the latest developments of air pollution levels as a consequence of implementing measures and actions are not considered, such as the policies to reduce air pollution in China or India starting

around 2010. First of all, the study is based on the newest and most consistent datasets available. Second, it establishes the period from 1996 to 2015 as the baseline of the study, which results from the overlapping period of the WSF evolution (1985 to 2015) on the one hand and the beginning of tropospheric composition mapping in 1995/1996 with GOME and successors to present on the other hand. Third, evaluating past air pollutant trends in urban areas gives insight into the effectiveness of previous or recently implemented policies (Goldberg et al. 2021). The results attained here can serve as the baseline to examine the effectiveness of current policies to control air pollution and urban sprawl, and possibly support measures for future policies (Baklanov, Molina & Gauss, 2016). The policy in China for example has reversed the NO_2 trends partly from a strong increase into a decrease in the meantime, thanks to the implementation of mitigation measures, i.e., NOx emission reductions. This reversal cannot be appropriately acknowledged by neglecting the situation in the decades before. A thorough evaluation of trends is a prerequisite for studying the effects of policy implementations like the introduction of emission taxes on pollution levels and economic value (Erbertseder et al., 2023). In general, however, the study will be performed from 2016 to the present as soon as the update of the WSF evolution will be available. This will provide a better statistical basis to identify trend reversals as initiated in (Georgoulias et al., 2019).

5.7. Megacity development and global economic landscape

As early as in 1992, Shukla and Parikh (1992) addressed the need for an empirical examination of differences and similarities between cities from the Global North and the Global South concerning air pollution, urbanization, and city size. At that time, based quite on speculative terms, cities from less developed countries were thought to hold more potential for much greater pollution than it was anticipated for developed economies. This assumption was based on the fact that they had to go a long path of industrialization and had more limited means to mitigate environmentally detrimental concomitants of this process. Therefore, Shukla and Parikh (1992) made an initial step toward assessing the magnitude of one aspect of the problem that cities from developing countries are facing, air pollution. In doing so, they conducted an empirical examination of differences and similarities between the two contexts. Cities of the Global South are characterized by greater population pressures, more rapid urbanization rates, the unbalanced growth of megacities beyond limits, and the elevated use of resources predicated by more appropriate technologies to higher development levels (United Nations, 2018). As presented in the current work, we can demonstrate what EO has accomplished in the meantime, by providing global datasets for comprehensive documentation and studies. As we neither use the concept of developed and developing countries nor Global North and Global South, we apply the global income classification of the World Bank that gives a much more detailed view. Furthermore, different groups are in a different stage of the economic development and as such in different technological and socioeconomic cycles.

Given the EO data records covering 20 years, we can take the opportunity to revisit development trends. In particular, by analyzing EO data records that have now become available, we retrospectively contribute to answering the hypothesis whether the sizes of

megacities are dynamically growing in less developed countries and whether the growth rate in major metropolitan areas in more developed countries is declining. In general, we confirm the hypothesis, but the systematic analysis of all megacities allows us to identify exceptions. As was shown in this study, Seoul is measured with exceptionally high settlement growth rates for a city in an HIC. On the contrary, Kinshasa and Lima show the lowest relative settlement growth rates although the cities are from LICs and MICs. Despite the low settlement growth rate in the time period considered, Kinshasa is predicted by the UN to become home to 58.42 million people by 2075. Furthermore, Chicago is supposed to become the fastest developing megacity from the HICs. Other emerging megacities are Dar es Salaam, Luanda, Baghdad, Chennai, and Bogota. The continuous monitoring of the tropospheric composition and the settlement footprint will be inevitable to track the development, urbanization, and environmental impact of existing and emerging megacities worldwide.

When analyzing relative settlement growth rates and interpreting them within the global economic landscape, we need to keep their relation to the size of the FUAs under consideration in mind. Due to the large size of the FUA of Los Angeles (10 407 km^2) for example and its built-up area in 1996 of 5091 km^2, a relative increase in settlement area of 0.6%/year corresponds to an increment of built-up area in 2015 by 571 km^2. This exceeds the increase in the settlement area of, e.g., Lagos that exhibits a higher relative growth of 3.4%/year corresponding to an increase in built-up area of 455 km^2. However, for reasons of comparison, we justify the concept of relative numbers, since it is a measure of the progress of settlement growth and its impact on the megacity.

5.8. Transformative aspects

Social, economic, and environmental justice and substantiating findings on air inequality are in demand. The study so far neglects transformative aspects since it currently does not consider social or socioeconomic aspects in detail. This has not yet been included in the current work for the following reasons: these aspects cannot (yet) be derived directly from or tackled with EO data. As such, the data available for megacities are derived by varying methods and are not available in a complete or consistent manner for the FUAs across the globe. The social or socioeconomic data currently available are provided for different administrative units, often only on aggregated scales such as national levels thus resulting in global data heterogeneity and uncertainty which hinders reliable conclusions. Also, the statistics of the World Bank applied in this study are based on national aggregates and should be interpreted carefully to capture or disentangle developments of an individual city. However, the significance of megacities in a country and their large share to most socioeconomic and environmental factors justify the assumptions made in this work.

5.9. Perspectives on integrated air pollution assessments

However, in order to consider the total air pollution situation in the megacities and to assess the associated health risks, the integration of other air pollutants, such as fine particular matter and O_3, are in demand. Therefore, the comparison of megacities, as presented here, needs to be interpreted carefully. We do refer to tropospheric NO_2 only. While the

distributions and trends of NO_2 predominantly reflect anthropogenic emission behavior, variability and changes in meteorological parameters may also play a decisive role. However, the time series analysis method used in this article should result in robust estimates and cope with seasonal and annual fluctuations (Weatherhead et al. 1998). Air pollution levels and the resulting (long-term) trends may also be subject to influences from other parameters and processes, such as changes in precipitation, the number of fog days or the concentration, and type of aerosols. Ultimately, more detailed and systematic studies are needed to investigate possible changes in the chemical regime of the megacities.

Current satellite missions and instruments, such as Sentinel- 5P/TROPOMI and GEMS, and the upcoming missions (Sentinel-4, Sentinel-5, and TEMPO) will provide unprecedented opportunities for studying the tropospheric chemical composition. In particular, by combining LEO and GEO orbits, which will resolve diurnal cycles of the atmospheric composition, we will be able to much better observe and analyze air pollution variability in megacities and megaregions worldwide.

6. Conclusion

A first systematic analysis of megacities in relation to tropospheric NO_2 pollution trends and settlement area growth was presented. The results advance beyond a pure presentation of NO_2 trends by systematically incorporating variables on urbanization, socio-economy, and demography in the analysis. The study covers a period of 20 years (1996 to 2015). By consistently applying the spatial concept of FUAs, we find that the megacities on average exhibit an increase in tropospheric NO_2 burden of $5.06 \pm 0.83\%$/year. The average relative settlement growth rate for the megacities is 2.87%/year. Among the 38 megacities under investigation, a wide range from increasing to decreasing NO_2 pollution levels was derived and, with respect to settlement area growth, cities with stagnating to highly dynamic growth rates. In terms of absolute growth of the built-up area, we report increases from 16 km^2 (Kinshasa, Democratic Republic of the Congo) and 46 km^2 (Lima, Peru) to 3511 km^2 (Guangzhou, China). Despite the variety, we found that the results exhibit a pronounced relation to the income groups following the world's economies classification of the World Bank. While the strongest increase of NO_2 pollution and settlement growth as well as per-capita pollution occur in megacities of UMICs, megacities of LMICs are centered in the midfield of the spectrum, but with the lowest per-capita pollution levels. Megacities in HICs are predominantly characterized by the smallest relative settlement growth rates and a decrease in NO_2 air pollution. Conversely, despite the decreasing trends observed in megacities from HICs, indicating advancements in air quality, the per capita pollution levels in these cities persist at some of the highest levels. This dichotomy underscores the complex interplay between economic development, urbanization, and environmental quality.

We conclude that the combination of tropospheric NO_2 composition records and the WSF evolution provide essential components to develop a better understanding of the drivers and impacts of urbanization and air pollution. Regression analysis indicated a statistical relationship between relative settlement growth and relative NO_2 increase among the megacities.

Furthermore, we conclude that the presented approach is suitable for identifying and comparing megacities concerning their socioeconomic development and transformation stages, such as the green transition and world economies. By a simple classification of EO-based parameters, we can label the megacities in their transformation stage. We can infer a dynamic cycle pertaining to urbanization and economic development levels.

Finally, we conclude that the findings give insight into the effectiveness of previous and current urban policies and developments, and possibly support future measures to transform existing cities according to the United Nations Sustainable Development Goal SDG 11 ("Make Cities Inclusive, Safe, Resilient and Sustainable").

Acknowledgements

The authors would like to thank Tropospheric Emission Monitoring Internet Service (https://www.temis.nl/) for the free use of tropospheric NO_2 data.

Part IV

Urban dimensions and rankings

CHAPTER XII	The urban morphology on our planet – global perspectives from space	283
CHAPTER XIII	A new ranking of the world's largest cities – do administrative units obscure morphological realities?	303
CHAPTER XIV	Which city is the greenest? A multi-dimensional deconstruction of city rankings	327

Chapter XII

The urban morphology on our planet

Global perspectives from space

Remote Sensing of Environment, 2022, vol. 269, 112794

Xiao Xiang Zhu, Chunping Qiu, Jingliang Hu, Yilei Shi, Yuanyuan Wang, Michael Schmitt & Hannes Taubenböck

© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.1016/j.rse.2021.112794

Received 23 June 2021; Accepted 3 Nov 2021; Published 24 Nov 2021

Abstract

Urbanization is the second largest mega-trend right after climate change. Accurate measurements of urban morphological and demographic figures are at the core of many international endeavors to address issues of urbanization, such as the United Nations' call for "Sustainable Cities and Communities". In many countries – particularly developing countries –, however, this database does not yet exist. Here, we demonstrate a novel deep learning and big data analytics approach to fuse freely available global radar and multi-spectral satellite data, acquired by the Sentinel-1 and Sentinel-2 satellites. Via this approach, we created the first-ever global and quality controlled urban local climate zones classification covering all cities across the globe with a population greater than 300,000 and made it available to the community (https://doi.org/10.14459/2021mp1633461). Statistical analysis of the data quantifies a global inequality problem: approximately 40% of the area defined as compact or light/large low-rise accommodates about 60% of the total population, whereas approximately 30% of the area defined as sparsely built accommodates only about 10% of the total population. Beyond, patterns of urban morphology were discovered from the global classification map, confirming a morphologic relationship to the geographical region and related cultural heritage. We expect the open access of our dataset to encourage research on the global change process of urbanization, as a multidisciplinary crowd of researchers will use this baseline for spatial perspective in their work. In addition, it can serve as a unique dataset for stakeholders such as the United Nations to improve their spatial assessments of urbanization.

Keywords

Remote sensing; Sentinels; big data; data fusion; deep learning; local climate zones; urban morphology; global urban LCZ dataset; global inequality

1. Introduction

1.1. Motivation

Urbanization is one of the most important trends in global change. Although a near-perfect correlation is verified between urbanization and economic prosperity of societies (Glaeser et al., 2013), this correlation does not automatically lead to a golden future: instead, the unprecedented dynamics and dimensions of natural growth of and migration into cities pose fundamental challenges to our human societies across the globe. Many international endeavors addressing issues of urbanization, such as the United Nations' call for "Sustainable Cities and Communities," are based on accurate measurements of urban morphological and demographic figures. Such measurements provide key scientific foundations for the allocation of valuable resources for a wide range of stakeholders and form the basis of global efforts to understand and track progress in improving human livelihoods.

Although urban geographic information has significantly improved in recent decades, the required geo-information is still not available in many countries. For example, the United Nations population figures, on which our understanding of urbanization at these scales is currently primarily based, do not provide information on the distribution, pattern, and evolution of the built environment (Zhang and Seto, 2011). Remarkably, in 2014, for the first time, the IPCC report included a specific chapter on urban areas, making note that cities of sufficient density and spatial scale can influence their local micro-climate (Revi et al., 2014). This highlights the importance of cities to climate action, yet there are significant gaps in knowledge and strategic directions. Global information on inter- and intra- urban variability, urban dynamics, 2D and 3D urban form, and links between urbanization and well-being are overdue (Chai and Seto, 2019; Mahtta et al., 2019). In Arellano Ramos and Roca Cladera (2016), the data of nighttime lights was utilized for detecting and classifying urban areas. Recently, the global land cover map (GLC), which includes urban and non-urban classes, is transferred from 30 m to 10 m (Gong et al., 2019).

In the past decades, satellite remote sensing has been the foundation of data collection and the development of knowledge about our Earth. During that period, new initiatives in global urban mapping have advanced the quality of spatial knowledge. On a global scale, there is a handful of global urban land use classification maps. For instance, the Global Urban Footprint (GUF) (Esch et al., 2013) provides a binary mask of urban vs. non-urban surfaces, while the World Settlement Footprint Evolution (WSF-Evo) (Marconcini et al., 2020) documents their evolution over time. In addition, the World Urban Database provides intra-urban local climate zone classification maps (Stewart et al., 2014; Stewart and Oke, 2012) – however only for about 100 metropolises and to a best resolution of a few hundred meters. Despite their significant contribution to large scale urban mapping, these approaches have not addressed the variations of intra-urban morphology. On the other hand, the potential of Earth observation to describe the complex, small-scale morphology of cities has been shown in many approaches, e.g. (Taubenböck et al., 2013; Bechtel et al., 2015a). However, global urban mapping approaches still lag behind. Consistent and resilient global spatial data in high geometric, thematic and temporal resolutions necessary to address the challenges

described above is still nonexistent. This is mainly due to the lack of high-resolution global satellite data and computationally and methodologically efficient algorithms.

In the meantime, Earth observation (EO) has irreversibly arrived at a golden era of big data. The game changer is the Sentinel satellites of the European Copernicus programme which provide continuous, reliable, quality-controlled acquisition of big EO data that are free and open. To date, tens of petabytes of satellite data from complementary sensors have already been acquired and shared. In addition, the programme offers a long-term perspective of guaranteed data acquisition and sharing until 2030. Recently, the European Space Agency (ESA) awarded contracts for the development of six new Copernicus missions. Beyond that, plans for 2040 are already under discussion. To effectively retrieve global urban geo-information from such a massive data source requires not only new technological approaches to manage large amounts of data, but also new analysis methods. Here, methods of data science and artificial intelligence (AI), such as machine learning, become indispensable. Deep learning in particular has led to a revolution in AI in recent years. Since taking off several years ago, deep learning in remote sensing has become a blooming research field (Zhu et al., 2017; Tuia et al., 2021; Camps-Valls et al., 2021). Its huge potential in global urban mapping using EO data is ripe for discovery.

1.2. Review of local climate zones classification algorithms

Local climate zones (LCZs) were originally developed for metadata communication of observational urban heat island studies (Stewart, 2011). But soon they also showed potential in urban morphology mapping. A significant part of the existing development of LCZ classification is community-based large-scale LCZ mapping using freely available Landsat data and softwares (Mills et al., 2015; Bechtel et al., 2015a,b; Hidalgo et al., 2019), known as the World Urban Database and Portal (WUDAPT). These community-based efforts mark the first step towards a more synergetic cooperation among researchers. Yet, (Bechtel et al., 2017, 2019) discovered that the quality of these maps is heavily dependent on individual producers. Ren et al., (2016, 2019) also argue that the standard WUDAPT mapping approach cannot fulfill quality demands of practical usage. Therefore, researchers also attempt in parallel to improve the classification methodologies. In terms of data sources, Bechtel et al. (2016) and Hu et al. (2018) demonstrate that synthetic aperture radar (SAR) data can be beneficial as an additional data source. Unger et al. (2014), Zheng et al. (2018) and Zhang et al. (2019) conclude that geographic information system (GIS) data are also beneficial to this classification, subject to the availability of GIS data. Regarding the classification algorithms, similar to WUDAPT, many studies utilized Random Forests to accomplish LCZ classification (Danylo et al., 2016; Demuzere et al., 2019; dos Anjos et al., 2017; Yokoya et al., 2017) due to its easy implementation and robust performance. However, Yoo et al. (2019) and Rosentreter et al. (2020) recently found that convolutional neural networks (CNNs) outperform Random Forests on LCZ classification with fair evidences. Sukhanov et al. (2017) fuse the predictions of random forest- and CNN-based classifiers to deliver LCZ maps. Qiu et al. (2019) propose a recurrent network to utilize multi-temporal Sentinel-2 data for LCZ classification of European cities. Qiu et al. (2020) report the performance of many

standard CNN architectures for LCZ classification in a practical and comprehensive manner. The authors also introduced a CNN which utilizes multi-scale representations.

1.3. Contribution of this paper

Here we demonstrate a novel deep learning and big data analytics approach in order to fuse freely available global radar and multi-spectral satellite data, acquired by the Sentinel-1 and Sentinel-2 satellites. Via this approach, we achieved the first-ever global and quality controlled urban LCZs classification covering all cities with a population greater than 300,000, i.e., 1692 cities in total, and made it available to the community. We validated our approach on 52 urban agglomerations across all five inhabited continents. To do so, we used different strategies, giving upper and lower bounds to the accuracy of our global classification result. This dataset, named "So2Sat Global Urban LCZ" (So2Sat GUL; data: https://doi.org/10.14459/2021mp1633461, demo: https://github.com/zhu-xlab/So2Sat-LCZ-Classification-Demo), will boost research on the global change process of urbanization, as a multidisciplinary group of researchers may use this baseline for the spatial perspective of their work – including planners, demographers, sociologists, economists, climatologists, and many others. On top of that, So2Sat GUL can serve as a unique dataset for stakeholders such as the United Nations to improve their spatial assessments of urbanization.

To stimulate the dataset's usage, we carried out a study to showcase urban geographic analysis to better quantify and understand the global urban morphology. We found that more than 30% of the human settlement area is sparsely built. Similarly, about 25% of human settlements are lightweight or large low-rise buildings, which are often slum areas in cities of the Global South. By combining the urban morpho logical map with a global population map, we discovered that the 30% sparsely built area contains less than 10% of the total population of the 1692 cities, whereas the 25% low-rise area contains 26% of the total population. In addition, we were able to show that the influence of geographical and cultural regions on the physical design of the urban landscape also empirically stands up to observations and theories. This quantifies the very uneven distribution of population and settlement types across the world.

The remainder of this paper is organized as follows. The next section briefly introduces the study area and data. Section 3 explains our method for the mapping of global local climate zones. Section 4 presents the experimental results. Section 5 showcases a pilot application of the resulting LCZ map by analyzing the relation between LCZ patterns and urban morphology. Finally, in Section 6 we share our perspectives on possible usage of our results for the analysis of global urbanization and on further improvements on our results.

2. Study area and data

2.1. Study area

According to the United Nations (United Nations, 2018), one-third of the world population will live in cities of at least 500,000 inhabitants in 2030. Those cities are facing a challenge of meeting the needs of numerous inhabitants for improving their livelihood while only

limited resources are available for a sustainable development. To support those cities with geographic information, our study focuses on producing maps of urban morphology for 1692 cities whose population is greater than 300,000 according to the census of the United Nations (United Nations, 2014). The location of the 1692 cities is shown in Fig. 1.

We used the existing GUF binary urban mask and the center coordinates of the 1692 cities to adaptively determine the extent of each city. We grow a rectangle centered at the coordinate of each 1692 cities, until half of the area is not built up anymore according to GUF. This serves as a basis to define the region of interest (ROI) of each city. In addition, since GUF and the population figure may be out of date for certain cities, to take into account the rapid urbanization, we expand each side of the rectangles by a factor of two (i.e. factor of four in area). In this way, we believe most built areas of cities of interest shall are included. A zoom in of the ROIs in Europe can be seen in Fig. 2. As one can see, the size of ROIs are adaptive according to the city size. Because of our extension, the ROI is much larger than the actual city boundary. As a result, some ROIs also partially overlap.

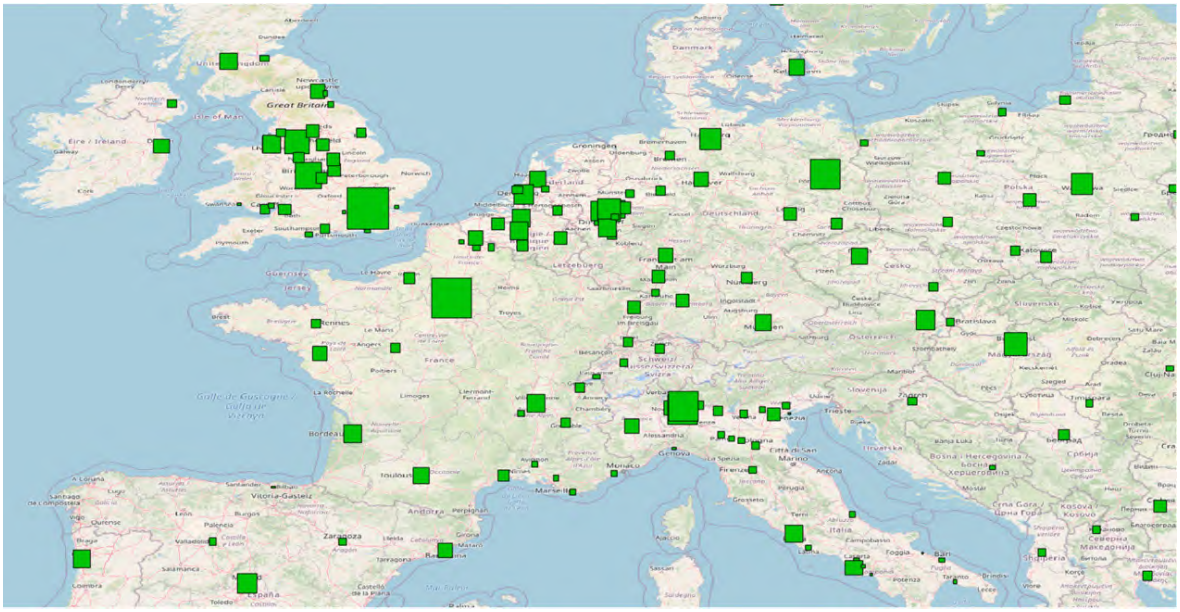


Fig. 1. A close up of the ROIs of our cities of interest in Europe. We grow rectangles centered at the coordinates of the 1692 cities, until half of the area is not built up anymore according to GUF. Since GUF and the population figure may be out of date for certain cities, we expand the each side of the rectangles by a factor of two (i.e. factor of four in area), to take into account the rapid urbanization. Because of this extension, the ROI is much larger than the actual city boundary. As a result, some ROIs also partially overlap.

2.2. Data source and preprocessing

We chose the freely available Sentinel-1 and Sentinel-2 satellite data as the data sources for global urban morphology classification. These two satellite missions provide highly complementary data. Sentinel-1 provides SAR imagery in the microwave range, whereas Sentinel-2 provides multi-spectral images. Nonetheless, it is not a trivial task to extract global urban morphology from these free and open global data. The first challenge is the complex relation between the measured satellite data and the urban morphology, but also the global diversities caused by differences in cultures, environments, locations, topography,

and climates. The second is the variability of the tremendous volume of satellite data for global mapping. To this end, we have developed a complete workflow of creating analysis-ready images, model training, and large-scale inferencing. Since the scope of this paper is more on the methodology and the final datasets, we briefly listed the pre-processing procedures for creating the analysis-ready data.

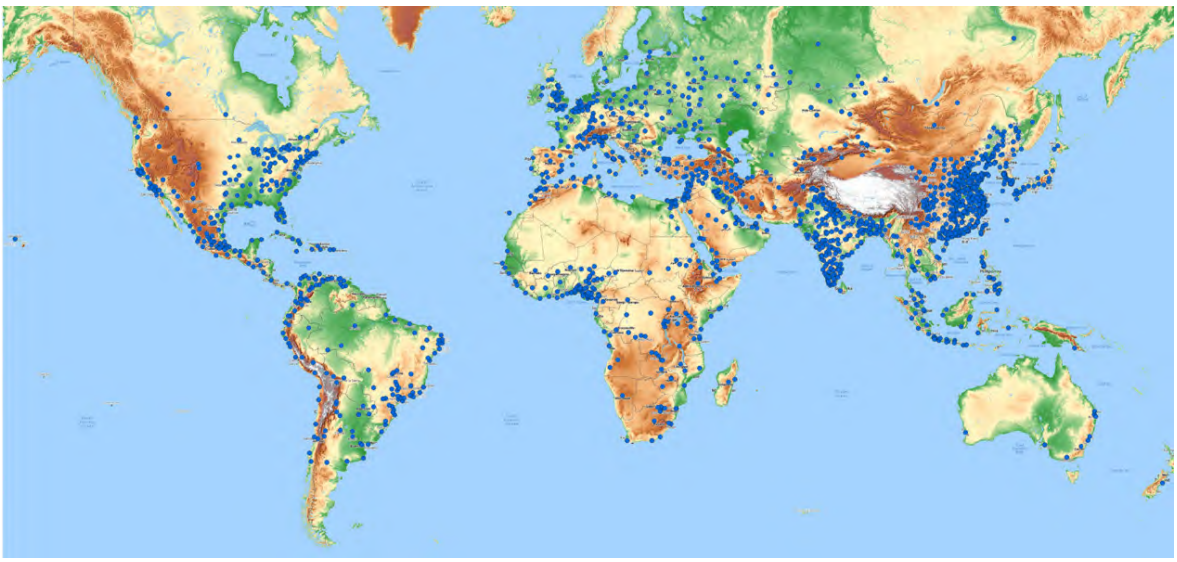


Fig. 2. Location of the cities included in this study. The blue dots on the upper figure indicate the location of each city. The background image shows the topographic height of the Earth surface (source: OpenTopoMap (CC-BY-SA)). The height increases from green to brown to white.

2.2.1. Sentinel-1

Sentinel-1 level-1 single look complex (SLC) dual-pol (VV-VH) Interferometric Wide swath (IW) data of 2017 summer were collected for the 1692 cities in this study. A series of processing steps using the ESA SNAP toolbox were applied to prepare an analysis-ready dataset. These processing steps are listed as follows:

• *Apply orbit profile*: This module downloads the latest orbit profile so that a precisely geocoded product can be achieved.

• *Radiometric calibration*: This step is to provide imagery in which the pixel values can be directly related to the radar backscatter of the scene. To do this, the output scaling applied by the processor is undone and the desired scaling is reintroduced according to a look uptable.

• *TOPSAR deburst*: For each polarization channel, the Sentinel-1 IW product has three swaths. Each swath image consists of a series of bursts. The TOPSAR deburst merges all these bursts and swaths into a single SLC image.

• *Polarimetric speckle reduction*: Speckle reduction was conducted by using the SNAP-integrated refined Lee filter. Unfiltered data can be preserved by skipping this module.

• *Terrain correction*: This step geocodes the range-azimuth SAR image to a geographic coordinate. SRTM DEM was used to provide the height information. The data was re-

sampled to a 10m GSD by the nearest-neighbor interpolation. The data was geocoded into the WGS84/UTM coordinate system.

After processing, seven real-valued bands are contained in the Sentinel-1 analysis-ready data. They are listed in Table 1.

Table 1. Description of Sentinel-1 analysis-ready data.

Band	Description
1	the intensity of unfiltered VH channel, in decibels
2	the intensity of unfiltered VV channel, in decibels
3	the coherence between unfiltered VV and VH
4	the intensity of the refined LEE filtered VH channel, in decibels
5	the intensity of the refined LEE filtered VV channel, in decibels
6	the coherence between the refined LEE filtered VV and VH
7	the phase difference of the refined LEE filtered VV and VH

2.2.2. Sentinel-2

Similarly, Sentinel-2 L1C data (top of atmosphere reflectance) were collected for our study area. The preprocessing Sentinel-2 images is less demanding than that of Sentinel-1. The main challenge is creating cloudfree mosaics. We achieved this by exploring an engineering approach relying on pixel-wise cloud detection and the combination of multitemporal images within short time periods (Schmitt et al., 2019). The Sentinel-2 L1C top of atmosphere reflectance data were also scaled by a factor of 1/10,000 (Gatti and Bertolini, 2013). In the Sentinel-2 analysis-ready data, 10 out of the 13 bands were used. Specifically, the channels with ground sampling distance (GSD) of 10 m and 20 m were used. The 10 real-valued bands are listed in Table 2.

Table 2. Description of Sentinel-2 bands used in this study.

Band	Central wavelength [nm]	GSD [m]	Description
B2	490	10	Blue
B3	560	10	Green
B4	665	10	Red
B5	705	20, upsampled to 10m	Visible and Near Infrared (VNIR)
B6	740	20, upsampled to 10m	VNIR
B7	783	20, upsampled to 10m	VNIR
B8	842	10	VNIR
B8a	865	20, upsampled to 10m	VNIR
B11	1610	20, upsampled to 10m	Short Wave Infrared (SWIR)
B12	2190	20, upsampled to 10m	SWIR

In addition, considering the much larger seasonal variation in optical images in comparison to SAR images, we collected a four-seasonal set of Sentinel-2 images of our study areas. The seasons collected were the winter of 2016, and the spring, summer, and autumn of 2017. It

is however worth to mention that some of the cities in our study areas do not have four seasonal data due to an excessive amount of cloud or data corruption. For those cities, we only include seasons that have data available.

2.3. The training dataset: So2Sat LCZ42

In order to train a deep learning model with good generalization abilities for inferencing with global data, a large-scale and representative annotated dataset is crucial. For this purpose, we created a rigorously labeled reference dataset: the *So2Sat LCZ42* benchmark dataset (Zhu et al., 2020). Over one month, 15 domain experts carefully designed the labeling workflow, the error mitigation strategy, and the validation methods, and then conducted the data labeling. The dataset consists of manually assigned LCZ labels of 400,673 Sentinel-1 and Sentinel-2 image patch pairs globally distributed in 42 urban agglomerations plus 10 additional smaller areas covering all the inhabited continents and 10 different cultural zones. We conducted a rigorous quantitative evaluation of 10 cities in the dataset by having a group of remote sensing experts cast 10 independent votes on each labeled polygon, in order to identify possible errors and assess the human labeling accuracy. The "human confusion matrices" per polygon and per pixel were created, where the confident of individual classes can be seen.

In general, our human labels achieve 85% confidence. This confidence number can serve as a reference accuracy for the machine learning models trained on this dataset.

3. Method

3.1. Deep learning model

We employed a modified residual neural network as our base model, because multiple literatures have shown CNN greatly outperforms random forest in LCZs classification accuracy (Yoo et al., 2019; Rosentreter et al., 2020). In addition, a residual network has the advantage over many other networks in facilitating model training and encouraging feature reuse, enabled by the short-cut connections within the network. Our network consists of one convolutional layer in the beginning, six following residual blocks, and one dense layer at the end, as shown in Fig. 3. Each convolutional layer is followed by one batch normalization layer and one activation layer. Unlike the well-known ResNet 101 or 152, our modified ResNet has a reduced number of residual blocks, in order to fit the input patch size of a typical LCZ classification problem. LCZ classification maps usually have a GSD of 100 meters. We choose an input patch size of 32 × 32, which corresponds to 320 × 320 meters. This choice is a balance between the depth of the network and the definition of LCZ.

Similar to many CNNs for image classification, our model progressively reduces the resolution of the feature maps by 2. This progressive reduction of spatial resolution and increase of the number of the feature maps compress the spatial information into learned features of large dimensionalities. While optimizing with LCZ labels that emphasize the composition of land surfaces, the learned features would find a balance between the spatial correlation and the land surface composition for the classification purpose. One skip

connection is applied within each residual block, which is the core characteristic of residual networks. A global average pooling layer is applied to the final 8 × 8 ×256 feature map, resulting in a 1 × 256 feature vector. This feature vector is further input to the following dense layer with a softmax activation function, outputting a 1 × 17 vector of softmax probability in the end. The soft max probability indicates the LCZ label of the input image patch. The output softmax probability is used for loss calculation together with the corresponding input reference label during the training phase.

3.2. Fusion of Sentinel-1 and Sentinel-2

There are two main advantages of fusing the classification results from Sentinel-1 and Sentinel-2 data. First, Sentinel-1 SAR images and Sentinel-2 multispectral images have complementary information. For example, the multi-spectral information of Sentinel-2 images provides spectral reflectance of covering materials, and the side-looking Sentinel-1 SAR data provides information of building structure and orientations. Second, Sentinel-1 data is not affected by weather and clouds. It provides more consistent classification results than Sentinel-2 data for large areas. We carried out the fusion at the decision level rather than the geometric or feature level, because of the different modality and geometry of Sentinel-1 and Sentinel-2 data. Decision level fusion averages the softmax probability of corresponding Sentinel-1 and Sentinel-2 data patches. The final prediction is the class with the highest mean probability.

3.3. Training strategy

In order to demonstrate a full spectrum of global performance of our deep learning model, we consider three training-testing splits: “randomsplit”, “block-split” and “cultural-10”. Among them, random-split indicates a spatially random sampling of evaluation points, which is commonly used in remote sensing. Since the distributions of training and test data are similar, random-split defines the upper bound of achievable classification accuracy. In contrast, block-split is a deterministic data split, i.e., the data from each city is separated into nonoverlapping east-west blocks and the accuracy is evaluated on unseen blocks. The block-split gives a representative measure of accuracy for unseen cities whose data distribution is similar to the training cities. Last but not least, cultural-10 defines the lower bound of the achievable accuracy by evaluating completely held-out data in a cross-validation scheme. We have selected 10 held-out cities, which come from 10 different culturally-defined geographical regions of the Earth. For each city, we trained the model with all other 41 cities and then evaluated on this held-out city. Finally results of the 10 cities were averaged. In this way, “cultural-10” provides a completely unbiased view of the accuracy achievable on absolutely unseen data. Together, these three evaluation scenarios give a solid impression of the quality of the global LCZ maps produced through this approach.

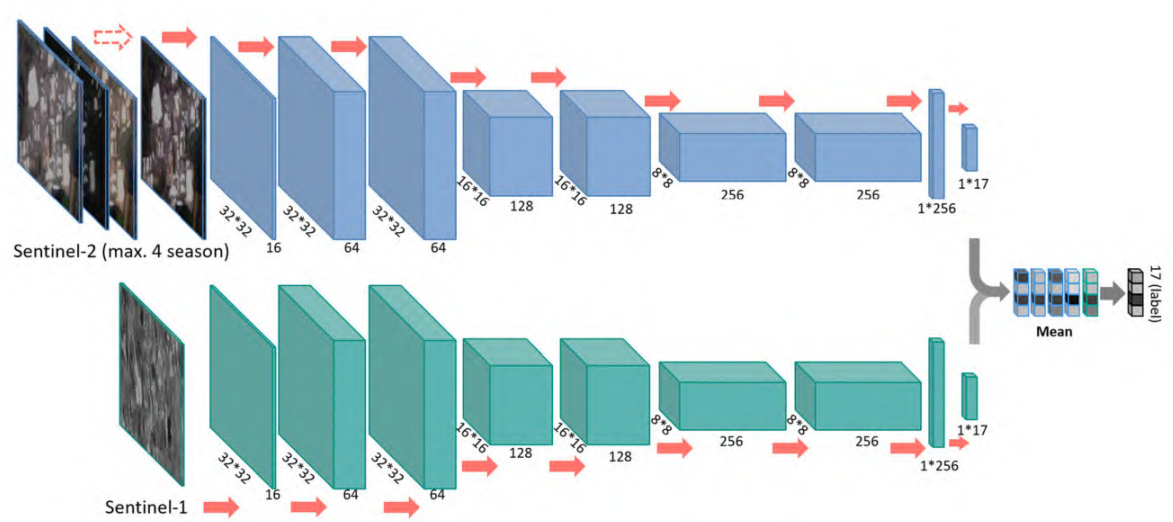


Fig. 3. The deep learning model architecture employed in this study. This is two streams of identical residual networks with 20 layers for Sentinel-1 and Sentinel-2 image patches respectively. The Sentinel-2 stream takes input from a maximum of 4 seasonal Sentinel-2 images of a city, whereas only one season was considered for Sentinel-1. The softmax layer of all the predictions are fused via averaging, yielding the final prediction.

3.4. Implementation details

The implementation of the network is under the framework of Keras with Tensorflow backend. For the training, the batch size was set to 32. We apply Adam (Kingma and Ba, 2014) as the optimizer with a learning rate of 1e-4. The number of training epochs was set to 300. Early stop was enabled with a patience of 30 epochs. The training was performed on NVIDIA Tesla P100 GPU.

4. Experimental results

4.1. Performance assessment

4.1.1. General assessment

We evaluated the performance of our deep learning model based on the three data splits mentioned in the previous section. Our deep learning model shows consistent and promising accuracies in the three scenarios, which are listed in Table 3. It shows the average accuracies of using Sentinel-1 image only, average of 4 seasonal Sentinel-2 images, and the fusion of Sentinel-1 and Sentinel-2 images. Firstly, we see that the three different evaluation strategies illustrate that our LCZ maps have an average accuracy between about 50% for cities that show characteristics disjunct from the training set, and more than 80% for cities that are fairly similar to the training distribution. This range of accuracies give us a realistic upper and lower bound of the performance of our model. Therefore, we are confident that the model will perform reasonably well on completely unseen cities. It should also be noted that average accuracy is a pessimistic metric compared to the more commonly used overall accuracy, as it is not weighted by the number of samples of a class. Secondly, it is obvious that 4 seasonal Sentinel-2 images (even the 1-season Sentinel-2 image, which are not listed)

consistently outperform the Sentinel-1 image. The reason is not only the averaging of the results from 4 seasons, but also the successful cloud removal performed on the Sentinel-2 images. Cloud-free optical images are more suitable for vision tasks such as LCZ classification. However, the importance of Sentinel-1 image shall not be downplayed, as cloudfree Sentinel-2 images are not always available. It is estimated that 60–70% of the global Sentinel-2 images are cloudy. Therefore, our algorithm includes the fusion of Sentinel-1 and Sentinel-2 results. Table 3 shows that the fusion of SAR and optical data consistently leads to the best results. Although, the improvement is not significant, the inclusion of Sentinel-1 image will play an important role in large scale or even global scale inference, in particular for cloudy areas such as western Africa and Amazon forest.

Table 3. Averaged accuracy of Sentinel-1 and Sentinel-2 data, and fused results, from three types of assessment: random-split, block-split, and held-out cultural-10 split.

	Random	Block	Culture-10
Sentinel-1	64.2%	51.6%	34.0%
Sentinel-2 (4 seasons)	82.2%	75.7%	51.0%
Sentinel-1 + Sentinel-2 (4 seasons)	83.4%	76.8%	51.3%

4.1.2. Per-class analysis

In order to get insight for future improvements, we studied the performance of our fused deep learning model in each LCZ class. The classification confusion matrix is shown in Fig. 4. We observed that certain LCZ classes are more easily confused in both Sentinel-1 and Sentinel-2 data than other classes. This can be seen both in the confusion matrices of our deep learning models as well as in the produced LCZ classification maps. Overall, we discovered three types of major confusions between the LCZ classes.

• Among the low-rise classes, including compact low-rise (class 3), large low-rise (class 6), and light-weight low-rise (class 7): About 45% lightweight low-rise (class 7) samples in our reference data were classified as compact low-rise (class 3). We also visually observe some confusion of light-weight low-rise with large low-rise in the predicted LCZ classification map of cities not in our reference data. These three classes have similar patterns of compactness. Their fundamental difference is the weight of their construction materials – information that is hard to retrieve from Sentinel-1 and Sentinel-2 images. Another reason for confusion is the limited amount of reference data in the lightweight low-rise class.

• Between heavy industry (class 10) and large low-rise (class 8): This is a reasonable confusion as large low buildings, such as factories and warehouses, are normally part of heavy industrial areas.

• Among bush, scrub (class 13), low plants (class 14), and bare soil or sand (class 16): The differences among these three classes mainly lie in the percentage of vegetation coverage, as well as on the vegetation species. Seasonal variation and the limited resolution of Sentinel-1 and Sentinel-2 data restrict their performance in discriminating above-mentioned differences.

Fig. 4. Confusion matrix of the Sentinel-1 and Sentinel-2 fusion model in the culture-10 split. Left: The value in each cell is the number of test samples. Right: the values are accuracies. A summary of the class-wise accuracy is also listed: on the right side is recall, on the bottom is precision. The classes are 1: compact high-rise, 2: compact mid-rise, 3: compact low-rise, 4: open high-rise, 5: open mid-rise, 6: open low-rise, 7: lightweight low-rise, 8: large low-rise, 9: sparsely built, 10: heavy industry, 11: dense trees, 12: scattered trees, 13: bush scrub, 14: low plants, 15: bare rock or paved, 16: bare soil sand, 17: water.

4.2. Global maps of urban morphology

4.2.1. Visual analysis

In total, we mapped 1.65 million square kilometers of urban areas of the 1692 cities. LCZ classification maps of those cities show that our deep learning model has a strong capability of classifying urban morphology. Figure 5 shows a few typical examples of cities with different morphological types in our study area. It can be seen that LCZ classification maps are capable of capturing different types of urban morpho logic compositions, e.g., many South American cities feature large compact high-rise areas, whereas European cities feature a mixture of open mid-rise and compact mid-rise areas.

Comparing to the state of the art, i.e., binary urban-nonurban maps, our classification provides fine grained intra-urban morphological variations. Figure 6 compares GUF (left) and our morphological map (right) of Delhi. Our morphological map makes it possible to localize the location of morphological variations within the settlement landscape, and further determine their exact morphological types. For example, the Delhi LCZ map shows there is no large compact high-rise district in Delhi as we see in Sao Paulo. The dense areas in Delhi are scattered across the city. We also observe large areas of large low-rise buildings, which could be related to informal settlements. Water resources such as rivers and lakes as well as vegetation are also clearly mapped. This is important information that can aid decision making on urban development.

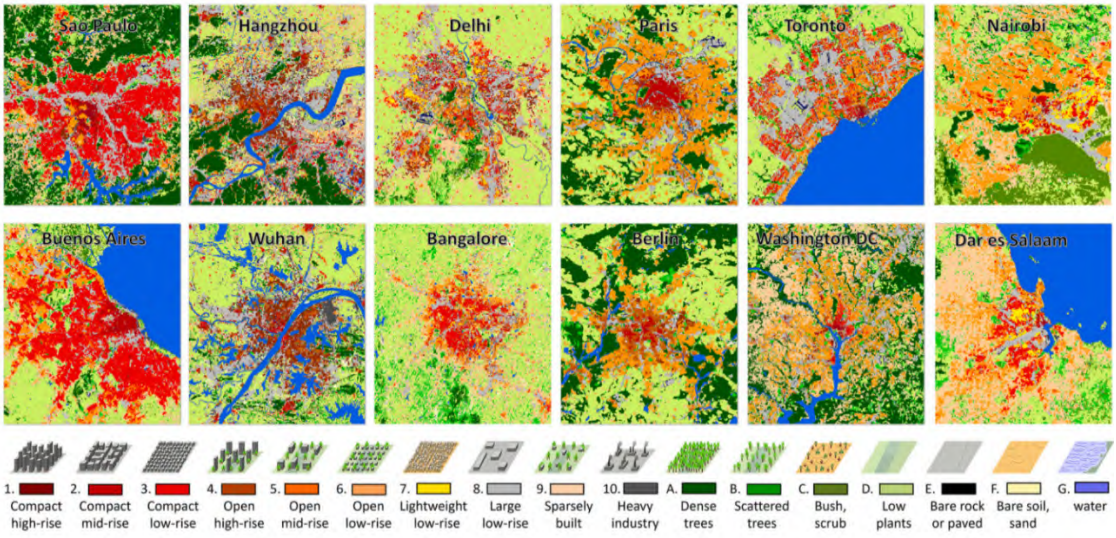


Fig. 5. Examples of the global urban morphological maps. We employed a deep learning model on Sentinel-1 and four-seasonal Sentinel-2 data to cover the largest 1692 urban agglomerations with population greater than 300,000 according to UN's World Urbanization Prospects 2014 (United Nations, 2014). A few examples of the produced LCZ maps are sorted here according to the morphological type of the urban area. From left to right, it shows a decrease of compactness and an increase of the mixture of different LCZ classes. Note the strong correlation between the morphological type and the geographic location. A more detailed analysis of the urban morphological pattern and culture can be found in the next section. At the bottom of the figure is the legend and schematic drawings of each LCZ class. Please note that the maps of a few cities in this figure are enlarged for better visualization. They are not true in proportion.

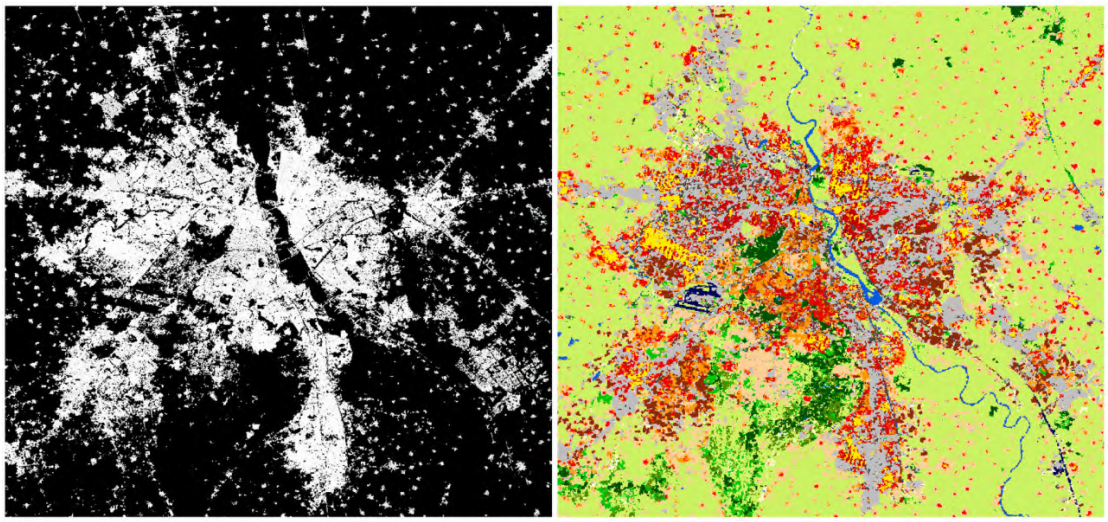


Fig. 6. Comparison of the state-of-the-art global urban map and LCZ classification. The left image is the urban and nonurban binary classification of Delhi from the global urban footprint. The right image is the corresponding LCZ classification.

Some LCZ classes also have strong correlation with certain settlement types, such as class 8 large low-rise vs. factory and warehouse, and class 7 lightweight low-rise vs. informal settlements. Therefore, LCZ maps could be employed to identify those areas. In Fig. 7, we listed a few examples of industrial areas identified in the LCZ maps. It shows the north of Munich, Germany, where the BMW factory and many historical factories are located, the

port area of Rotterdam, and the industrial area of Shenyang – a typical industrial city in China. This information provides a significant step in how we can perceive cities, with more ac curacy, to make better planning decisions in urban areas. We make those data available to the community.

4.2.2. Quantitative analysis

The LCZ classification maps of the largest 1692 cities provide the possibility to analyze the composition of global urban areas for the first time. In the 1.65 million km^2 we mapped, about 550 thousand km^2 are built-up areas, i.e. the LCZ class 1 to 10. Among these 10 classes, the sparsely built, large low-rise, open low-rise, and compact low-rise occupy the four largest proportions by area, as shown in Fig. 8 right. The largest built-up area is sparsely built, accounting for about 30% of the 1692 cities. More than 26% of the built-up area are large/lightweight low-rise buildings, often warehouses and informal human settlements. The open and compact classes occupy about another 25% and 14%, respectively, often taking the form of residential and commercial buildings.

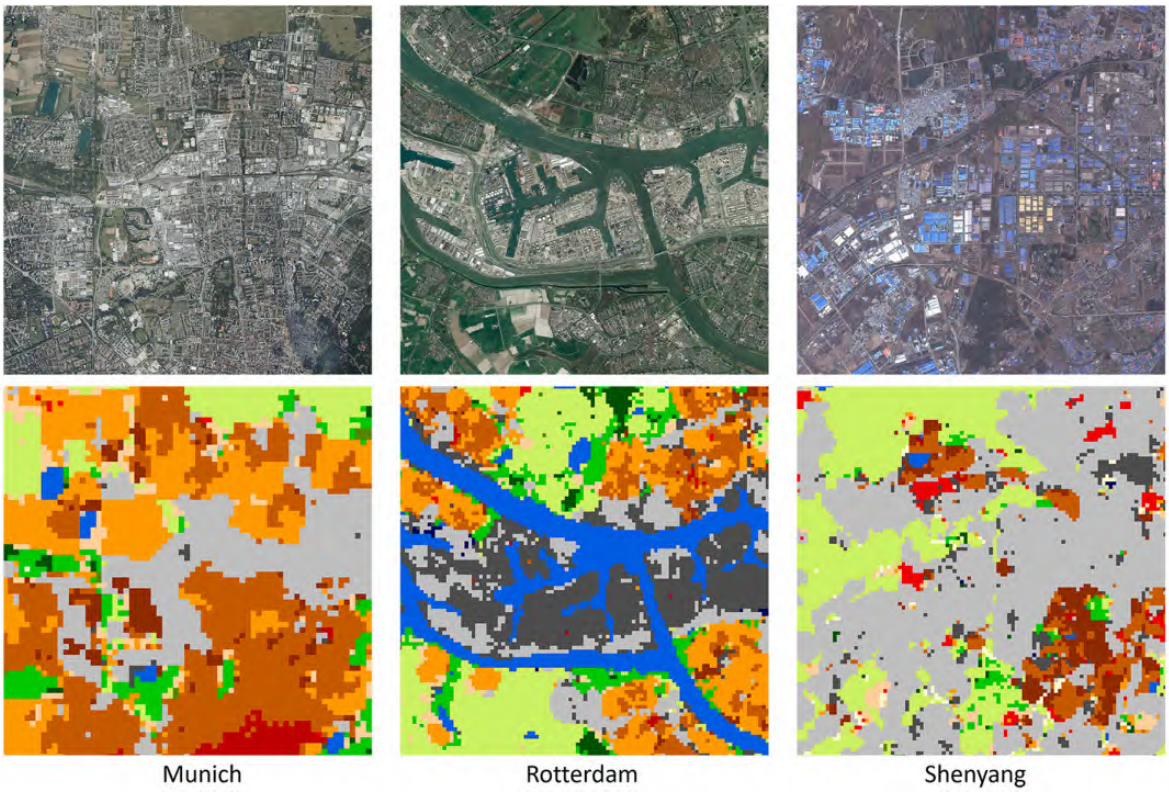


Fig. 7. Examples of industrial areas identified in LCZ maps. From left to right, it shows the north of Munich, Germany, where the BMW factory and many historical factories are located, the port area of Rotterdam, and the industrial area of Shenyang, a typical industrial city in China. The upper image is cited from Bing Map.

Alarmingly, using the population of Global Human Settlement Layer in year 2015 (Schiavina et al., 2021), we found that the 30% sparsely built areas only accommodates 10% of the total population of the 1692 cities. In contrast, more than 30% of the population lives in the 14% compact areas. The difference in population density in these two types of settlement is about seven-fold at a global scale. The population pro portion of each built up class can be seen in the left of Fig. 8. To see the full picture, we summarize the area and

population of the four aggre gated classes, i.e., compact, lightweight/large low-rise, open, and sparsely built, in Table 4. To summarize, 60% of the population live in compact and lightweight/large low-rise areas, occupying 40% of the built-up area. Another 35% of the population lives in open or sparsely built areas, occupying 55% of the areas. The remaining 5% of the population is distributed in the non-built-up LCZ classes. This huge difference in population density reflects a possible inequality in living conditions.

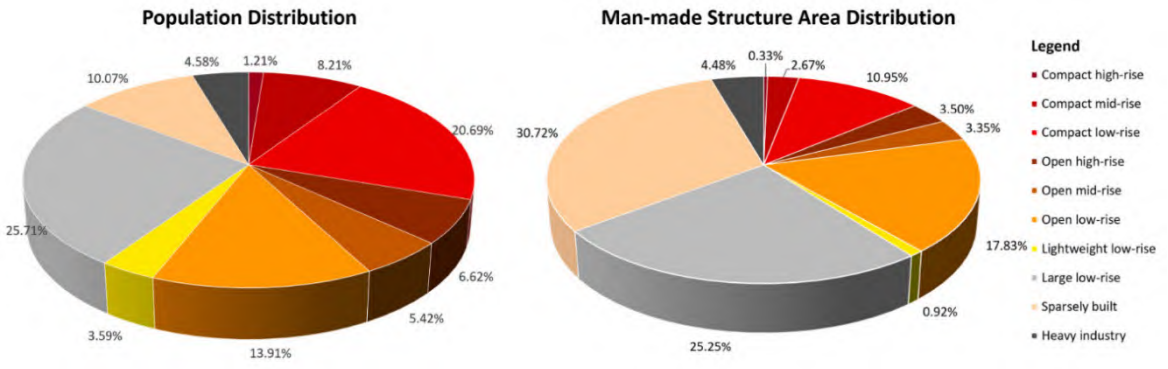


Fig. 8. The proportions of population (left) and area (right) of the ten LCZ built up classes (classes 1–10) in the 1692 cities.

Table 4. Area and population of human settlements of different compactness in the 1692 cities. From left to right, the compactness of the settlement decreases, and the population density increases.

	Compact	Lightweight / large low-rise	Open	Sparsely built
% in built up area	13.95%	26.17%	24.68%	30.72%
% of population	30.11%	29.30%	25.95%	10.07%
population per km^2	9122	4733	4445	1385

5. Pilot application: empirical study of LCZ patterns vs. culture

The new availability of intra-urban LCZ classifications describing the morphologic configuration of urban landscapes enables new empirical science on cities. It enables applications in a wide variety of thematic domains, such as urban heat islands, population estimation, or analyses directly related to building and open space structures and patterns. We illustrate this capability here, with one example, by presenting a morphologic categorization of city types.

Traditionally, urban structural models have been developed in urban geography for different geographical-cultural areas, such as for the Western or the Islamic world. These models have been based on theories and observations. Our LCZ classifications now allow us to test whether groups morphologically form according to geographical-cultural regions.

To test whether spatial patterns of LCZs in cities within the same geographic-cultural spaces are more similar in their quantitative appearance as well as in their spatial composition than compared to others, we conducted the following approach: a clustering based on

unsupervised methods (k-means (Hartigan and Wong, 1979) as well as expectation-maximization (Dempster et al., 1977) algorithms) is applied

on the quantity and spatial location of LCZs across all cities. The optimal number of clusters is determined by the gap statistics algorithm (Tibshirani et al., 2001). Various feature spaces are tested, combining the occurrence of the 17 LCZs in general and with respect to their locations in the city; the city size is integrated as a feature as well (Taubenböck et al., 2020). We rely on comparable spatial baselines of city extents based on a consistently applied morphologic delineation method (Taubenböck et al., 2019). As it is a priori unclear which feature space and which clustering approach fits best, we evaluate our resulting clusters in relation to geographical-cultural regions based on predefined regions, using those defined by Huntington (Huntington, 1997). Using the Simpson Evenness Index (SIEI) (Simpson, 1949), we define the clusters closest as the best result.

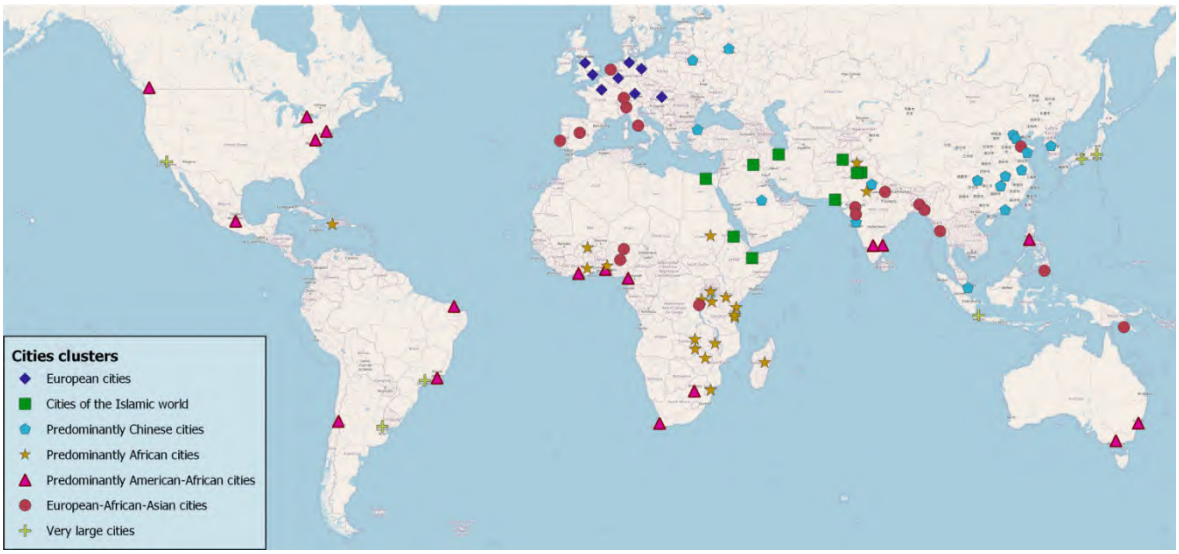


Fig. 9. Seven resulting clusters of morphological city pattern types. Findings are based on the k-means algorithm using an 18-dimensional feature space (consisting of 17 LCZs as well as city size as features) at the spatial unit of the morphological urban areas.

We find that the resulting clusters of cities of similar morphological patterns of the built environment do reflect geographic-cultural regions to a certain degree (Fig. 9). It can be seen that Central European cities form one cluster, and Islamic cities form another one. We also identify clusters that predominantly feature Chinese and African cities. However, we also find clusters that are more spatially complex, such as clusters containing predominantly cities in American and in African as well or European as well as African and Asian cities. One last cluster consists of very large cities, which obviously represent a specific pattern. In general, this confirms that certain aspects of geographic-cultural heritage have an influence on the spatial design of urban space. However, strict geographic-cultural models are not sufficient to describe urban settlement patterns: such models would negate the complexity of global interconnections in urban development.

6. Perspectives and discussions

The way we physically build our cities defines our life patterns, our mobility, the quality of life, the resilience of our cities, and much more. Across the globe, there are no two identical urban configurations and yet there are many groups of cities that are similar in their morphological characteristics. The measured differences in land consumption per person testify to the variability of cities worldwide and are an expression of inequality. A uniform and consistent classification is the basis for such new and better understandings of the status quo and thus also for improving urban planning in the future.

In times of global urbanization, it must therefore be of central interest to create such a data set for cross-disciplinary use. The global high resolution and consistent local climate zone classification map generated in this manner can be the basis for various applications: just a few examples include understanding the urban thermal environment to mitigate urban heat islands in human settlements, assessing population distributions, and simply analyzing urban morphological configurations. Our geostatistical result exemplifies the latter, demonstrating that patterns of urban morphology have a strong correlation to the geographical region and the related cultural heritage. However, we believe that extensive urban-related researches based on information extracted from this dataset can help stakeholders to develop appropriate policies and distribute resources more effectively. Of course, we are fully aware that the LCZ scheme is only one possible conceptual approach among many others to a systematic intra-urban description of urban morphology. To what extent these 17 LCZ classes can cover the urban landscape comprehensively and also in its manifestations worldwide must be investigated in future studies. However, we believe the LCZ scheme as a generic, culturally-neutral description of land-use and landcover (Stewart and Oke, 2012) relying on universal, standardized and measurable parameters of urban form allows for a consistent and sys tematic description of the built urban landscape.

Although the overall accuracy of the classification results we achieved in this work is high, the discrepancy of accuracy between different classes, particularly, height-related classes, is obvious. The complementary Sentinel-1 data can improve the result. However, the moderate resolution of Sentinel-1 limited significant improvement. We have investigated the potential of using TanDEM-X data, whose spatial resolution is much higher than Sentinel-1 data, for the same task. Our preliminary experiments show promising results, suggesting that higher resolution SAR data would significantly complement the lacking height information. Since the TanDEM-X bistatic data provides global coverage and is free from clouds, a higher resolution and better quality classification dataset could potentially be achieved in the future to enable more precise analysis in supporting different applications.

In general, the volume of open-access remote sensing data is increasing significantly. The proposed deep learning based multi-sensor fusion framework exemplifies the potential of developing advanced machine learning pipelines for extracting desired information from such a “gold mine.” Together with high performance computing (HPC), high-quality geo-information layers that currently do not exist can be efficiently generated at a global scale from peta bytes of EO data. This offers a great perspective toward filling the geo-information gap between knowledge and strategic directions, e.g., while addressing United Nations

Sustainable Development Goals, or monitoring mega trends such as climate change and global urbanization.

Funding

The work is jointly supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [ERC-2016-StG-714087], Acronym: *So2Sat* and grant agreement No. [957407], Acronym: *DAPHNE*, by the Helmholtz Association through the Framework of Helmholtz AI (grant number: ZT-I-PF-5-01) - Local Unit "Munich Unit @Aeronautics, Space and Transport (MASTr)", the Helmholtz Klimainitiative (HICAM) and Helmholtz Excellent Professorship "Data Science in Earth Observation - Big Data Fusion for Urban Research" (grant number: W2-W3-100)) and by the German Federal Ministry of Education and Research (BMBF) in the framework of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond" (grant number: 01DD20001).

Acknowledgements

The authors thank Nikolai Skuppin for supporting the organization of data set and code repository.

Chapter XIII

A new ranking of the world's largest cities

Do administrative units obscure morphological realities?

Remote Sensing of Environment, 2019, vol. 232, 111353

Hannes Taubenböck, Matthias Weigand, Thomas Esch, Jeroen Staab, Michael Wurm, Johannes Mast & Stefan Dech

© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.rse.2019.111353

Received 25 Feb 2019; Accepted 29 July 2019; Published 2 Aug 2019

Abstract

With 37 million inhabitants, Tokyo is the world's largest city in UN statistics. With this work we call this ranking into question. Usually, global city rankings are based on nationally collected population figures, which rely on administrative units. Sprawling urban growth, however, leads to morphological city extents that may surpass conventional administrative units. In order to detect spatial discrepancies between the physical and the administrative city, we present a methodology for delimiting Morphological Urban Areas (MUAs). We understand MUAs as a territorially contiguous settlement area that can be distinguished from low-density peripheral and rural hinterlands. We design a settlement index composed of three indicators (settlement area, settlement area proportion and density within the settlements) describing a gradient of built-up density from the urban center to the periphery applying a sectoral monocentric city model. We assume that the urban-rural transition can be defined along this gradient. With it, we re-territorialize the conventional administrative units. Our data basis are recent mapping products derived from multi-sensoral Earth observation (EO) data – namely the Global Urban Footprint (GUF) and the GUF Density (GUF-DenS) – providing globally consistent knowledge about settlement locations and densities. For the re-territorialized MUAs we calculate population numbers using WorldPop data. Overall, we cover the 1692 cities with > 300,000 inhabitants on our planet. In our results we compare the consistently re-territorialized MUAs and the administrative units as well as their related population figures. We find the MUA in the Pearl River Delta the largest morphologically contiguous urban agglomeration in the world with a calculated population of 42.6 million. Tokyo, in this new list ranked number 2, loses its top position. In rank-size distributions we present the resulting deviations from previous city rankings. Although many MUAs outperform administrative units by area, we find that, contrary to what we assumed, in most cases MUAs are considerably smaller than administrative units. Only in Europe we find MUAs largely outweighing administrative units in extent.

Keywords

City size; urban agglomeration; rank-size distribution; remote sensing; global urban footprint; urban morphology

1. Introduction

"Tokyo is the world's largest city with an agglomeration of 37 million inhabitants, followed by Delhi with 29 million, Shanghai with 26 million, and Mexico City and São Paulo, each with around 22 million inhabitants. Today, Cairo, Mumbai, Beijing and Dhaka all have close to 20 million inhabitants" (UN, 2018). This quote from the 2018 World Urbanization Prospects publication is a clear statement about the ranking of the largest cities in the world. This paper, however, calls this ranking into question.

The UN ranking is based on population numbers relying on administrative space units. Naturally, these spatial units are crucial as they determine a clear-cut city's boundary to establish jurisdictional competence of its municipal government (Parr, 2007). In a world, however, which is experiencing highly dynamic transformation processes through global urbanization (e.g. Angel et al., 2011; Taubenböck et al., 2012), conventional, administrative spatial units represent reality less and less. Scholars describe a continuous transformation from formerly compact, concentrated land use patterns into spatially extended and morphologically less clear-cut urban configurations (e.g. Anas et al., 1998; Batty et al., 2004; Siedentop and Fina, 2010; Small et al., 2011; Taubenböck et al., 2019). Terms such as 'peri-urban', 'urban fringe', 'leapfrog development', 'suburb', among many others describe a complex, indistinct, irregular, and dynamic transitional zone which is more an urban-rural continuum than a clear-cut boundary.

The implications, however, of the common administrative space units are critical: Beyond the political control competence, the question of city size is of fundamental statistical significance. Size is of relevance with respect to the ranking and hierarchical ordering of cities (Auerbach, 1913; Zipf, 1941; Eeckhout, 2004), to the relation to such variables as per capita income, unemployment rate, inequality, among many others (Parr, 2007) or to aggregated numbers such as total population or gross domestic product (UN, 2018; Openshaw, 1983). Without a comparable and meaningful unit of space, these numbers lead to distorted patterns of explanation, e.g. in multivariate analysis.

Let us take Tokyo as an example: The city of Tokyo has, according to official figures, 9.5 million inhabitants on an administrative area of 622 km^2. These numbers would rank Tokyo at number 31 in the list of the largest cities in the world (UN, 2018). The metropolitan region (more or less the continuous built urban area), however, is listed with 37 million inhabitants at 13,572 km^2, making it officially the largest city in the world. In global statistics, the latter number is commonly used and compared to other cities such as Guangzhou in China, which is reflected in this statistic with 12 million inhabitants, ranking it at 23. However, the continuous built urban area in Guangzhou is, as it has been shown by scholars, significantly larger (e.g. Florida et al., 2008; Taubenböck et al., 2014). Neighboring cities such as Foshan, Dongguan or Shenzhen have physically merged together with Guangzhou but make the list by themselves. So, Guangzhou is treated individually here. Surprisingly, these inconsistencies are largely hidden in the statistics; they are tacitly accepted and barely questioned.

In consequence, we argue a fundamental challenge for any empirical investigation is to define comparable spatial units in order to provide a realistic statistics of city sizes or above

mentioned downstream statistical indicators. Critics claim that much of existing urban research studies can be contested due to a data-driven approach and a rather arbitrary use of geographical boundaries (e.g. Lechner et al., 2013; Riitters et al., 1995; Berry & Okulicz-Kozaryn, 2012; Masucci et al., 2015; Taubenböck et al., 2016). In a non-academic sense, the size of the city contributes to its global perception. The opening quotation *"Tokyo is the world's largest city"* assigns this city to a certain global city network and implies, triggers, attracts and advertises economic and political spin-offs in terms of, for example, investments by international companies, attraction to high-skilled labor force or presence of a broad skilled labor force (e.g. Weber, 2001). Rankings have become a fundamental element in the contest of globalization.

In this paper we question whether and to what extent today's morphological urban areas (MUAs) differ from conventional administrative spatial units. To do this, we derive MUAs from globally consistent geodata on settlement structures derived from remote sensing data. In other words, we investigate whether the current ranking of city sizes based on the United Nations' statistics are strongly influenced by the artificial spatial units of administrative boundaries compared to the built-up extent of cities. But perhaps more importantly, we propose a new ranking of the largest cities in the world based on consistently calculated MUAs and related population figures.

The remainder of this paper is organized as follows. Section 2 provides more background and introduces the conceptual foundation. In Section 3 we introduce the data used and the developed methodology. In Section 4, we present the empirical results for 1692 cities across the globe, which are listed larger than 300,000 inhabitants in 2016. In Section 5, we critically discuss the capabilities and limitations of data and methods and foremost the implications of these findings, and in Section 6, we conclude the study.

2. Background and conceptualization

What defines the size of a city? Is it the administrative unit, the functional urban region, the physical extent of the built environment, the total population, the volume of trade, the economic power, the interlacing space, the functional move-in area, the perception of people, or other related factors?

These different concepts to approach the size of the city show, as Jessop et al. (2008) remark, that a city, metropolitan area or region can be imagined and constructed in varying ways with different methods and indicators for region-building: e.g., from tightly sealed areas with a territorially-embedded thinking of regions, to porous nodes in a networked space of flows (Castells, 1999; Harrison & Growe, 2014). As Castells (2000) discusses, space consists in the conflict between the '*space of flows*' and the '*space of place*'. The '*space of flows*' is defined as regular flows of people, goods, or information between separate but networked locations, while the '*space of place*' refers to the physical boundaries of locations. In consequence, there is neither an "all-purpose definition", nor is there a universal truth for the size of the city. Definitions and delimitations will differ depending on the specific purpose, the data and the methodologies used.

A fundamental basis for assessing the size of a city is to define what is meant by '*urban*'. In the scientific discourse on this subject no standard and internationally accepted definition of '*urban*' or '*urban population*' has yet prevailed. Each country uses its own definition, and collects data accordingly. The statistic that currently 55% of the world's population is urban dwellers (UN, 2018) relies on adding up these figures coming from often incomparable definitions, which are even based on arbitrary and thus inconsistent spatial admin units (Deuskar, 2015). Research studies delimiting urban space rely on various methods and geodata such as street networks (Masucci et al., 2015), remote sensing (e.g. Liu et al., 2016; Esch et al., 2014), demographic (Rozenfeld et al., 2011), among other data. Multi-criteria attempts using indicator combinations of minimum population size and density, travel times to the central place, among others are suggested and in use (e.g. World Development Report, 2009; Abed & Kaysi, 2003; Dijkstra & Poelman, 2014). Georg et al. (2018) take account of this by delimiting urban space using remote sensing, infrastructure, population and economic data constructing several possible spatial forms illuminating the fuzziness of the delimitation of territorial space. The definition and delimitation of the urban and the related urban-rural transition is subject to many years of scientific discussion, without, however, arriving at a uniformly accepted approach (Simon, 2008). Ross (2011) notes that boundaries are malleable as the transition is complex, indistinct, irregular, and dynamic and they should include some capacity for flexibility depending on the purposes. In spite of the large body of literature, Masucci et al. (2015) remark, that the very concept of cities remains obscure, hidden or assumed.

In consequence, any classifications of city sizes are not innocent. The concept, criteria and methods used might change details, and thresholds applied might be manipulated with often undocumented effect. In recommendations by the United Nations it is even formulated that due to different characteristics and understandings of 'urban' and 'rural' across the globe, a global definition is not possible (Dijkstra & Poelman, 2014). In this paper, however, we argue that, even though local forms of settlement, flows of goods, and life vary widely, a consistent methodological approach generating statistically comparable spatial units based on one consistent data source brings to light novel perspectives and facts about city sizes. Against this background we understand 'urban' in this work in reference to Schneider et al. (2009) and Taubenböck et al. (2012) as places dominated by the built environment. The built environment includes human-construct elements, roads, buildings, or the like. Non-built land (e.g. vegetation, bare soil) is not considered urban, even though they may function as urban space.

We are building on this definition because we approach the delimitation of 'urban' using a globally consistent data source of satellite-based earth observation (EO). This data has developed to a tool providing area-wide spatial information on the location, spatial distribution and characteristics of settlement features at global scale with high geometric resolution (e.g. Esch et al., 2012; Pesaresi et al., 2013) and high accuracies (Esch et al., 2018a, b; Klotz et al., 2016; Taubenböck et al., 2011). We rely on the Global Urban Footprint suite (GUF) (Esch et al., 2018a, b) capturing physical attributes of settlements allowing us to focus on urban form. Building on this, we construct territorially bounded city areas based

on similar characteristics of the built city. Thus, we use the '*space of place*' logic to construct physical boundaries of cities.

While many EO-based studies approached the global urban extent (e.g. Schneider et al., 2009; Elvidge et al., 2007; Pesaresi et al., 2013; Esch et al., 2012), only few studies, mostly relying on night light emission and/or gridded population data have approached the delimitation of city extents on a global scale. In these studies, the power law scaling of rank-size distributions is documented (Decker et al., 2007; Small et al., 2011; Small & Sousa, 2016) and consistency is confirmed over time (Small et al., 2018). Small et al. (2011) found spatial networks of urban development are vastly larger than the administratively defined cities they contain. At higher resolutions, but only for a comparatively small subset of the world's cities, Fragkias & Seto (2009) revealed oscillations of rank-size distributions over time due to the coalescence of settlements across administrative city boundaries. This scarcity of studies at global scales originates from the conceptual complexity for a consistent and comparable delimitation of city boundaries, any widely accepted definition of what constitutes a coherent urban area and the necessary globally consistent data set.

So, what does this imply regarding the delimitation of cities to determine their sizes? Unlike the above-mentioned studies, which used night-time light emissions or population data as proxy for settlement extents, this study is based on another proxy: the empirical relationship between spatial variability of X-band backscatter and assumed structure of built morphology mapped at a comparatively high geometric resolution. We assume that the urban-rural transition can be defined by the slope of the gradient of physical density of built-up structures towards the periphery. Although reality is often more like a fuzzy transition in the urban-rural continuum, we aim at a clear-cut, but compared to conventional administrative units, consistently derived boundary. In this way we try to spatially designate the built city and to re-territorialize the existing measure of the administrative spatial units more sensibly. Based on these new re-territorialized urban entities, we use WorldPop data (Tatem, 2017) for calculating related population figures. The effects of potential changes of re-territorialized city sizes (in extent and population) compared to administrative spatial units (in extent and population) are represented by rank size distributions and related statistics.

3. Data and methodology

3.1. Data

Our analysis bases on data from *remote sensing* (Esch et al., 2012; Esch et al., 2018a), geodata from *OpenStreetMap* (OSM, 2018), the *Global Administrative Areas* dataset (GADM, 2018), data from the United Nations' *populations statistics* (UN, 2015), and the global population dataset from *WorldPop* (Tatem, 2017).

The two remote sensing-based mapping products are from the Global Urban Footprint suite: The *Global Urban Footprint* (GUF) and the *GUF Density* (GUF-DenS) classification. The GUF is a binary raster layer presenting the worldwide distribution of human settlements in a so far unique spatial resolution of 12 m (Esch et al., 2012; Fig. 1b). Using the operational Urban Footprint Processor framework described by Esch et al. (2013), the vertical built-up

structures in urban and rural environments are mapped based on a global analysis of > 180,000 TerraSAR-X and TanDEM-X StripMap radar images collected between 2011 and 2014 (93% of images recorded in 2011–2012, 7% in 2013–2014). Spatial complexity of varying objects within small areas is characteristic for urban areas. This is represented in highly textured image regions of strong directional, non-Gaussian backscatter due to double bounce effects in radar data. This information is used in combination with the intensity information to delineate 'settlements' from 'non-settlements' using an unsupervised image analysis technique. The experimental GUF-DenS layer is a semantically enhanced version of the GUF that provides information on the built-up density or – as an inverse, the greenness – in form of the percentage of impervious surfaces within the area assigned as settlement in the GUF (Esch et al., 2018a). The GUFDenS results from a combination of the binary GUF (which is used as a mask) and a TimeScan dataset derived from > 450,000 Landsat images (Esch et al., 2018b). The TimeScan layer is used to actually model the imperviousness/greenness via the temporal characteristics of the Normalized Difference Vegetation Index (NDVI) (e.g., mean, max, min, stdev) over a 3-year period. Since the NDVI is only calculated for the "built-up" regions within the GUF, the potential uncertainties due to mix-ups with bare soil/sand (e.g. in semi-arid regions) are significantly reduced. Assuming a strong inverse relation between vegetated and impervious surfaces, the intensity of vegetation cover defined by the NDVI can be used as a proxy for the percent impervious surface following an approach introduced by Esch et al. (2009). The resulting GUF-DenS is a raster layer in 30 m spatial resolution that shows values between 0 and 100 as an indication of the percentage of impervious surfaces per grid cell (Fig. 1b).

We use *administrative units* for each city provided by the Global Administrative Areas (GADM) database (GADM, 2018). However, there is no internationally agreed definition of metropolitan areas, which are the spatial basis for the World Urbanization Prospects statistics (UN, 2018). The GADM database contains administrative units at different levels: from the national to the municipal or even to the district level. The supposed spatial units of the population figures of the World Urbanization Prospects statistics, however, are not simply mapped in the GADM database. To achieve congruent spatial units, we adjust administrative units to metropolitan areas where the World Urbanization Prospects statistics refer to this spatial entity (cf. example of Tokyo in the Introduction). Beyond, for application of our monocentric sectoral city model (cf. Section 3.2.), we need to define the location of the city center. Here, we rely on the spatial definition of the central points for each city as provided by the United Nations (2014) (Fig. 1b).

For our comparative spatial city analysis, we exclude all areas containing water bodies such as oceans, rivers and lakes as they represent non-buildable areas. To do so, we use the water bodies' dataset from the OpenStreetMap project (OSM, 2018).

For the selection of the cities under investigation we rely on population statistics from the United Nations (2015). We perform the analysis for all cities on our planet with > 300,000 inhabitants, i.e. a total of 1692 cities are included in this study.

For the calculation of population numbers for the newly re-territorialized cities we rely on *WorldPop data* (Tatem, 2017). WorldPop data are gridded population counts at spatial scales finer than the administrative unit level of census data using a suite of geospatial layers (e.g.

Sorichetta et al., 2015). The approach relies on a random forest-based dasymetrically disaggregation of the census counts from administrative units into grid cells (Stevens et al., 2015).

3.2. Spatial delimitation of morphological urban areas (MUAs) from rural environments

Where do patterns of settlements change from urban to rural? There can be no simple answer to this rather philosophical question. To address this question, however, we use a conceptual framework that is tailored to our available global data. We assume that the urban-rural transition is represented somewhere along a decreasing gradient of physical built-up density with rising distance to the urban center. Thus, we measure the built-up urban extent of cities in their physical form, as opposed to metropolitan areas, which are defined in their economic or functional form.

For measurement and mathematical representation of the gradient of built-up density from the center to the periphery we choose a monocentric, sectoral city model as spatial entity. The monocentricity assumed here is, of course, a simplification of the urban spatial patterns. This traditional, monocentric model, known as the "Alonso-MillsMuth" model, refers to the equilibrium distribution of land use within a city as a function of land rents, which decreases with distance from the center (Paulson, 2012). Due to its empirical traceability and a proven explanation of a substantial portion (at least 80%) of the variation in urbanized land area across cities (e.g. Spivey, 2008), it continues to serve as a theoretical and empirical core concept of urban land development (e.g. McMillen, 2006). Beyond, it can also be conceptually reconciled with our assumption of an urban-rural transition along a decreasing gradient of built-up density with rising distance to the center. We use ring models with commonly used bandwidths of 1 km to determine the location relative to the defined urban center point. As the urban-rural gradient may vary around the city and urban or rural islands may occur within the cityscape (Simon, 2008), we additionally subdivide the urban area into 16 corridors (sectors) to take account of the local spatial context. Each sector opens radially outwards at 22.5°, starting at the defined city center (cf. Fig. 1a)

For the measurement of the urban-rural transition, we suggest a morphological settlement index (MSI) consisting of three indicators for defining the Morphological Urban Area (MUA): 1) The settlement area, 2) the settlement area proportion, and 3) the density within the settlements (Fig. 1d).

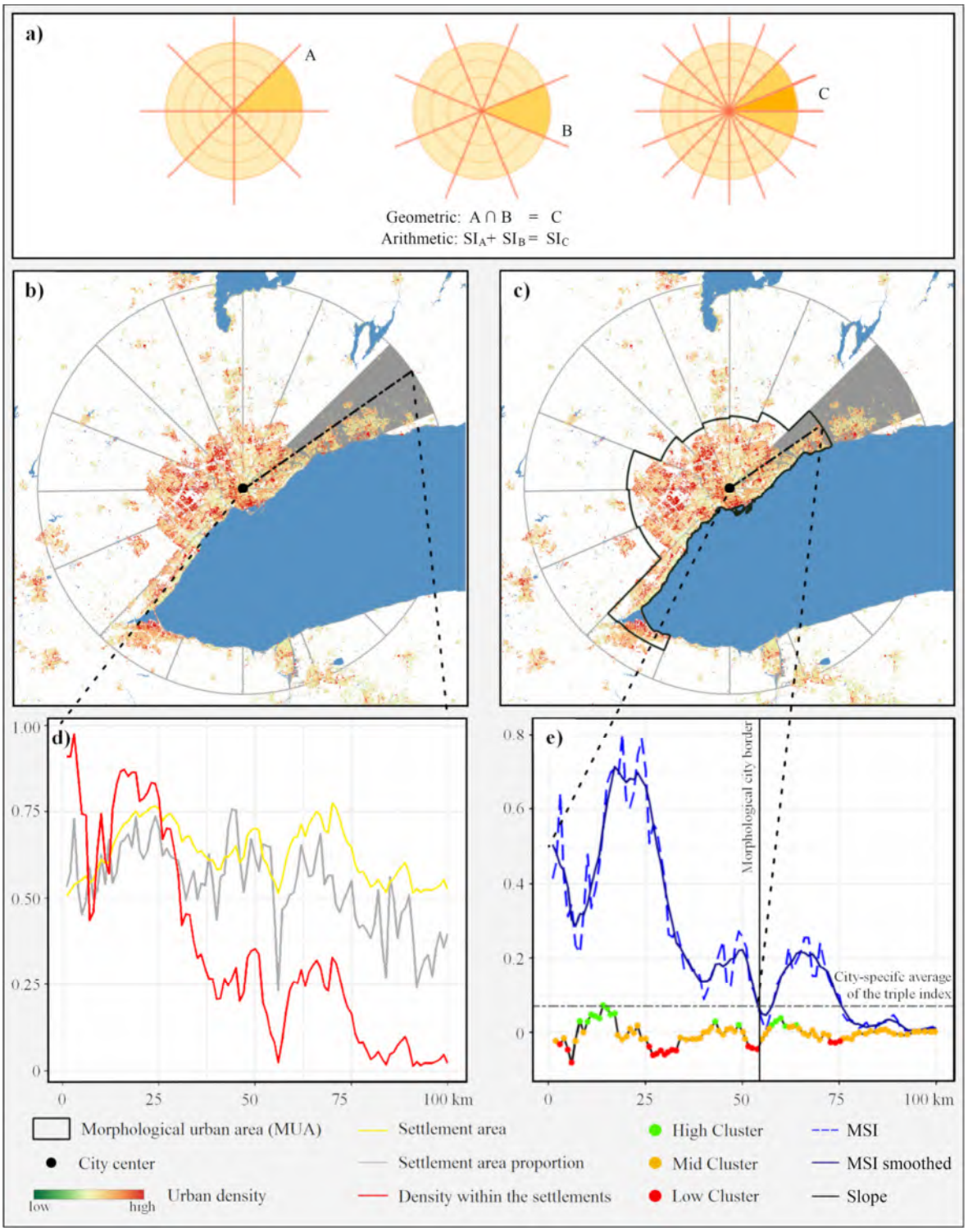


Fig. 1. Workflow for delimiting MUAs: **a)** The monocentric, sectoral city model with the two 8 sector models and the 16 corridors derived from it as well as bandwidths of 1km form the spatial entities for analysis; **b)** The sectoral, monocentric model projected around the defined city center, onto the settlement area from the GUF and the calculated densities of the GUF-DenS classification within the settlements – the example pictures the city of Toronto, Canada; **c)** The delimitation of the MUA based on the morphological settlement index (MSI); **d)** The gradients of all three indicators –settlement area, settlement area proportion and density within settlements– for a sample sector from center to periphery; **e)** The MSI and the slope allowing to classify the three clusters. The first low point of the low cluster below the MSI average of the low cluster is used as cut-off value.

1) The sectors from the city center to the periphery contain a certain settlement area. Along these sectors, the available area per ring, i. e. per 1 km bandwidth, increases with rising distance to the center point and so does the potentially available settlement area. We assume that as long as the absolute settlement area per ring is increasing towards the periphery, the sector still belongs to the MUA of the city.

2) The settlement area proportion corresponds to the extents of the settlement areas per respective ring per sector. We assume that a decreasing density hints at a transition area from urban to rural.

3) The settlement area proportion may in some cases be equal at two different locations; however, the built-up density within the settlement areas might differ. We assume decreasing densities hint at a transition towards rural environments. Here we calculate the mean builtup density within the respective settlement area per ring and sector.

The Pearson product-moment correlation between the three indicators is low (between settlement area and settlement area proportion r = 0.14; between settlement area and built-up density within the settlement areas r = -0.12; and between built-up density and settlement area proportion within the settlement areas r = 0.26); therefore, these indicators are permissible for our model. As it has been shown, we formulate a hypothesis per indicator which points to the urban-rural transition. However, a simple decrease in settlement area proportion from one ring to another – to take one example – is obviously given in most cases and, thus, not meaningful for finding an urban-rural boundary. In consequence, we combine all three normalized indicators to a MSI (Fig. 1e). We assume if all three indicators in combination suggest a decreasing gradient the probability for an urban-rural transition rises.

We calculate the MSI per sector (✶) and per respective ring area to display the density gradients. For the derivation of the MSI, however, we would also like to add the larger urban context to each sectoral gradient. The following train of thought is the basis for it: if a low dense island occurs within one 22.5° sector embedded in highly dense urbanized areas in the surrounding sectors, the localization of a morphological city boundary should be less likely than if a low dense island in the central sector is flanked by undeveloped open spaces in the surrounding sectors. To account for this consideration, we integrate MSI values of neighboring sectors for the derivations of the final MSI value. To do so, we calculate the MSI values for two separate (+ and ×) sectoral models consisting of 8 sectors each (see Fig. 1a). Since both sectors are rotated by 22.5° (cf. in Fig. 1a the sectors 'A' and 'B' illustrate this), their geometric superposition results in the sector 'C' (cf. Fig. 1a). The final MSI value for sector 'C' is a function of both, the MSI values of sector 'A' and 'B' by addition. Thereby, the central sector 'C' is given a higher weight, since its density values contribute to both sectors ('A' and 'B').

For a reasonable classification of the MUA that only defines a city boundary if there is a real morphological change towards low dense, rural structures, we calculate three clusters based on the distribution of MSI values. We establish clusters by initializing MSI values using a k-means approach. The three clusters are the high cluster indicating a rising morphological density gradient, the mid cluster indicating a generally constant density gradient, and the low cluster indicating a decreasing density gradient. The high and mid clusters testify to a

continuous settlement area. The low cluster, however, suggests decreasing morphological density. In Fig. 1d the three morphologic indicators are illustrated along the rising distance from the city center. In Fig. 1e the MSI and the related increase or decrease of the gradient is visualized. Also indicated are the three different clusters, with the low cluster (in red dots) suggesting potential candidates for urban-rural cutoff values. For delimiting the MUA, we define the cut-off value as the first dot towards the periphery belonging to the low cluster which features a MSI below the particular city average of all low clusters. This in turn, allows that –as we argued above– urban or rural islands can occur within the cityscape. As an example, our approach allows a park area (i.e. very low built-up density) with a continuous built urban landscape before and after the open area to be classified as part of the city as long as the identified low cluster is above the city average of the MSI at all cut-off candidates. Compared to other studies, which are working with minimum fixed distances between settlements to be still considered continuously urban (e.g. 200 m in the study of Weber, 2001), our approach is independent from such strict thresholds.

Many urban regions across the globe have experienced a coalescence of multiple, once morphologically separate cities, but remain jurisdictionally separated (e.g. Taubenböck & Wiesner, 2015). Although we chose a monocentric approach based on the city centers of the still jurisdictionally separated cities, we aim to capture morphologically merged cities as one MUA. If MUAs from two (or more) neighboring cities overlap, we combine the MUAs from both (or more) cities into one. We then count them as one urban agglomeration in our statistics and attribute this 'city' as a metropolitan region (M.R.).

3.3. Spatial statistics for analyzing the effects of re-territorializing cities by MUAs

We present the effects of changes in city extents by our morphologic approach compared to the commonly used administrative spatial units at two spatial entities: at *individual city level*, and at *continental level*.

At *individual city level*, we compare all 1692 cities across the globe based on the rank in the rank-size distribution. Rank-size distributions show the relation between rank $\tilde{N} = \tilde{N}(S)$ (Auerbach, 1913). The rank indicates the position of a city if you arrange the cities according to their size. Scholars observed that the rank is proportional to the reciprocal size of the city (e.g. Auerbach, 1913; Nitsch, 2005; Soo, 2005).

At *continental level*, we analyze trends between administrative space units and our re-territorialized MUAs. To do so, we propose the Normalized Difference Area Index (NDAI). It is calculated as follows:

$$NDAI = \frac{MUA - AUA}{MUA + AUA}$$

where MUA is the morphological urban area and AUA is the administrative urban area. The NDAI ranges from -1 to 1. Negative values indicate a significantly larger AUA in reference to the MUA. Positive NDAI values indicate that the MUA exceeds the AUA. The NDAI allows projecting all cities onto a world map to capture regional differences. In addition, we aggregate the results of all cities belonging to a continent using boxplots for comparison.

3.4. Accuracy assessment of input data

The validity of our intended results depends on the accuracy of the input data: For the GUF and the WorldPop data we do not perform an accuracy assessment here, but rely on measures presented in other studies (Stevens et al., 2015; Taubenböck et al., 2011; Klotz et al., 2016). For the experimental GUF-DenS layer, we perform an accuracy assessment. We relate the density values to aggregates of built-up densities based on independent, highly resolved geodata: We apply OSM data (OSM, 2018) for Paris, France and New York City, Dallas, and Las Vegas, USA. We use building footprints alongside features for rails, roads and industry. For the city of Munich and the entire state of Bavaria we substitute the building footprints from OSM data by Level-of-Detail 1 building models provided by the German Federal Agency for Cartography and Geodesy (BKG) (www.bkg.de). From a geographic point of view, we chose the latter example to encounter for a landscape consisting of highly dense urbanized areas and low dense, rural environments. These samples are picked as for them a complete set of built-up features is available. Using these data bases, we calculate built-up density for four different scenarios: (1) buildings, (2) buildings and roads, (3) buildings, roads and railways, (4) buildings, railways, roads and industrial facilities. For these scenarios, the built-up density is aggregated to square kilometers. The GUF-DenS product is then compared against all four scenarios, to identify the most corresponding thematic relation and to provide a validation of the input data set.

4. Results

4.1. Mapping results: morphological urban areas and administrative urban areas

The proposed approach allows a demarcation of the main morphological urban entity from low dense peripheral and rural areas in a globally consistent manner. In Fig. 2 cartographic results of the derived MUAs are projected onto the input data, i.e. onto the GUF classification and the density values of the GUF-DenS layer within the settlement areas. Administrative boundaries and defined urban center points derived from data from the United Nations (2014) are visualized as well. The following cases have been identified:

a) The resulting MUAs related to different urban center points spatially overlap and, based on our methodology, are merged to one polycentric spatial entity (in our abbreviation M.R.) consisting of more original cities by UN definition (Fig. 2a).

b) The resulting MUA is smaller than the administrative unit (Fig. 2b).

c) The resulting MUA is larger than the administrative unit (Fig. 2c).

d) The resulting MUA and the administrative unit are comparatively similar and can be considered a 'true-bounded' city (Fig. 2d).

e) The resulting MUA and the administrative unit are comparatively similar in size; however, the spatial extents are not congruent (Fig. 2e).

The newly calculated re-territorialized spatial boundaries of MUAs are available for download in vector format in supplementary A-1.

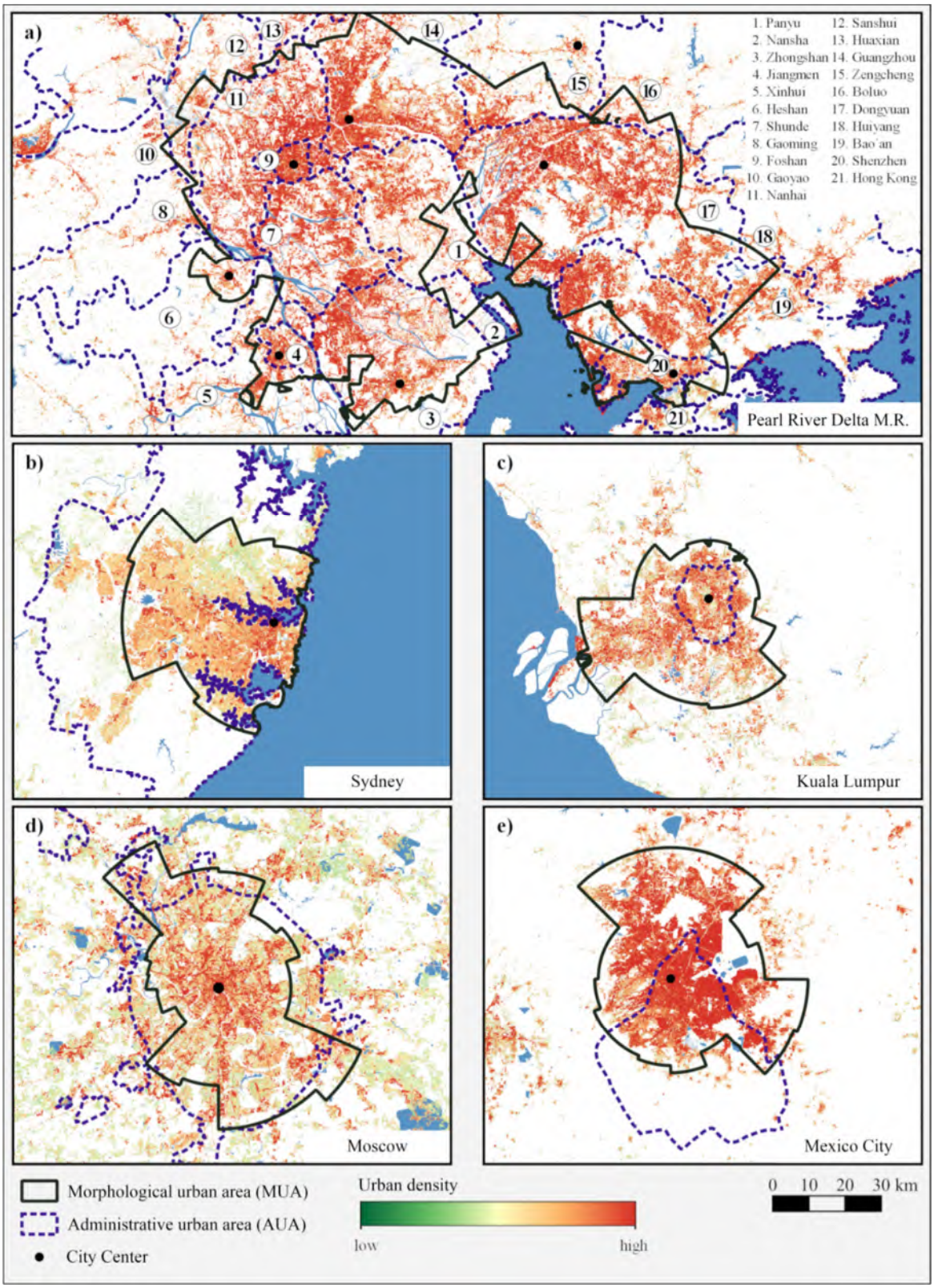


Fig. 2. a) The largest city in the world by areal size and population: the Pearl River Delta (PRD) Metropolitan Region (M.R.) in China and the many individual administrative units of cities commonly used for statistics; b) Sydney, Australia resulting in a smaller MUA than the administrative unit; c) Kuala Lumpur, Malaysia resulting in a larger MUA than the administrative unit; d) Moscow, Russia with a comparatively small discrepancy between MUA and administrative unit, and e) Mexico City, Mexico resulting in a comparable MUA and administrative unit; however, spatially both units are displaced.

The mapping results rely on multi-sensoral remote sensing data and do not provide accuracies of cadastral data. However, the studies of Klotz et al. (2016), Taubenböck et al. (2011) and Mück et al. (2017) clearly show the improvement of map accuracy of the GUF layer over other global urban mapping products. Especially for cities, areas characterized by high settlement densities, the GUF features high accuracies of about 90%. In turn, we assume this input data set provides a reliable basis. Also, the GUF-DenS values show high agreement with very high-resolution reference data from OSM or from BKG. At an aggregated grid of 1km^2, we find the density values correspond best to the built-up density calculated by the combination of the thematic classes 'buildings', 'streets' and 'railways'; in turn, this defines basically the thematic definition of our GUF-DenS input layer. Median absolute errors of 1.67 for Paris, 4.92 for Munich or 3.29 for Dallas prove the validity of the input layer. However, we also observe tendency of slight overestimation of density values in this data set in general, and a tendency of higher overestimation for areas of higher density within cities in particular. This is e.g. true for Las Vegas where the aridity of the location leads to open soil as dominating land cover beyond impervious surfaces. Thus, these two classes seem to get mixed up spectrally causing overestimation with a median absolute error of 29.71 (Fig. 3). However, we also believe that it is highly probable that the measured overestimation is in reality lower because, unlike the BKG data in Germany, we cannot ensure that the OSM are indeed complete. In general, we conclude that our input data set still provide highly even if variable accurate density values for different areas or landscape types across the globe and represents the built-up density of 'buildings', 'streets' and 'railways'.

For the WorldPop data high accuracies have been measured for national-scale population distributions presented by Tatem (2017) or Stevens et al. (2015). In turn, we also assume this input data set provides a reliable basis for population assessment.

4.2. Global rank-size distributions at individual city level

With the first sentence of this paper – quoting the 2018 World Urbanization Prospects publication – a clear statement about the ranking of the largest cities in the world is given: Tokyo is the largest city of the world measured by population based on administrative units (Fig. 5a). Our results, however, suggest a different truth. If one determines the MUA of cities and calculates the populations related to them, the currently largest urban agglomeration in the world is the continuous urban landscape in the Pearl River Delta (PRD) metropolitan region (M.R.) in China (Fig. 2a; Fig. 4).

In this PRD M.R. the once physically (and still administratively) separate cities of 21 administrative units (among them are such large cities as Guangzhou, Dongguan, Foshan or Shenzhen which are entirely or partly within the new MUA) morphologically coalesced. The PRD M.R. is home to 42.6 million inhabitants. With 5.6 million more than the city of Tokyo, which has been listed as the largest city to date, this is 15.1% more compared to Tokyo in the United Nations ranking. The dimension of such an urban agglomeration becomes clearer when one compares it with population figures of countries. In a ranking of countries, this single urban agglomeration would rank at 34, larger than populations of e.g. Canada, Poland, or Australia. Considering the large but unknown informal population in the PRD M.R. (assumptions suggest about 20 million (Liang et al., 2014)), the PRD M.R. would be home

to about 62.6 million which would rank the city even at 22, in the range of Great Britain, Italy or South Africa. Based on our approach, Tokyo ranks now at number 2; however, its population is with 31.9 million inhabitants far behind the PRD M.R.

If one focusses on the newly calculated MUAs, the PRD M.R. is also measured spatially the largest agglomeration of today. The second largest city by MUAs is the coalescent M.R. of Los Angeles, USA; interestingly, its population with 13.7 million inhabitants ranks L.A. only at 16. At rank three in MUA is the M.R. of Changzhou (consisting of the cities of Changzhou, Jiangyin, Jingjiang, Suzhou, Wuxi and Zhangjiagang). However, this city is with 14.5 million inhabitants only ranked at 14. Vice versa, the fourth largest city by MUAs is Tokyo, Japan, but ranked second in population. In this context these relationships reveal indirect statements about the particular density of built structures.

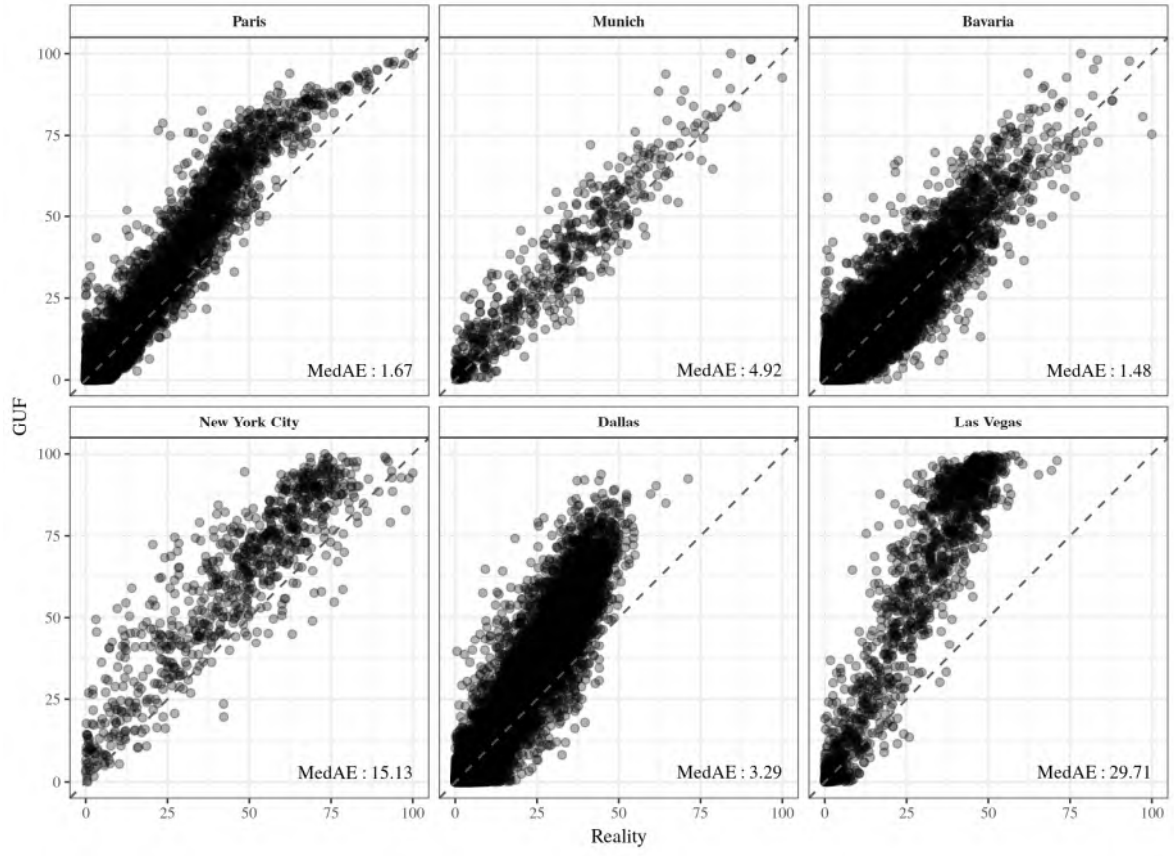


Fig. 3. Validation of the GUF-DenS layer at an aggregated grid of 1km² for different areas: Mean absolute errors (MAE) calculated for Paris, France; Munich, Germany; and the State of Bavaria in Germany consisting of cities and large rural environments and New York City, Dallas und Las Vegas, USA.

We have also specifically included the city of Ad-Damman (SaudiaArabia) in Fig. 4, as it leads the AUAs size ranking (cf. Fig. 5b). Ranked at 383 by population based on MUAs and ranked 173 using MUAs also reveals clearly how arbitrary spatial units are and how they may differ to structural characteristics.

In general, we find the three largest and 14 out of the largest 30 MUAs are, by our definition, polycentric metropolitan regions where once separated cities coalesced. This testifies to the fact that formerly morphologically separated urban areas have merged into continuous urban landscapes of a new dimension exceeding current spatial control units of AUAs. In our

analysis of MUAs, the PRD M.R. is the largest urban agglomeration in the world with the largest population. In previous statistics the PRD M.R. is not listed, but the cities that have grown together in the meantime are still counted individually. As examples, the city of Guangzhou is listed at rank 19, the city of Shenzhen at 26 by population in UN statistics (Fig. 5b). By administrative spatial units Guangzhou ranks only at 699 revealing the skewness of the statistics (Fig. 5a). In Fig. 5 we reveal that the largest extents are found in Saudi Arabia; cities, as indicated above, far from being the largest cities in the world. These areas carry political history in them and we should disregard them in geographical comparisons. Let us take as one example the city of Kuala Lumpur in Malaysia: the city is ranked 40 based on population figures relying on MUAs. The MUA ranks it even on 33. For administrative units Kuala Lumpur ranks at 49 for population, but for administrative space units it is only at 1253. This finding reveals how the common, accepted but somehow artificial and arbitrary AUAs and the related statistics obscure reality and challenge any ranking system due to an unequal and thus generally incomparable denominator.

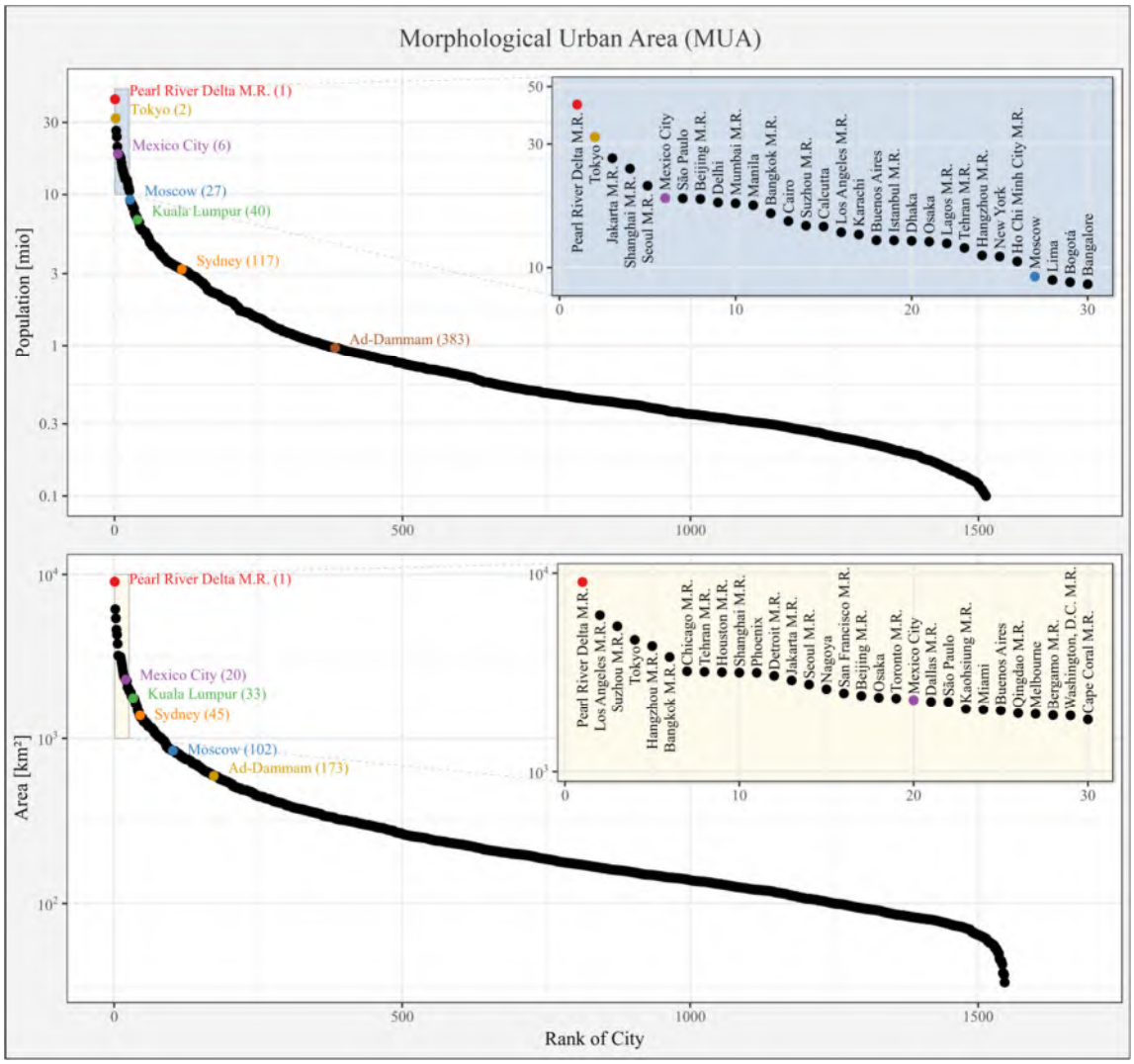


Fig. 4. Rank-size distributions of the largest 1,692 cities across the globe: For population derived from re-territorialized MUAs (on top) and for AUAs (below); the details in the boxes present the largest 30 cities.

In general, this alternative approach using MUAs as spatial baseline reveals a striking spatial difference to the extents of the existing AUAs. Considering a city 'true-bounded' if spatial deviation is within 10% between MUAs and AUAs, we find this only in 3.7% of all cases. This seems to be an alarmingly high value that should make us re-think existing administrative units.

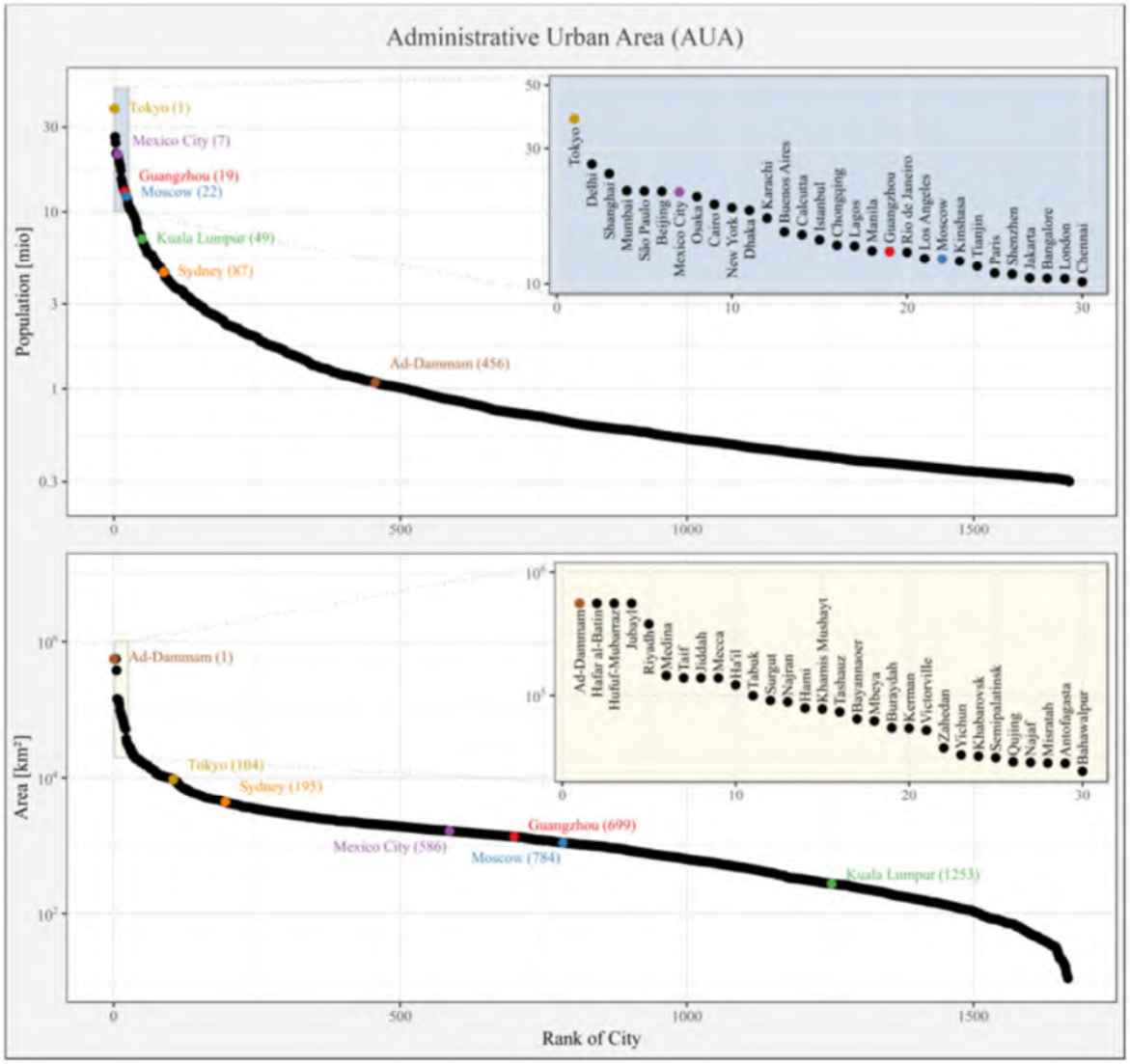


Fig. 5. Rank-size distributions of the conventional administrative space units for the largest 1692 cities across the globe: for population related to AUAs (on top) and for spatial extents (below); the details in the boxes present the largest 30 cities.

Although according to this study a new largest city in the world has been identified, which is with 42.6 million measured larger than the previously assumed 37 million of Tokyo, overall fewer cities achieve mega-city status (mega cities are defined as cities with > 10 million inhabitants (UN, 2018)). Compared to the currently 30 led by the United Nations statistics, only 26 are identified by our approach. Perhaps even more interesting is the fact that these two lists are very different from each other: 10 mega-cities listed by the United Nations – Bangalore, Chennai (Madras), Chongqing, Moskva (Moscow), Tianjin Kinshasa & Brazzaville, Lima, Rio de Janeiro, London, Paris – are not considered mega-cities based on our newly designed MUAs. In turn, six metropolitan regions do now feature mega-city

status based on our approach (Ho Chi Minh City and Bien Hoa M.R. in Vietnam; Seoul, Incheon, Suweon, Seongnam, Goyang Bucheon, Ansan, Anyang, Uijeongbu, Siheung and Gwangmyeong M.R. in South Korea; Suzhou, Wuxi, Changzhou, Jiangyin, Zhangjiagang and Jingjiang M.R. in China; Hangzhou, Shaoxing, Cixi, Yuyao and Shangyu M.R. in China; Tehran, Karaj, Eslamshahr, Malard and Qods M.R. in Iran; Bangkok, Samut Prakan and Nonthaburi M.R. in Thailand). What is striking here is that based on our approach all newly identified mega-cities are located in Asia, while 6 out of 10 mega cities that fall out of this statistic are outside Asia.

4.3. Geographical patterns with regard to the different spatial units (MUAs vs. AUAs and related populations) at continental level

If one looks at the discrepancies between MUAs and AUAs in terms of geographical distribution, some things stand out especially: The previous analysis focused on the largest cities in the world, where some MUAs far exceed their administrative boundaries (as shown e.g. for the PRD M.R.). It might be surprising that in most cases the opposite is measured. The actual tendency reveals that the calculated MUAs are more often much smaller than AUAs. This tendency is particularly pronounced in Asia (Fig. 6). The exception, however, is Europe (Fig. 6). In Europe the spatial units of MUA and AUAs match on median, the distribution, however, indicates that a majority of AUAs is smaller than the calculated MUAs. This reveals a spatial land organization on comparatively small entities. It is also interesting that a gradient from west to east is emerging for Europe. In the west there are many cities that have grown beyond their administrative borders, while in the east this is the other way round. Another example for regional gradients: In China, there are only a few cities that go beyond established administrative boundaries; however, these are found almost exclusively on the east coast. A few other regional peculiarities are the following: in Indonesia, MUAs are measured consistently larger than AUAs, whereas in India, the Arabian Peninsula, as well as in the Middle East, it is predominantly the other way around.

Table 1. Geographic trends at continental level: Quantity of cities per continent belonging to the largest 100 cities based on MUAs and AUAs as well as related population figures.

Continent	Population (MUAs)	MUAs	Population (AUAs)	AUAs
Africa	12	5	10	8
Asia	63	48	58	73
Europe	5	6	4	0
North America	13	32	16	16
South America	7	5	10	3
Oceania	0	4	2	0

If we take the 100 largest cities for comparison, we find that independent from the method of measurement (population or extent based on either MUAs or AUAs) the large share of the largest cities is located in Asia (Table 1). 48 cities based on MUAs and even 63 for population related to MUAs out of 100 are located there. For the AUAs, 73 cities in spatial extent and 58 in population are located in Asia. At the other end, Oceania features only 4 for the MUAs, or even 0 respectively for population; for the AUAs it is 0 or just 2. Remarkable is the difference in city extents in North America: For the MUAs, 32 cities are counted

among the largest 100, while for the respective population only 13 belong to this list. This indicates to the extensively large and continuously low dense sprawling cities in the USA with comparatively low dense populations. The detailed list of the 100 largest cities and their attributes are available in supplementary A-2.

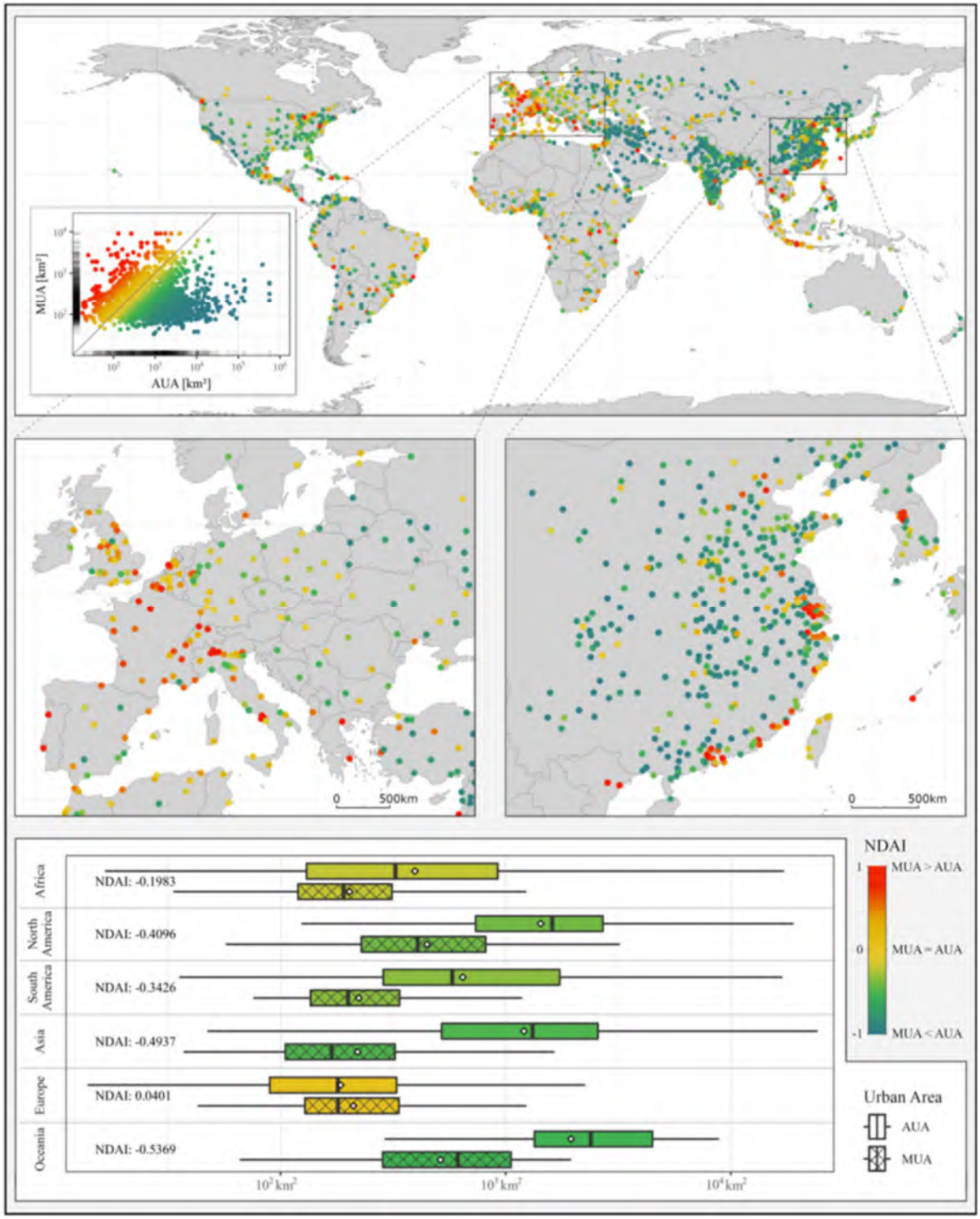


Fig. 6. Global map of cities classified based on the NDAI indicating the relationship of MUAs to AUAs. Boxplots illustrating the MUA distributions versus the administrative units aggregated to continental level. Be aware of a non-linear nature of the NDAI.

5. Discussion and interpretation: Shaky truths and morphologic realities

The morphologically coalesced polycentric metropolitan region in the Pearl River Delta is currently the largest urban agglomeration in the world. Its population is suggested at 42.6 million. Tokyo, however, usually considered the largest city in the world, is ranked second. With 31.9 million it is suggested with fewer inhabitants than in UN statistics. So, in general we see a shake-up in the city sizes and populations and their respective rankings at global scale comparing this alternative, but methodologically and spatially consistent approach using MUAs vs. the common UN statistics relying on administrative units.

So, is in consequence our global urban landscape defined by larger urban entities than we have assumed? There is no simple answer to this question and this must be considered in a differentiated way: On the one hand, conceptual approaches such as 'mega-regions' or 'urban corridors', which see themselves as an integrative cluster of cities with their surrounding suburban hinterlands forming far larger urban landscapes, have been identified a long time ago (Gottmann, 1957; Whebell, 1969). In this context, Small et al. (2011) revealed spatial urban extents vastly larger than the administratively-defined cities. These evolving new city patterns have been documented as an interweaving space as well as in their physical shape and are already acknowledged for a new dimension of urban landscapes (e.g. Florida et al., 2008; Taubenböck et al., 2014). On the other hand, within these mega-regions or urban corridors individual cities (e.g. New York City within the Boston to Washington corridor) remain individual lighthouses regardless of being part of a larger urban constellation for the global perception of the city, its economic success, creativity, cultural diversity, and much more. It therefore remains relevant to shed light to these physical dimensions of individual cities within these larger, more loosely bounded urban landscapes. The statistical shift of city sizes measured by MUAs, as done in our study, reveals that we are de facto dealing in parts with larger urban entities than we have assumed in traditional statistics (PRD M.R. vs. Guangzhou within a region as well as PRD M.R. vs. Tokyo in a global ranking). In large parts, however, the statistical shift of city sizes directs us towards smaller entities in MUAs than in administrative units.

From a methodological point of view, this approach aims to translate the agglomeration concept understood as continuous physical builtup landscape into statistical practice. Naturally, the city sizes based on MUAs are indicative rather than definitive. It is clear that there is not a single regional logic, nor a single dimension defining connected spaces, nor is a constructed territorial space "correct" or "incorrect" in absolute measures (Taubenböck et al., 2017). Rather, the approach is sensitive to our assumptions, i.e. the usage of the spatial entity of a monocentric city model, the defined center points, the corresponding indicators and the thresholds set. The monocentric city model is documented empirically robust and analytically tractable (Paulson, 2012), the defined center points are uniformly derived from one database (UN, 2014), the corresponding indicators only show a low correlation and their combination to the MSI makes it more robust and allows for a factual hypothesis, and the thresholds set are also based on a factual approach that produces results that are reasonable in the spatial domain (cf. Figs. 1 and 2). These explanations may be more or less comprehensible; however, we argue here that the decisive added value does not lie in the choice of these specifics, since our results are plausible spatial delimitation of the physical

city extents, but in the globally consistent application of these specifics and thus a comparable basis is created. Moreover, it must also be clear to one that the analysis also relates very much to the spatial and thematic scale. The abstract representation of settlements used in our EO-mapping products influences the results; if we consider this mapping product in comparison to lower resolution data such as night-time lights as used by Small et al. (2011) or to geometrically and thematically higher resolved data sets such as three-dimensional city models, which might even be enriched with usage types, different diversities of morphological and functional details allow for different concepts, methods, perceptions and results on delimiting urban from rural. However, the latter higher resolved data sets are not consistently available at global level, if they are available at all and thus do not allow for a global study as performed here. The results of our MUAs also depend on the accuracy of the input data – GUF and GUF Density. We need to be aware that the GUF does not feature a consistent high accuracy across the globe. However, if one considers the studies on this subject, for urban areas, the accuracies are documented consistently high (Taubenböck et al., 2011; Klotz et al., 2016) even in landscape types such as arid regions (Mück et al., 2017). The accuracies of the GUF Density layer are documented also high, however, also with varying precisions for different location. This inconsistency in accuracy for different areas on our planet must remain an unknown, since good and complete reference data are largely unavailable. And even if, as the Las Vegas example shows, the data set tends to be overestimated in certain areas, our approach for delineating cities based on local context information such as the threshold relative to the city average is adaptive and can compensate for misclassification. So in consideration of these aspects, we think the decisive contribution of this study is not 'general truth', but 'a truth through methodological and as far as possible data technical consistency'.

Geographically, the physical approach of consistently delineating cities using the local context produces MUAs that still contain a large morphological variability in it. It is also clear that measurement of physical contiguity of a city is one-dimensional and does not necessarily imply a high degree of interaction and interdependence within urban landscapes. This means, our analysis focusing on the "space of place" is only one perspective; however, it is in line with the perspective taken by ranking lists on city sizes. For a more comprehensive understanding of the physical and virtual sizes of cities, it needs to be complemented by the "space of flows" within and across city areas. Compared to all these theoretically to be considered data, conceptual and methodical possibilities, the strength of this approach here is to be seen in its global consistency regarding input data, concept, methods, and in consequence geographical results.

So in conclusion, is 42.6 million inhabitants the new correct number for the largest morphologically contiguous urban agglomeration in the world? Unfortunately, this must be questioned, too. First, as just discussed, the MUA approach is sensitive to the choice of certain indicators and the input data has its own errors. Manipulating the algorithm would allow us to expand or shrink the MUA resulting in larger or smaller population figures. Let's take the example of the PRD: Our approach separates the MUA of the PRD from the city of Hong Kong, as the topography and the ocean there constitute a natural barrier of low settlement density (or no settlement at all). Our method takes this circumstance into account.

Many urban geographers, however, would argue that Hong Kong is functionally part of the PRD metropolitan region. By separating these morphological units, we measure the PRD with 42.6 million and Hong Kong with 4.5 million. With a different conceptual and methodological approach the PRD could therefore also be measured at 47.1 million by adding these numbers. So, we have to recognize, there is no 'absolute truth' to it; we argue if you look closely at the results in Fig. 1c and 2, you see that the derived MUAs capture the morphological urban space well and thus our approach is reasonable, transparent and consistent without claiming to be the only truth. Second, while the largest MUA according to our approach comprises 42.6 million people, we find in the extended Shanghai region a very high density of MUAs in close proximity to each other on a comparatively small area of 250 × 300 km. 17 MUAs – among them are the Shanghai, Kunshan, Taicang MUA (with 24.1 million the fourth largest in the world), the Suzhou, Wuxi, Changzhou, Jiangyin, Zhangjiagang, Jingjiang MUA (with 14.5 million the 14th largest) and the Hangzhou, Shaoxing, Cixi, Yuyao, Shanggyu MUA (with 11.1 million the 24th largest) – add up to 61.2 million inhabitants (the entire stretch is even home to 96 million). While we list these 17 individual MUAs in our statistics separately, this clustering of so many large MUAs indicates a larger urban agglomeration developing within the above mentioned conceptual framework of a mega-region. From this conceptual point of view, this mega-region (which is about the size of Austria) could be described with 96 million inhabitants as the largest urban agglomeration in the world. With it the urban agglomeration would be ranked 16 among countries with a population larger than Germany or Turkey. Third, as shown in other studies, it is highly probable that today's population figures especially in such dynamically growing large cities are still rather underestimated, mainly due to the difficulty to record informal population groups (e.g. Taubenböck and Wurm, 2015). Ultimately, therefore, the absolute population figures determined must be regarded as uncertain (just take the possibly 20 million informal workers in the PRD M.R. (Liang et al., 2014)). However, the relative population figures appear to be much more consistent than any previous data. This is due to the comparable spatial basis of consistently derived MUAs and the globally consistent input data. In summary this means, 42.6 million might not be the new correct number for the largest city in the world, but if you take a comparable spatial baseline the PRD M.R. is de facto the currently largest individual urban agglomeration in the world.

Consequently, the globally consistent and harmonized approach mapping MUAs provides a comparable basis for geographic research. It thus allows scrutinizing administrative units whether they indicate close to reality statistical information and/or whether they form meaningful areas of political competence. As we find only 3.7% of the 1692 cities 'true-bounded' we can now clearly state that the usual statistics rely on for comparisons actually inadmissible spatial entities and thus obscure morphologic reality. The newly derived MUAs may allow overcoming the sometimes arbitrary effects caused by AUAs, they may reveal misjudgments based on previously accepted statistics or they will make us re-think about whether existing spatial units should or should not be reformed. Ranking Los Angeles M.R. at number 2 (by MUA size) or at 16 (by MUA population) is both correct but global perception is, depending on the particular list, fundamentally different. The results discussed thus make clear that one must critically question every ranking list – knowing that one single truth may not exist.

6. Conclusion and outlook

Yes, administrative units obscure morphologic reality and significantly influence statistics and perceptions of cities. We find our planet already consists of larger individual city entities than generally accepted. The metropolitan region of the Pearl River Delta is currently the largest urban agglomeration in the world instead of Tokyo. We need to re-think current ranking lists on the spatial and demographic dimension of urbanization in a critical manner.

This work is a plea to overcome historical or arbitrary spatial units for generating statistics, and more importantly for managing our living environments from jurisdictional, political and planning perspectives. As we could show, in most cases the administrative spatial units and the true morphologic extents of cities do not match. Our newly generated spatial entities derived in consistent manner may form an admissible spatial baseline for better and more comparable statistics and urban research studies in the future. We propose that this approach needs to be extended systematically by analyzing the influence of different data and methods for city delineations as well as by a multi-temporal analysis of city size development over time for further geographic findings.

Isn't it interesting how old questions that seemed answered can now be re-examined in a more objective way, but remain unanswerable? By this re-examination, however, we believe that based on our analysis the opening quote of this article would be closer to the truth if changed to *"the Pearl River Delta M.R. is the world's largest city with an agglomeration of 42.6 million inhabitants, followed by Tokyo with 31.9 million, Jakarta M.R. with 26.5 million, Shanghai M.R. with 24.1 million and Seoul M.R. with 20.7 million Today, Mexico City, Sao Paulo, Delhi, Mumbai, Beijing M.R., and Manila have all close to 20 million inhabitants"*.

Acknowledgements

We would like to thank our supporters for their commitment to this project: Kevser Basdas, Carla Madueno, Nicole Osterkamp, and Thomas Spieß.

Appendix A. Supplementary data

Supplementary data to this article can be found online at

https:// doi.org/10.1016/j.rse.2019.111353

Chapter XIV

Which city is the greenest?

A multi-dimensional deconstruction of city rankings

Computers, Environment and Urban Systems, 2021, vol. 89

Hannes Taubenböck, Marina Reiter, Fabian Dosch, Tobias Leichtle, Matthias Weigand & Michael Wurm

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.compenvurbsys.2021.101687

Received 17 Mar 2021; Accepted 30 June 2021; Published 13 July 2021

Which city is the greenest?

Abstract

The question "*which city is the greenest*" sounds trivial, but in reality, this question contains statistical ambiguities. In this study, we approach this issue by ranking cities by green space shares. However, we do not base our ranking only on one green parameter and the commonly used administrative boundaries. Instead, we broaden access to rankings through several approaches: First, we calculate two parameters, i.e. green space shares and green space per capita. Second, we apply these parameters for two cases: for all green areas as well as for green areas with a minimum size of one hectare. The latter are considered to have an impact on near-home recreation and the local climate. Third, we relate these parameters on the one hand to administrative spatial units constituting the entity 'city', but juxtapose these on the other hand with two alternative spatial reference units: a morphological spatial unit that closely encompasses the built-up pattern of the city, and a standardized buffer unit around the city centers. The variability of these manifold rankings obtained by this study makes clear: the rank of one city in a relational system to other cities depends strongly on these parameters and spatial units applied. In our experiments we rank and compare the 80 major cities in Germany. The diversity of results allows to discuss the susceptibility of spatial statistics to ambiguities that may arise from the use of different concepts. By integrating these multidimensional concepts into one final ranking, we propose a strategy for a more holistic and robust approach while revealing uncertainties.

Keywords

Urban green; remote sensing; spatial metrics; city ranking; comparative urban research

1. Introduction

Cities have always been in competition with each other. In times of globalization, positioning a city in this competitive environment has become increasingly important (e.g. Begg, 1999; Hall, 1995). Cities compete and seek to attract global businesses, investors, tourists and capital (Giffinger, Haindlmaier & Kramar, 2010). One popular tool, which gives expression to this competition, are city rankings. Cities are compared, for example, according to certain key figures: e.g. by population (UN, 2018) or economic turnover (Sassen, 2019), by spatial specifications of key figures (e.g. by urban populations (Melchiori et al., 2018) or by areal extents (Taubenböck et al., 2019)), or by multiindicator analytics (e.g. done by consulting companies such as Resonance Consultancy, 2020; The Economist Intelligence Unit, 2019; Mercer Consulting, 2019). Rankings are in the vernacular and practices quasi constitutive of the rules of the game of urban competition, to which players must inevitably adhere (Bok, 2021). It has become an advantage in the competition for perception to give the city a trademark that stands for a definable quality or for uniqueness (Löw, 2008). These attributes often aim to capture the particular character of a city, framing it as dynamic, cosmopolitan, traditional, ecologic, green, among others.

However, city rankings are not without controversy. McManus (2012) urges to critically question who produces rankings, what goal they serve, which cities are addressed, which indicators are used and how they are calculated and weighted, and how results are interpreted. A veritable industry with well over 500 different urban benchmarks has more or less taken on a life of its own (Acuto, Pejic & Briggs, 2021). City rankings, however, are conceptually and methodologically demanding, and often lack transparency, empirical basis or data appropriate for comparison (White & Kitchin, 2021). In this paper, we want to follow Acuto et al. (2021) call to engage conceptually, methodologically and empirically with city rankings, as critical urban research must bring more than "*limiting itself to criticism*". We aim to respond to Derudder & van Meeteren's (2019) call to engage with 'urban science', to experiment with methods to better assess current urban conditions by the use of technology-based quantitative analysis. In this sense, our goal is to objectify a ranking through a multi-layered approach based on remotely sensed data, without denying the ambiguities of statistics.

One strategy to give the city a positive image is to highlight statistics of spatial indicators. Green spaces, as an example, have widely positive connotations in society. In Germany, for instance, the cities of Berlin, Bonn, Halle (Saale) or Hanover take advantage of this and use the adjective *'green'* in relation to their city to build a positive image. Hanover even promotes itself as Germany's greenest city. This branding, how ever, remains statistically unquestioned, as long as this interpretation is accepted as plausible by society. It is precisely this unquestioned aspect that we examine in this study and which we underpin with quantitative statistics on urban green space shares. For this purpose, we present multiple approaches of city rankings of the proportion of green spaces in major German cities. On the one hand, this allows to show to what extent the described perception of these four exemplary cities is also statistically reliable in spatial terms. On the other hand, this allows to systematize and discuss statistical ambiguities, which is created by the choice of parameters and their conceptualization, and the spatial reference units.

In this study, we pick the parameter 'urban green' as example. Green spaces, of course, have vital relevance for cities beyond perception and branding. They play a crucial role in urban ecosystems: they allow rainwater infiltration refilling groundwater (e.g. Bolund & Hunhammar, 1999; Wolch, Byrne & Newell, 2014) which improves resistance to flooding (Banzhaf & De La Barrera, 2017). Cooling effects (evaporation, shading, fresh air corridors, etc.) could be empirically demonstrated (Reis & Lopes, 2019) and thus the potential to mitigate the effects of climate change to which urban flora, fauna and humans are increasingly exposed (BMUB, 2017). And, they contribute to biodiversity within cities as habitats for flora and fauna (Aronson et al., 2017; BMU, 2019). Green spaces also play a crucial role for urban citizens: In this context, the material resources of a city can be seen as a foundation through which society is specifically constituted (Löw, 2008). In particular, green areas provide spaces for recreation, social interaction and physical activity that have a positive impact on the mental and physical health of humans (Bertram & Rehdanz, 2014; TEEB, 2011). They are known to reduce mortality and the risk of cardiovascular and respiratory diseases (Gascon et al., 2016; Weigand et al., 2019). And, the design and quantity of greenery impacts on the visual character and perception of the urban landscape (Chmielewski et al., 2020). With it, they contribute to the improvement of life satisfaction (reduction of stress, aesthetic experiences, spiritual enrichment) (WHO, 2016).

In a world that is urbanizing at highest dynamics (Taubenböck et al., 2012; UN, 2018), a balanced combination of built and natural urban landscapes is increasingly important. With rising numbers of people living in cities, the pressure on urban ecosystems and the environment increases. Land consumption is rising, for instance, as more space is required for living, commerce or traffic. Exposure to noise or air pollution increases (Gomez-Baggethun et al., 2013; Grunewald, Xie & Wüstemann, 2018). The more important cities become as a living space, the more important become urban living conditions. Initiatives, such as "*Die Grüne Stadt*" (transl. "*The Green City*") and public petitions (e.g. "*Grünflachen erhalten*" (transl. "*Preserve green spaces*") in Munich) show how not only decision-makers, but also the general public are becoming increasingly aware of these issues and the protection of green spaces is a key concern (Bürgerbegehren Grünflachen erhalten, 2019; Die Grüne Stadt, 2019).

In theory, a city ranking could be based on manifold parameters or methodological approaches. Qualitative studies compare or evaluate the 'urban green' e.g. by visual assessments on greening, questionnaires on subjective perception or parameters such as green quality (e.g. Ellaway, Macintyre & Bonnefoy, 2005; Giles-Corti et al., 2005; Hoehner et al., 2005). Other approaches rely on quantitative spatial parameters. Data are applied from citizen science approaches, e.g. for the assessment of resilience of green spaces (Pudifoot et al., 2021), from official governmental data or by remote sensing (Shahtahmassebia et al., 2021). A variety of measures are applied such as green space accessibility, green volume, green space proportion or green space provision (e.g. BBSR, 2018; Fuller & Gaston, 2009; Grunewald et al., 2018; Gupta et al., 2012; Morgenpost, 2016; Richter et al., 2016).

Overall, however, it has to be stated that spatial knowledge about the green stock is still often insufficient which is an obstacle for Derudder & van Meeteren's (2019) call for experimenting with technology-based spatial analyses. On the one hand spatial data are often

neither available in the necessary spatial extent, the thematic or spatial detail, the needed accuracy and consistency, nor the desired up-to-dateness. In Germany, as example, the federal area statistics based on the real estate cadaster do show parameters for green areas, but their temporal and spatial consistency and nationwide availability is currently limited. On the other hand, the spatial indicators or measurement methods, which permit an evaluation, are neither obvious nor unambiguously defined, i.e. the comparison of spatial green indicators is not trivial.

In recent years, remote sensing has become a crucial data source for the classification of green areas and it allows to work over large areas in a consistent manner (e.g. Lamchin et al., 2020; Richards & Belcher, 2020; Shahtahmassebia et al., 2021). Methods for deriving green spaces based on remotely sensed data have advanced significantly even in the complex and small-scale urban landscapes (e.g. Parmehr et al., 2016). Approaches on very high-resolution (VHR) data with spatial resolutions of 1 m and better combined with three-dimensional data (e.g. LIDAR) allow a highly detailed recording of vegetated areas, green volume estimates or different vegetation species (e.g. Tooke et al., 2009). The limitation for these VHR applications is primarily the limited possible spatial coverage, due to data cost and availability. In contrast, large-scale approaches with free-of-cost data are more limited in their spatial and thematic resolution. Nevertheless, very good classification accuracies have been achieved with sensors such as Sentinel-2 (e.g. Weigand et al., 2020) or Landsat (e.g. Pflugmacher et al., 2019) using shallow machine learning techniques. Lately, new image classification methods such as semantic segmentation from convolutional neural networks have proven very high accuracies for the detection of small-scaled structures in complex urban environments (Wurm et al., 2019), and have also been successfully applied for urban green spaces (Albert, Kaur & Gonzalez, 2017). In this study, we therefore use Sentinel-2 satellite data instead of indicators from the area statistics, which are prone to errors due to changes in the survey method. In this way, we aim to ensure comparability across cities.

In the field of spatial parameters for analyzing and comparing the green stock across space or over time, concepts, methods, thematic content, spatial reference units, among others vary. We want to illustrate these methodological particularities with the example of spatial reference units: Whatever the intended statement in studies analyzing urban green is (as done by Bertram & Rehdanz, 2014, Fuller & Gaston, 2009; Grunewald et al., 2016; Larondelle & Haase, 2013; Richter, Behnisch & Grunewald, 2017 or Richter et al., 2016; Kabisch, Strohbach, Haase & Kronenberg, 2016), in most cases indicators are developed on administrative reference units. While these units form the basis for political decision-making, in a geographic sense their historically and politically drawn boundaries do not make them an admissible basis for consistent comparisons. The modifiable area unit problem (MAUP) testifies to aggregation effects on zonal statistics as well as on size effects (Openshaw, 1983). These have been found to be scale dependent and nonstationary over space (Margulies et al., 2016) and are prone to distort or even obscure reality (Taubenböck et al., 2016). To address these challenges in part with respect to the example of spatial reference units, approaches have been developed to spatially separate cities from their surrounding areas in a consistent manner to provide an admissible spatial unit for cross city

comparisons (e.g. Dijkstra & Poelmann, 2014; Melchiori et al., 2018; Taubenböck et al., 2019). Nevertheless, these approaches also carry conceptual challenges (e.g. border effects); a universal approach is non-existent.

Against this background, we aim in this study to systematically compare the green stock in major German cities. The focus of our interest is on the structural differences and commonalities between cities. To do so, we generally rely on classified green spaces derived from remote sensing data. In contrast to other studies, we aim to develop a city ranking that is fed by different urban green parameters and different spatial statistical perspectives. Based on a spatially variable multisubject concept, sufficient empiricism and a consistently conducted methodology, we aim to contribute to a more holistic, reliable comparative urban research. In this way, we want to reduce the lack of empirical knowledge about the positioning of different cities in a relational field, as demanded by Löw (2008).

2. Conceptualization of the study

Green spaces are one (of many) important indicator to describe and evaluate urban living conditions and quality of life. Related statistics often serve as a basis for societal debates or political decision-making. Spatial statistics, however, are always prone to ambiguity – which parameters are used, which scale of analysis is applied, how accurate or consistent are input data, or to which spatial reference unit is the measurement referred to. And with that, we run the risk that ill-considered spatial statistics either lead to a random result or are selected specifically to achieve certain results, which are then to serve as the basis for political decisions.

In this study we build a relational reference system of cities with respect to green spaces. It is intended to offer more resilient results through multiple perspectives. Our focus is on two common parameters in this field: The *proportion of green space* and the *provision of green space* per capita. Both parameters are calculated relative to area or number of inhabitants and are thus basically suitable for comparisons of cities of different sizes. We choose the parameter ‘proportion of green spaces’ because a comparable study by data analysts from Morgenpost (2016) used exactly this parameter for their ranking. And we rely on the parameter ‘provision of green space per capita’ as it is recognized and used by the World Health Organization (WHO, 2010). Our scale of analysis is *the city level*. However, in our conception, we do not want to conceive the city (or its spatial extent) here as a given territorial unit. Rather, we vary the spatial reference units defining the city extent in order to gain a deeper understanding of green space proportions. Our starting point is on administrative reference units, which is the spatial entity that is commonly used and the legal basis for decision-making. In addition, we elicit how consistent results are when we change the spatial reference unit. For this purpose, we use a morphological spatial unit that encompasses the built space of a city and a standardized radius-based buffer unit around the center. Our input data are based on one consistent data source: *remote sensing data*. Thus, we rely on a consistent and full survey without distinguishing between public and private green spaces. The data set has a transparent assessment of accuracy and the original data sets have a high spatial resolution.

Our study focuses on the systematic comparison of cities in terms of spatial green shares. We relate cities to each other as we believe that they are a conceptually similar target entity. For this purpose, we use *city rankings* as a means of choice. In general, we imply higher spatial shares and larger entities of green spaces are related to a higher likelihood of positive effects (e.g. Arlt et al., 2005; Gill et al., 2007). On the one hand, this means that we set the city with the highest share in the ranking to one. On the other hand, we consider that the size of an urban green area is of great importance for their effects: The larger the green space, the higher its potential benefits in terms of provisioning, regulating, habitat-shaping and cultural services. With respect to climate-relevant effects, some re searchers state that regulating services, such as cooling the environment, are only provided by green spaces of one hectare or more (Arlt et al., 2005; Finke, 1994). Thus, we conduct our analysis for *all green spaces per city* and in addition also only for *green spaces larger than one hectare*.

3. Study areas, geodata and spatial parameters

3.1. Study areas: major German cities

The selection of the study areas is based on the need for an up-to-date, large-area, consistent geodatabase of spatially high and thematically appropriate resolution. Since this is available for Germany (cf. Section 3.2), we include for our analysis all German cities which are 'large cities' by definition (Fig. 1). These are 80 cities which have at least 100,000 inhabitants (BBSR, 2018). The data of the German Federal Statistical Office served as a basis for the selection and they refer to December 31st 2017 (DeStatis, 2020). This population information was assigned to grid cells with a side length of 100 m in an INSPIRE compliant grid (BKG, 2020). For the spatial definition of the respective city centers, we rely on the data sets provided by the German Federal Agency for Cartography and Geodesy (BKG) (BKG, 2020).

3.2. Geodata: green spaces

In this study we rely on green spaces classified from Sentinel-2 satellite data. This data is free of cost and features area-wide coverage with a high spatial resolution of 10 m ground sampling distance. The acquisition period of the images for the classification was between 2015 and 2017, i.e. the classification is based on data that were not recorded at the same time. The classification algorithm was trained for seven thematic land-cover classes: artificial land, open soil, high seasonal and perennial vegetation, low seasonal and perennial vegetation as well as water. Artificial land relates to built-up areas, low seasonal vegetation com pares to croplands, low perennial vegetation features pastures, vine yards, and orchards, high seasonal vegetation subsumes deciduous tree cover including forests and fruit tree crops, and high perennial vegetation relates to evergreen tree cover (cf. nomenclatures in alignment with Anderson et al. (1976)). Figure 2 illustrates the classification for two sample cities, Karlsruhe and Dresden. The classification features an overall accuracy of 93.07%. Details on data, methods and results are shown in Weigand et al. (2020).

Green spaces are heterogeneous not only in their spatial distribution, but also in their phenological characteristics, as already indicated in the classification scheme. This means that a green space can be low seasonal vegetation (such as shrubland), high perennial

vegetation (such as deciduous forest), among others. In our analysis, we do not differentiate here. We aggregated the various vegetation classes into one thematic class: green areas. This abstracts our input data but also makes them consistent, so that a comparison is feasible.

3.3. Spatial reference units

Typically, city rankings are based on administrative boundaries. These spatial units are crucial as they determine a clear-cut city's boundary to establish jurisdictional competence of its municipal government (Parr, 2007). Therefore, we use them as a starting point in our analysis. We rely on *administrative boundaries* as provided by BKG (BKG, 2020).

However, these administrative units form an artificial unit as they have been created by very different political and historic developments. Although often applied, they do not form a consistent and thus com parable spatial entity for city comparisons from a geographic point of view (Taubenböck et al., 2019). It has been shown that the zonation effect, i.e. a re-arrangement of a spatial reference system into different units, may lead to – although based on the same initial values – different results and conclusions (Jelinski & Wu, 1996; Madelin et al., 2009).

Against this background, we apply two other spatial reference units that are supposed to also mark the entity 'city': a *buffer* and a *morphologic unit*.

The *buffer unit* is understood as a uniform unit, independent of city sizes or spatial patterns. Here we apply a standardized buffer, using a 10 km radius around each city center. The selected circle and radius are here representative for other possible spatial units like squares or rectangles, or other radii. The point here is simply to incorporate a uniform size and shape as an example, as one possible variant of consistent geographic comparison. Since there are cities included in the analysis which are located in close proximity in highly agglomerated areas, such as the Ruhr area (Fig. 1), in some cases the buffer units of different cities overlap. However, in our analysis we treat each city individually, i.e. green spaces in overlapping situations are part of the analysis for both cities.

Both spatial units, the administrative and the buffer unit, do naturally not capture the morphologic settlement extent of cities in a perfect sense. Therefore, we introduce a *morphologic unit* as a third spatial entity. It is our basic assumption that boundaries of the city are conceptually difficult to determine (Sievers, 1997), but in principle they exist. To address this concretely, we apply an approach that delineates cities as a spatially coherent, comparatively dense built-up landscape. Using remotely sensed classifications on built-up structures and their density, the city boundary can be determined along an urban-rural gradient in a data-driven way (cf. for methodological details Taubenböck et al., 2019). Based on a monocentric city model, it is assumed that with increasing distance from the city center, the transition from urban to rural is along a decreasing built-up density (Fig. 2). As it is impossible to distinguish the urban from the rural according to a universal truth, the strength of this approach lies in its consistency, i.e. for all test cities the boundaries are set in a mathematically unambiguous, consistent and thus comparable manner. For some cities located in close proximity, the morphology units also do overlap. As for the buffer units, the

same assumption was made, that some green spaces are part of the statistical analyses for two cities at the same time.

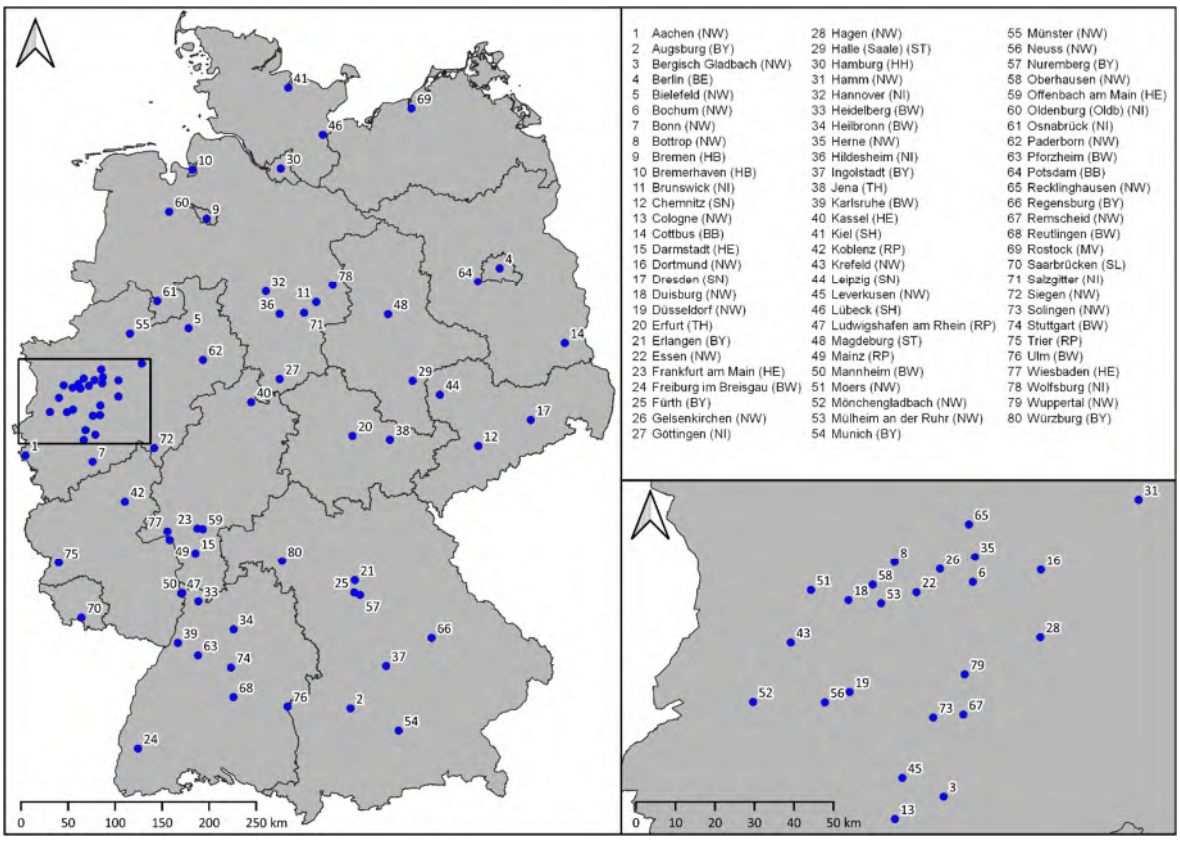


Fig. 1. Selected German cities with corresponding Federal state: BB = Brandenburg, BE = Berlin, BW = Baden-Wuerttemberg, BY = Bavaria, HB = Bremen, HE = Hesse, HH = Hamburg, MV = Mecklenburg-Western Pomerania, NI = Lower Saxony, NW = North Rhine-Westphalia, RP = Rhineland-Palatinate, SH = Schleswig-Holstein, SL = Saarland, SN = Saxony, ST = Saxony-Anhalt, TH = Thuringia.

Fig. 2 illustrates the three spatial units that variably constitute the entity city in spatial comparison. In the example of Dresden, the administrative unit includes the Dresdner Heide, a large forest area in the north-east of the city center. The morphological boundary, in contrast, draws the border tightly around the built city and thus the Dresdner Heide is not part of the spatial reference unit – with corresponding consequences for the green share.

3.4. Spatial parameters on green space shares

In this study we apply a quantitative approach, since we rely on consistent data sets for green space shares as well as for city populations for all cities. We base our city ranking on two quantitative parameters: (1) *Green space proportion*, i.e. the share of the total green space relative to the total reference area calculated in percent. Although this metric merely reflects the quantity of green space and not its quality, it is a parameter often applied as proxy for evaluating the quality of the urban life and the ecosystem (BBSR, 2018; Faryadi & Taheri, 2009; Morgenpost, 2016). (2) *Green space provision* per capita, i.e. the total green space available for each citizen calculated in m^2 per person. This measure is often used in the

political decision-making process. The WHO, for instance, calls for a minimum of 9 m^2 of urban green areas per inhabitant and defines 50 m^2 per capita as desirable (WHO, 2010).

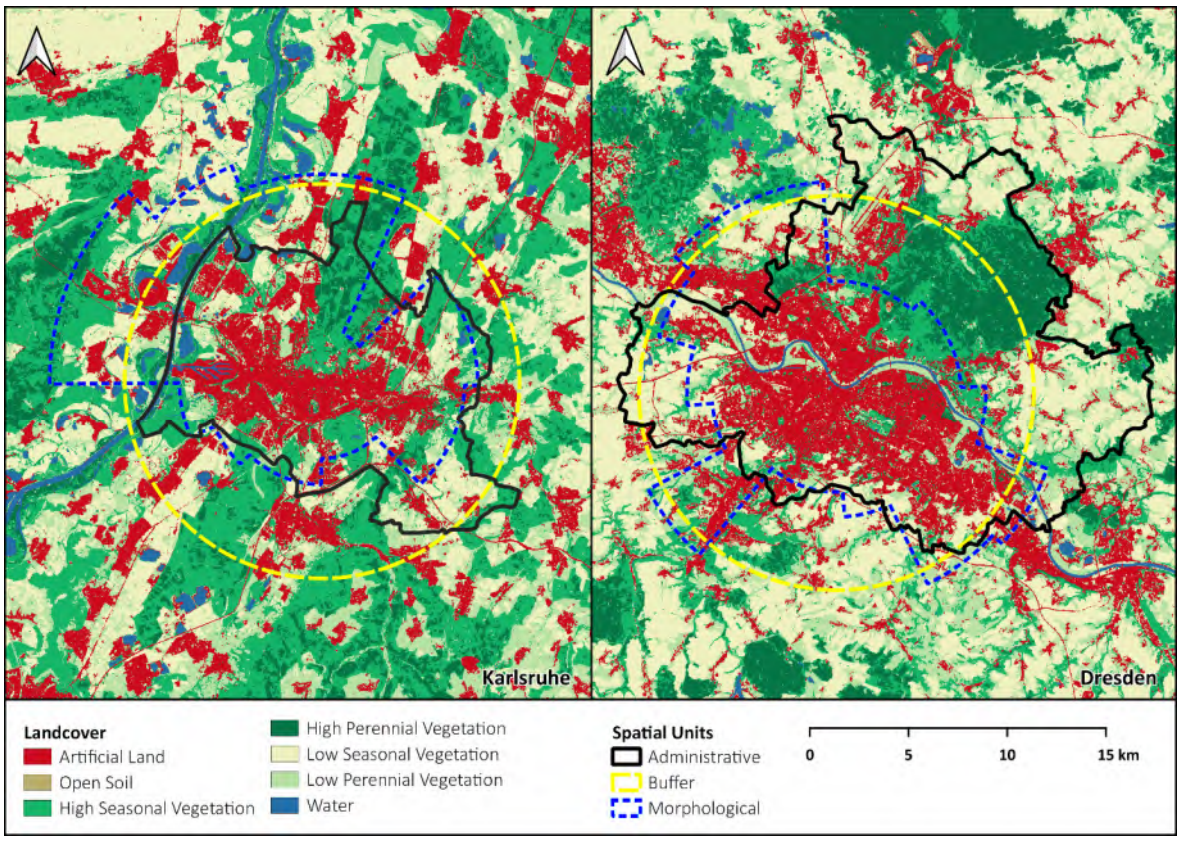


Fig. 2. Land-cover classification and three different spatial reference units delineating the entity city differently for the examples of Karlsruhe and Dresden.

3.5. City rankings and their deconstruction

With the two parameters on green space shares, the two-dimensional distinction according to spatial extent between all green spaces and the climate-relevant green spaces larger than 1 ha, and the three spatial reference units, overall twelve different city rankings are possible. This alone shows the diversity of spatial statistics.

In deconstructing this variability, we adopt the following rationale: We first rank the cities based on their green space shares (total and per capita), i.e. the highest share is number one. We use the city ranking based on the administrative spatial units as a starting point and analyze the deviations in the other spatial units. To show the differences be tween the results statistically, we determine the correlations between the approaches and describe their relationship through the coefficient of determination. We illustrate city sizes in these correlations in order to relate trends in green space shares to city sizes.

In order to identify cities with higher rank changes across spatial units, we classify them according to the comparison to administrative ranks, i.e. in those that feature above average rank losses in comparison to administrative ranks (indicated in red in Figs. 4 and 5 as well as in Figs. A and B in the Appendix), and vice versa the ones with above average rank gains (indicated in green in Figs. 4 and 5 as well as in Figs. A and B in the Appendix). The

classification is based on the average of the changes in ranks for morphological and buffer units. All mean values above the respective median are classified as cities with higher changes in rank.

Rankings are naturally also based on the desire to be able to make simple and clear statements and classifications. However, the diversity of the conceptually and methodologically reasoned variety of results contains a complexity that makes a simple answer to our guiding question "*which is the greenest city?*" difficult. Thus, we generate a final result to meet both challenges: to abstract and simplify this complexity to make a clear statement by benchmarking cities in one final ranking, but at the same time not without quantifying the uncertainties in the process. Therefore, we merge all twelve rankings into one single ranking. To do this, we use all twelve ranks per city and list them descending over their mean value, with the lowest total in first place and the highest total in last place. We use standard deviations for the ranks to indicate the uncertainties. This way, we base the final result on all views in equal parts.

4. Results

Depending on the green parameters (green space share, green space share >1 ha, green space per capita, green space >1 ha per capita) and the spatial units (administrative, buffer, morphologic) used, results vary. First, we introduce some general results in an overview and secondly, we present the city rankings.

4.1. General results

For the aggregated results from our sample of 80 major cities in Germany, we highlight the following key statistics: For the three *spatial reference units* applied, we measure differences in size and extent. In comparison to the common administrative units, the morphological boundaries draw the city boundary more tightly around the built-up space, i.e. this reference unit is on average smaller. Specifically, 56 of our 80 sample cities are smaller, by an average of nearly 33 km^2. In contrast, the standardized buffers of 10 km radius predominantly encompass the administrative city units: 75 of the 80 city entities are larger, by an average of almost 142 km^2 more area. These differences in extent, of course, have an impact on the population figures covered by these spatial reference units. On average, over 138,000 more people live in the buffers than in the administrative units. It is remarkable that morphological boundaries, although smaller on average, still hold nearly 50,000 more people on average than administrative units. This effect is mainly related to the cities in close proximity with a continuous built landscape, whose morphological units cover areas that belong to administratively different cities.

If we now look at the parameters, i.e. the (1) green space shares and (2) *green space* per capita aggregated across all cities, we measure the following differences to administrative units: (1) morphological boundaries with their smaller extents result in green space shares that are on average 6.8% lower (- 7.2% for green areas >1 ha). 67 of the 80 cities have a lower green space share on this reference unit. This reveals that administrative units usually integrate more natural and close-to-nature areas around the actual built space and therefore

green pro portions are higher. Buffer units, although spatially very different from administrative boundaries, have similar green space shares on average (- 0.4% and - 0,2% for green areas >1 ha), but basically it is distributed indifferently here. 35 cities have higher proportions and 45 have lower proportions. This shows that this spatial unit is neither adapted to the shape nor to the size of the particular city. (2) For the aggregated green space shares per capita, we see fundamental differences in the statistics: For the morphological units, the supply of green space per capita is decreasing on average by 140 m2 per capita (139 m^2 for green areas >1 ha) (and 71 of the 80 cities have falling figures) compared to administrative units (on average 331.7 m^2). For the buffer units (on average 475.4 m^2), on the other hand, the provision per capita increases by 144 m^2 per capita (+141 m^2 for green areas >1 ha) (and 64 of the 80 cities have higher proportions).

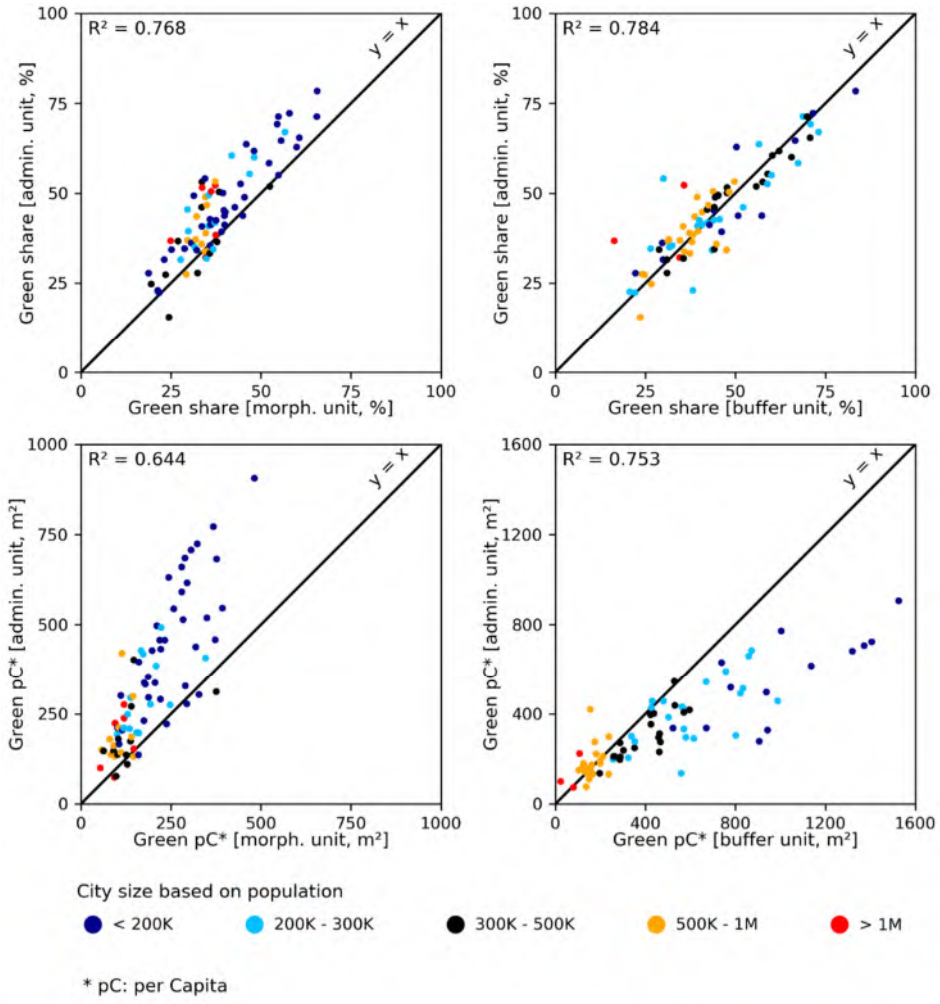


Fig. 3. Relationships of spatial green shares and green spaces per capita for morphologic and buffer units to the reference values based on administrative units. All cities are classified using population numbers.

The statistical relationship of the different parameters and spatial units generally reveal good linear fits with coefficients of determination between 0.644 and 0.784 (Fig. 3). While this may be considered a good correlation in principle, it shows that the spatial units do have a strong influence on our result. The different spatial units all have similar spatial bases, i.e. the city center and more or less the contiguous built-up areas, but they are drawn very differently at the edges of the cities, sometimes narrower, sometimes wider. And the fact

that this alone is enough to have coefficients of determination of only 0.644 shows how fragile spatial statistics can be. Especially in the case of green space per capita and the morphological spatial units, we see immense deviations to lower values compared to administrative spatial units.

With regard to available green spaces, a certain relationship to the size of the city can be identified. For cities that have very high green space shares, we exclusively see low population numbers within our sample. Vice versa, the cities with large populations such as Berlin or Munich (larger cities are indicated in red and orange colors in Fig. 3) show low green shares. This inverse relationship is even more pronounced when looking at green space per capita.

4.2. The city rankings

As variable as the parameters on urban green in relation to the different spatial units are, one thing remains indisputable according to our city ranking analyses: The city of Siegen can claim to be the greenest large city in Germany (cf. Figs. 4 and 5). In every ranking except for one, this city is ranked number one (green space shares for administrative/ buffer/morphologic units: 78.4%/83.4%/65.6%; green areas >1 ha: 77.3%/82.6%/63.9%; green area per capita: 906 m^2/ 1526 m^2/ 482 m^2; green areas per capita >1 ha: 894 m^2/ 1511 m^2/ 470 m^2). Siegen is second only in the ranking of the morphologic units on green areas >1 ha, with Bergisch-Gladbach slightly overtaking Siegen with 64.0% compared to 63.9%. capita.

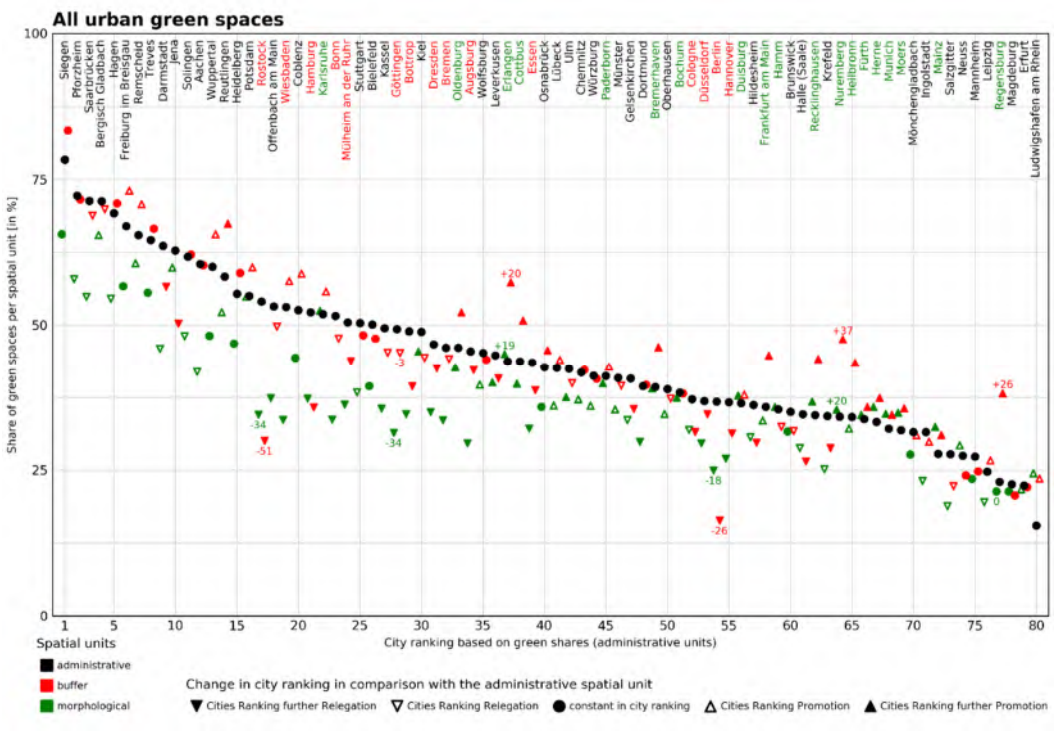


Fig. 4. Proportion of green spaces in German cities at three different spatial units: administrative, buffer and morphological. The cities are ranked along the x-axes for administrative units and deviations in the ranking are marked for the buffer as well as the morphological units. City names are classified by above average rank loss (red) and above average rank increase (green) relative to administrative units.

The situation is more ambiguous with regard to the last place in the ranking. Here, the cities of Ludwigshafen (administrative units: green space shares 23.6%; green areas >1 ha 13.5%; green areas per capita >1 ha: 67 m^2), Munich (administrative units: green areas per capita: 74 m^2), Berlin (buffer units: green space shares: 16.4%; green area per capita: 24 m^2; green areas >1 ha: 13.6%; green areas per capita >1 ha: 20 m^2; morphologic units: green area per capita: 54 m^2; green areas per capita >1 ha: 48 m^2), and Salzgitter (morphologic units: green area shares: 18.8%; green areas >1 ha: 16.6%), alternate.

In general, we can state the range of green space shares in German cities is very uneven. To pick just two examples: the range of green area shares between 1st and 80th place is 62.9% at the administrative level, on the morphologic unit it is less with 46.8%, but still remarkably high. The tangible effects on the individual person become clear in the second parameter: the range between 1st and 80th place for green space shares per capita are 832 m^2 at the administrative level, on the morphologic unit it is with 428 m^2 again less. While this is an indication that the figures of the morphological delineation are more comparable due to the data-driven approach, we can state that green space provision across German cities varies immensely.

Figure 4 illustrates the rankings among all 80 cities for the green space shares at administrative reference units. It reveals deviations in the ranking based on the alternative morphologic as well as on the buffer units.

Basically, it can be stated that the ranking for the different spatial units varies. It is interesting to note that 34 cities register strong changes in the ranking – 15 cities generally fall back in rankings, while 19 cities rise. For example, the city of Wiesbaden ranked 19 for green space shares on administrative spatial units (53.1%) drops by 35 ranks on morphologic units (33.6%). Or the city of Recklinghausen ranked 62 on this parameter on administrative units (34.4%) gains 27 ranks with morphologic units (36.7%). Without judging or evaluating what is true, these examples illustrate that statistics can make cities rank much better or worse depending on the calculation method.

We observe the same effects regarding the ranking based on the green spaces per capita (Fig. 5), here 35 cities register strong changes in the ranking: An extreme example is the city of Erfurt. On an administrative spatial unit, it offers over 300 m2 of green space per capita. On a morphological unit, however, it only has 114 m^2 and falls 25 places (from 37 to 62) in the ranking. Vice versa, Regensburg is an example in the opposite direction: on morphological units this city gains 31 ranks (from 74 to 43) and increases its per capita green share from 136 m^2 to 160 m^2 compared to the statistics on the administrative level.

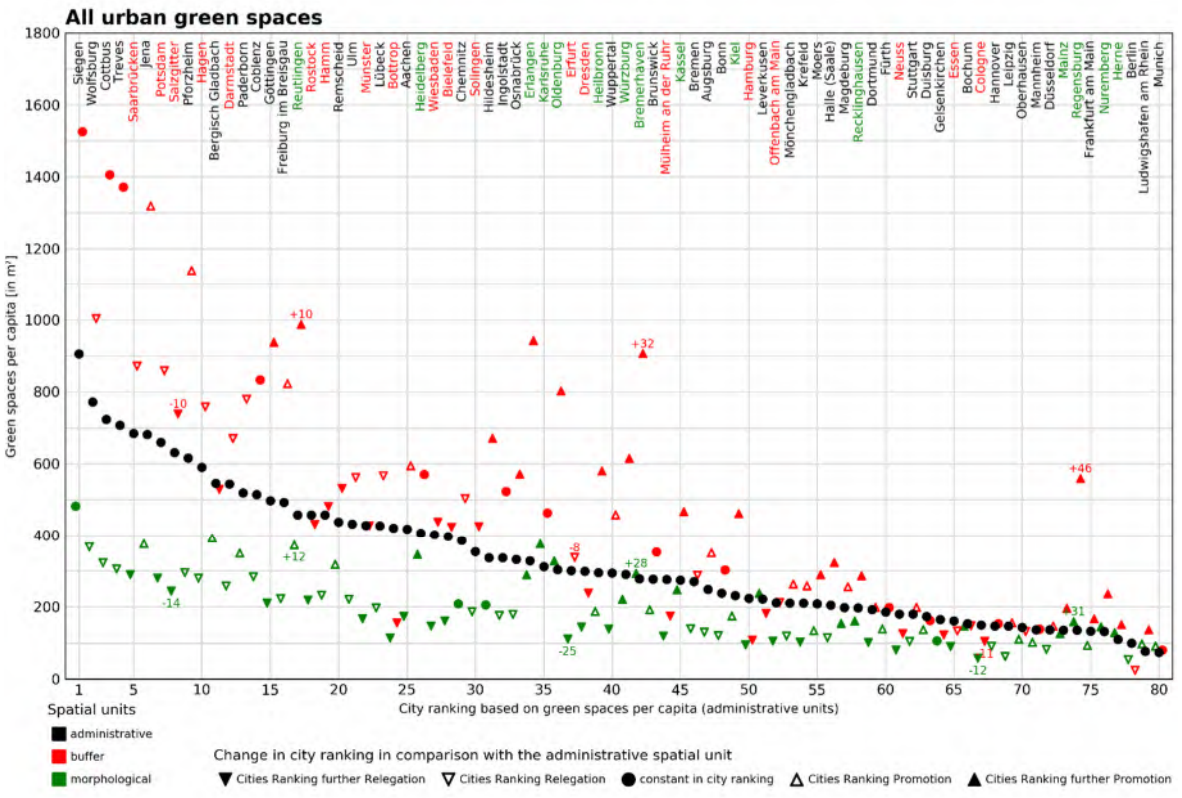


Fig. 5. Proportion of green spaces per capita in German cities at three different spatial units: administrative, buffer and morphological. The cities are ranked along the x-axes for administrative units and deviations in the ranking are marked for the buffer as well as the morphological units. City names are classified by above average rank loss (red) and above average rank increase (green) relative to administrative units.

In the Appendix, Fig. A analogously presents the ranking for green spaces larger than 1 ha and Fig. B illustrates the ranking for green spaces larger than 1 ha per capita.

In a final attempt to create one final ranking that is as holistic as possible to abstract and simplify this complexity in rankings, we combine all the ranks of the 12 individual city rankings (Fig. 6). Thus, with regard to the two parameters ‘green space shares’ and ‘green space provision per capita’, the two variants according to spatial extent ‘all green spaces’ and ‘green spaces larger than 1 hectare’, as well as the three spatial reference units, it can be said that Siegen is undoubtedly and unambiguously the greenest of all large cities in Germany. This position is also reflected in the gap to the second place (Pforzheim). Ranks two to eleven are occupied by cities such as Treves, Saarbrucken or Potsdam – all of which tend to have smaller population numbers in our sample. At the lower end of the ranking we find many of the larger cities such as Berlin, Munich, Cologne or Hamburg.

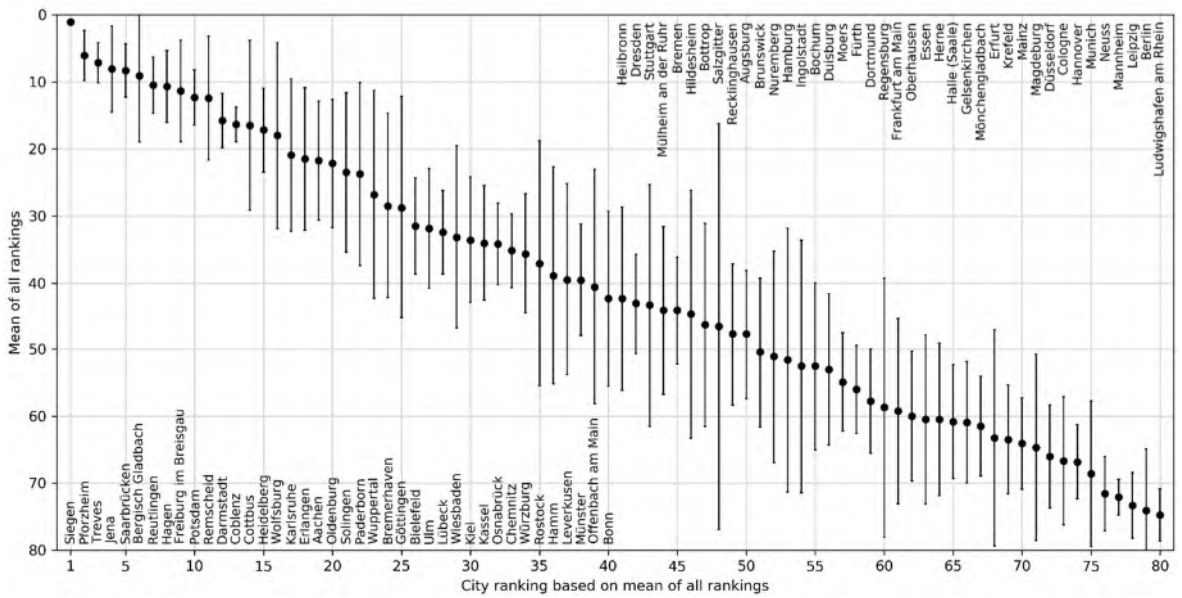


Fig. 6. Final ranking of the 80 largest German cities with respect to green area shares from the combination of all rankings and their standard deviations.

Within this relational field generated from twelve rankings the probability of being subject to method-specific coincidences is lower, but the standard deviations indicate related uncertainties. For some cities the standard deviations are low or relatively low (especially at the top as well as at the lower end of the ranking). With it their positioning is generally stable and less guided by spatial statistical coincidences (e.g. Siegen, Coblenz or Mannheim). In contrast, some cities (e.g. Salzgitter, Rostock or Regensburg) feature very high standard deviations and their positioning is prone to methodological specifics. This can be taken as an indication that the local specifics, i.e. e.g. spatial patterns of the city or city sizes, must be particularly considered here for a precise estimation.

5. Discussion

In order to see one's own, comparison is required. One way of assessing oneself is within a relational system by means of a ranking. So, whether a city is well or poorly equipped with green spaces is not easy to assess in absolute terms, but to classify in comparison. However, there are a wide variety of challenges and pitfalls to permissible and meaningful comparisons. We now critically review methods and data used to compile our rankings, we reflect on rankings in general and suggest a strategy for more scientifically sound rankings, and finally we discuss the geographical results of this study.

5.1. Reflections on methods and data

As we have seen in this study, the choice of methods and data in fluences the resulting rankings to a not insignificant degree. In this context, we discuss the applied parameters, spatial units and accuracies of the input data.

In terms of *parameters*, we have relied on two accepted measures, 'green space shares' and 'green space shares per capita', each also ac cording to 'all green spaces' and to 'green spaces

larger than 1 hectare'. This decision is not based on the fact that these parameters are the most relevant, but rather due to the need for consistent, comprehensive and appropriate geoinformation. We are aware that these parameters can be expanded almost arbitrarily (e.g. green volumes, green space types, or quality), if the data situation allows, and that this would even create more variability in rankings. It also remains to be considered that the quantity of green spaces measured in this study, does not equal quality. In this sense, we must be aware that it is not only quantity but also the respective spatial patterns which have effects on ecological or social functions, i.e. whether green spaces are more spatially clustered, large, and contiguous or more distributed, small and dispersed.

In terms of *spatial units*, the administrative unit is crucial, as it forms the spatial basis for planning and political decisions. At the same time, the administrative entity distorts the statistics because these spatial reference units are not consistent across cities. For this reason, we have introduced morphological units. These are based on a data-driven delineation of the built environment from a remotely sensed classification. Therefore, we consider them to be a consistent spatial baseline for comparing cities. The built urban space is here predominantly more narrowly delineated around the built urban environment than administrative units and, hence, the actual amount of green space available within the built city can be determined. However, it is also clear that this approach simultaneously negates green spaces in the close vicinity of the built landscape. And, as a third, supplementary spatial unit, we have added a constant area in terms of a buffer unit of 10 km around the city center. We believe, this unit allows to determine green space shares beyond the built space and to compare them permissibly on the basis of the same extent for all cities. However, the buffer unit is very artificial because it does not take the spatial pattern or size of the city into ac count. Nevertheless, this unit adds an additional perspective on green space shares in space that can contribute to the overall view as done in the merged ranking. In general, we should also be aware that these two spatial units, morphologic and buffer units, could also be permuted by other radii or shapes, or by alternative spatial definitions of city centers. Effects of these variants have been shown (e.g. James et al., 2014); however, the effects to our rankings have not been tested in this study and must remain unknown. This once again demonstrates in general that the operational measurement of green space shares is subject to conceptual and methodological specifics and can therefore be manipulated.

From the perspective of our *input data*, the remotely sensed data are a consistent database and feature a generally high accuracy of 93.07%, which makes our results robust. Nevertheless, we have to note that the spatial resolution of 10-m pixels does lead to challenges in specific structural environments. In areas of large parks or forest areas, our classification even achieves higher accuracies. However, in small-scale, complex structures (e.g. detached houses with small gardens), we come up against limits with spatial resolution. Here we also measure lower accuracies. Accordingly, since many areas in cities have this complex structural composition, it can be assumed that certain measurement inaccuracies show through and have an unknown, although probably not significant, effect on our ranking.

5.2. Reflections on the rankings

Our results of various city rankings with respect to green spaces show, first of all, that there is *no absolute truth*. There is no one right parameter or no one right spatial reference unit. We argue here that achieving a higher reliability in benchmarking a city within a relational system is only possible through manifold methodological variants. And still, although the various empirical approaches presented are systematic, it is not possible to produce a simple or universally valid ranking. And it is precisely this that exposes pitfalls of rankings. In scientific literature, city rankings are often criticized, from an epistemological point of view as well as from the ideology behind them (White & Kitchin, 2021). Therefore, as critically discussed by Acuto et al. (2021), any ranking must be critically questioned in terms of how it came about, who makes it, what it represents, and who it influences.

The challenge that despite a uniform data basis and consistent calculations, rankings (in our case on green shares) vary due to the different choice of parameters or spatial reference units is shown in this work. This explains to some extent why there is neither a scientifically approved benchmark for urban green spaces nor are there generally accepted survey and measuring methods. In Germany, only some orientation values, parameters and indicators are defined for individual urban districts without general validity (BBSR, 2018). And yet, ac cording to Kitchin, Lauriault & McArdle (2015), rankings are attractive to practitioners because they offer a semblance of technocratic objectivity and satisfy the need for political expediency. And, rankings have become a central tool of intercity competition.

Therefore, this work intends to contribute to the objectification of rankings. We believe that the presented multi-layered ranking allows a more objective and stable positioning of cities within a relational comparison system. Through conceptual and methodological diversity, it makes rankings less susceptible to corresponding randomness and the representation of uncertainties allows to overcome absoluteness in rankings. Thus, this multi-layered approach can indicate, via un certainties, whether local specifics need to be considered when develping interventions strategies. But we must be aware, that even this multi-layered approach cannot create complete unambiguity in the ranking. The combination of twelve variants in this work is thus intended to provide a blueprint that can be extended at any time to include additional green parameters.

The combination of the various rankings allows us to undoubtedly designate Siegen as *the greenest city in Germany*. This is first of all a confirmation of a similar result of a methodologically different study a few years ago (Morgenpost, 2016). However, only one parameter and one measurement method were used in their study. While the same study provides an equally clear result regarding the last place (Ludwigshafen), we paint a more complex picture here due to the un certainties (Ludwigshafen, Munich, Berlin, Salzgitter). This means that secondly, in our study the challenge and the potential lies rather in the detail of data interpretation. Whether Siegen has 65,6% (morphologic) or 83,4% (buffer) green space shares or whether it provides 482 m^2 (morphologic) or even 1526 m^2 (buffer) green spaces per capita – so in fact very different values that can be attributed to a city – depends on the spatial reference unit applied. These variabilities in numbers reveal the complexity of a reasonable assessment of the situation and a good strategy of evaluation cannot emerge by a more or less randomly chosen approach, but only in a differentiated view: an evaluation

that reflects conceptually and methodologically and co-evaluates local specifics in individual cases to derive appropriate recommendations for action.

And in this sense, benchmarking can be a valuable tool to develop and address constructive interventions and to contribute to the strategies of local governments (Robin, 2021). Whether and how local actors can translate this knowledge into action and to what extent rankings in general and this one in particular help to define interventions must, however, remain open at this point.

5.3. Geographical reflections

From a geographical point of view, we can generally state from our analysis that German cities are well equipped with green spaces. When we relate our results to the suggestions of the WHO calling for a mini mum of 9 m^2 of urban green spaces per capita and defining 50 m^2 per capita as desirable (WHO, 2010), we see that even the cities at the bottom of the ranking are (at the administrative unit: Munich 74 m^2 and at the morphological unit: Berlin 54 m^2) equipped with even more green space per capita than desirable. For the buffer unit, the green area space per capita for the ranking tail end Berlin is 24 m^2, which is about half of what was suggested as desirable. Herein this particular case we have to consider that the 10 km radius for Germany's largest city only covers the core city. Based on its green space shares on the morphologic (54 m^2 per capita) as well as the administrative unit (100 m^2 per capita), we see the proposed minimum is far exceeded.

With respect to the four example cities from the introductory section – Berlin, Bonn, Halle and Hanover – which have branded themselves by the adjective 'green', we observe something somewhat surprising: none of the three cities is in the top 10 of our ranking or even close to it. Bonn is at rank 41 in the combined ranking, Halle at 65, Hanover, the self-proclaimed greenest city in Germany, ranks 74th, and Berlin even at the penultimate position 79. Thus, the question arises whether this attribute for branding is by any means justified. These claims can be attributed to the practice of '*greenwashing*' (e.g. Seele & Gatti, 2017), i. e. a form of marketing in which e.g. green statistics are used to convince the public that, in our case, the city stands out in particular. When relating our results to the marketing use, it seems that whether a city is accepted as green in the public perception seems to depend less on objective spatial shares and more on what people accept as plausible. In this sense, Bok (2021) points to the powerful influence of rankings in defining and narrating the normative conditions of discourse and practice. This is, of course, accompanied by the need for objectification.

In our rankings, we have basically assumed that a higher share of green spaces is related to a higher likelihood of a positive effect (Arlt et al., 2005; Gill et al., 2007). In this rationale, the purely quantitative approach in this study is justified, but of course, reality is more complex. We should be aware, that the distribution of green spaces reflects the importance of the local. The impact of green areas is on the space that encompasses everyday life or on cooling effects in the immediate vicinity, and that is something quite different than measured at administrative (or other) units defining the city (Löw, 2008). In addition, we should be aware that quantitatively fewer green areas do not necessarily have fewer positive effects:

with suitable spatial arrangement or different types of green areas, climate-relevant effects may be achieved with less spatial quantity. For compensation in the social sphere, high quality and accessibility can be more important than quantity in certain cases. Last but not least, it must also be borne in mind that in the struggle for urban space, never only one parameter (in our case green areas) must be considered whether a city is liveable. This is always linked to the need for housing, infrastructure, culture, and much more. The fact that many of the larger cities such as Berlin, Munich or Hamburg are at the bottom of the ranking also makes the following clear: these cities are highly sought-after residential locations in Germany with highest influxes or highest (and increasing) housing prices. Their attractiveness and the perception of qualities of place and life do not seem to be influenced in the least by the lowest proportion of green space among German large cities.

6. Conclusion

The question "*what is the greenest city of Germany?*" seems to be trivial. Nevertheless, behind this supposedly simple question lies a deeper complexity. True to the motto "*don't trust statistics that you haven't faked yourself*", our different experiments on the proportion of green spaces in cities are not 'fakes', but transparent in their parameters, data and methods, and still nothing is black or white. Based on conceptual definitions, applied parameters or spatial units of measurement, a myriad of statistics which are all mathematically correct, can be produced to steer the impression one or the other way.

With this study we show how fragile spatial statistics are, but how they can then be (mis-)used for political decisions without considering the complexity of reality. However, with manifold different perspectives – according to parameters and spatial units – we believe we can contribute to a more holistic understanding for systematically approaching these issues. We argue here in the line of Derudder & van Meeteren's (2019) call to engage with '*urban science*': for expanding this multi-layered ranking including related uncertainties in future scientific work by experimenting with methods and technology-based quantitative spatial analysis to better assess current urban conditions. The addition of alternative spatial approaches such as green components in the immediate surroundings of households (Wüstemann, Kalisch & Kolbe, 2017) seems to be just as relevant as systematically integrating further parameters such as urban green accessibility, size, edges, volumes, spatial patterns or connectivity, among others to extend the suggested multi-layered concept of rankings.

As spatial thinking forces to think in difference, we believe being able to describe and understand these differences between approaches as well as between cities in their complexity in a more holistic sense. We see this as a starting point for thinking about urban developments. With this basis of good empirical knowledge within a relational system, the necessity for constructive interventions can be identified, examples of others can be studied and planning attempts can be made to overwrite or to adapt the existing structures of a city.

Acknowledgements

This study was partly funded by the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) (project “Monitoring des Stadtgrüns), the German Federal Ministry of Transport and Digital Infrastructure (BMVI) (projects “meinGruen” and “SAUBER” (funding codes 19F2073B and 19F2064B, respectively)), and by the German Federal Environmental Foundation (DBU). We would also like to thank Julia Tenikl and Lennart Imberg for their support.

Appendix

Supplementary data to this article can be found online at

https://doi.org/10.1016/j.compenvurbsys.2021.101687

Part V

Urban patterns

CHAPTER XV	Seven city types representing morphologic configurations of cities across the globe	349
CHAPTER XVI	Planned, unplanned, or in-between? A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe	375
CHAPTER XVII	Patterns of Eastern European urbanisation in the mirror of Western trends – convergent, unique or hybrid??	405
CHAPTER XVIII	Measuring the spatial hierarchical urban system in China in reference to the Central Place Theory	427

Chapter XV

Seven city types representing morphologic configurations of cities across the globe

Cities, 2020, vol. 105, 102814

Hannes Taubenböck, Henri Debray, Chiunping Qui, Michael Schmitt, Yuanyuan Wang & Xiaoxiang Zhu

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.cities.2020.102814

Received 22 Jan 2020; Accepted 31 May 2020; Published 9 June 2020

Abstract

What we understand by the simple term ‘city’ is in fact describing highly diverse domains: different economies, demographics, ways of living, land uses, built-up morphologies, among other things. The built landscape alone ranges from low-density, one-storey suburban settlements to high-density accumulations of skyscrapers. Models have repeatedly attempted to describe these various ‘city’ manifestations and to understand the processes that shape these spatial appearances and patterns. In this paper we analyze the morphological-spatial configurations of urban landscapes. We empirically examine 110 cities distributed around the globe. By using the Local Climate Zones (LCZs) classification scheme, we quantitatively describe morphologic variances of the built landscape within cities. We find seven city types (clusters) that capture the global diversity of spatial urban configurations. These seven types testify in parts to common geographic-cultural spaces. Some are largely congruent with wellknown spatial units such as Europe or the Islamic world. In contrast to theoretical city models, however, we also find clusters that are more spatially complex such as African-American or Asian-African clusters. On the one hand, the study confirms that similar cultural, socio-economic, demographic or political conditions in fact do produce similar morphologic-spatial urban configurations. On the other hand, it also shows that there exist similar morphological configurations across geographic-cultural spaces.

Keywords

City models; remote sensing; local climate zones; urban morphology; patterns; comparative urban research

1. Introduction

We perceive the city first and foremost through its physical form. When visiting a new city, its spatial structure and its interrelationships appear at first inscrutable. Nevertheless, we expect an underlying spatial order which follows a more universal structural configuration of urban space. A configuration we can understand within the framework of previous experiences. An order that adheres to certain basic principles – a dense, multifunctional city center, an arrangement of subcenters of commerce and industry spread over the city, different densities of living forms, and a generally decreasing density with increasing distance to the center.

Urban geography has long endeavored to uncover rules and general principles of order (e.g. Bettencourt, 2007) and to develop ideal structural models to describe the spatial patterns of cities. As early as the Chicago school, the general spatial order of cities has been conceptualized by observations and theories. Cities were spatially modelled in concentric zones (Burgess, Park & McKenzie, 1925), sectors (Hoyt, 1939), or multiple nuclei configurations (Harris & Ullman, 1945). The various models created over the years with a claim to universal validity generally refer to social-ecological space, to transport costs, land market and land use, to population density, to functional interdependence, central-local network theory or to diffusion (Hofmeister, 1991). The 'Alonso-Mills-Muth' model (Alonso, 1964), as example, has been established with a monocentric character referring to the even distribution of land use within a city as a function of land rents, which decreases with increasing distance from the center (Paulson, 2012). This monocentric character, however, has been questioned by scholars, since a constantly decreasing density with increasing distance to the center no longer adequately reflects today's complex polycentric city patterns (e.g. Adolphson, 2009; Siedentop et al., 2003). In this regard the "Urban Realms" model tried to capture the simplified representation of the spatial, social, economic and cultural pattern in monocentric models by realms representing more complex polycentric patterns (Vance, 1964). Models that reflect this higher complexity refer to, for instance, effects of economic externalities such as traffic congestion (Anas & Kim, 1996; Fujita & Ogawa, 1982) or indirect utility functions with economic activity to cluster in several interacting centers (Anas, Arnott & Small, 1998; Roca Cladera, Marmolejo Duarte & Moix, 2009; Solow, 1973). Edge cities have been described, reflecting spatial concentration of office and retail space, often in conjunction with other types of development, including residential, at the nodes of major express highways (Bontje & Burdack, 2005; Garreau, 1991). The unbridled dispersion of economic activity outside the centers as a whole has continued leading to patterns described even as 'beyond polycentricity' (Gordon & Richardson, 1996), often resulting in forms of low dense urban sprawl (Nechyba & Walsh, 2004).

In addition to these more or less universally valid urban structural models, attempts were made to reflect the relationship between global economic forces and the local cultural context. Spatial models are intended to reflect the geographical-cultural diversity by regionally specific features of the urban configuration at the inner-city level (Pacione, 2009). Models for the European (e.g. Lichtenberger, 1972), the US American (e.g. Hahn, 2014), the Ford-Griffin-Model for Latin America (Griffin & Ford, 1980), the oriental Islamic model

(Ehlers, 1993), the African model (UN, 1973), the Chinese city (Gaubatz, 1998), among many others have been conceptualized. Further spatial refinements to urban forms within these large, often continental regions have been developed: For Africa, as example, models for the indigenous city (Krapf-Askari, 1969), the colonial city (Pacione, 2009), the apartheid city (Davies, 1981), among others were conceptualized. The basis for such intercultural comparisons of cities is the division of the earth into cultural areas or cultural earth parts, as it has been done e.g. by Kolb (1962) or Huntington (1997).

Modeling approaches tried to better understand the complex underlying processes shaping urban patterns. The complexity is also reflected in the fact that, like DNA, each city developed its own unique configuration. The most obvious configuration is the morphologic-spatial structure. It is the direct physical implementation of a complex interrelation of economic, social, demographic, legal, and political systems (Tonkiss, 2013), historical path dependencies (Kostof, 1991) and location factors such as topography or climate (Taubenböck, 2019). This 'physicalism' (Batty, 2013) is, on a very generic level, composed of buildings, streets, plots, and open spaces (Kostof, 1991; Salat, 2011). This urban form is designed so differently across the globe that two ends of a morphologic-spatial continuum may be marked by a planned, geometric, regular, low dense arrangement of suburbs with high vegetation fractions and open spaces on the one hand, and informally shaped, organic, irregular, high dense arrangements of shacks in slums basically without open, public space, on the other hand.

However, empirical studies comparing the urban spatial configuration of cities across the globe are few due to data scarcity or data inconsistency. An urgent need for a comprehensive database worldwide that goes beyond an urban footprint analysis with the need to capture the internal structure of cities is formulated (Bechtel et al., 2015). Remote sensing from space is one data source that has the capability to increasingly reduce these data gaps. In the recent past, global mapping products based on Earth observation data with ever better spatial resolutions have emerged (e.g. Esch et al., 2012; Small & Sousa, 2016) and related data-driven urban analysis have been conducted predominantly on urban growth (e.g. Angel et al., 2011; Taubenböck et al., 2012). First models have been developed that allow evaluating the spatial dispersion of urban landscapes on this new empiricism (Taubenböck et al., 2019). However, at continental or even global scale these analyses are still limited in a thematic sense due to their binary representation of the settlement space. Existing works that analyze and compare the internal morphologic structure of cities using EO-data are so far mostly based on a comparatively small number of cities (e.g. Seto & Fragkias, 2005; Cai, Huang, & Song, 2017; Taubenböck et al., 2017; Wurm et al., 2019) and the specifications of the needed or used input data rarely allow, for several reasons, a global approach. Thus, studies that analyze the intra-urban structure of cities or the conceptualized models of cities in a global empirical comparison on the basis of a consistent data set do not exist so far.

Another challenge for comparative urban research derives from a rather arbitrary use of geographical boundaries for cities or urban landscapes which makes many comparisons inadmissible (e.g. Lechner et al., 2013; Openshaw, 1983). It has been shown that the commonly used conventional, historical administrative spatial units do represent cities in a world characterized by highly dynamic urbanization less and less (Taubenböck, et al., 2019).

As a result, geographical knowledge is often distorted by those incomparable spatial baselines.

Against these backgrounds, we aim in this paper for a global empirical analysis on the intra-urban structure of cities using consistent data with comparable spatial baselines. To do so, we characterize the intra-urban morphologic-spatial composition of urban landscapes across the globe. We have selected 110 representative cities. This selection is proportional to the relative global share of urban population per continent, generally selecting the largest cities on each continent and at the same time ensuring a balanced spatial distribution. We classified them according to the classification scheme of the Local Climate Zones (LCZs) (Stewart & Oke, 2012) using remotely sensed data (Qiu, Schmitt & Zhu, 2019) and we rely on a comparable spatial baseline of city extents (Taubenböck et al., 2019). With it, we specifically ask the following research questions:

• Are intra-urban morphologic configurations of cities across the globe similar or different?

• And, should they be different, do groups form according to geographical and cultural aspects?

2. Conceptualization of the study

The general idea of this study is to quantify intra-urban morphologic-spatial configurations of cities across the globe and to identify groups consisting of similar spatial configurations. This requires a clear definition of the components for analysis and comparison, i.e. we must ensure that we do not use a random *selection of cities* for our experiments, that we do not compare apples with oranges in terms of *thematic content* (i.e. structural classes defining the different morphologic compositions within the urban landscape) and, since space plays a decisive role, that the *spatial units* across cities are comparable.

For a representative *selection of cities,* we select cities whose cumulative populations are proportional to city populations on continental level. We base this on statistics from the United Nations (UN, 2018). Beyond, we ensure a balanced spatial distribution across the respective continent. Our object of investigation from the perspective of the *thematic content* is the intra-urban morphologic configuration of an urban landscape. As elements of this configuration we understand homogeneous parcels that contain a similar equipment of space of the built and non-built (non-natural and natural) environment. For this purpose, we adapt the Local Climate Zones (LCZs) scheme as it is a generic, culturally-neutral description of land-use and land-cover (Stewart & Oke, 2012). Although originally developed for studies on urban climate, the classification scheme relies on universal, standardized and measurable parameters of urban form. The classes are based on the characteristics ‘density’, ‘building size’ and ‘building height’ for the built landscape. For the non-built landscape, they are based on the theme (e.g. trees or open spaces). These “*regions of uniform surface cover, structure, material, and human activity span hundreds of meters to several kilometers in horizontal scale*” (Stewart & Oke, 2012). Figure 1 introduces the 17 thematic LCZs split into ten LCZs defining the built types of the urban landscape and seven LCZs defining the non-built types of the city. In our study we define that different LCZs reflect a different morphologic urban configuration. In consequence, we understand the morphologic urban configuration of cities

being similar or different, if the share of specific LCZs is similar or different and/or if the spatial distribution of LCZs within cities is similar or different.

For the *spatial units* defining the city extents, we rely on morphological urban areas (MUAs). These encompass the built space of cities independent of administrative/political borders. Since critics question much of existing comparative urban research because of a rather arbitrary use of geographical boundaries (e.g. Lechner et al., 2013; Masucci et al., 2015; Taubenböck et al., 2016), we do not use administrative or standardized spatial units, but a consistent delineation of urban from rural areas. We apply the MUAs as provided by Taubenböck et al. (2019). Decisive here is not whether these MUAs capture the urban space in a universal true manner, but that they delimit the urban spaces in a globally consistent manner, which provides the spatial basis for a valid comparison.

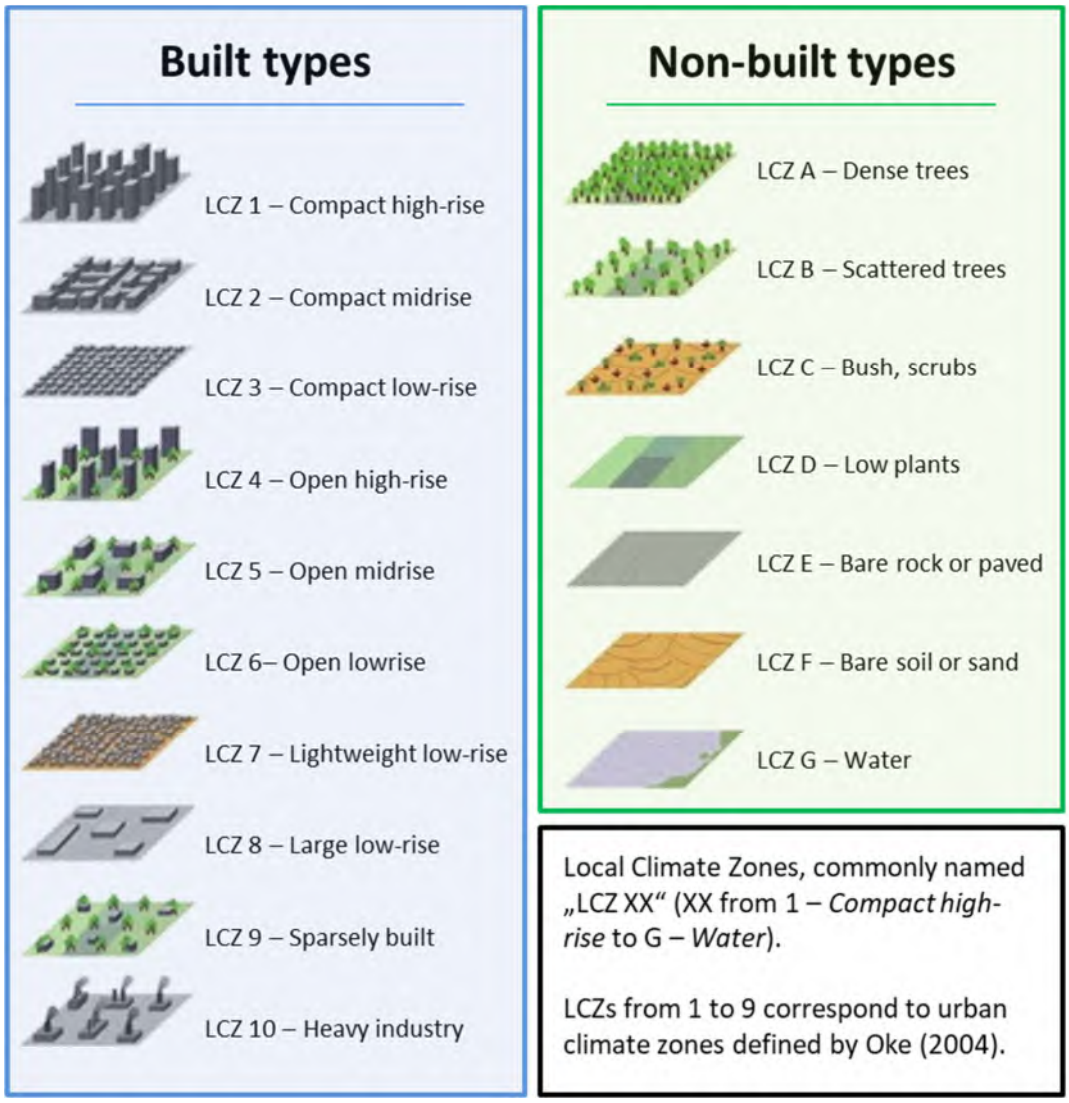


Fig. 1. The 17 Local Climate Zones (LCZs) as a generic, culturally-neutral description of land-use and land-cover (based on Steward & Oke, 2012).

3. Study sites, data and methodology

3.1. Study sites

In order to compile a sample of about 100 cities that allows a globally representative evaluation of urban configuration, we have made use of the following considerations: We calculate the relative global share of the urban population per continent using population statistics from the United Nations (UN, 2018). Then we select cities per continent, which in their cumulative number of inhabitants correspond to the relative continental share. In general, we select the largest cities of the continents; however, for a reasonable spatial distribution that ensures all regions are covered, we do not strictly follow the population numbers. However, any selection of cities is naturally not fully proportional to the urban population per continent. Thus, we define an error tolerance of 3% and choose cities that keep the error tolerance low. Based on these considerations, we selected an ensemble of 110 cities. All 110 cities are displayed in Fig. 3 and stated in Table 2.

3.2. Satellite data and classification of LCZs

For the classification of LCZs we rely on multi-seasonal Sentinel-2 satellite data (ESA, 2012). The sensor acquires multi-spectral optical imagery at high spatial resolution of 10, 20 or 60 m with 10 bands in the visible, near infrared, and short-wave infrared part of the spectrum.

A recurrent residual network (Re-ResNet) architecture capable of learning a joint spectral-spatial-temporal feature representation within a unitized framework has been exploited and trained on the LCZ42 dataset. The labels for training have been generated manually by 15 domain experts following a carefully designed labeling workflow and evaluation process over a period of six months. The labels are with an overall confidence of 85% based on a rigorous quality assessment (on technical details we refer to Zhu et al., 2019). Afterwards, the trained network is applied to perform the LCZ classification at a resolution of 100 × 100 m per city from seasonal Sentinel-2 images, followed by a majority voting on the multi-seasonal predictions (for details on the classification procedure we refer to Qiu et al., 2019). In Fig. 2 we visually exemplify the LCZ classification.

3.3. Spatial units of analysis

A fundamental challenge for any comparative urban study is the spatial unit to base the empirical investigation on. As introduced above, we apply in our study MUAs as provided by Taubenböck et al. (2019). They understand MUAs as a territorially contiguous settlement area that can be distinguished from low-density peripheral and rural hinterlands along a gradient of decreasing built-up density from the urban center to the periphery. Figure 2 underpins the spatial discrepancy between the arbitrary administrative units and the MUAs. The exemplified administrative area of Madrid and the respective MUA are very similar around the center. The administrative area, however, includes large non-built forest areas in the northwest that are not considered part of the city in the MUA. At the same time, large suburbs are not considered part of the city by the administrative area in the southwest, while

the MUA integrates these built settlement landscapes. By applying the same method to all larger cities worldwide, the provided MUAs are re-territorializing commonly used urban extents. These MUAs constitute a consistent and thus permissible spatial unit for comparative urban research by capturing the built urban space, rather than arbitrary administrative city boundaries or standardized extents.

For analyzing the intra-urban spatial distributions of LCZs, we apply a monocentric city model. Although, as discussed above, this model certainly does not correspond to the many spatial characteristics of the cities investigated here, it allows us to investigate a basic assumption of decreasing densities with a greater distance to the center. And, this model still proves empirical understandability and an explanation of an essential part (at least 80%) of the variation in urbanized land area (e.g. Spivey, 2008) that functions as a theoretical and empirical basis to evaluate intra-urban locations (e.g. McMillen, 2006). As a spatial starting point of the models we rely on the defined center points for each city as provided by the United Nations (2014). We use two different spatial approaches: For a standardized comparison we apply ring models with bandwidths of 100 m to determine the location related to the defined urban center point. For a normalized comparison we apply 100 rings of uniform bandwidths relative to the specific MUA of a city. Figure 2 exemplifies some sample rings around the defined urban center.

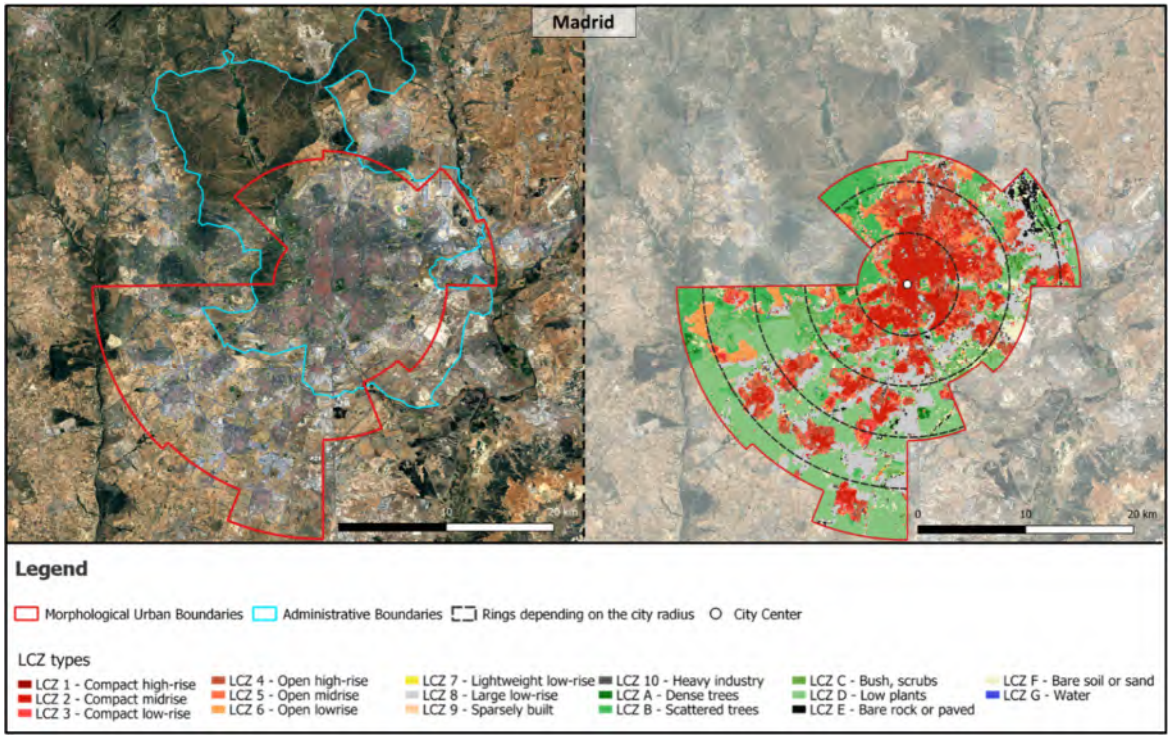


Fig. 2. The morphological urban areas (MUAs) in comparison to administrative unit for the sample city of Madrid (left) and the related classification of LCZs with few sample rings around the defined city center to visualize the spatial concept of the monocentric city model (right).

3.4. Methods for comparing landscape configurations of cities

3.4.1. Clustering of cities

We understand the LCZs as the individual building blocks that in their entirety define the morphological-spatial configuration of a city. Against the background of these considerations, we try to identify similarities and differences between urban configurations using two indicators: The *spatial proportions of LCZs* and the *spatial proportions of LCZs depending on their locations within the city*. We want to identify groups for both indicators.

For the *spatial proportions of LCZs*, we count the occurrence (in pixels) of each of the 17 LCZs per city at the spatial unit of MUAs. We apply minimum-maximum normalization to them relative to all cities. With it we produce a 17-dimensional feature space per city (we refer to it as '17f'). To take specifically account of the various city sizes, we derive the spatial share of the 17 LCZs per city from the 17-dimensional feature space and we additionally integrate the 'city size' derived from the MUAs (Taubenböck et al., 2019) as further variable. We apply minimum-maximum normalization to the 'city size' relative to all cities. With it we produce an 18-dimensional feature space per city ('18f').

For the *spatial proportions of LCZs depending on their locations within the city*, we count the occurrence (in pixels) of each of the 17 LCZs per city at the spatial unit of MUAs for each of the 100 concentrical rings of relative width using our monocentric city models. As above, we apply minimum-maximum normalization to them relative to the basic population, i.e. all cities. With it, we produce a 1700-dimensional feature space per city ('1700f'). In addition, we derive the spatial share of the 17 LCZs per ring from the 1700-dimensional feature space and we also integrate the size of the city based on the size of MUAs. We apply minimum-maximum normalization to the 'city size' relative to all cities. With it, we produce a 1701-dimensional feature space per city ('1701f').

With it four features spaces are at our disposal. This allows us to find various clusters of similar proportions of the various LCZs in the first place, and to find various clusters of similar proportions of the various LCZs with respect to their spatial location within the urban landscape in the second place. For clustering we apply two unsupervised methods: k-means (Hartigan & Wong, 1979) and expectation-maximization algorithms (Dempster, Laird & Rubin, 1977), where we assume a model of type "EII" (Spherical distribution, Equal Volume, Equal shape, without specification of orientation). Both algorithms apply centers of clusters to model the data; however, k-means tends to find clusters of comparable spatial extent; the expectation-maximization allows clusters to have different extents. Expectation-maximization can detect clusters partially mixed in the feature space; k-means, in comparison, tends to find clusters with more neighborhood-oriented logics.

In unsupervised clustering methods, the number of clusters is not known a priori. For determining the optimal number of clusters, we rely on the gap statistics algorithm (Tibshinari, Walther & Hastie, 2001). We apply 100 random launches with different seeds. To define the number of clusters for all four feature spaces (17 and 18-dimensional as well as 1700 and 1701-dimensional) we calculate the average from the 100 launches. As a result, we plot the determined particular clusters on a world map.

3.4.2. Correlation of identified clusters to geographical-cultural spaces

The clustering results (based on the four feature spaces and the two clustering methods) differ in composition and thus their spatial expression. It is not a priori clear which clusters are our target. As in this study we want to focus on whether groups are emerging that correspond or come close to predefined geographical-cultural spaces we evaluate our resulting clusters in relation to geographical/cultural regions based on those defined by Huntington (1997). He divided the world basically into the following geographical/cultural regions – Western, Latin American, Orthodox, Islamic, Sinic, Buddhist, Japanese, Hindu, African (cf. Fig. 1A, Appendix). It is perfectly clear to us that this division into regions is problematic. Distinct cultural boundaries do not exist in the present day and are oversimplifying the complexity of our globalized world (e.g. Berman, 2003). However, the underlying argument of using these geographical/cultural regions in our study is that, as assumed in many studies, similar cultural, demographic, socio-economic and political conditions produce similar spatial patterns. This is reflected by specific city models designed for different cultural circles assuming they influence the morphologic-spatial configuration of cities (e.g. Hahn, 2014; Griffin & Ford, 1980). Against this background, we want to take a closer look at the clustering that comes closest to this geographical and cultural division.

To identify this clustering, we apply the Simpson Evenness Index (SIEI) (Simpson, 1949), which defines the probability that two sub-units (in our case the cities) of the landscape (in our case the geographiccultural region), selected at random, belong to different types (in our case the clusters). The SIEI is defined as

$$SIEI_{geographic/cultural\ zone_i} = \frac{1 - \sum_{k=1}^{7}\left(\frac{number\ of\ cities\ of\ cluster\ k}{number\ of\ cities\ in\ i}\right)^2}{1 - \left(\frac{1}{7}\right)}$$

producing values from 0 to 1; with 0 indicating there is only one cluster present in the respective region and 1 indicating there is an even repartition of all clusters in the respective region. In other words, the lower the value, the higher the congruency.

3.4.3. Morphologic descriptions of clusters

Based on the analysis above, we continue this study with only one clustering result having highest congruency to geographical/cultural regions as defined by Huntington (1997). The idea now is to characterize these resulting clusters in their morphologic-spatial configuration. To do so, we apply four elements: 1) The *average size of the cities*, a parameter which was not tackled by theoretical models; 2) the *share of built* versus *non-built up LCZs*, a parameter to assess the intensity of used urban space; 3) *distributions of LCZ shares* within the respective MUAs as the main information on the structural configuration of the cities; we illustrate the internal variability per cluster in boxplots; 4) and, the *volume density gradients* from the center to the peripheries, a parameter to describe the distribution of the usage of space within the urban landscape. For the latter, we transfer the LCZs into measures of three-dimensional built-up density. We extract the building surface fraction and the height of roughness element (as provided in Stewart & Oke, 2012) per LCZ. For each grid, we calculate the approximated built-up volume by multiplying the building surface

fraction with the provided height of roughness elements. We aggregate all resulting built-up volumes from the grid level to the corresponding ring of the monocentric city model. From it we calculate the average built-up volume along the gradient from the defined city center to the peripheries of the MUAs for each of the resulting clusters. We display the resulting density gradients per cluster in two different ways – absolute and normalized. In absolute manner, we derive for each city and every ring (100 m of width) of the monocentric city model the average built-up volume. For the clusters we rescale the resulting values of the built-up volumes to 0 (minimum value) and 100 (maximum value) and plot the clusters' depending on the absolute distance to the center (x-axis from 0 to the maximal radius of the city). In the normalized manner, we derive for each city 100 individual rings varying in widths depending on the particular city radius as defined by the MUAs. From it we calculate the average built-up volume for each of the clusters and rescale the values as done above. We then plot the clusters' built-up volume depending on the relative distance to the center (x-axis from 0 to 100% of the maximal radius of the city).

We illustrate one representative city per cluster in a LCZ map, we plot the distribution of LCZs within the clusters, we describe the main morphologic-spatial characteristics of this cluster and we name the cluster according to urban geographical features.

4. Results

4.1. City clusters and their geographic-cultural expression

The gap statistics algorithm suggests in principle, for all four features spaces, the 17 and 18-dimensional and the 1700 and 1701-dimensional feature space, that on average seven clusters best group the data. As a consequence, the input data of LCZ classifications, the spatial baseline of the MUAs, the introduced methodological approach, and application onto these sample cities indicate that seven types of morphological-spatial city configurations generally allow capturing the built appearance of cities on our globe.

Based on this, we have systematically tested the different methods for clustering (k-means and expectation-maximization) onto the different feature spaces (17 and 18-dimensional as well as 1700 and 1701- dimensional) for the a priori defined seven clusters. Of course, the resulting groups are not so unambiguous that the results remain the same regardless of the feature space or method. Since we want to find out in this study whether similar groups are emerging that are similar to predefined geographical-cultural regions as proposed by Huntington (1997), we continue from here with the result that shows the greatest possible similarity to the geographical-cultural region. By weighting the Simpson Evenness Index by the number of cities in each region, we find the 18-dimensionsal feature space using the k-means algorithm provides highest congruence to the suggested geographical-cultural regions (Table 1).

Table 1. Results of the weighted Simpson Evenness Index based on the different feature spaces (17-, and 18-dimensional features (17f, 18f), and 1700-, and 1701-dimensional features (1700f, 1701f)) and the different clustering methods (k-means, expectation-maximization).

Method used	Weighted SIEI
k-Means & 17f	0.81
k-Means & 18f	**0.77**
k-Means & 1700f	0.82
k-Means & 1701f	0.82
Expectation-maximization & 17f	0.79
Expectation-maximization & 18f	0.83
Expectation-maximization & 1700f	0.87
Expectation-maximization & 1701f	0.82

This empirical analysis according to physical equipment and spatial distribution for cities on our globe shows that whether we consider only the proportions of the LCZs or whether we consider the proportions of the LCZs as a function of their relative position relative to the center, or whether we apply k-means or expectation-maximization as clustering methods, we find slightly different spatial compositions of the seven clusters. Nevertheless, we generally also see the geographical-cultural clusters remain comparatively constant. Relative to our best result (k-means, 18f), we observe the greatest match of classes if we leave the feature space the same. Applying the expectation-maximization algorithm on the 18-dimensionsal feature we have an 80% match in the class assignment of the cities (cf. Fig. 2A, Appendix). But even if we vary the feature space (k-means, 1701f), compared to our best result (k-means, 18f), there is still a 60.9% agreement in the class assignment of the cities. For the latter example, still 66 of our 110 sample cities remain in the same cluster constituting similar geographical-cultural regions.

If we plot the best results on the world map (18f and k-means), we can see certain spatial distributions of the clusters, which obviously represent in parts geographical-cultural spaces (Fig. 3). This is a first clear indication that the geographical regions, their cultural and historical background do influence the morphologic-spatial appearance of cities.

In a first step, we describe the geographical distribution of the identified clusters. The morphological specifics of the identified clusters are introduced below in Section 4.2.

Cluster 1: mainly **Asian and African cities**

Of the 22 cities in this cluster, 19 are located in Africa and Asia (10 in Africa, 9 in Asia with 5 of them in India). Many cities have a colonial or dominion heritage. It should also be mentioned here that also two cities from South America and one from Europe are belonging to this cluster.

Cluster 2: mainly **Asian and American cities**

Of the 15 cities in this cluster, 9 are from Asia and 5 are from America. The only exception here is Melbourne in Australia.

Cluster 3: dominated by **European cities**

21 out of 29 cities belonging to this cluster are on the European continent. Of the 22 European cities that are part of our entire sample, 21 belong to this cluster (the only exception is Bucharest belonging to Cluster 1). It is also worth mentioning that contrary to our expectations some cities like Addis Ababa in Ethiopia or Wuhan in China also belong to this cluster.

Cluster 4: mainly **Eastern African and Eastern Asian cities**

Of the 20 cities in this cluster, 10 are located in Africa. It is striking that all these cities are located in (south-)eastern parts of Africa. 8 cities are located in Asia, mainly in far Eastern parts and one city (Port-auPrince) belongs to America and one (Sydney) to Australia. It is interesting to note that most of the cities here do have a colonialist legacy.

Cluster 5: dominated by **cities of central Africa**

8 out of 10 cities of this cluster are located in Central Africa - distributed from west to east. Just one from Oceania (Port Moresby) and 1 from Asia (Poona) add to it. Again, most of the cities here do have a colonialist legacy.

Cluster 6: Cities of the Islamic world

It is noteworthy that all cities in this cluster are exclusively located in the Islamic region from Ouagadougou to the Hindu Kush. 9 cities (5 in northern parts of Africa and 4 in the Middle East) define this cluster. In this region, two cities part of our sample – Bagdad and Riyadh – do not belong to this cluster.

Cluster 7: consisting of **very large cities**

This cluster contains only 5 cities. All of these five cities are mega cities with four of them located in Eastern Asia. The remaining one (Los Angeles) is from America. They all are very large cities and assume that the share of LCZs could be characteristic for these mega cities.

All in all, we draw the following interim conclusion from this empirical analysis according to physical equipment and spatial distribution for cities on our globe: In many theory-based studies, city models are strictly constituted according to more or less accepted geographical spaces. This empirical analysis, however, shows that these geographical spaces represent a unit that is too simplified and too generalized. Here historical path dependencies such as colonial influences or economic situations overlap with many other factors and influence the constitution of the built city. But it remains undisputed that geographical-cultural spaces are reflected in the morphological configuration of cities (and yes, outliers do exist). The *European city model* (Cluster 3) and the *Islamic city model* (Cluster 6), as presented in many theoretical models, are clearly identified by this empirical investigation. City configurations that apply mainly to *Central Africa* (Cluster 5) could be localized. However, other clusters are found that cannot be assigned so clearly to only single geographical-cultural area.

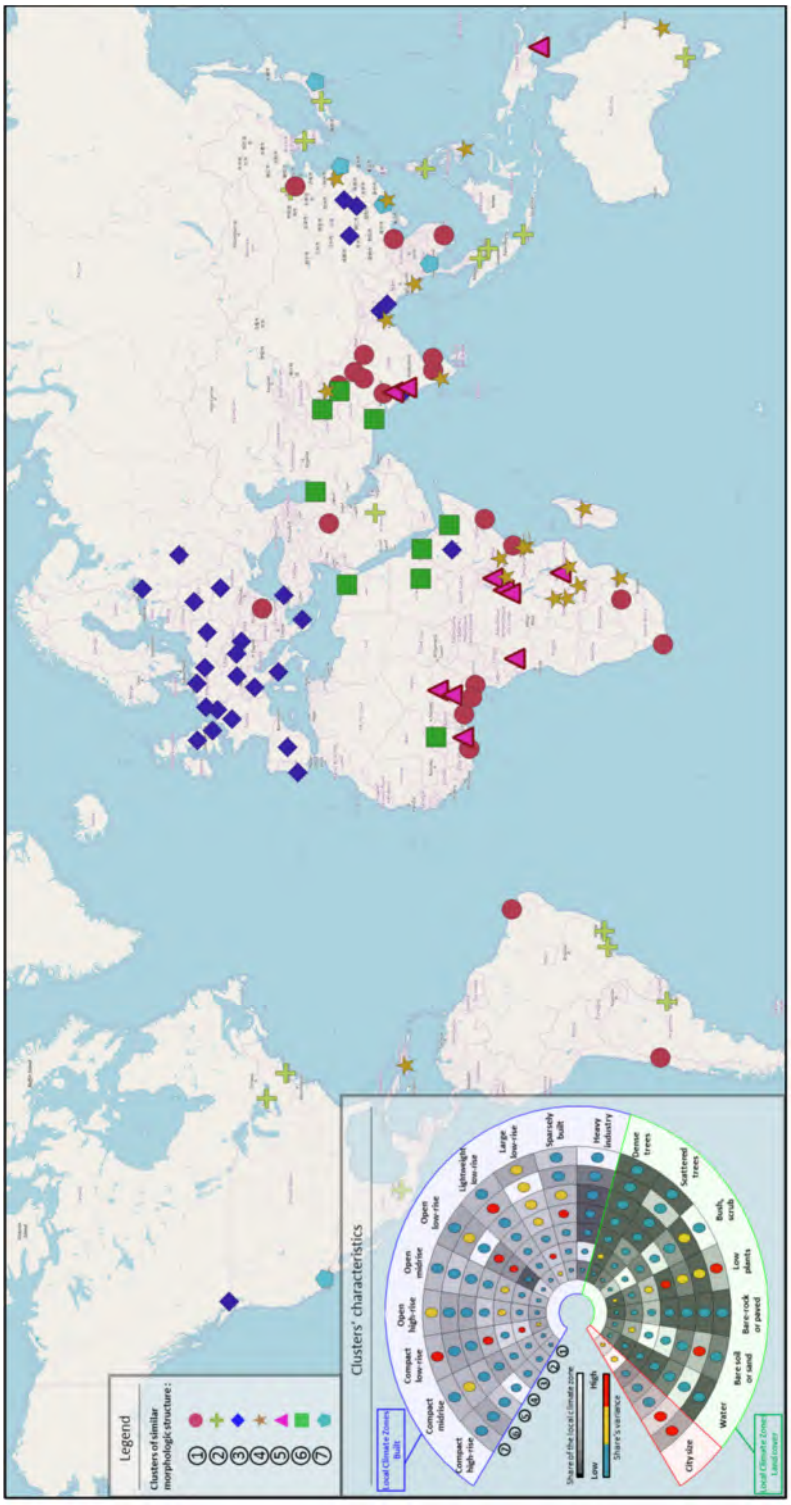


Fig. 3. Seven resulting clusters based on the k-means algorithm using the 18-dimensional feature space (consisting of 17 LCZs as well as city size as features) at the spatial unit of the morphological urban areas. The seven clusters are presented in their global distribution and by their cluster characteristics: For the latter, we compute average and standard deviation of the share of the 18 features. We divide the diagram in seven rings, one for each cluster. We divide the rings in 18 radials sectors, one for each feature. Thus, we obtain 18*7 cells, each related to one cluster and one feature. On each cell we plot a gray shade related to the share of this individual feature in the cluster: white is high, black is low. On each cell we plot a colored pastille related to the standard deviation across the cities in the respective cluster of the share of this individual feature. Red is high, blue is low. We arrange the clusters in this diagram so that the closest neighbors in the diagram are also the closest neighbors in the feature space. This means fewer differences between clusters that are neighbors in the diagram.

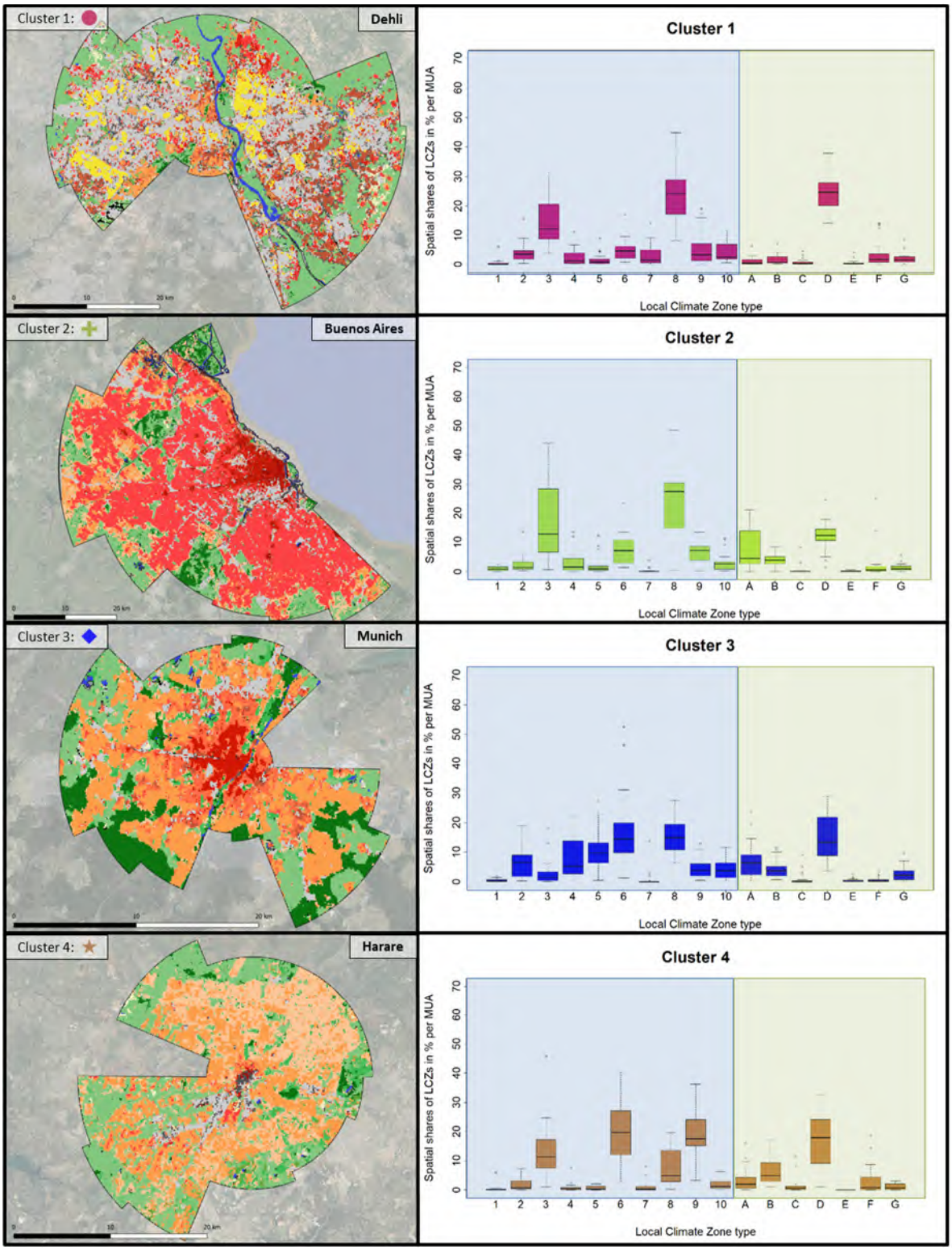


Fig. 4. The seven resulting clusters based on the k-means algorithm using the 18-dimensional feature spaces consisting of 17 LCZs and city size at the spatial unit of the morphological urban areas.

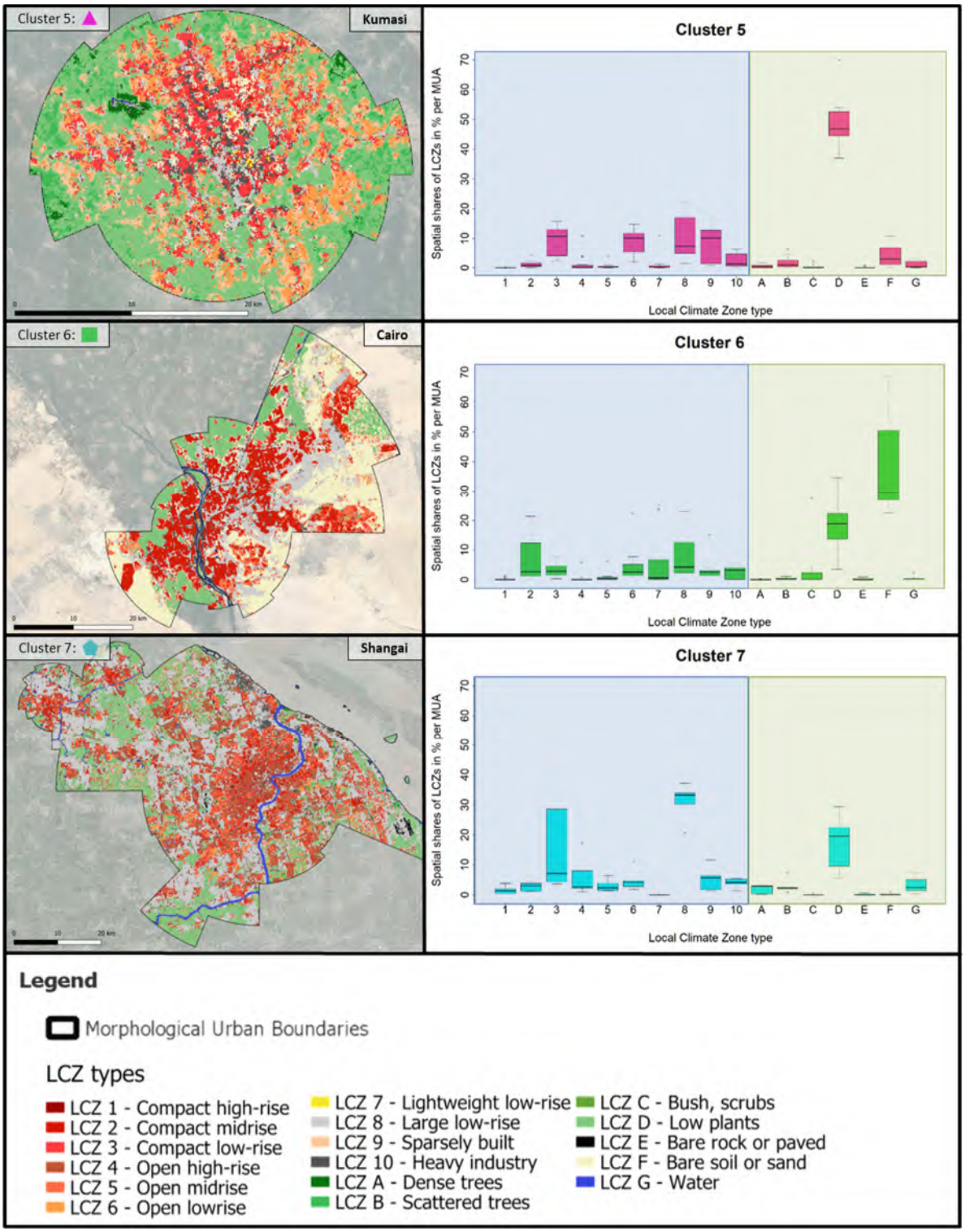


Fig. 4. continued.

Clusters are identified whose similar morphologic-spatial urban characteristics appear predominantly in cities located at different continents: *Asian and African* (Cluster 1), *Asian and American* (Cluster 2) cities. For the first case –Asian and African–, a further cluster is identified which takes a clear spatial concentration on the eastern parts within the particular continents (Cluster 4). Despite some geographical outliers of these clusters, we find there are cultural spaces that seem to have a clear influence on the morphologic-spatial configuration of cities (Figs. 3 and 4).

In the following we present the morphologic specifics of the seven clusters based on the 18-dimensional feature space and the k-means algorithm. For the morphologic-spatial characterization we apply the four elements, size of the cities, the share of built versus non-built up LCZs, the general shares of LCZs and, the volume density gradient.

In Fig. 4 we illustrate the cartographic results of the LCZ classification for one representative city per cluster as well as box plots showing the distributions of LCZs. We also illustrate the three-dimensional volume density by density gradients from the urban center towards the peripheries (Fig. 5). We do so for the standardized and for the normalized distances.

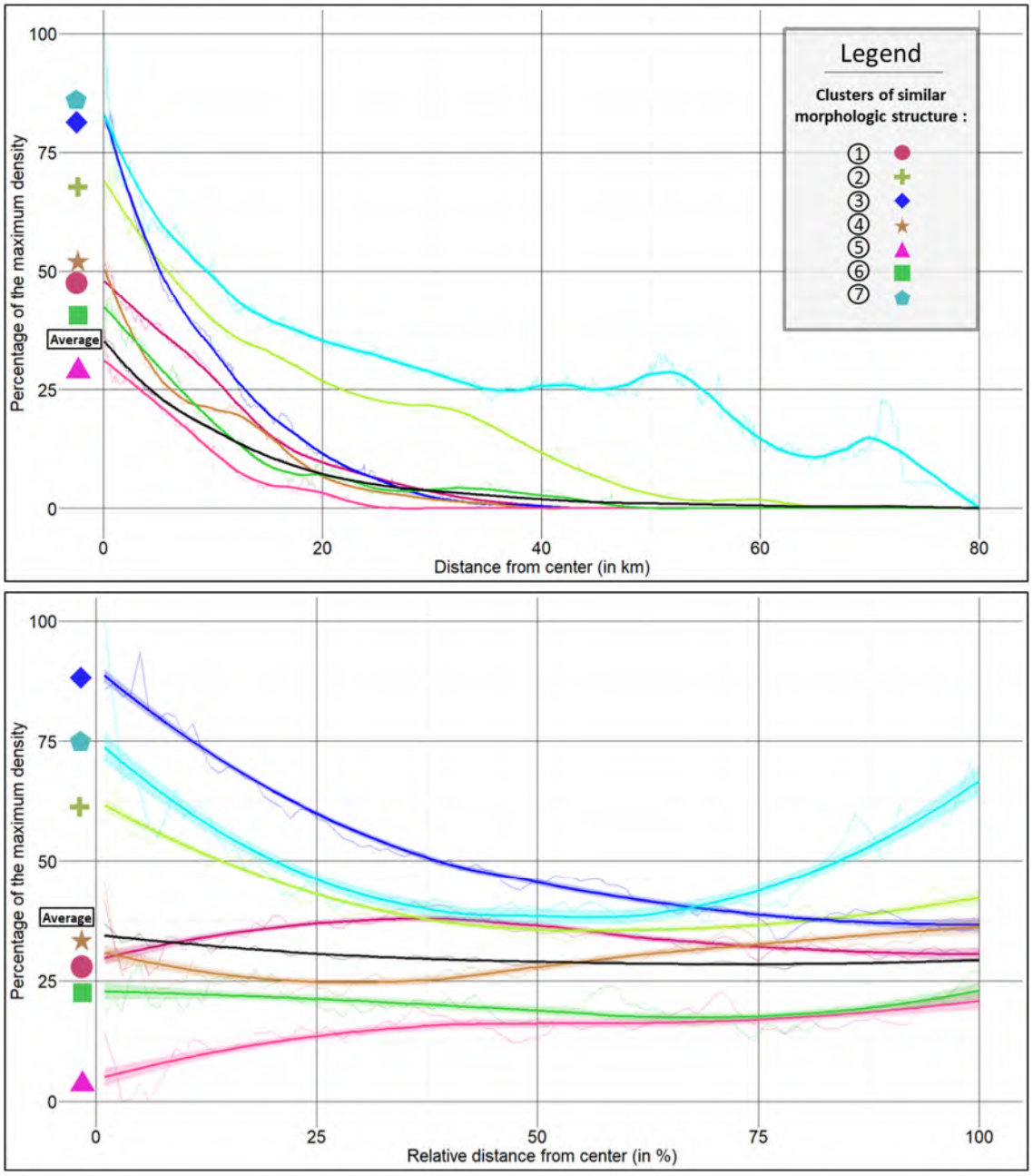


Fig. 5. Volume density gradients of the seven resulting clusters (k-means algorithm; 18f feature space, spatial entity MUAs) standardized distances (top) as well as normalized distances (bottom) from the city center.

Cluster 1: (mainly **Asian and African cities**)

Medium large cities of low structural variability, medium compact, low-rise

Cities of this cluster have on average a radius of 14.9 km. With it, this clusters features cities in a medium size among the seven clusters. Structurally, these cities show an average share of 35% of non-built up land. This ratio seems to be rather typical for cities across the globe as four out of the seven clusters show values between 31 and 35%. The structural variability of these cities is low. Three LCZs, which together account for about 61% of the MUAs, dominate the structural configuration: low plants (LCZ-D with 25% on median), large low-rise (LCZ8; 24%), and compact low rise (LCZ3; 12%). The remaining LCZs cover without exception less than 5% each of the urban area. From the point of view of spatial statistics, this class has the greatest similarity with cluster 2.

Uniform density decrease (abs.) and convex gradient (normalized)

For cluster 1 we measure the expected decreasing built-up volume density gradient with increasing distance to the center. The general volume density for these cities, however, is just above average among all clusters (Fig. 5). It is remarkable that when we evaluate the normalized approach, we do not get a flattening curve towards the urban periphery, but a rising curve outside the city center which eventually decreases again on the edges of town. In other words, in comparison to other clusters, the drop in density is less pronounced and the urban morphology remains relatively constant within the MUA.

Cluster 2: (mainly **Asian and American cities**)

Large cities of medium structural variability, medium compact, low-rise

With 25.1 km radius on average, this cluster contains large cities compared to other clusters. The share of non-built up is with 31% still in a medium range. While this cluster features one LCZ that stands out in terms of spatial shares (large low-rise structure (LCZ8) accounting for 28% of the MUAs), more other LCZs have higher spatial proportions compared to cluster 1 (compact low rise (LCZ3; 13%), low plants (LCZD; 12%), open low rise (LCZ6; 7%), sparsely built low rise (LCZ9; 7%), dense (LCZ-A; 5%) and scattered trees (LCZ-B; 4%)) (Fig. 4). From the point of view of spatial statistics, this class has among the remaining clusters (3–7) the highest similarity with cluster 3.

Concave decreasing density gradient (abs. and normalized)

The volume density of the built-up in the center is on average comparatively high in this cluster and we measure a continuously decreasing, concave density gradient towards the periphery (Fig. 5a). A specific characteristic here is the comparatively high volume density at a long distance from the center (even at 25 km at over 20%), which is an indication that there are many very large cities in the cluster. This is also represented in the normalized density gradient showing a relative increase in the periphery (Fig. 5b).

Cluster 3: (dominated by **European cities**)

Medium-sized cities of high structural variability, medium compact, mid-rise

With 13.7 km radius on average, this cluster contains medium sized cities compared to other clusters. This morphological configuration consists of a generally medium share of non-built up (31%). This cluster features no dominance of one or few LCZs, but more balanced shares of LCZs: Open low rise (LCZ6; 15%), large low-rise (LCZ8; 15%), low plants (LCZ-D; 13%), open mid-rise (LCZ5; 10%), compact mid-rise (LCZ2; 6%), dense (LCZ-A; 6%) and scattered trees (LCZ-B; 4%) show the high structural variability. From the point of view of spatial statistics, this cluster has the among the remaining clusters (4–7) highest similarity with cluster 4.

Strongly decreasing, concave uniform density gradient (abs. and normalized)

The volume density gradient of this cluster comes closest to the generalized assumption that a continuous decrease from the center to the periphery exists. The volume density in the center is very high in this cluster and we observe the steepest decrease among all clusters with increasing standardized distances from the center (Fig. 5a). And, we observe an even and continuous decrease in density using relative distances to the center, too (Fig. 5b).

Cluster 4: (mainly **Eastern African and Eastern Asian cities**)

Small cities of high structural variability, low compact, lowrise

With 12.2 km radius on average, this cluster contains small sized cities among our clusters. The share of non-built up is with 32% again in a medium range. This cluster features no dominance of one LCZ but a high structural variability: open low rise (LCZ6; 20%), low plants (LCZ-D; 18%), sparsely built (LCZ9; 18%), compact low rise (LCZ3; 11%), scattered trees (LCZ-B; 5%) and large low-rise (LCZ8; 5%) are the six classes featuring highest shares. Statistically, this cluster has the among the remaining clusters (5–7) highest similarity with cluster 5.

Strongly decreasing, non-uniform density gradient (abs.) and change from decreasing to increasing gradient (normalized)

The standardized volume density gradient of this cluster features a high drop in density outside the center with a lower drop in the peripheries. The relative density gradient reveals a comparatively high density in the periphery and an unconventional increase relative to all clusters towards the periphery (Fig. 5).

Cluster 5: (dominated by **cities of central Africa**)

Small-sized cities of medium structural variability, low compact, low-rise

With 11.8 km radius on average, this cluster contains the smallest sized cities among our clusters. The share of non-built up is here with 57% large, i.e. the general structure of these cities differs significantly to clusters 1–4. The large amount of undeveloped land is generating a less intense usage of urban space. This cluster features an exceptional proportion of one non-built LCZ: low plants (LCZ-D; 47%) defines the pattern. In the built domain, the shares of individual LCZs are exceptionally small: large low-rise (LCZ8; 8%), compact low rise (LCZ3; 11%), open low rise (LCZ6; 10%), sparsely built (LCZ9; 10%). Many LCZs do not exist at all. Statistically, this cluster has among the remaining clusters (6, 7) highest similarity with cluster 6.

Uniform density decrease (abs.) and increasing gradient (normalized)

This cluster features lowest volume densities across all clusters independent from the location within the city. The standardized volume density gradient features the expected continuous decrease towards the periphery starting at the lowest measured rate of 32% of the maximum density measured across all cities. Interestingly, the relative density gradient reveals an increase relative to other clusters towards the edges of towns.

Cluster 6: (dominated by **cities of the Islamic world**)

Medium-sized cities of high structural variability, medium compact, low-rise

With 15.8 km radius on average, this cluster contains medium sized cities. The share of non-built up is with 62% the largest of all clusters; a structural configuration that leaves a large part of the urban landscape undeveloped. So, this cluster shows large shares of non-built up LCZs: Bare soil or sand (LCZ-F; 29%) and low plants (LCZ-D; 19%) are dominating. In the built types individual LCZs feature on median comparatively small shares. Compact mid-rise (LCZ2; 3%), compact low rise (LCZ3; 3%), open low rise (LCZ6; 3%), large low-rise (LCZ8; 4%), sparsely built (LCZ9; 3%) and heavy industry (LCZ10; 3%).

Uniform density decrease (abs.) and basically constant gradient below average (normalized)

The volume density gradient features a continuous and uniform decrease in density towards the periphery. The relative volume density gradient reveals a cluster consistently below average densities across clusters.

Cluster 7: (consisting of **very large cities**)

Very large cities of low structural variability, medium compact, mid-rise

With 37.4 km radius on average, this cluster contains by far the largest cities among our clusters. The share of non-built up is with 26% the lowest of all clusters and reveals highly developed urban spaces. In this cluster two LCZs cover more than 50% of the entire MUAs. In the built domain large low-rise (LCZ8; 33%) dominates the physical configuration of the cities. In the non-built domain, low plants (LCZ-D) cover 20%. Other LCZs (3, 4, 5, 6, 9 and 10) are in the range of 2.5–8%. It is noteworthy that even for these very large cities, the compact highrise (LCZ1; 1%) occupies only a small proportion of the urban area (albeit more than in the other clusters).

Concave density gradient (abs. and normalized)

Volume density is highest in the center among all clusters and due to the city sizes in this cluster the gradient is, in comparison to others, featuring a lower decrease with remarkable densities even 50 km from the city center. Due to this, the relative density gradient reveals a strongly concave shape with significant relative increases in density compared to other clusters.

Summarizing, Table 2 provides a detailed overview of the seven clusters with highest congruence to geographical-cultural regions. The associated cities, the geographical allocation and their physical characteristics in terms of LCZs and density are provided.

We conclude that based on these specific data, these methods, these spatial units, we find *seven clusters* (or city types) represent the morphologic-spatial configuration of cities worldwide. And these seven clusters do relate in parts to geographical-cultural regions.

Table 2. Overview of the Clusters (C), the containing cities depending on the feature space (18f), and their geographical allocation and spatial features.

Cluster	Cities (k-means, 18-dimensional feature space)	Geographic-cultural allocation	LCZ features	Density features
1	Abidjan, Ahmadabad, Baghdad, Bangalore, Bucharest, Cape Town, Chennai, Delhi, Douala, Fortaleza, Ha Noi, Ho Chi Minh City, Jaipur, Johannesburg, Lagos, Lahore, Lucknow, Mogadishu, Mombasa, Port Harcourt, Santiago, Tianjin	Mainly **Asian and African** cities	Low structural variability, medium compact, low-rise, medium shares of open space	Uniform density decrease (abs.) and convex gradient (norm.)
2	Beijing, Buenos Aires, Jakarta, Kuala Lumpur, Manila, Melbourne, Mexico City, New York, Osaka, Rio De Janeiro, Riyadh, Sao Paulo, Seoul, Singapore, Toronto	Mainly **Asian and American** cities	Medium structural variability, medium compact, low-rise, medium shares of open space	Concave decreasing density gradient (abs. and norm.)
3	Addis Ababa, Amsterdam, Athens, Berlin, Brussels, Budapest, Changsha, Chittagong, Chongqing, Dhaka, Hamburg, Istanbul, Kiev, Lisbon, London, Madrid, Manchester, Milan, Minsk, Moscow, Mumbai, Munich, Paris, Rome, Saint Petersburg, Vancouver, Vienna, Warsaw, Wuhan	Dominated by **European** Cities	High structural variability, medium compact, mid-rise, medium shares of open space	Model-like, strongly decreasing, concave uniform density gradient (abs. & norm.)
4	Antananarivo, Blantyre, Calcutta, Calicut, Dar Es Salaam, General Santos City, Harare, Hong Kong, Islamabad, Kitwe, Lusaka, Maputo, Mwanza, Nairobi, Nanjing, Port Au Prince, Sydney, Yangon, Zanzibar	Mainly **Eastern African & Eastern Asian** Cities	High structural variability, low compact, low-rise, medium shares of open space	Strongly decreasing, non-uniform gradient (abs.) and change from decreasing to increasing gradient (norm.)
5	Abuja, Bujumbura, Kampala, Kano, Kigali, Kinshasa, Kumasi, Lilongwe, Poona, Port Moresby, Surat	Dominated by **Central Africa** cities	Medium structural variability, low compact, low-rise, high shares of non-built space	Uniform decrease (abs.) and increasing gradient (norm.)
6	Asmara, Cairo, Faisalabad, Hargeysa, Kabul, Karachi, Khartoum, Ouagadougou, Tehran	Dominated by cities of the **Islamic world**	High structural variability, medium compact, low-rise, very high shares of non-built space	Uniform decrease (abs.) and basically constant gradient below average (norm.)
7	Bangkok, Guangzhou, Los Angeles, Shanghai, Tokyo	Consisting of some **Very large cities**	Low structural variability, medium compact, mid-rise, lowest shares of open space	Concave density gradient (abs. and norm.)

5. Discussion

Cities are diverse – in their economic, social, demographic, administrative, political structure, but also in their atmospheres, in their daily routines, in their sound volumes or smells; and, of course, also in their physical appearance, i.e. the built and non-built urban landscape. The morphologic-spatial appearance is naturally only one part reflecting this diversity; yet, it defines many routines and modes of operation of a city (e.g. Glaeser, 2010). In this study we have filtered out similarities of built landscapes of cities across the world and formed groups that represent different morphologic-spatial types.

City models, as mentioned in the Introduction, are based mainly on observations and theories. They rely comparatively little on truly consistent empirical data, which are suitable for comparison. Our study is based on consistent, structural knowledge of the built landscape based on a uniform concept. The structural knowledge is derived from remote sensing data. This enables a new empirical approach to analyze city theories and models.

It may not come as a surprise, but this study shows that the theoretical spatial model concepts of cities and the seven spatial city types (clusters) determined from our empirical work have common features; however, differences also become apparent.

With respect to our *first research question* – are intra-urban morphologic configurations of cities across the globe similar or different – we empirically find that different morphologic-spatial configurations of cities across the globe exist. Against the often-stated narrative of 'homogenization' (e.g. Gordon & Cox, 2012), we found here a large structural variation. In general, seven clusters of morphologic city models have been statistically identified representing the various morphological appearances of cities across our globe.

The *second research question* – do groups form according to geographical and cultural aspects (with the underlying assumption that similar processes produce similar spatial patterns) – cannot be answered so unambiguously. On the one hand, we find clusters of cities that show structural similarities in our global sample and correspond to the often proposed geographical-cultural aspects. The European, the Islamic and the central African model are the representatives of this group. On the other hand, we identify groups in this study that surprise us with regard to a geographical-cultural classification: The 'Asian and African', the 'Asian and American', and the 'Eastern African and Eastern Asian' clusters that have, to the knowledge of the authors, never been identified explicitly with respect to similar morphological-spatial configurations. These clusters suggest that there are indeed urban configurations in which cities are more similar across continents than to cities in geographically and culturally closer regions. City models that generate themselves exclusively from these geographical-cultural spaces thus seem to represent reality in an over-simplified way.

However, our results are, and this must be discussed here, also fraught with uncertainty. The uncertainties are derived from the following points: 1) inaccuracies in the classification products, 2) the applied classification scheme of LCZs, 3) the spatial units of measurement, 4) the general spatial entity of MUAs, and 5) the selection of cities: 1) The input LCZ classifications have an overall accuracy of 87.3 and a kappa value of 0.65. Overall accuracies among the thematic classes vary significantly. 65.1% for the 'LCZ-C – bush, scrub' up to

99.3% for 'LCZ-G – water' in the natural domain and, as examples in the built domain 0.58 for the 'LCZ-6 – open low-rise' to 0.87 for 'LCZ-8 – large low-rise'. Although these are fair numbers, the effects of misclassifications to the derivation of clusters is unknown. Furthermore, it must be stated that the LCZ classifications from Sentinel-2 data are two-dimensional. The derivation of built-up volumes is purely based on a statistical approach derived from the concept of Stewart and Oke (2012) and thus contains further unknown uncertainties. 2) The applied classification scheme of LCZs is a concept aggregating the structures to a larger spatial unit, in our case a standardized raster of 100 × 100 m. In this spatial generalization, small-scale, structural transitions or invariances are lost, of course, due to an averaging effect in the classification process demanding for a discrete LCZ class. Beyond this, we must be aware that although LCZs are generic, culturally-neutral descriptions of land-use and land-cover, we cannot rule out that morphologically similar areas exhibit different meanings in different cities or areas. 3) The spatial units of measurement we applied rely on a monocentric city model. While it is still a well-accepted approach (e.g. Spivey, 2008), it is once again a generalized assumption that structural peculiarities, especially in the periphery, are masked by averaging. 4) Beyond, although the MUAs are consistently derived and thus form a comparable basis (Taubenböck et al., 2019), they are only one way of delineating urban from rural. A different setting allowing larger or smaller MUAs, while still being designed consistent, might have effects on the resulting clusters. We do have tested in our study also administrative city units as well as standardized circles as spatial extents; the resulting clusters, however, were geographically not as congruent with the geographic-cultural regions used, and thus we disregarded them in this paper. And, our analysis is based on the fact that the extent used for the cities, except for the class 'water', which is eliminated, is flat and therefore the same. However, topographic, economic-specifics, historic path-dependencies such as colonial influences or any other related possible effects shaping the configurations of cities have not been considered in this analysis. These things may counteract geographic-cultural processes, but the effect is unknown to us. 5) Last but not least, our analysis relies on 110 cities. While this is a large sample, with a proven representative distribution across the globe relative to continental urban population shares, a smaller or a larger sample, might also affect the resulting clusters. All in all, we conclude that based on our particular data and methods, we think the resulting seven clusters are one reasonable possibility of describing the city configurations across the globe in a transparent way that reveal that geographical-cultural aspects influence their spatial configuration; but, these seven clusters are neither the only possibility nor the universal best fit for a reality probably more complex.

When we close our eyes and think of big cities on our planet, the often-appearing feature in our mind is a high-rise skyline. This often feeds the feeling that cities are becoming more and more homogeneous. However, the related morphological class (LCZ1 – compact highrise) plays almost no role in the morphologic-spatial composition of cities, since their spatial proportions usually make up only about 1% of the MUAS. And this focus on the most conspicuous morphological manifestations, but extremely underrepresented in the urban landscape as a whole, probably also obstructs the feeling that there is a wide range of urban configurations worldwide. On our global analysis seven clusters with still a large variation in themselves are the result. Even if we, for instance, find the European cluster relatively

homogeneous within our global sample, other studies reveal, looking only at a sample of European cities, that a large variety exists also within Europe (e.g. Schwarz, 2010; Siedentop & Fina, 2012; Taubenböck et al., 2019). This clearly shows that the generalization of the clusters here as done by our study is on a highly aggregated level. Cities like Mumbai and Hamburg fall into one cluster. However, this is legitimate as geographically motivated city models have always adopted this high abstraction level. The fact that similar morphological structures produce very different atmospheres, life routines or many other things (as Mumbai and Hamburg testify) is not the issue here and goes without saying. This study is exclusively about morphologic-spatial manifestations. Given these variations, the identified seven clusters allow us to take a novel look at city configurations and their geographic distributions across our planet. We must not forget that the generalized models conceal regional specifics of urbanization processes that often materialize in highly site-specific urban outcomes that are incompatible with the model.

6. Conclusion and outlook

Science, but also reports or popular rankings have always tried to fit the diversity of our urban landscapes into generalizing and thus simplifying models. This study shows with novel data from remote sensing an improved empiricism to identify morphologic-spatial characteristics for such models across the globe. And the realization that geographic-cultural characteristics can be empirically proven with regard to morphological urban configurations is a confirmation of a long-held observation. At the same time, however, it also shows that morphological urban configurations also have similarities across geographical-cultural spaces. Understanding and fathoming these will be part of future research tasks. With a large database of 110 globally distributed cities, selected in proportion to the urban population per continent, we can principally make generalizable statements at global level. Nevertheless, an expansion of the empirical basis remains necessary. In times of ‘Big Data’ this study is only a beginning to master this complexity with more precise classifications, more cities, or alternative methods. It may be a basis to more systematically analyze which forms and models work better or worse.

Firmly convinced that the spatial-morphological configuration of cities makes a decisive contribution to whether cities offer quality of life, whether they are socially acceptable, ecologically sustainable or economically successful, this study attempts to create a spatial-morphological basis to examine these questions more systematically: In the hope of better understanding urban space as the theatre of life and thus being able to make it more livable.

Acknowledgements

This work has received also funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No [714087]-So2Sat).

Appendix

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cities.2020.102814

Chapter XVI

Planned, unplanned, or in-between?

A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe

Landscape and Urban Planning, 2023, vol. 233, 104711

Henri Debray, Nikolas J. Kraff, Xiao Xiang Zhu & Hannes Taubenböck

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/j.landurbplan.2023.104711

Received 28 June 2022; Accepted 29 Jan 2023; Published 14 Feb 2023

Abstract

The physical appearance of the built urban landscape is the result of multiple, intertwined processes. The relationship between the existing morphology and the diversity of planning processes, however, has been little studied empirically at a global scale. In this study, we develop an ontology of planning intensity: conceptualizing intemediate categories of whether an urban structure is planned single-handedly or constantly updated by myriads of participants. Thus, we move away from the ’planned/unplanned’ dichotomy and develop a continuum of Intensity of Plannedness (IoP). The focus of research is whether these conceptualized categories of IoP show demonstrable differences in morphology. Hence, we operationalized the urban structure by three structural elements: buildings, morphological units and streets. Curating geodata on 381 study sites across the globe, we empirically investigate the relation of the IoP to the structural complexity of the urban fabric. Tests of significance of difference and post hoc analyzes are performed on the statistical distribution of structural complexities of the categories of IoP. This study proves empirically that the distinct IoP has significantly contrasting structural complexities. From this, we conclude that there is indeed a relationship between both, the intensity of the process of planning and the resulting urban morphology and that this relationship is non-linear.

Keywords

Intra-urban morphology; global study; intensity of plannedness; structural complexity

1. Introduction

How is the built city shaped? Sometimes, the form of the city emerges from a totally unplanned process. Occasionally, on the opposite, its design is subject to a completely planned operation. However, staying at this simplistic dichotomy does not give its credit to the rich diversity of urban processes, as a given city or settlement does not necessarily fall easily into the categories fully '*planned*' or totally '*unplanned*'. In this research, we investigate a concept that we coined the Intensity of Plannedness (IoP). The Intensity of Plannedness is a scale characterizing how acutely and completely an urban space is designed. It follows the idea that there exists a conceptual continuum between two archetypical extremes of city-making that we could call '*total design*' and '*total self-organization*' (Fig. 1).

Prior to the works of C. Alexander and J. Jacobs, modern and premodern approaches on urban making theory were attached to explain the strive of cities and the collective qualities of urban life by focusing almost exclusively on the planning of the materiality of the built-up environment and the services offered in the city (Batty & Marshall, 2012). The shift proposed in (Alexander, 1965; Jacobs, 1961) was to consider these qualities of urban life in their relation to the subjective and inter-subjective aspirations of individuals and groups (Cozzolino, 2018). These new paradigms are giving their due importance to the social structure in the design and control of space (Cozzolino & Moroni, 2021; Cozzolino, 2020; Kostof, 1991b; Kropf, 2009; Lynch, 1984), the relevance of the contribution of the ordinary inhabitants (Cozzolino, 2021; Jacobs, 1961; Noizet & Clemençon, 2020; Tonkiss, 2013), because of their influence on the resulting diversity and complexity (Batty & Marshall, 2012; Batty, 2005, 2009; Bettencourt, Lobo, Helbing, Kühnert, & West, 2007; Jacobs, 1961; Moroni & Cozzolino, 2019; Portugali & Stolk, 2016; Salat, Bourdic, & Labbe, 2014), and in the ensuing enabling of incremental adaptivity (Akbar, 1988; Cozzolino, 2020; Habraken, 1998; Hakim, 2014).

In this work, we draw from this theoretical background, especially from the importance it attaches to the actions of individuals, the interactions between them, and the social processes emerging from it (Moroni & Cozzolino, 2019). More specifically, we see the IoP as both, an emerging result of the social processes, but also as the social frame to which the making of the built-up environment is conditioned to.

By acknowledging the contribution of the ordinary city dwellers to their direct environment, this shift of paradigms enabled to look past a too rigid dichotomy between '*planned*' and '*unplanned*'. In their book '*Collage City*', Rowe and Koetter (Rowe & Koetter, 1978) developed the idea of '*total design*' or '*total planning*'. This can be understood, per negative, as the acknowledgment of the existence of '*non-total planning'*: a state below '*total planning*' still pertaining to the notion of '*planning*' thus, an in-between '*planned*' and '*spontaneous*'. Later, and more conclusively, Habraken was among the first to propose to see it as a combination of top-down and bottom-up processes in relation to hier archical overlapping of '*fields of control*' of different scales (Habraken, 2016, 1987, 1998); Kelso (Kelso, Portugali, & Stolk, 2016) described this intermediary state as the synergetic combination of the complementary pair '*self-organization – design*'.

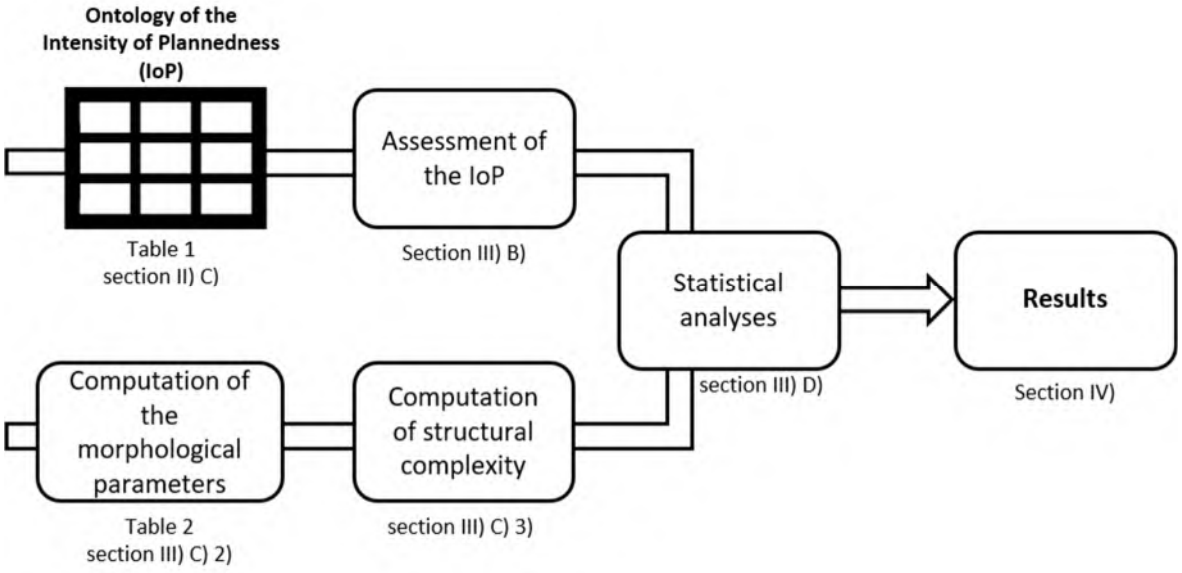


Fig. 1. Workflow of the study.

Among the research fields conceptualizing the spectrum between '*spontaneous*' and '*planned*', it is interesting to mention the contribution of the study of informal settlement. Noticeably with Kuffer, Barros, & Sliuzas (2014) proposing an '*unplanned settlement index'* to describe morphologies of informal settlements and Dovey et al. with their work (Dovey & Kamalipour, 2018; Dovey et al., 2020) on the morphogenesis of informal settlements proposing to consider their morphogenesis as indexed to a gradient of formality both, at the urban design level but also at the architectural level.

Congenially, a few urban thinkers and historians employed them selves to unfold the multi-dimensionality of this spectrum: Ikeda (Ikeda, 2017) proposed to take in account the '*scale of design*' and the '*passage of time*'; stepping on these spatial and temporal extent, Cozzolino (2018) added that the '*level of spontaneity*' in urban context is tributary to the initial physical dimension and, more critically to the social dimension. More recently (Cozzolino, 2020) proposed a richly multi-factorial set of attributes characterizing the spectrum between '*anti-adaptive neighborhoods*' and '*adaptive neighborhoods*' including the evolution in time of the attractiveness, the scale of design, the amount of designers, the construction time, the types of planning rules, the diversity and interrelations of functions, the amount of public open spaces and the ownership system (further developed in Cozzolino & Moroni, 2021; Cozzolino & Moroni, 2022). Focusing on the French context, Noizet & Clemençon (2020) proposed to conceptualize a typology of '*lotissement*' (allotment) ranging from '*a minima*' to '*a maxima*' planning. This typology presents six levels based on the inter-relation of buildings, parcels, streets, "morphological urban islands", showcasing, as in the works of Ikeda and Cozzolino et al., connections with the concepts of ownership and asynchrony.

Despite this acknowledgment of intermediary states between '*planned*' and '*spontaneous*' in the literature, there exist, to the best of our knowledge, no comprehensive operationalization pertaining to the spectrum of these states fitting global scale analysis. Therefore, the first aim of this study is to develop an ontology of the IoP.

The second aim of this work is to investigate the hypothesis that form of the built-environment relates to the IoP. On one side, the assumed main tool of urban planning is, in itself, carving, in a top-down manner, a specific design in the intra-urban morphology (Batty & Marshall, 2009; Kostof, 1991b). On the other side, the self-organization as well, has its way of inscribing itself spatially in the urban morphology, this time in a bottom-up manner that is often described as presenting "organic" or "instinctually-grown" patterns (Batty & Marshall, 2009; Kostof, 1991b; Larkham, 1992). In fact, if both paradigms shape morphology, the resulting layouts cannot be more strikingly different (as exemplified in Fig. 2). As such, our hypothesis is that the continuum between both will, in turn, propose a wide range of morphological manifestations. There fore, we want to address in this study the question whether the here conceptualized IoP is incrusted in the intra-urban morphology in a such a way that it is reflected in a spatially measurable way?

To answer this question, we follow a qualitative and quantitative approach to study intra-urban morphology. Following studies such as (Boeing, 2019; Dibble et al., 2019; Rashid, 2017; Fleischmann et al., 2021; Pont and Olsson, 2018; Taubenböck et al., 2018b), we evaluate the statistical dependencies between quantitative morphological fea tures and our qualitative characteristic, the IoP. Our empirical investigation is based on 381 study sites around the world.

Fig. 2. (From left to right, from category Extremely spontaneous to Completely planned) Caracas, Kyoto, Paris, Chicago, Cairo.

Summarizing, the scientific contribution of this paper is twofold: 1) Develop and propose an ontology of IoP that we use to 2) empirically investigate at a cross-cultural, global scale to which extent the IoP in scribes itself into the urban physical form.

To do this, we *first* develop a conceptual basis to define the IoP of settlements in section II). We ground this concept on a literature review that allows to lay a theoretical foundation and to elaborate the practical elements needed for an empirical investigation. We collect common traits and differences of generic urban processes among cities and we abstract them to specific levels of IoP. We operationalize this by three structural elements: the buildings, the morphological units and the streets (Table 1).

Table 1. Ontology of the 'intensity of plannedness' using five structural categories and three defining structural element.

		Characteristic urban processes for the different categories of *intensity of plannedness*				
		Extremely spontaneous	Predominantly spontaneous	Evenly spontaneous and planned	Predominantly planned	Completely planned
Structural elements	Buildings	Buildings are mostly self-constructed individually by its future user. As such, there is no official general or organized plan for the building disposition and no explicit concerted decision to build in a certain way. As work force for building is limited, circumstances mean that the choices of structures and building materials to be used is limited. Therefore, the buildings of the overall settlement will be rudimentary, often of makeshift character, and quite limited in their diversity.	Buildings are either self-constructed by the main users with the help of other members of the local community or the construction process can be delegated to more specialized people. Both ways provide more work force and provide more elaborated structures that are disposed in an, at least implicitly, accepted manner by the community. As the workforce is typically drawn from a small pool of local people, who often share same skills, hence producing some norms of building in a manner tributary to local resources and culture. This enhances only little variety in the buildings' type.	Buildings are constructed in the frame of explicit rules such as implemented cadastral parcels division of the land and norms of constructions. The owners acquire the parcels and have the buildings renovated or built anew to their idea. As such, the owners rely on significant enough financial funds and are likely to rent the edifices for at least recovering the investment. Therefore, the users of the buildings are unlikely to be the owners. The variety of buildings can be high to the measure of the diversity of funds of each agent and of the diversity of dates of building, but is nuanced by the constraints related to the norms, already pre-existing parcels, and the homogeneous skills of different local building companies.	Buildings are constructed typically in planned zones where cadastral parcels are already systematically structured in their layout by the local planning authorities. Parcels are of middle size or composed of small sized ones that can be combined easily into single ones. This allows private individuals or real estate companies to construct buildings at a larger scale. This is reflected – by either large and/or high structures built by a single company, large low-rise residential areas commissioned to multiple building companies, or medium-scaled medium-rise residential areas built by a single company. As such, there is a tendency for standardization of building types.	Buildings often originate from large-planning policies and legally-binding land use plans. The cadastral system is generally designed at a coarse scale and the parcels are designed to host big estates. Thus, only local governances, big real estate companies or rich private individuals can afford to own them and to build on them. As a result of the dimension of the programs, a small pool of companies has the technical capacities to develop these buildings. This can lead to either single or few monumental buildings int the case of limited scale programs. Alternatively, in large programs, the economy of scale in the design, in making, in logistic, etc. tailors a masterplan based on replication, and tends to enhance a strong standardization of the buildings.
	Morphological Units	Morphological units emerge haphazardly, by conglomerating buildings for practical reasons, being mostly reducing distances for rapid access to outbuilding and annexes. People develop these units in free range, over time and in a non– or loosely concerted manner. Therefore, the morphological units exist, in the settlement, in a large variety of types of conglomeration, each specifically meant to fit the need of its direct local users.	The community implicitly, or explicitly, designed or regulated specific places as commons over time. Morphological units are constrained to form outside such areas. Those common places, being non-buildable or public places (e.g. streets or squares), define the outermost extension the private space can take in the morphological units. The inner part of the morphological unit is not constrained in the exception of possible inner streets -and squares.	The relation of the limits of the morphological units to the public places in the district or the city is explicitly regulated to ease the design of streets, for rights of way, for the access to the sunlight or for esthetic reasons. This regulation implies definitions of how the frontage of the lots of land sets back from the public place. This control of the buffer between the outer built -or owned -perimeter of the morphological units and the public places, in turn, means that the outer perimeter of the unit is defined by the lots.	Rules are implemented that seek to respect theorized optimal densities for the morphological unit as a block. The rules can concern the width, the spacing of the parcels or the depth by which they can be built. Those policies concerning the density can range from taxation system to enforcement of specific buildable space in the lots. This process often benefits from an already laid out street network hosting surveyed plots ready to be bought rendering the policies easier to apply.	The morphological unit and its constitutive buildings are the result of a single design. Due to the monetary costs of such projects, agents planning such unit might be ones having enough fund to design multiple units on larger area. The morphological units follow strictly local norms and rules such as the density and the height-distance to road ratios. As consequence of those rules, the streets are not necessarily anymore an element constraining the extent and the morphology of the morphological unit.
	Streets	The streets emerge from frequently used non-built trails and pathways that people use to connect buildings or places. In dense situations they are defined by the surrounding private spaces and thus are tributary from their shapes. They tend to follow a logical path of least energy or least commutation time making them prompt to follow the topology and avoid obstacles.	The streets of the settlements are subject to common rules of habits aiming at facilitating their overall use by granting access to the buildings and to improve the basic freedom of movement. The street network is usually renovated in this sense but it still inherits from the shapes of former patterns or is impacted by the topology.	The street network is planned according to a system. This system is thought to improve the quality of the network in regard of specific aspects through regulations (e.g. street geometry standardization). The system can concern the entire network or focus on a specific level of hierarchy.	The street network is designed to accommodate an area as large as a district. It is often planned prior to the existence of the settlement or else, in a massive renovation project. The planning follows systematically norms and a goal of optimization (most of the time to ease the circulation of vehicles).	The street network is completely designed. It originates from a single entity being in charge of deciding the layout of a new district, or a complete city with the street network being its structuring backbone. To comply with accepted theories (on the 'ideal city' or on circulation for examples), a strong hierarchy will be implemented between the different streets. This will often support the argumentation for specific designed patterns.

Second, in section III) A), we present how we select the largest possible number of study sites across the globe where we find both, reliable specific information on the undergone urban processes' strata and where appropriate geodata is available.

Third, in section III)B), we describe how, we assess qualitatively the IoP of the structural elements for each settlement.

Fourth, we summarize in section III)C) how we estimate the structural complexity of the intra-urban morphology of our study sites on vector geodata. We first give a summary of the vector geodata we use in section III)C)1). Subsequently, in section III)C)2) we present the computation of a set of morphological features indicative of urban planning on the different structural elements. In section III)C)3), we explain how we compute, the joint entropy of the value's distribution for sets of morphological features for a settlement as a proxy for the *structural complexity* of the morphological patterns of the settlement.

Finally, in section III)D), we describe how we statistically analyze the obtained diversity measure across the categories of IoP in order to investigate whether correlations between the conceptualized IoP and physical appearance are significant.

With the concept of the IoP, we aim to contribute to the development of a theoretical foundation of the notion of intermediate degrees of plannedness of the urban form. By showing that the intra-urban morphology is charged with social values generated by different types of politic, economic and social structures, from the most monopolized to the most dispersed, we aim to highlight the relevance of the IoP for policy making. Although the concept of IoP may not have a direct impact on local planning decisions, we believe that this conceptual tool could be useful for studies and discussions on policy design.

2. Concept and ontology of the intensity of plannedness

2.1. Intensity of plannedness and intra-urban morphology: Theoretical background

In this study we propose a conceptual framework in which we consider the morphological configuration of a city as a result of the liberty, or lack of it, with which inhabitants shape their living environment over time.

The physical layout of cities is the artifact of cumulative processes of construction, destruction, renovation, replacement, and replenishment (Conzen, 1960; Kostof, 1991a, 1992). Those transformative actions, those modifications, in their great majority, are sought by agents. The type of those agents and their intentions come in great diversity. It could be a speculative investor transforming open lands into marketable plots in Los Angeles (Kostof, 1991b); a centralized socialist state curating the city of Moscow with large housing estates to propose a solution to an accommodation crisis in compliance with its ideals during the Socialist Realism movement (Kostof, 1991b); or a Peruvian farmer leaving his village in the Selva, informally building a brick house in the foggy hills of San Juan de Lurigancho in north of Lima (Inostroza, 2017; Riofrío, 2003; Turner, 1967). All those actions, exemplified here, leave a trace in the morphology of a settlement with their diversity of processes, technical solutions and scales. In turn, the physical layout is the cumulative result of a plethora of such processes. To unveil the specific variations of the exact process that formed a settlement as a whole, to say who built what, how and in which purpose, is a tedious task that one could claim to achieve only with great care and abundant contextual knowledges (Conzen, 1960; Kostof, 1982; Kostof, 1991b). This in-depth morphogenetic analysis is a challenging goal for a single city, let alone a consistent cross-cultural global investigation.

In our study, the goal is not to achieve such a detailed investigation, rather, we propose a more abstract approach, suitable for a large-scale global cross-cultural study. As developed by (Lefebvre, 1974), cities are the projection of the sociological interactions on the ground, or in other words by (Conzen, 1980), writing about the form of cities: "*Few social values and actions are so abstract that they fail to be reflected in material forms*". Here, we aim to find in settlements the morphological reflection of the two main relations to the space defined by (Lefebvre, 1974), i.e. the appropriation of space and the domination of space.

What Lefebvre defines as '*domination of space*' are the collective measures concerning planning (urbanism ordinances, rules and codes, hygiene and security rules, planning projects, administrative and management norms, norms of real estate, police, among others), over lapping, in our case with the notion of '*design control*' (Cozzolino, 2020; Habraken, 1987, 1998). In summary, this comprises top-down planning rules, norms and ordinances (Portugali, 2016). The effects it has on the physical layout can be stereotyped by the extreme example of master plans that strictly design the urban environment, but also by the less extreme example of ordinances regulating the distance of buildings to the road, conferring specific aesthetic character to street corridors.

On the other side, the bottom-up initiatives such as building, renovation, decoration, occupation, cleanness, co-security, among others are what (Lefebvre, 1974) describes as '*appropriation of space*'. The way this '*appropriation*' inscribed itself in the urban layout can be grasped in its cumulation when considering the uncoordinated apparition of informal settlements, shaped by the accretion of individually built buildings in a seeming haphazardly arrangement. Or, more discreetly, by the way specific buildings are renovated, making them standing out of its neighbors, despite having been originally built in the exact same way at the same time.

In our conceptual frame, we argue that the physical layout of the city is the evolving compromise of the transformative actions originating from those two social relations to the urban space, the complex dialectic relation between bottom-up and top-down processes.

As the relative importance of both types of processes exist in many variations and equilibria, cities, quarters, neighborhoods, or plots can vary from extremely spontaneous to completely planned, although a large number of urban areas falls in between these two extremes. Our goal is to formulate a categorization that covers the breadth and variety of morphological manifestations and that establishes the relation to their '*planned*' or '*spontaneous*' origin, by highlighting the characteristics of their formations. In this scope, we focus on actions that have a direct impact on the physical layout of cities, ranging from bottom-up, appropriation processes of ordinary urbanism (Cozzolino, 2021; Noizet & Clemençon, 2020; Tonkiss, 2013) to top-down, control sought actions and incentives of urban planning (Habraken, 2016; Kelso et al., 2016).

Both processes shape cities, each in their own ways. Literature shows that different policies (orientation of urban planning process) (Kostof, 1991a; Taubenböck, Murawski, & Wurm, 2018b) or lack of policies (room for spontaneous urban process) (Dovey, 2020; Kohli, Sliuzas, Kerle, & Stein, 2012; Kraff, Wurm, & Taubenböck, 2020a; Kraff, Wurm, & Taubenböck, 2020b; Taubenböck & Kraff, 2014) lead to different patterns. We additionally

formulate the hypothesis that different nuances of mixtures lead to different patterns as well. Although the combinations of spontaneous and planning process are maybe as numerous as the number of cities on Earth through its whole history, we propose an abstraction to few categories, i.e. we conceptualize an ontology enabling the comparison between mixtures. To do so, we define a gradient scale ordering these mixtures based on the relative importance of spontaneous urban processes against urban planning processes. We draw on the categorizations and scales from the works of Cozzolino (2020), Dovey et al. (2020), Kuffer, et al. (2017), Noizet & Clemençon (2020) and Taubenböck, Kraff, & Wurm (2018a). In the following, this scale is what we refer to as the *'intensity of plannedness' (IoP)* of cities.

We do not envisage *a priori* the IoP to be a quantitative measure, rather, we consider it as a qualitative assessment of a level on the gradient between unplanned and planned. To be able to assess the IoP of a settlement could be, alone, a challenging work of research in the vein of the works of scholars such as Conzen (1960) or Kostof (1991b, 1992). Their work should be a *memento* to us, not to forget the complexity to endeavor understanding the rich social dynamics of a place across the epochs down to the last iota of detail. Nonetheless, in the sake of producing data on the IoP, we have to resort to alternatives. Here, we used a hybrid approach, using literature on contextual and general urban planning history.

Along with our conceptual considerations, one can envision the IoP as a rough estimate of the balance between the number of people whose actions created and shaped the city (Moroni & Cozzolino, 2019) and the number of people living and performing activities in the city. As such, the layout of a city with a low IoP is understood as the physical result of the cumulative actions of a large part of its inhabitants. On the opposite, the design of a city with a high IoP will originate from the intervention of one single, or very few agents (local governances, big real estate companies, etc.) (Cozzolino, 2020; Moroni & Cozzolino, 2019). Although we believe the scale between low and high IoP to be potentially continuous, we aggregate to five main categories covering this continuum. Thus, we are able to conceptualize and describe these categories. Each IoP cate gory is representative of a relative importance of spontaneous urban processes vs urban planning in their combination. The IoP is presented in Table 1. The conceptualized categories are: *Extremely spontaneous*; *Predominantly spontaneous*; *Evenly spontaneous and planned*; *Predominantly planned*; *Completely planned*. These five categories are exemplified in Fig. 2: From '*organic and complex*', i.e. extremely spontaneous, to '*geometric and ordered*', i.e. completely planned.

2.2. The intensity of plannedness of structural elements.

The five settlements shown in Fig. 2 exemplify stereotypizations of the five conceptualized categories of the IoP. However, among the large variety of urban processes across the globe, a lot of settlements stray from these stereotypes. For instance, it is not rare that a street network is *predominantly planned* in a gridiron, whereas the buildings are *evenly spontaneous and planned* or even are *predominantly spontaneous*. Indeed, different structural elements of the settlements' physical layout do not result necessarily from the same processes. As such, this non-correlation of the IoP across different structural elements is depending on the asynchrony between them, and on the hierarchical nesting of the power of the different urban agents (Cozzolino, 2020; Noizet & Clemençon, 2020).

For the operationalization of measurability in the physical appearance of cities, we detail our ontology further, so these differences be tween structural elements can be addressed. Following the Conzenian tradition, the physical layout of a city is sketched by its buildings, streets, and the arrangement of its plots and open spaces (Conzen, 1960; Kostof, 1991b; Kostof, 1992; Kropf, 1993; Salat, 2011). Hence, we propose to use three major structural elements, the *buildings*, the *morphological units* and, the *streets* as illustrated in Fig. 3. We consider *buildings* being any type of built structure purposely sheltering people, wares or activities. We understand the *morphological units (MUs)* as the structural element formed by the aggregation of plots (Kropf, 1993), i.e., the buildings and their bounded open spaces. Hence, we define the MUs as the surface on which a collection of buildings in direct vicinity sits, sharing the same type of topological relationship. The MUs are under stood as an extension of the concept of '*street blocks*' to acknowledge a specific type of building conglomeration that emerges in contexts of a low IoP, where building conglomerations may not be defined by streets. Last, we understand the *streets* in their large definition of ways where the public circulation of people and goods is enabled. The computation of the structural elements is further described in the data section III) b).

We believe that the interplay of these three structural elements provides an exhaustive overview of the main observable physical layout of settlements. Beyond, we see these structural elements as direct targets of transformative actions of urban processes, top-down and bottom-up alike (Cataldi et al., n.d; Kostof, 1992; Mangin & Panerai, 1999; Pan erai, Demorgon, & Depaule, 1999).

Fig. 3. Structural elements illustrated for the example of Trier (Treves), Germany. (From left to right) buildings; morphological units; streets.

2.3. Ontology of the intensity of plannedness

Based on all these considerations, we conceptualize, describe and summarize the mixture of processes at stake for each of the three structural elements and for each of the five categories of IoP. The proposed ontology is presented in the following Table 1.

3. Experimental set-up

3.1. Study sites

We understand study sites here as samples representative of a type of urban process. As a city generally is composed of multiple areas tributary from different (intra-)urban processes, we select and delineate as samples, areas where the urban process shaping the physical appearance is as much as possible homogeneous, close to the definition of *Morpho logical* or *Morphogenetic Regions* (Conzen, 1988; Gu, 2019; Whitehand, 2001), or *Sanctuary Areas* (Dibble et al., 2019). These sample settlements are typically of the size of a district (typically ranging between 0.5 km^2 and 10 km^2), often following actual district borders, having their singular histories and, as a result, presenting a unique IoP.

We are aware that in different cultural regions, similar urban processes might shape different physical urban appearances (Kostof, 1991b). Thus, we attach importance to the fact that the different categories of plannedness are also covered at different cultural regions across the globe.

Another factor that we take in account for selecting the study sites is that we need to access appropriate data covering the different structural elements.

The selection of study sites is, therefore, based on the following criteria: We aim at reaching the highest possible number of settlements, that represent examples from all across the world; that are representative of the large variety of intra-urban morphologies; that high resolution vector geodata are accessible, depicting accurately buildings and roads and that the Mus can be derived from it; and, that literature references allow us to infer the category of the IoP for the selected settlements.

Fig. 4. Map of the aggregated number of study sites and their geographical distribution (points without number signify only one study site at this location).

Following these criteria, our sample selection are 381 study sites with consistent information to infer the IoP of the settlement (see Fig. 4). Among these 381 study sites, a subset of 260 is used for the *buildings* and the *morphological units*, and a different subset of 296 is used for the *streets*. Out of the 381 study sites, 69 are located in Africa, 107 are located in Asia, 36 in Central and South America, 98 in Europe, 55 in North America and 16 in Oceania. In this selection, some cities provided multiple study sites, given the distinctly different properties of their morphologies or their process of planning. Overall, this sample covers more than 150 independent cities. Additional maps of the geographical distribution of the study sites used specifically for *buildings*, *morphological units* and the *streets* can be found in Appendix A.

3.2. Assessment of the 'intensity of plannedness' of a settlement.

There does not exist yet a comprehensive guideline to perform the assessment of the IoP in to the categories. Thus, we have developed a systematic workflow to qualitatively perform this categorization on the base of literature and other exegetical data. This workflow is a decision tree that can be applied for each of the three structural elements. The develop decision tree is introduced in Fig. 5 and the steps are described in detail below.

• For each of these study sites, we *first* gather literature and other exegetical data, pertaining to the history of the urban development phases and processes that affected the study site considered.

Ideally, we would rely on information for each and every specific individual transformative action involved in the creation and the transformation the built-environment underwent, with knowledge on which agent instigated and contributed to each transformation. Unfortunately, data with this type of information is almost never openly available, is based on knowledge from many different actors, or does not even exist. Therefore, we build the assessment of the IoP on the inference of urban processes at a macro-level based on aggregated information sourced in literature.

With this regard, drawing from (Cozzolino, 2020; Noizet & Clemençon, 2020), specific attention is given to the search of information on planning rules, ownership systems, amount of designers, observable scale of design, mix or segregation of functions, and to the construction time and changes. To gather this broad range of nature of information, the search is performed across multiple types of source (academic literature, official documents, planning survey, historical maps, historical aerial pictures and to some extent grey literature and local knowledge).

• *Second*, using this information, we identify the decisive characteristics of the major original urban process at stake in the study site.

Namely these characteristics are: the type of planning rules and codes (prescriptive or proscriptive (Cozzolino, 2020)), the frame of these rules and codes (formal, *para*-formal, informal (Dovey & Kamalipour, 2017; Tonkiss, 2013)), the type of participants to the urban process (institutional or private entities, organizations, distinct agents) and amount of participants in relation to the scale of the area (Cozzolino, 2020) (in this study, we propose the four following tiers: Single (one participant responsible of the whole area); Few

(typically less than 5 for 1 km^2); Medium (typically be tween 5 and 50 for an area of 1 km^2), High (typically more than 50 per km^2)).

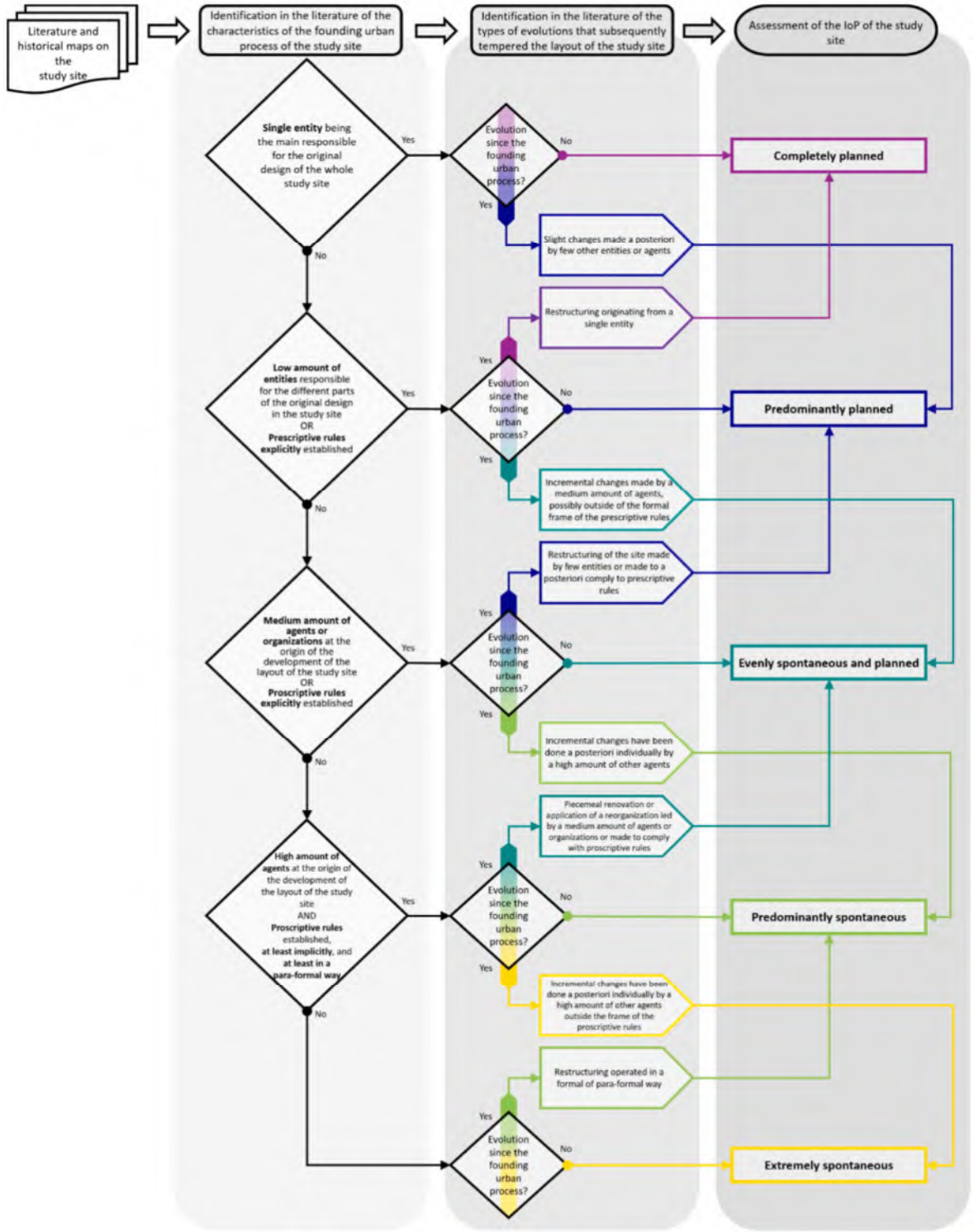


Fig. 5. Workflow of the assessment of the IoP of a study site.

Then, following a descending chain of *a minima* conditions on the combination of the characteristics, we can evaluate the IoP of the founding urban process of the study site.

• *Third*, we identify the evolutions the study site underwent since the founding urban process. This crucial step helps us to navigate from the IoP of the original state to the current one, considering the layers of history the study site went through.

Following (Cozzolino, 2020), we distinguish two major trends of evolution, namely: "incremental changes" done by additional participant being, in amount, at least as many as the original participants and over a longer term than the founding process; "restructuring led by comparatively less participants and happening comparatively in a more sudden way than during the founding process.

• *Last*, using the evaluation of the original IoP and the identification of the subsequent phases of evolutions in the study site, we formulate the assessment of its current IoP into the categories '*Extremely spontaneous*', '*Predominantly spontaneous*', '*Evenly spontaneous and planned*', '*Predominantly planned*' and '*Completely planned*'.

Among academic sources used to infer urban processes, none were as instrumentally valuable as the works of Kostof (1992, 1991a, 1991b). As much for his encyclopedic knowledge of cities across the world as for his deep understanding of urban processes from the most motivated ones to the most ordinary ones. For the systematic overview of this information, which is the basis for the qualitative classification of the sample to an IoP category, we refer to Supplementary Material 1.

3.3. Estimation of the intra-urban morphology by the structural complexity

3.3.1. Geodata base for analysis of the urban layout

We predominantly rely on geodata from the collaborative worldwide mapping project OpenStreetMap (OSM) using the python library developed by Boeing (2017). As OSM is based on volunteered contributions, the completeness of its data relies on the involvement of map pers, local or not, that chose their working site. As such, places of secondary interest for worldwide mappers and places with few or none voluntary mappers often are omitted or only superficially covered (Herfort et al., 2021; Seto, Kanasugi, & Nishimura, 2020). This omittance mostly happens in poorer places or less densely populated places (Herfort et al., 2021). Conversely, settlements with complex morphologies like slums, or of small size such as villages are less represented. Further, the incompleteness of the OSM project is more important for buildings than for road data (Herfort et al., 2021). Overall, this creates an imbalance going against our goal of multi-dimensional representativity in data, study sites and thus, our categories of IoP. Therefore, we complete our dataset by data from the study of (Taubenböck et al., 2018a) covering deprived places across the globe on building level and by manual digitization when Open Street Map do not provide consistent or comprehensive enough data. When digitizing, we remain consistent with the OSM format.

Overall, our 381 study sites reflect a selection of settlements where we could confirm visually that in our geodata not more than ~10 % of the structural elements were missing or that the missing data would influence drastically the quantitative estimation of the intra-urban morphology.

For the empirical examination of the intra-urban morphology, we use the structural elements of *buildings, morphological units*, and *streets*, as outlined in section II) B).

For the structural element *buildings*, we gather solely building's footprints as the building height is yet not consistently present in OSM data. The rooftop represents a single building and is used as proxy in formation for a building's ground area. For details about manual mapping in complex urban areas we rely on (Kraff et al., 2020a).

As for the *streets*, we extract them as geographical network, composed of nodes and edges as done by (Boeing, 2019). As no information on width could be persistently used, we limit ourselves, here, to consider streets as linear objects.

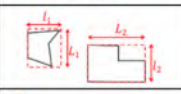
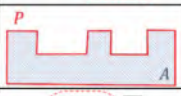
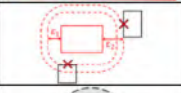
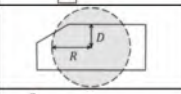
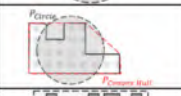
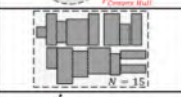
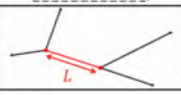
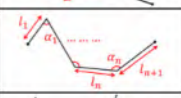
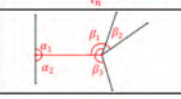
For the *Morphological Units* (*MUs*), as discuss in section II)b), we see this structural element as a generalization of the '*aggregate of lots*' (Kropf, 1993). As discussed in Fleischmann et al. (2020) retrieving individual lots, or plots (beyond considerations on their existence outside of formal estate economy) is no trivial task as the data is mostly unavailable. Therefore, we extract the MUs as the polygons constrained by street meshes when the streets form such a mesh, following the street-dependent acceptation of the definition proposed in Kropf (1993). When streets do not form meshes (mostly in *Extremely Spontaneous* settlements), we propose the production of the outlines using an aggregation algorithm based on an adapted spatial version of the OPTICS clustering (Ankerst et al., 1999). The process of production of the outlines is introduced in Appendix B.

3.3.2. Morphological parameters

We aim at investigating in quantitative manner whether the IoP has a relation on the three structural elements considered here: *buildings, morphological units* and *streets*. We therefore rely on measurements of specific characteristics of these structural elements. Furthermore, following (Batty, 1974, 2016; Boeing, 2019; Cozzolino, 2021; Habraken, 2016; Jacobs, 1961; Kostof, 1991b), we specifically believe that the diversity of the urban morphology is negatively corelated to processes of norms and standardization. Therefore, to probe the IoP within settle ments, we propose to compute the diversity of sets of three parameters for each of the structural elements. Later, we refer to this diversity as the *structural complexity*.

The morphologic features considered here are mostly geometric measurements, or *morphometric characters* (Dibble et al., 2019). We choose them because they encompass features directly or indirectly impacted by planning ordinances and norms. We describe these parameters, their computation and their relevance with respect to urban planning in Table 2.

Table 2. Description of parameters of structural elements and of the specifics for measurement. (See below-mentioned references for further information.)

Structural element	Parameter	Description	Illustration	Equation	Bin size
Buildings	Building dimensions	*Building dimensions* is a feature measuring the width and the depth of buildings footprints. We define the buildings dimension as the double of values (L, l), L being the long side of the bounding box of the building and l its small side. Dimensions of buildings give information on their functions and on the financial context as well as land availability when the buildings were constructed.		(L, l)	(5m; 5m)
	Compacity	*Compacity* is a geometrical feature describing the shape of buildings. We compute the geometrical compacity of buildings by considering the ratio between the perimeter of the building and the square root of its surface. Here we consider the compacity of a building as a proxy for the specific design of the building's footprint, from simple to complex ones.		$\frac{P}{\sqrt{A}}$	0.1
	Wall-to-wall distance to neighbors	*Wall-to-wall distance to neighbors* is a parameter defined as the mean distance to the two closest neighbors of the building. The wall-to-wall distance to neighbors sheds light on the types of mutual distances the buildings have in their vicinities with each other. It is a very local measure of building density.		$\frac{\varepsilon_1 + \varepsilon_2}{2}$	0.5m
Morphological units	Girth index	*Girth index* is a ratio describing the nominal thickness of a shape. It is computed as the ratio of the thickness of the layer insulating the innermost core of the shape from its periphery and the radius of a circle with the same area as the shape. This indicator proposed by Angel (2010) translates the relation between the core of the morphological unit and its boundaries. In other words, this reflects how deep the morphological unit is.		$\frac{D}{R}$	0.01
	Detour index	*Detour index* is a measure computed as the ratio between the perimeter of an equal-area circle and the perimeter of the convex hull of a shape. This index proposed by Angel (2010) reflects the typical distance one has to walk to skirt a shape. Here, we use it as a proxy for the design of the morphological units.		$\frac{P_{Circle}}{P_{Convex\ Hull}}$	0.01
	Order of magnitude of number of buildings	*Order of magnitude of number of buildings* is computed by taking the logarithm of its raw number of buildings. This parameter indicates how many buildings are located within a morphological unit while acknowledging that there are more differences between small headcounts (e.g. of 5 and 6 six buildings) than between large headcounts (e.g. of 100 and 101 buildings).		$\log_{10}(N+1)$	0.1dit
Streets	Length of street section	*Length of street section* is a geometric feature measuring the length between two crossroads. This parameter is a proxy of the regularity of the mesh size in the street network.		L	5m
	Curviness	Curviness is a geometric measurement of how bent and counter-bent a section of a road is. We compute it as the mean of the angles between the segments of the road, weighted by the length of the segments. This measure quantifies the design of the curvature through the settlement.		$\sum_{i=1}^{n} \frac{\alpha_i * (l_i + l_{i+1})}{2 * \sum_i l_i}$	0.1°
	Symmetricity of crossroads	*Symmetricity of crossroads* is measured as the doublet (α, β) which is respectively the mean angle of the paths protruding from both ends of a section of road. Here we use this couple of values as a proxy for the design of crossroads emerging from a road section.		$(\overline{\alpha_i}, \overline{\beta_i})$	(5°; 5°)

3.3.3. Structural complexity as joint entropy of morphological parameters

For operationalization of the *structural complexity*, we compute for each settlement the normalized Shannon entropy (Shannon, 1948) of the individual parameter distributions and the normalized joint-entropy (Cover & Thomas, 2005) of the set of three parameters for each structural element (see Table 2). The entropy and joint entropy are computed by approximating, in each individual settlement, the parameters' distribution by their histograms as in (Purwani, Nahar, & Twining, 2017). The bin-widths for each parameter is chosen with respect to the scale of its physical reality and is noted in Table 2. Shannon entropy is a measure of a system's diversity, or information complexity. Here, we use it assess the dispersion of values computed for both individual parameters and for sets of parameters. This is motivated by its presumed information related to convergence to, or divergence from, uniformity, fostered by planning, ordinances, rules, or contextual constraints. Respectively,

the joint entropy mathematically represents the individual entropies of its variables reduced by their mutual information (Cover & Thomas, 2005). In our case, we understand it as the specific diversity for a structural element of the combined triplet of morphological parameters in a given settlement. We further refer to the joint entropy of a *structural element* in a settlement as its *structural complexity*. Equations on normalized entropy and normalized joint entropy can be found in the Appendix C.

The introduced parameters in Table 2 are common spatial metrics in this domain. However, it is clear that any selection of spatial metrics cannot be comprehensive. To make sure that

these selected metrics do not produce statistical randomness in the measurement of the urban layout, we calculated 139 additional metrics and analyzed to what extent the results obtained with these metrics hold up (see Supplementary material 2).

3.3.4. Statistical analyses on the relation between the Intensity of Plannedness and the structural complexity

The crucial point in our study is to demonstrate the relation between the categorized IoP derived from the literature and the output of the *structural complexity* computed on the structural elements.

To demonstrate this, we analyze whether the values of *structural complexity* differ significantly between the five categories of IoP. To do so, we analyze the variance among the distributions of the different categories. As the distribution of the *structural complexity* measures is not assumed to trivially follow normal gaussian distributions, we resort to a non-parametric version of the analysis of variance (ANOVA) (Fisher, 1925): the Kruskal-Wallis test (Kruskal & Wallis, 1952). The KruskalWallis test indicates whether the H0 (null hypothesis) is respected, i.e., whether the distributions of each category of IoP are all equivalent to each other. As a threshold to decide whether H0 is respected, we consider the generally accepted value of $\alpha = 0,\ 05$ (Glantz, 2012). H0 is respected if the p_{KW} value of the Kruskal-Wallis test is above this threshold. If the H0 is not verified, we assume that at least one statistical distribution is different from the rest. To observe which statistical distribution differs from the rest, we proceed with a post-hoc pair-wise comparison of significant differences among the individual categories' distributions using the Mann-Whitney test (Mann & Whitney, 1947; Wilcoxon, 1945). Here as well, we consider a pair of distributions to be significantly different if $p_{MW} \leq 0{,}05$.

4. Results

In general, our findings empirically show, that the elements of the built landscape result in a physical form that reflects the intensity of the underlying human planning, i.e. it relates to the conceptualized IoP. In this section, we first illustrate the distribution of the entropy per individual parameter as well as for each structural element in general. Beyond this, we reveal the result of their statistical significance tests.

In the following paragraphs, we first present the results with Figs. 6a and 6b. This figure is composed of snippets of 500 m × 500 m of figure ground diagrams, each displaying, the morphology of *buildings* per settlement. The color of the snippet's border relates to its IoP category. The snippets are ordered by increasing *structural complexity* of the element *buildings*. The equivalent figures for, respectively the structural element *morphological units* and the structural element *streets* can be found in the Appendix, respectively Appendixes D1 – E2.

Subsequently, we illustrate the statistical distribution of the structural complexity in Fig. 7. For each structural element, the statistical distribution of the values of our metric per settlement grouped by categories of IoP is displayed. In this figure, as a supplementary information, we also plot the statistical distribution of the entropy for each individual

parameter which the structural complexity is based on. Finally, we discuss the statistical relationships between the categories of IoP and the *structural complexity* for each structural element.

For the structural element *buildings*, as observed in Fig. 6 and Fig. 7, the different categories of IoP tend to overlap, however, still distinguishable trends are measured. For the parameter *building dimension*, a consistent increase of the entropy with increasing IoP is found. In comparison, *compacity* presents a slightly different profile. The entropy is quite consistently low for a low IoP. It increases to a relative maximum with the category *Evenly spontaneous and planned* and a decrease again with the category *Completely planned.* This category shows a large spectrum of entropy possible but, on the majority, as low as the *Extremely spontaneous* settlements. For the third parameter, the *wall-to-wall distance to neighbors,* there is a clear pattern of, on average, increasing entropy with increasing IoP, with a slight recess for the category *Evenly spontaneous and planned.* When combined, the set of the three parameters concerning the *buildings* show an increasing structural complexity along with the ranks of categories of plannedness.

Statistically, the distributions of the structural complexity for the structural element *building* reveal differences. With 260 settlements, the result of the Kruskal-Wallis analysis shows a p_{KW} value of 2 e^{-17}, which is far below the threshold of 0,05. Hence, we conclude that categories of IoP feature at least one category significantly different from the others for the *buildings*.

This is generally confirmed by the Mann-Whitney analysis (Table 3). With the exception of the two categories with the highest IoP – *Pre-dominantly planned* and *Completely planned* – we measure significant differences in term of structural complexity for all other categories. Thus, we conclude that for the structural element *buildings*, settlements of different IoP categories are discernable based on their structural complexity.

Table 3. Symmetric matrix of the pairwise results of the Mann-Whitney significance test for structural element Buildings. A value of $pMW \leq 0,05$ indicates that the two categories are significantly different in term of structural complexity. For ease of understanding, we crossed the values above the threshold.

Structural element: **Buildings**	Extremely spontaneous	Predominantly spontaneous	Evenly spontaneous and planned	Predominantly planned	Completely planned
Extremely spontaneous		…	…	…	…
Predominantly spontaneous	$6,9 \times 10^{-3}$		…	…	…
Evenly spontaneous and planned	$3,9 \times 10^{-8}$	$5,5 \times 10^{-3}$		…	…
Predominantly planned	$5,7 \times 10^{-11}$	$7,0 \times 10^{-7}$	$7,9 \times 10^{-3}$		…
Completely planned	$3,0 \times 10^{-10}$	$4,1 \times 10^{-6}$	$1,9 \times 10^{-2}$	~~1,0~~	

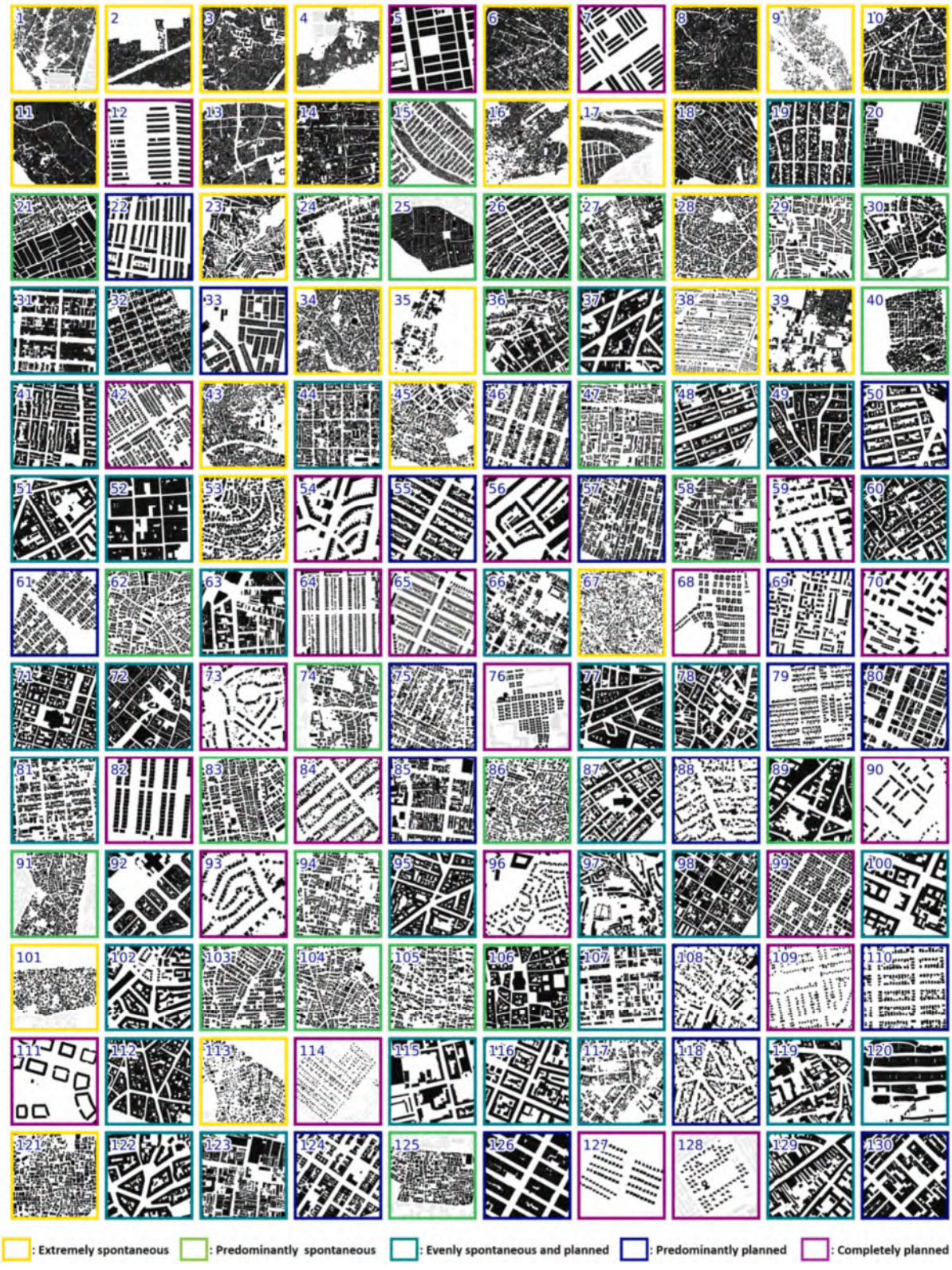


Fig. 6a. Figure-ground diagram snippets of a 500 m × 500 m subset of each study site used for the analysis of the structural element 'buildings'. The snippets are ordered by the measured increasing structural complexity and framed in color with respect to their IoP. Buildings in black are part of the study site, the ones in faded grey are not. List of the study sites below.

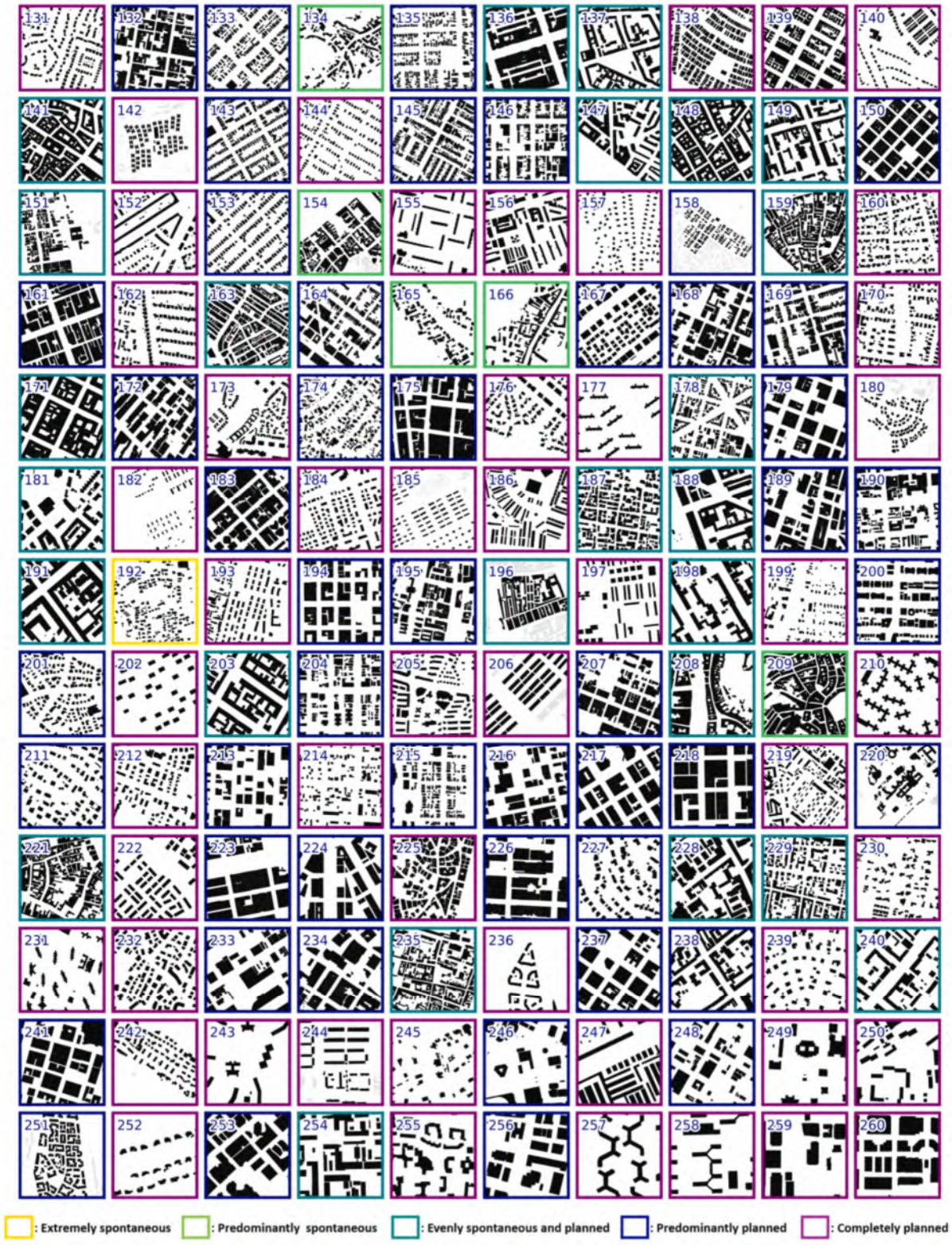


Fig. 6b. Figure-ground diagram snippets of a 500 m × 500 m subset of each study site used for the analysis of the structural element 'buildings'. The snippets are ordered by the measured increasing structural complexity and framed in color with respect to their IoP. Buildings in black are part of the study site, the ones in faded grey are not. List of the study sites below.

1. Manila North Cemetery
2. Manila Tondo
3. Mumbai Dharavi
4. Mumbai Bharat Nagar
5. Khartoum Al-Sahafa
6. Lagos Makoko
7. Chandigarh Sector 59 East
8. Caracas Petare
9. Afgooye Lafoole Refugee Camp
10. Mexico City Magdalena Contreras
11. Nairobi Kibera
12. Qazvin Navvab Safavi
13. Mumbai Santosh Nagar
14. Beijing Dashilanr
15. Cape Town Silver Town
16. Rio De Janeiro Turano
17. Cape Town Victoria Merge
18. Cairo Manshiet Nasser
19. Santiago De Chile Lo Prado
20. Cairo Imbaba
21. Tunis Cite Ettadhamen
22. North Philadelphia
23. Lima Comas
24. Bucharest Tei Toboc
25. Dhaka Islambagh
26. Teheran Eslamshahr
27. Ho Chi Minh Khu pho2
28. Dar es Salaam Tandika
29. Kyoto West
30. Lisbon Cova
31. Rio de Janeiro Morro do Cantagalo
32. Sliven Nadezhda
33. Brasilia Paranoa
34. Dar es Salaam Keko
35. Madrid Canada Real
36. Rio de Janeiro Morro Do Valongo
37. Brussels Molenbeek
38. Al Zaatari Refugee Camp
39. Dhaka Kadamtali
40. Shenzhen Baishizhou
41. Athens Agios Panteleimonas
42. Calcutta Sector 1
43. Dar es Salaam Lugalo Golf
44. Kyoto Nakagyo Ward
45. Izmir Kadifekale
46. Los Angeles West Addams
47. Kyoto Yamashina Ward
48. Amsterdam Jordaan
49. Madrid Centro Madrid
50. New York Lower Manhattan
51. Brussels Anneessens
52. Budpest Monserrat
53. Istanbul Gulsuyu Gulensu
54. Edinburgh Greenbank
55. Manhattan East Village
56. Edinburgh West End
57. Hyderabad Boudha Naga
58. Hyderabad Gandhinagar
59. Edinburgh The Grange
60. Merida Old Center
61. Tokyo Urayasu
62. Tokyo Minamimagome
63. Sao Paulo Centro
64. Riverside North Lawndale
65. Forest Hill Forest Hill
66. Ho Chi Minh Binh Tho
67. Lusaka Misisi
68. Bangkok Panthiya Village
69. Bogota Polo Club
70. Madrid Moratalaz
71. Pavia Old Center
72. Barcelona Ciutat Vella
73. Edinburgh Trinity
74. Kyoto Minamieganosho
75. Dar es Salaam Kiwalani
76. Seoul Hankuk University
77. Paris Central Arrondissements
78. Boston North End
79. Obihiro Jominami
80. Tokyo Shinbashi
81. Nagoya Naka Ward
82. Brasilia Lago Sul
83. Seoul Mangwon-dong
84. Sydney Padstow
85. Porto Alegre Menino Deus
86. Katmandu Anamnagar
87. Naarden Naarden
88. Paris Gagny
89. Trier Hauptmarkt
90. Columbia Maryland Owen Brown
91. Katmandu Kalopul-Ratopul Rd
92. Barcelona Eixample
93. Edinburgh Craiglockhart
94. Calcutta Lake Market
95. Paris 9 Arrondissement
96. Toulouse Le Mirail
97. Lisbon Alfama
98. Bogota La Canderlaria
99. Delhi Karol Bagh
100. Berlin Kreuzberg
101. Dar es Salaam Kigamboni Edge
102. Munich Marienplatz
103. Calcutta Lake Garden
104. Tokyo Shimokitazawa
105. Nagoya Arakocho
106. Rome Municipio I
107. Tokyo Sumida
108. Paris Adamville
109. Gaborone 8/10/11/12 Extensions
110. Miami Little Havana
111. Madrid Horcajo
112. Paris 6 Arrondissement
113. Lusaka Bauleni
114. Lusaka Luburma Central
115. Dublin Temple Bar
116. Munich Maxvorstadt
117. Dar es Salaam Kigamboni
118. Paris Clamart
119. Paris Belleville
120. Bern Innere Stadt
121. Dhaka Motijheel
122. Prague Prague
123. Historic Centre of Mexico City
124. New York Bronx
125. Katmandu Bagbazar
126. Manhattan Midtown
127. Qazvin Artesh Hospital
128. Calcutta Sector 1 Block CC
129. Bratislava Old Town
130. San Francisco South Of Market
131. Brie Comte Robert
132. Philadelphia Center City
133. Denver Five Points
134. Trier Im Falschen Biewertal
135. Miami Allapattah
136. Edinburgh New Town
137. Berlin Europa City
138. Delhi Greater Kailash
139. Manhattan Washington Heights
140. Charlottesville Fry's Spring
141. Vienna Innenstadt
142. Antananarivo Antsahamamy
143. Ottawa Downtown
144. Charlottesville Belmont
145. Perth Center
146. Washington Downtown
147. London Westminster
148. Budapest Inner City
149. Edinburgh Old Town
150. Cape Town Central Business District
151. Phnom Penh Sangkat Srah Chak
152. Munich Harlaching
153. Dallas Junius Heights
154. Antananarivo Central Station
155. Sarcelles HLM
156. London Tower Hamlet
157. Oslo Nordberg
158. Lusaka Bauleni West
159. Venice San Polo
160. Munich Obermenzing
161. Melbourne Central Business District
162. Edinburgh East Craigs
163. Nairobi Central
164. Richmond Downtown
165. Suloszowa Linear City
166. Pavia Borgo Ticino
167. Abu Dhabi AL Danah
168. Wellington Te Aro
169. Toronto Old Toronto
170. Munich Bogenhausen Nord
171. Oslo Sentrum
172. Downtown Los Angeles
173. Evry Evry-Courcouronnes
174. Dar es Salaam Masaki
175. Sydney Central Business District
176. Trier Petrisberg
177. Teheran Ekbatan South East
178. Belo Horizonte Santo Agostinho
179. Downtown Portland
180. Antananarivo Cite Belle Vue
181. Oslo Frogner
182. Dar es Salaam Mikoroshoni
183. Brisbane Central Business District
184. Pavia Parco Leopardi
185. Dar es Salaam Mikoroshoni (West)
186. Delhi Laxmi Bai Nagar
187. Bairut Hamra
188. Warsaw Srodmiescie
189. Auckland Central Business District
190. Pretoria Central
191. St Petersburg Admiralteysky District
192. Nakatsugawa Ochiai
193. Pavia Via Umberto Olevano
194. Downtown San Diego
195. Colombo Fort
196. Venice Castello
197. Brasilia North Wing
198. Saint-PBurg Vasileostrovsky District
199. Glendale Glendale
200. Vancouver Fairview
201. Rabat Administrative Sector
202. Brasilia Lago Sul Lake
203. Saint-Petersburgh Smolninskoye
204. Denver North Capitol Hill
205. Shanghai Tianping Res District
206. Brasilia ASA Sul
207. Charlottesville North Downtown
208. Ljubljana Center District
209. Sienna Old Center
210. Manhattan StuyTown
211. Budpest Gellerthegy
212. Ferrara Via Comacchio
213. Kansas Downtown
214. Bishkek Park
215. Indianapolis Mile Square
216. Chicago Central Business District
217. Downtown Seattle
218. Chicago Loop
219. Ferrara Ippodromo
220. Antananarivo Lake Anosy
221. Venice Giudecca
222. Reston North West
223. Durban Central
224. Detroit Downtown
225. Cairo Garden City
226. Calgary Downtown Commercial
227. Oslo Holmenkollen
228. Moscow Tverskoy District
229. Ferrara University District
230. Johannesburg Robin Acres
231. Manhattan Williams Bridge
232. Sienna San Prospero
233. Cleveland Downtown
234. Downtown Atlanta
235. Ferrara Doro
236. Paris Amerique
237. Houston central business district
238. Minsk City Center
239. Oslo Ulvoya
240. Kiev Pechers'kyi
241. St Louis Downtown
242. Sienna Ravacciano
243. Minsk Syerabranka
244. Seoul Bupyeong 1
245. Delhi Meena Bagh
246. Tokyo Shinjuku Center
247. Saint-Petersburgh Municipal Okrug
248. Minneapolis Downtown West
249. Minsk Kamennaya Gorka
250. Minsk Kastrycnicki Raion North
251. Venice Castello West
252. Madrid La Finca
253. Charlotte City Center
254. Saint-Petersburgh Moskovskaya Zastava
255. Munich Arabella park
256. Tokyo Chiyoda
257. Teheran Ekbatan Phase 2
258. Teheran Ekbatan Phase 1
259. Columbia Maryland Town Center
260. Reston Town Center

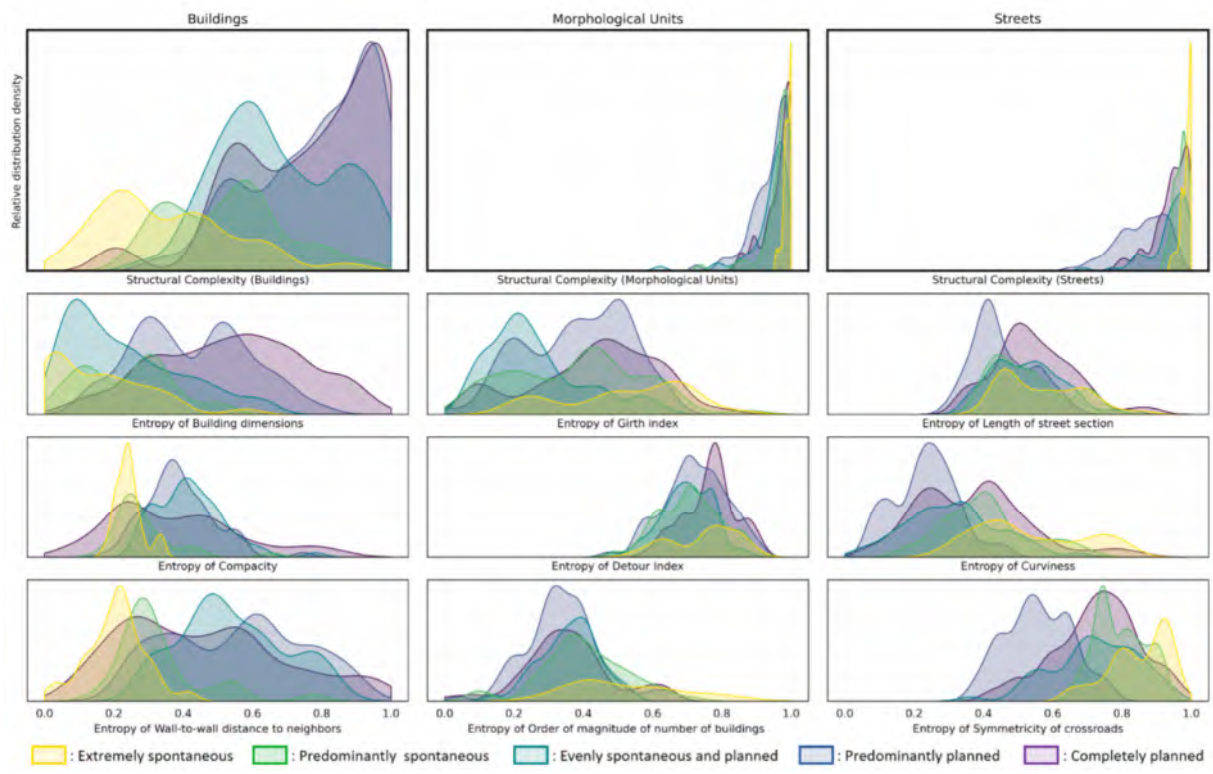


Fig.7. Distribution of the structural complexity (top row) and distribution of the individual entropies of the three parameters, grouped and colored by category of IoP for each of the three strucural elements.

For the *morphological units*, the two first parameters seem to present more compact distributions of entropies for the settlements of different categories of IoP. For the *girth index*, even though clumped together, we discern a trend of decrease and then increase of the entropy with higher IoP, with a minimum for the *Evenly spontaneous and planned* category. The *detour index,* on the other hand, presents less discernable differences in terms of entropy, except for an ever so slight decline with increasing IoP. The last parameter of the *morphological units* distinguishes itself and offers a less packed profile: The entropy of the *order of magnitude of number of buildings* is decreasing and then re-increasing with ascending categories of IoP, just as the *girth index* but in a more discernable way. The structural complexity of the *morphological units*, being the joint entropy of all three parameters, shows for all categories a high entropy indicating high variability of the individual parameters with regard to each other's. We observe that for the MUs as well, the different categories of IoP tend to overlap. but some trends exist. Here, we find a slight decrease and then a secondary increase with a minimum for the category *Evenly spontaneous and planned*.

With the test being led on 260 individual settlements, the result of the Kruskal-Wallis analysis shows a p_{KW} value of 1 e^{-08}, which is below the threshold of 0,05. However, this value is higher and closer to it than for the *buildings*. We conclude that categories of IoP present at least one category significantly different from the others. The Mann-Whitney analysis results in the following matrix (Table 4):

We conclude for the structural element *morphological units* that most of the categories of IoP are discernable based on their structural complexity. Still, two pairs of categories that are relatively close, '*Evenly spontaneous and planned/Predominantly planned'* and '*Predominantly planned/Completely planned'* cannot be considered as significantly different. Further, for this structural element, two categories with high differences of IoP '*Predominantly spontaneous/Completely planned'* feature the same types of structural complexity and '*Predominantly spontaneous* and *Predominantly planned'* are not significantly different.

Table 4. Symmetric matrix of the pairwise results of the Mann-Whitney significance test for structural element MUs. A value of $pMW \leq 0,05$ indicates that the two categories are significantly different in term of structural complexity. For ease of understanding, we crossed the values above the threshold.

Structural element: **Morphological Units**	Extremely spontaneous	Predominantly spontaneous	Evenly spontaneous and planned	Predominantly planned	Completely planned
Extremely spontaneous		…	…	…	…
Predominantly spontaneous	$1,9 \times 10^{-3}$		…	…	…
Evenly spontaneous and planned	$3,1 \times 10^{-8}$	$1,3 \times 10^{-2}$		…	…
Predominantly planned	$2,3 \times 10^{-6}$	~~$6,7 \times 10^{-2}$~~	~~1,0~~		…
Completely planned	$3,1 \times 10^{-2}$	~~1,0~~	$2,4 \times 10^{-2}$	~~$6,3 \times 10^{-2}$~~	

For the structural element *streets*, we observe again a different statistical trend for the different individual parameters. First, for the *length of street section*, although presenting slight but still discernable differences of entropy between the different categories of IoP, there is no major trend identifiable. Rather, the distribution fluctuates back and forth around the main mode of the entropy distribution of the *Extremely spontaneous* category. For the *curviness*, the profile gives a clearer trend: the entropy decreases constantly with ascending IoP and only increases again for the category *Completely planned*. For the *symmetricity of cross roads*, the categories of IoP reveal mostly distinctive entropies with again the same trend as previously observed: Values decrease progres sively with increasing IoP, except for the *Completely planned* category where we observe a re-increase. In summary, we observe for the struc tural complexity of the *streets*, the same type of trend as for the *morphological units*, an overall clustering of the categories in the high values of the spectrum.

The result of the Kruskal-Wallis analysis led on 296 settlements shows a p_{KW} value of 3 e^{-19}, which is even further below the threshold of 0,05 than for the *buildings*. Hence, we conclude that categories of IoP present at least one category significantly different from the others. The results of Mann-Whitney analysis are presented in Table 5.

For the structural element of *streets*, we generally ascertain significative differences in term of structural complexity of the different categories of IoP. The exception is the category

Completely planned being similar to the category *Evenly spontaneous and planned* and not significantly different from the category *Predominantly spontaneous*. Note worthy is the fact that these two last categories are, between them, significantly different. Further, contrary to the two previous structural elements, the difference between *Predominantly planned* and *Completely planned* is highly significant.

To summarize, from the analyzes on the three structural elements, it is empirically clear that the built landscape does not correlate linearly or one-to-one with the degree of planning. In general, however, our results prove a difference in the particular distributions of the parameters. In other words, despite overlapping in the distributions, a definite trend is emerging that reveals empirically that the built landscape reflects the degree of planning.

Table 5. Symmetric matrix of the pairwise results of the Mann-Whitney significance test for structural element Streets. A value of $pMW \leq 0,05$ indicates that the two categories are significantly different in term of structural complexity. For ease of understanding, we crossed the values above the threshold.

Structural element: **Streets**	Extremely spontaneous	Predominantly spontaneous	Evenly spontaneous and planned	Predominantly planned	Completely planned
Extremely spontaneous		…	…	…	…
Predominantly spontaneous	$5{,}0 \times 10^{-4}$		…	…	…
Evenly spontaneous and planned	$1{,}0 \times 10^{-7}$	$2{,}7 \times 10^{-2}$		…	…
Predominantly planned	$7{,}7 \times 10^{-17}$	$7{,}8 \times 10^{-18}$	$8{,}0 \times 10^{-10}$		…
Completely planned	$1{,}8 \times 10^{-7}$	~~$6{,}2 \times 10^{-2}$~~	~~1,0~~	$2{,}5 \times 10^{-13}$	

5. Discussion

In most cases, the physical structure of cities repeatedly undergoes a chain of transformation processes over time. The resulting morphology is an artifact that we have to observe through the scope of these successive modification sought by people. This scope allows us to under stand the intra-urban morphology as the crystallization of urban processes whose main drivers are societal in nature. Despite the knowledge of this relationship between urban forms and undergone urban processes, to unveil the later from the first is a challenge.

In this paper, we proposed a conceptual frame of bottom-up and top-down planning to examine the intra-urban morphology in this regard and to empirically reveal the intertwined relation of urban processes and built forms. We developed a categorical scale of the IoP for urban processes and we tested empirically whether this scale is adequate to observe significant physical differences in the measured *structural complexity* operationalized by three main structural elements of the urban fabric: the *buildings*, the *morphological units* and the *streets*. We used an unprecedented large sample of 381 settlements from as many different cultural zones of vernacular architecture as we could find reliable information on.

In our experiments, we were able to empirically prove what has often been described in the literature in a theoretical manner: There exist generally and globally significant morphological appearances arising from distinct types of urban processes, here understood as different categories of IoP. Additionally, the pairwise comparisons of *structural complexities* of different categories of IoP expose, in most cases, largely statistically significant differences in their morphologies, as estimated by our set of parameters. For the buildings, 9 out of 10 pairwise comparisons show significant differences in the *structural complexity* of the categories of IoP. For the Morphological Units, 6 out of 10 and for the streets, 8 out of 10 present significant differences.

Interestingly, no manifest co-evolution between the *structural complexity* and the IoP can be showcased here. As a rule of thumb, one might have expected the *structural complexity* to be at its lowest when the process of standardization is at its strongest. In consequence, the *structural complexity* was expected to decrease with an increasing IoP. What we observed, on the contrary, is that the *structural complexity* of the *buildings, morphological units* and *streets* evolve very differently with regard to their IoP. Although anecdotical for this study, this specific secondary result seems interesting and is suggested e the for further studies.

Despite our approach demonstrating the connection between the IoP and our estimator of the urban morphology – the *structural complexity* – this study also highlights some nuances: For each structural element, a few pairs of distinct categories of IoP present hardly distinguishable *structural complexities*. Also, in the case of every structural element, the juxtaposition of the *structural complexity* with the individual entropy of each constitutive parameter brings forth a complex situation of mutual information between parameters that was not quantified here.

Although these nuances are not impeding our demonstration of the relation between the IoP and the intra-urban morphology, they raise interesting limitations of the use of the *structural complexity* as an un equivocal estimator of the intra-urban morphology outside of the scope of this study.

We acknowledge that the specific set of morphological parameters we have chosen, despite being based on known urban processes, is non-exhaustive. We do not see this set of parameters as being necessarily the only one possible for the demonstration intended here. Therefore, beside the *structural complexity* and its constituent morphological parameters presented in the main body of this article, we investigated the relation between IoP and other morphological parameters. We tested 58 additional parameters (resp. 60) for the buildings and the Morphological Units proposed in (Fleischmann, 2019) and consolidated with the additions presented in (Fleischmann et al., 2020). For the streets, we computed 21 additional parameters issued in (Boeing, 2017, 2019). On these additional parameters, we computed the individual normalized entropy when fitting, and then proceeded to the Kruskal-Wallis test and, when required, the Mann-Whitney test. The results of this additional analysis concur with the empirical relations achieved with the *structural complexity* and are compiled in Supplementary material 2. That is, among the pairs of categories of IoP, in general, significant differences are showcased by most of these parameters. With it, we show that the literature-based selection of the spatial parameters has not led to space-statistical coincidence. These additional parameters paint a more complete picture of what are relevant

parameters showing contrast in the morphology of different categories of IoP (see Supplementary material 2). This was not investigated here and could be at focus in further studies. In this article, we focus specifically on the *structural complexity* as its constitutive parameters are clearly linked to the urban processes, simple enough in their formulation and they have established them selves to be suitable to answer our research question.

On a more general ground, we believe that shallow statistics, such as the parameters used here, are limited in their capacities to unveil, in their uttermost depth, the interdependencies between urban process and resulting morphologies. Partially or completely overlapping distributions of *structural complexities* established this. Therefore, it is not our claim here that the IoP can be automatically classified using this *structural complexity* metric or other morphological parameters out of com parable geodata.

This lack of direct prediction power from our quantitative measure was to be foreseen in the face of the complexity of the intertwined path dependencies of urban forms and processes but does not deter any qualitative prevision regarding the morphology displayed by these processes (Moroni & Cozzolino, 2019).

Beyond these considerations on the validity of the *structural complexity* as an unequivocal estimator of the intra-urban morphology, we do want to acknowledge observed limitations within our study. The origins of these limitations were kept, by the authors, as circumscribed as possible not to mitigate the overall analysis.

The first limitation is tied to the nature of the data used. Despite being of really high spatial resolution, OSM data, and other manually digitized data are prompted to be inconsistent across settlements. For example, the way buildings are digitized can vary from one polygon for a continuous built street block, to multiple polygons for a single-family house as they can vary from interpreter to interpreter (Kraff et al., 2020a). We paid attention to this while curating the data to eliminate gross inconsistencies. Still, residual inconsistencies remained in our dataset. This type of inconsistencies occurring across our samples tend to blur the computation of the *structural complexity* and, as a consequence, impact the analysis by an unknown factor. Datasets of large coverage in the likes of (Microsoft, 2020) would be of great interest in future studies to help eliminate further residual inconsistencies.

The second limitation is related to the residual uncertainties in the qualitative assessment of the IoP. While a great deal of effort has been made in assessing the categories of IoP for each type of structural element per settlement based on an extensive literature research, we acknowledge that an irreducible part of uncertainties remains. Levels of certainty in the assessments of the IoP can originate both, in the quality of sources available and in the clarity of our ontology. E.g., whereas categories at the ends of the spectrum of our IoP concept, i.e. *Extremely spontaneous* and *Completely planned* are easy to identify, intermediary categories – from our experience – are less unambiguous when classifying along our ontology. In the Supplementary Material 1, we propose a self-evaluation of our level of certainty with respect to the assessment of the IoP. In addition, while in that scope of this explorative study, this relative uncertainty was not a deterrent, we encourage future research to further refine and detail the ontology and the operationalization of the assessment of levels of the IoP to curb as much as possible the part of uncertainty.

Despite the existence of these limitations within our study, we argue that the validity of the quantitative measure of the *structural complexity* and the qualitative assessment of the IoP is transparent according to a clear literature-based approach. Moreover, we argue that the high amount of sample used for each statistical test, further add to the reliability of the global analysis and of our empirical demonstration.

The primary scientific aim of this study is to elaborate the concept of the IoP and to empirically investigate its relation to the built-up form.

With the concept of the IoP, we aim to bring forward the under theorized political and societal component of the making of the intraurban morphology. Therefore, the IoP shall serve as a stable concep tual ground to renew the study of a large body of topics where the built-up form stands at the interface with societal matters.

In his work, Christopher Alexander (Alexander, 1979) put forth "*a quality without a name*" of prosperous urban spaces. Although we do not claim here that the IoP of an urban space explains altogether the actual existence or not of this "*quality without a name*" within it, we nonetheless do believe that it is indeed one of the underlying tenets to this quality. As such, a field of study, trailblazing the "*unnamed*" quality as suggested by Alexander and where we do believe in the relevance of the integration of the ontology of the IoP is the study of the "*urban quality of life*" (Higueras, Roman, & Farina, 2021).

This field of study has already proven relevance in the analysis of urban fabric's form to investigate multiple facets of individual and collective qualities of urban life (McDonald, Wise, & Harris, 2008; Mouratidis, 2018; Sapena et al., 2020, 2021). It has already been shown that the form of the built-up environment share a relation with aspects driving the urban quality of life such as (among others) well-being, physical health, mental health (Frank & Engelke, 2001; Hajrasoulih, del Rio, Francis, & Edmondson, 2018; Krefis et al., 2018; McDonald et al., 2008; Mouratidis & Poortinga, 2020; von Szombathely et al., 2017), and social relations (Legeby, Berghauser Pont, & Marcus, 2015; Mouratidis & Poortinga, 2020; Mouratidis, 2018). Nonetheless, this field of study displayed so far some limitations in fully explaining how the material features of the built-up environment do influence these aspects of the urban quality of life (Berghauser Pont et al., 2021; Hajrasoulih et al., 2018). As such, the body of literature on this topic ac knowledges the great complexity of it that fail to be captured by physical features alone (Hajrasoulih et al., 2018; von Szombathely et al., 2017).

Therefore, numerous arguments could be made in favor of integrating the IoP ontology in this field of study, on the base of its unique characteristic to be at the interface between the material and the social part of the built-environment. Yet, one decisive argument for the relevance of the IoP to this field of study is related to one of the important aspects of the urban quality of life proposed in (Higueras et al., 2021), namely the levels of freedom of the inhabitants. We argue that this aspect relates greatly to the notion of self-agency of the population in the making and adapting of their urban environment that we argue to be a key factor of the ontology of the IoP.

Although we see the study of the urban quality of life as a privileged field of research where the ontology of the IoP could be key to operationalize this multi-faceted, complex phenomenon, we do not see a limitation of its relevance to this field only. We also argue

firmly that its benefits extend to the wider scope of study of the empirical signs of the collective, social, economic or even esthetical prosperity urban spaces at large.

Therefore, we suggest to integrate the use of the concept of IoP to the study in a large range of urban topics such as, the capacity of communities to form in cities (Mouratidis & Poortinga, 2020; Tonkiss, 2013), the possibility for informal micro-economy to emerge (Salat & Bourdic, 2012; Thai, Stevens, & Rogers, 2020; Thai, Stevens, & Rogers, 2019), capacity to adapt to the lack of surplus of housing (Zhang & Zhao, 2018), or the esthetical appreciation of cities (Cozzolino, 2021), or even the emergence of land-use patterns (Conzen, 1960) among many more.

We ultimately regard the ontology of the IoP to be an adequate conceptual frame to compare different urban planning policies and their subsequent impacts. We consider that such investigations would nourish the discourses and theories on urban planning policies and therefore, we deeply encourage these fields of study to seize for themselves and utilize the concept of the IoP.

6. Conclusion

For a long time, urban geography was rich in theories and poor in data, at least at a global context. Developments in remote sensing and crowd-sourcing approaches, however, now make it increasingly possible to support or reject theories with empirical findings – at least for the built urban landscapes.

Physicalist approaches of town and city planning theory outline the deep impact that the configuration of built environment has on the communities it hosts (Batty & Marshall, 2009). The built-up density, the height of buildings, the amount of green space per capita, the distance to amenities, among others are applied for the assessment of the quality of the built-up environment through the perspective of a physicalism (Batty & Marshall, 2009). And still, statistical approaches solely focusing on the material part of urban morphology display conflicting evidences of its impact in a comprehensive manner (Hajrasoulih et al., 2018; Krefis et al., 2018).

We argue that a new approach can focus on taking a look at the ef fects of the physical environment on communities through the prism of the IoP. This is motivated by the fact that the IoP is a synthetic infor mation of the brawls and struggles the society went and is going through. We have the intuition that our concept could reveal itself particularly fitting to the study of individual and collective urban qualities that have yet been hard to grasp through shallow-statistic ap proaches, such as social well-being and individual quality of life, economical and communities thrive and resilience.

Based on these considerations, we suggest empirical work analyzing in which way specific types of the IoP are statistically related to these urban qualities. From such work, we expect new insights and new tools for the practice of the city-making to allow for the desired balance between bottom-up democratic expression of the inhabitants and users' needs, and top-down control of a protection of commons (Cozzolino, 2020; Ostrom, 1990; Shipley & Utz, 2012).

With this research we attempt to lay a foundation – conceptually and empirically – to measure the built morphology and to relate to what can be read in it with respect to the IoP. As data on built morphology are becoming more and more available, this is intended to be a conceptual as well as empirical basis for large-scale or global studies of cities, their structure, their planning processes and their comparisons to understand the regularities, similarities and specifics in more detail. With these considerations, we encourage further studies willing to investigate the *Intensity of Plannedness* to enrich the ontology we propose in Table 1.

Funding

This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No [714087]-So2Sat).

Appendix

Supplementary material to this article can be found online at

https://doi.org/10.1016/j.landurbplan.2023.104711

Chapter XVII

Patterns of Eastern European urbanization in the mirror of Western trends

Convergent, unique or hybrid?

Environment and Planning B: Urban Analytics and City Science, 2019, vol. 46(7)

Hannes Taubenböck, Christian Gerten, Karsten Rusche, Stefan Siedentop & Michael Wurm

© The Author(s) 2019. Published by Sage. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

https://doi.org/10.1177/2399808319846902

Received 28 Aug 2018; Accepted 05 Apr 2019; Published 19 Aug 2019

Abstract

This paper presents a cross-sectional analysis of the urban land use patterns of 230 city regions in 34 European countries and an in-depth longitudinal analysis of 10 selected regions. The guiding question is whether the post-socialist transformation of urban spatial structure in Eastern European regions can be interpreted as an adaption process to Western-style urbanization and how far a process of 'mimicry' has reached. Our empirical approach is based on a model designed to measure binary urban land use patterns with respect to spatial dispersion. As cities and city regions vary in spatial pattern and size, we calculate the dispersion index for three different standardized extents: squares of 25 and 50 km around the defined urban centers as well as city adjusted accessibility isochrones. Our input layers are binary settlement classifications derived from multi-temporal Earth observation data. For the cross-sectional analysis, we cover entire Europe, and for the longitudinal analysis, we cover a sample of 10 cities for Western and Eastern Europe of predominantly capital cities of different sizes at four-time steps – 1975, 1990, 2000 and 2010. We found significant differences between Western and Eastern European city regions as they have entered different stages of urbanization. Eastern city regions are less populated, less urbanized, more dispersed and denser than regions in the West. Processes of post-socialist urban restructuring have definitely resulted in a change in land use patterns similar to that of Western Europe since the late 1950s. We nevertheless do not think that a 'catch-up' growth, leading to full convergence with Western-style urbanization, will be a realistic future scenario. Eastern European urbanization can best be characterized as hybrid: cities and city regions simultaneously manifest characteristics of convergent adaptation and path-dependency; they prove typical features of capitalist urbanization, but relics of the socialist past are still omnipresent.

Keywords

Post-socialist transformation; urban sprawl; compact city; remote sensing; comparative urban research

1. Introduction

Even 30 years after the collapse of socialism in Central and Eastern Europe (CEE), the region can still be seen as a laboratory allowing us to learn about the change of urban spatial structure in response to far-reaching shifts in political and social organization (Stanilov & Sykora, 2014). The transition from a socialist economy to free-market capitalism did not only result in a political, economic, social and cultural restructuring process but it also implied a large-scale transformation of urban structures.

Looking back at the past three decades, post-socialist transition is usually associated with the introduction of a democratic political system, with economic liberalization, the privatization of state assets and opening of local markets to international economic forces, the decentralization of political power to local governments and the predominance of liberal, rather permissive regulatory planning frameworks (Sykora & Bouzarovski, 2012). Alongside cultural and social changes, these large-scale restructuring processes have reshaped the institutional framework of urban and regional development as well as the built environment of cities and city regions themselves. The centralized, hierarchically organized planning system of the socialist era was abruptly replaced by a fragmented, local self-government structure where individual communities competed for tax income and economic stakeholders gained influence over land use decisions (Sailer-Fliege, 1999; Smigiel & Brade, 2011; Sykora & Stanilov, 2014). Whereas socialist urban policies aimed at 'balancing' development and reducing spatial disparities (mainly by distributing development over a larger number of industrial hubs and central places), post-socialist development agendas were instead competitive and growth oriented (Hamilton, 2005; Sailer-Fliege, 1999; Smigiel & Brade, 2011). Key features of urban transformation in the 1990s and 2000s include, but are not limited to, the commercialization, regeneration and densification of inner-city areas and a dynamic expansion of built-up space, largely driven by private actors on the re-established land markets (Sykora & Bouzarovski, 2012).

A predominant interpretation is that the adoption of Western institutions, cultures, policies and regulative mechanisms has led to a convergent adaptation of spatial structures, implying a profound change in functional, morphological and socio-spatial patterns (Haußermann, 1996). One essential spatial outcome of this '*catch-up*' process (Ott, 2001) is suburbanization, a phenomenon that was largely suppressed in CEE countries before 1989 (Smigiel & Brade, 2011; Sykora & Stanilov, 2014). The relatively compact urban form of the socialist city was, thus, gradually replaced by a more decentralized and dispersed urban structure (Ianos et al., 2016; Sykora & Bouzarovski, 2012). From this perspective, urban restructuring in post-socialist countries follows a '*common logic*' that is deeply rooted in institutional, political and social change (Sykora & Bouzarovski, 2012). Proponents support the '*convergence*' hypothesis with the observation of similar trends in urban land use change, motorization and mobility, housing preferences, spatial patterns of retail and office decentralization, or social segregation.

In contrast, other voices emphasize the path-dependent and unique nature of urban change in CEE countries after the fall of the Iron Curtain (Großmann et al., 2015; Schmidt et al., 2015; Sheppard, 2000; Tosics, 2005). First, scholars claim that post-socialist societies do not follow exactly the same transition track. Differences in privatization (or more specifically:

restitution) policies (Sykora & Bouzarovski, 2012) and housing policies (Sailer-Fliege, 1999) or in the pace of decentralizing the control over land use (Tosics, 2005), as well as the abolishment of rent controls, can be taken as examples. Although all post-socialist nations are moving away from their socialist past, a certain variation in the individual pathways of managing the transitional process might explain different spatial outcomes in terms of land use and urban form. Second, it is self-evident that Eastern European cities still exhibit physical, social and cultural remnants from the pre-capitalist era or even pre-socialist past of urban development. Sykora & Bouzarovski (2012, see also Großmann et al., 2015), therefore, took a middle position by stating that Eastern European cities are not yet fully developed capitalist cities: '*Looking at their morphology, land use and social segregation, we can document typically capitalist city areas and districts, while sections of urban landscapes resemble frozen mirrors of socialism*'. The reconfiguration of the built environment is – following Sykora & Bouzarovski – far from being completed. Third, commentators have also considered the specific demographic processes in many CEE countries as an exceptional circumstance for urban development. The massive outmigration from CEE countries, especially in the economically troublesome 1990s, and negative natural population balances have resulted in a strong population decline, affecting in particular economically lagging secondary and small cities. Schmidt (2011) found the so-called sprawl without growth pattern characteristic of East Germany as well as other post-socialist states.

In this paper, we ask whether the transformation of urban spatial structure can be interpreted as an adaption process to Western-style urbanization and sprawl and – if the answer is 'yes' – how far this process of 'mimicry' has reached. If the answer is 'no', we ask for the kind and intensity of demonstrable differences. Thus, we aim to locate the urban development trajectories of Eastern European countries in a broader context of European and global urbanization trends by analyzing urban land use patterns from both cross-sectional and longitudinal perspectives. Our focus here is on the intensity and form of urban sprawl and dispersion.

For this purpose, a new metric – the '*dispersion index*' (DI) – was utilized (see Taubenböck et al., 2019 and section 4.2. for details). Within its conceptual framework, this metric allows for an unambiguous ranking of any land use pattern with respect to the dispersion (or compactness) of built-up areas. In accordance with the literature, we understand dispersion as a specific spatial pattern of urban land use. Disperse land use patterns are typically characterized by a less concentrated and less spatially clustered or clumped configuration of urban functions (Schneider & Woodcock, 2008). Against this background, we consider a given land use pattern as being more or less dispersed if the number of contiguous urban patches is higher or lower and the size of the largest (central) patch is lower or higher, respectively. In other words, more compact urban forms are characterized by a larger continuously urbanized core area and a lower number of urban patches. In this paper, we approach the physical configuration of cities using the spatial patterns of urbanized land and assume that they reflect different historical trajectories influenced by political, economic, social and demographic factors as well as topographical factors.

As research on urban restructuring in CEE countries has so far largely focused on a few major cities such as Budapest or Prague (Sykora & Bouzarovski, 2012; Tosics, 2005), it is

our intention to take all cities above a certain population threshold into account. By doing so, we present an analysis of urban land use patterns in 230 European cities and city regions (among them 119 East European cases) and changing land use patterns over time in 10 selected regions. Since conceptually equivalent delineations of functional urban areas are not available for a comparative European study, we develop a spatial concept based on static as well as city-adapted spatial units. First, we employ static 25 and 50 km squares spanned around the center of the dominant city core as a consistent spatial analysis frame (see section 4.1.). We are fully aware that the size of these areas (625 and 2500 km^2, respectively) is in many cases larger or smaller than the 'real' functional urban area of a particular city (indicating commuter belts or housing market regions); however, these consistent spatial units provide permissible comparative entities. Second, we employ a city-adapted regionalization in dependence on accessibility by car.

The comparative perspective – taken in this study – demands consistent data sets. Using a spatial approach, remote sensing is a valuable source for providing consistent mapping products across space and time. Recent advancements in data and processing power have resulted in large areas or even global settlement mapping products at resolutions between 30 and 12 meters (Esch et al., 2012).

The remainder is structured as follows. In the next section, we survey the literature on urban restructuring in CEE countries. In the section 3, we introduce data and materials used. In section 4, we present the spatial concept, and we review and introduce spatial metrics applied for measuring urban form, and in section 5, the empirical results are presented. Section 6 offers a critical discussion on the results and the capabilities and limitations of these data and methods, and the final section 7 concludes this paper with a perspective.

2. Urban transformation in CEE countries: Convergent or path-dependent?

Suburbanization and urban sprawl are often considered global phenomena, expressing a universal pattern of land use change under liberal market regimes (Dielemann & Wegener, 2004; Gordon & Cox, 2012). Key features in this sense are growing physical footprints of cities, the deconcentration and spatial separation of urban functions, and a declining urban density as well as flattening density gradients. Although the 'European city' has often been portrayed as a distinct morphology, rooted in a long history and a rich culture (Dielemann & Wegener, 2004 or McNeill, 1999 for a critical debate), European cities and city regions have also witnessed a deep transformation of spatial structures, resulting in a less dense and less concentrated urban form. Salvati & Carlucci (2015) view suburbanization as '*the common trait*' of spatial patterns in European cities, '*regardless of their geographical, economic, or administrative characteristics*'. They argue that common drivers, such as rising living standards, growth-oriented development policies, housing preferences in favor of less dense urban environments or changes in mobility behavior, manifest themselves in convergent land use trends towards less dense and dispersed urban structures.

Against this popular narrative of '*homogenization'* or *'standardization*', researchers have found a tremendous rate of variation in both structures and structural change across Europe (Kasanko et al., 2006; Schwarz, 2010; Siedentop & Fina, 2012). Based on an analysis of

more than 200 European cities, Schwarz (2010) concluded that neither the European compact city nor a clear pattern of urban form in different regions of Europe can be found. Siedentop & Fina (2012) found '*substantial differences*' in terms of the scope of urbanization, the dynamics of urban growth or spatial patterns of land take in their comparison of 26 European countries. Given this diagnosis of variation, regional '*vernaculars*' of urbanization processes can be assumed, materializing themselves in highly place-specific urban outcomes.

With a view to the CEE region, there is significant evidence that CEE city regions are still less urbanized and decentralized than their Western European counterparts (EEA, 2016; Sailer-Fliege, 1999; Salvati & Carlucci, 2015; Siedentop & Fina, 2012). Especially for south-eastern countries (Bulgaria, Romania or Albania), low values of dispersion and land consumption were measured when taking Europe as a whole as a benchmark (EEA, 2016).

The diagnosis of a relatively higher concentration and level of compactness can be seen as a legacy of socialist state policies favoring large-scale housing projects at the urban fringe (Musil, 1993; Sailer-Fliege, 1999). Moreover, the concentration of economic and urban development in selected cities prevented the more land-consuming spread of urbanization across larger territories that is typical for many Western European countries. Finally, the absence of stronger suburbanization and the reliance on public transport in the socialist era explain the still higher compactness of CEE regions.

Nevertheless, the erosion of the once-compact urban form in CEE cities is an empirical fact, but this transition did not happen overnight. Rather, the first years of transition were marked by economic decline and deindustrialization, collapsing housing production, massive outmigration (Stanilov & Sykora, 2014) and, thus, a stagnation of urbanization. With the emerging economic recovery in the late 1990s, however, a rising demand for new housing and commercial space was obvious, and suburbanization became the 'predominant mode of urban growth' in CEE countries (Sykora & Stanilov, 2014). The relatively high rates of land consumption (Schmidt et al., 2015; Siedentop & Fina, 2012) implied a further decrease of urban densities (Schmidt et al., 2015). The data presented in Wolff et al. (2018) demonstrate a significant East–West divide in the emergence and dynamics of this phenomenon. Many CEE countries have experienced a much stronger reduction of urban densities than have Western European countries. The main reason for this difference lies in the massive population decline in the majority of cities in CEE countries, driven mainly by outmigration and a negative natural population balance (Haase et al., 2016; Mykhnenko & Turok, 2008). At the same time, scholars have found tremendous regional variation between both Western and Eastern countries and within the group of CEE countries (Siedentop & Fina, 2012; Wolff et al., 2018).

Another major characteristic of urban change in CEE countries is the tremendous contrast in urban size and growth rates between the capital and major industrial cities, on the one hand, and smaller secondary cities on the other hand (Hirt, 2007; Musil, 1993). This development '*asymmetry*' within the urban system (Sailer-Fliege, 1999) can also be seen as a legacy from the socialist era: based on the idea of economies of scale, the concentration of industrial production in growth poles should gain higher levels of efficiency. The prioritization of state-led investments in favor of the capital cities and some secondary cities resulted in

uneven economic and demographic development. The gaining cities had much stronger employment growth than other cities, and they also became hubs of in-commuting from surrounding areas since housing lagged behind demand (see also the debate about the '*under-urbanization*' phenomenon in Tosics, 2005). In accordance, Siedentop & Fina (2012) showed that urban sprawl in CEE countries during the 1990s and early 2000s occurred mainly around a few major cities and growth corridors, whereas a significantly divergent pattern was visible in many Western European countries, where spatially extended processes of suburbanization resulted in a much more decentralized and land-consuming pattern of urban growth. We, therefore, expect structural differences of compactness according to the size of cities and city regions.

In sum, we assume that the transformation of urban spatial structure in CEE countries is similar, but not necessarily fully convergent, with Western European processes. In other words, urban restructuring in Eastern Europe is expected to demonstrate characteristics of convergence and path-dependency. Suburbanization and urban dispersion might resemble Western-style features but, at the same time, significant divergence can be hypothesized. For example, there are no signs of a '*post-suburban*' development in CEE (Sykora & Stanilov, 2014). Scholars still characterize suburbanization in the CEE context as a rather unplanned, piecemeal decentralization of housing, commercial and retail developments in a dispersed spatial arrangement. Moreover, in contrast to many Western European cities, urban areas in CEE countries have not yet experienced a significant regrowth in their city centers (Haase et al., 2017) and sub-center formation in suburban areas. Eastern European urban spatial structure is, therefore, presumably less polycentric than that of Western regions.

3. Data sources

Our main data source for analyses is multi-source satellite data. All mapping products understand the term '*urban*' as the land directly occupied by a particular built physical structure. On the one hand, we capture binary patterns of urban versus non-urban land use on a continental scale for the year 2012. We use the Global Urban Footprint mapping product derived from TanDEM-X data (for methodological details, see Esch et al., 2012; for data access cf. supplemental material-1). On the other hand, we refer to multi-temporal mapping products based on Landsat sensors available since the 1970s. We monitor the evolution of settlement patterns using classifications at the time steps 1975, 1990, 2000 and 2010 for specific areas of interest (for methodological details, see Taubenböck et al., 2012).

We use residential population estimates for the target years 1975, 1990, 2000 and 2015 provided by the Gridded Population of the World (GPWv4data) database (SEDAC, 2018). The input data provide the distribution of human population (counts and densities) on a continuous global raster surface. The data rely on census information, which is disaggregated to grid cells, informed by the distribution and density of built-up areas as mapped in the Global Human Settlement Layer per corresponding epoch (https://ghsl.jrc. ec.europa.eu/).

4. Methods

4.1. Selection of study sites and spatial entities of measurement

Our study site is the European continent. We select all cities that feature a population larger than 250,000 inhabitants at administrative level for the year 2015 (based on an EU Urban Audit for European countries and the World Population Review for Russia and Ukraine). Overall, 230 cities fulfil this criterion, with 111 cities in the West and 119 cities in the East (geographically we divide Europe in West and East along the political system boundaries before 1990). For the multi-temporal analysis, however, we reduce our sample to 10 cities because settlement classifications with documented high accuracies are not available for the entire continent. We chose an even distribution of five cities each for the Western and Eastern regions.

For a consistent and valid comparison of the settlement patterns across time and space, the spatial units of measurement need to be defined. Administrative boundaries are artificial, non-uniform spatial units that are not appropriate for meaningful comparative geographic analysis. So, for reasons of comparability, we develop a two-fold spatial concept based on (1) static and (2) city-adapted spatial units: (1) In the static approach, we use fixed spatial entities: (a) We apply a 25 x 25 km square around the respective city center, with which we aim to capture the core city pattern; (b) we utilize a 50 x 50 km square around the respective city center, with which we aim to capture the core city pattern as well as the functional urban area. The city centers are defined as centroids based on the ESRI World boundaries and places. We argue that although these two spatial entities are artificial and are unlikely to fit the specific context of the local urban structure, they are – compared to artificial administrative entities – consistent and thus a permissible comparative entity. (2) In the second analysis, we employ city-adapted extents based on accessibility isochrones to also account for specific regional conditions. For each city, a size-dependent travel time threshold is used. With regard to gravitation theory, we assume that larger cities have larger catchment areas. The resulting isochrone geometry represents a flexible spatial geometry, taking functional interrelations between core cities and their catchments into account. Travel times were derived from network analysis based on the ESRI road network at free flow conditions. We set 20 minutes as the minimum travel time for the city with the smallest population in our sample, and 60 minutes as the maximum for the city with the largest population. The values in between have been transformed proportionally to this range and have subsequently been standardized to values from 0 to 1. In order to reduce impacts of outliers (Moscow, London), we modified the resulting attraction curve with a simple square root function, so that maximum travel times for city catchments were spread more evenly around the value range (see supplemental material-4).

4.2. Measurement of settlement dispersion

The main idea is to evaluate the urban land use patterns of cities and city regions. For measurement of settlement dispersion, we rely on a model-based conceptualization of spatial pattern evaluation suggested by Taubenböck et al. (2019). This model has been developed € for a simplistic case of spatial patterns – two-dimensional patterns constituted by two

thematic classes, 'settlement' and 'non-settlement' – and is based on two spatial metrics: (1) the largest patch (LP) of a landscape and (2) the number of patches (NP) of a landscape (see details in supplemental material-2).

Argued from a geographical perspective, the LP is assumed to be a proxy to evaluate whether or not it is – in relation to all other urban patches – dominating an urban landscape. This proxy is inspired by monocentric city models where the LP relates to a dense core city surrounded by a less dense suburban area (Anas & Kim, 1996). Regarding the NP of a landscape, larger numbers around the dominating LP indicate that the landscape is less compact and more dispersed.

These two spatial metrics span a two-dimensional feature space that is defined as the DI. Both parameters are normalized (LPn, NPn) to equal ranges (0 to 100) and weighted equally. Every possible two-dimensional binary (settlement) pattern can be projected in this feature space. One end in this model is marked, if the landscape consists of only one patch (which is then, naturally, the largest one), and the DI is 1 indicating a perfectly compact pattern. Any additional urban patch (increase in NP) around this LP transforms the pattern into a more fragmented, less compact and thus more dispersed one. On the other end, if the LP is minimal and the complete class area is represented by the maximum possible number of non-coalescent individual patches, the pattern is ranked with a DI of 100 (perfectly dispersed). It has been shown that this conceptualization allows for an unambiguous ranking of patterns in relative and absolute terms between compact and dispersed layouts (for more details, see Taubenböck et al., 2019).

4.3. Comparing trajectories of urban land use patterns

In general, we approach the evaluation of urban land use patterns in cities across Europe spatially (cross-sectional) for the year 2012 and temporally (longitudinal) for decadal points in time since the 1970s:

• Cross-sectional: We evaluate trends using the DI. We calculate the DI for the static two units of measurement (25- and 50-km rectangles) as well as for the city-adjusted accessibility isochrones. For evaluation, we use boxplots as well as their distributions in the feature space of our model. Furthermore, we test some hypotheses in terms of explaining variation in DI values using regression modelling. We conduct multivariate regressions to allow a deeper understanding of relationships between the DI and relevant covariates. For all spatial units, we employ separate ordinary least squares (OLS) regressions that use a small set of explanatory variables. In both cases, this involves the settlement density as the share of urban land uses within the specific analytical unit. The reference area of the spatial unit is reduced by large water areas such as oceans.

• Longitudinal: We first evaluate trends across Eastern and Western cities by relative spatial growth rates. Using the spatial extent of settlements per city for the year 1975 as 100%, we calculate relative growth rates until 1990, 2000 and 2010. For the analysis of spatial trajectories of settlement patterns, we use the DI and its evolution over time. Longer paths within the model reflect higher dynamics in settlement patterns. As the multitemporal

analysis street network data for past time steps are not available, we disregard the city adjusted spatial units by travel times.

5. Empirical results

5.1. Cross-sectional analysis of urban land use patterns

Figure 1 illustrates the land use patterns found for three groups of least, most and moderately dispersed city regions according to the conceptual logic of the DI. The LP is displayed in red in all cases. DI values range from 0.7/9.3 (London) to 48.5/55.3 (Kryvyi Rih, Ukraine) for the 25/50 km square. The fascinating diversity of European urban landscapes becomes intuitively apparent. We consider topographical conditions (coastal location, mountains, major rivers), historical settlement structures, economic and cultural factors as well as regulatory frameworks for land use planning as the main factors accounting for this tremendous variation. Supplemental material-3 gives an overview of all 230 city regions and their respective DI values for the three spatial units.

Beyond this impression of uniqueness of each individual pattern, our analysis reveals significant differences in land use patterns between the East and West. When considering the individual DI values, we see that among the 20 European regions with the lowest DI, only two Eastern European cases are found (Moscow and Krasnodar for the 25 km square); for the 50 km square, there are five regions (Moscow, Krasnodar, Makhachkala, Bucharest and Cherepovets). By using city-adjusted accessibility isochrones, six Eastern European regions belong to the 20 most compact ones. At all three spatial levels, higher mean values of the DI for Eastern European city regions were measured. This holds true for all city size classes (0.1–0.5, 0.5–1.0, >1.0 million). The Eastern regions are thus considered less compact than their Western counterparts.

This result, unexpected at first glance, can be explained by a lower degree of urbanization in Eastern European cities or city regions (Tosics, 2005). In general, low DI values are found in larger, densely populated urban regions with an extended core area and a lower NP due to the merging of formerly isolated patches. In contrast, high DI values are typically found in regions with a smaller core (LP) and a larger number of small patches. The latter is characteristic of many Eastern European regions that are still in a relatively early phase of suburbanization. Such regions are often shaped by a '*leapfrog*' mode of urban growth and/or the remnants of rural landscapes characterized by smaller cities and villages. This characteristic especially applies to more remote parts of the region – the peripheries of commuter belts – that remain less affected by urban growth and sprawl. The emergence of large conurbations due to the physical merging of settlements is a characteristic feature of Western urbanization, leading to compactness at higher spatial scales. This trend is confirmed by the fact that the mean LP in Western European regions is more than twice the size of the Eastern regions (114 km^2 versus 49 km^2 for the 25 km square).

We therefore interpret our results as an indication of a considerable West–East gradient of urbanization. This appraisal is further confirmed by the widely varying settlement density in Eastern and Western European regions. Whereas the proportion of urban land use to the total

land area is only 17.3% (25 km square), 7.2% (50 km square) and 18.2% (city-adjusted isochrones) in Eastern regions, the respective values of Western regions are 32.8, 16.7 and 26.8%.

Another interesting fact is that Western European regions show a much larger range of DI values than Eastern regions (Fig. 2). Extremely compact regions with DI values lower than 10 (25-km rectangle) are mainly located in Western Europe, with the notable exception of Moscow (and some smaller regions). At the same time, we identified some 'high DI' regions in Western Europe (values > 50). As these regions are located in quite different countries (Spain, Germany, Belgium, the Netherlands), we cannot assume a '*systematic*' effect of nationally specific regulatory regimes or housing preferences. We rather believe that topographical features as well as historical settlement patterns account for these outliers.

A further result of our analysis is a correlation of the DI and urban density, here measured as the number of inhabitants per km^2 of urbanized land. In general, the more dense regions are also the more compact ones. However, a closer view reveals that Eastern European regions are denser by population while showing a lower compactness. We measured a mean urban density of 6780 inh./km^2 (25 km), 5640 inh./km^2 (50 km) and 6422 inh./km^2 (city-adjusted isochrones) for the East and 5880, 4680 and 5185 inh./km2 for the West. This result clearly demonstrates the legacy of socialist development policies favoring compact urban extensions and housing at higher densities (see section 2). In contrast, long-lasting suburbanization has led to a strong dedensification of urban areas in most Western European regions (see also Wolff et al., 2018). Thus far, high urban density and a low degree of compactness in a single region do not pose a contradiction in these terms. Both indicators measure different urban attributes and should, therefore, be used in combination.

When differentiating the DI statistics according to nation states, some interesting observations can be made. The lowest DI values were measured in countries with a less decentralized planning system (as in Denmark, France, Greece and the UK; the latter one exemplified in Fig. 2). In contrast, considerably higher values can be found in countries like Germany (exemplified in Fig. 2), the Netherlands or Switzerland, where local governments have a stronger competence for urban development and communities have a fiscally motivated interest in urban growth. In Eastern Europe, Poland is an example of a country with a rather permissive land use policy. This has led to sprawling suburban landscapes indicated by significantly higher DI values than, for example, their eastern neighbor Ukraine. This variation might be an indication of the general relevance of land use governance and planning systems for the dynamics of urban dispersion (see also OECD, 2017a, 2017b).

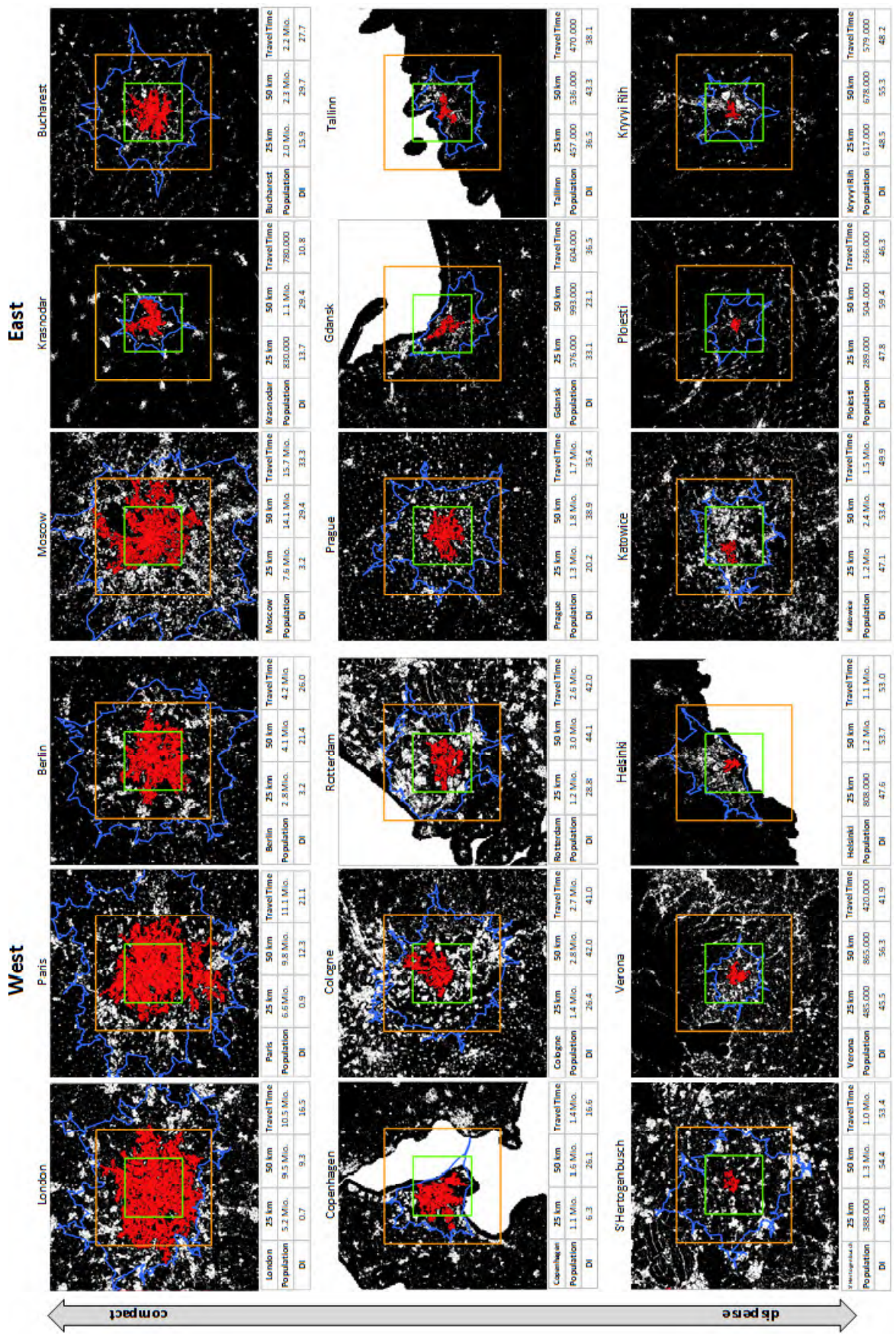


Fig. 1. Urban land use patterns in Western and Eastern European city regions (25-km square: green; 50-km square: red; city-adjusted isochrones: blue), grouped by compact to disperse patterns (LP in red).

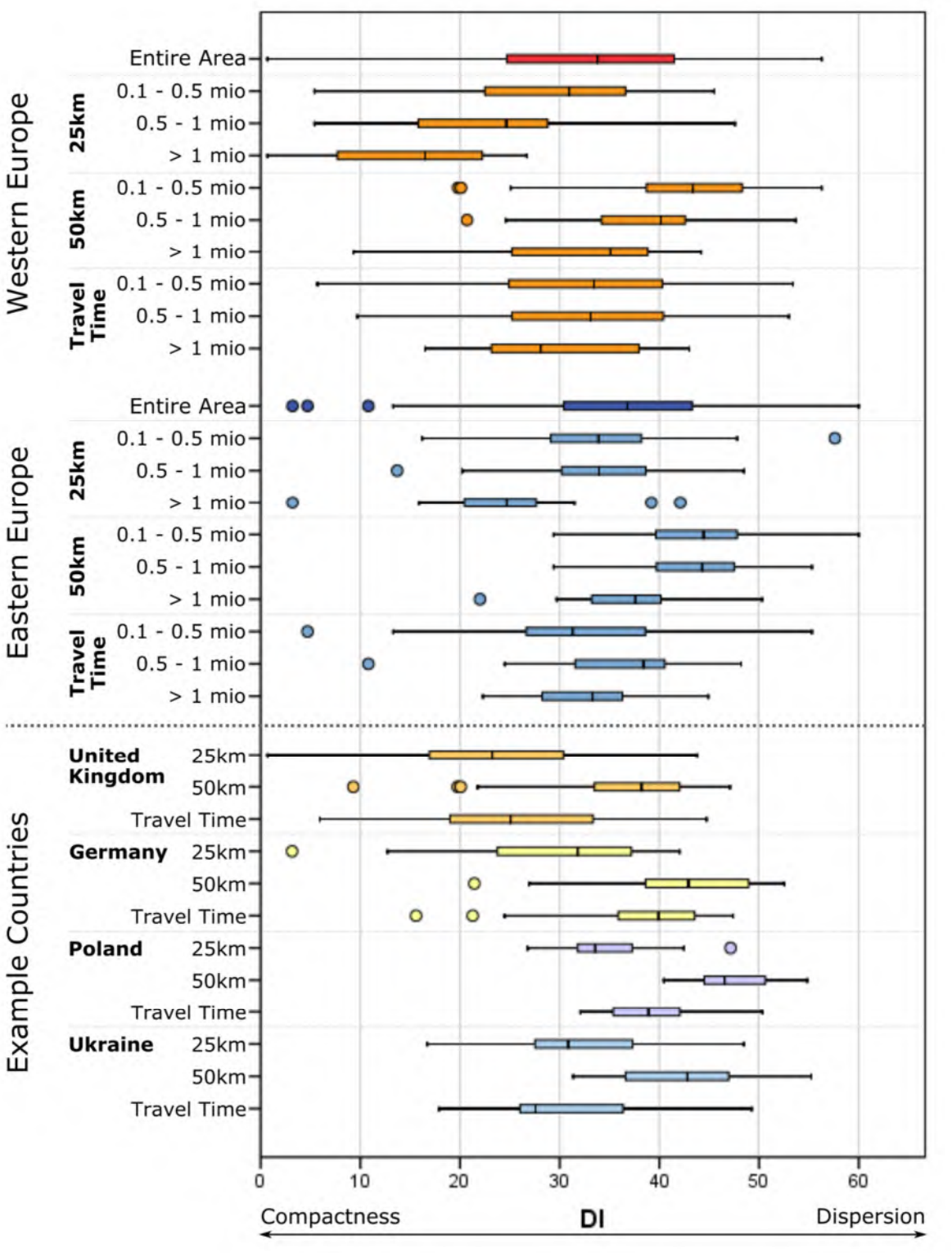


Fig. 2. Boxplots illustrating the medians and distribution of the DI across Eastern and Western European cities as well as for sample countries at three spatial units and according to population size.

For a deeper understanding of relationships between the DI and relevant covariates, we conduct multivariate regressions. We expect that regions with a higher share of built space will showcase lower index values. Furthermore, regressions include urban density (as the number of people per km^2 of urbanized land). Density can be seen as a measure of the degree of urbanization of a region. We expect that city regions with a higher urban density will

show a smaller DI value. In addition to these two main covariates, we added a dummy variable that indicates which cities are located in CEE countries. If there were a significantly different pattern in the form of a clear-cut West–East divide, this dummy would capture this effect.

In the course of the specification setup, other relevant variables were tested (e.g. population size, geographical coordinates) but they did not add relevant information to the regression results, i.e. the coefficients of the other variables remained significant, and newly included variables did not raise the explanatory power of the regressions. We tested level-level, log-level and log-log specifications, among which the log-level setup delivered the most robust results.

Table 1 lists the results of the OLS calculations. For all models, the regression diagnostics are similar. The degree of multicollinearity in both cases is low (condition number under 10, variance inflation factors all under 2), the variable setup is significant as a group according to the F tests and both regressions have an average level of explanatory power. Nevertheless, the adjusted R^2 for the 50-km unit and the city-adjusted accessibility isochrones is considerably lower than for the 25-km unit. Standard errors and significance levels are based on white standard errors because the basic OLS setup in both cases had problems of remaining heteroscedasticity in the regression residuals. Altering the estimation type for the variance– covariance matrix did not considerably change the values for the heteroscedasticity-consistent results. In addition, residuals are non-normal due to a skewness caused by a few extreme values of some city regions. Based on central limit theorems, asymptotic normality can still be assumed.

Table 1. Results of the OLS calculations.

	Dependent variable: log (dispersion index)		
	OLS model 1 (25 km)	OLS model 2 (50 km)	OLS model 3 (city adjusted isochrones)
Settlement density	−0.027*** (0.004)	−0.015*** (0.003)	−0.018*** (0.004)
Urban density	−0.00005*** (0.00002)	−0.00003*** (0.00002)	−0.0001 (0.00005)
Dummy East	−0.027 (0.0577)	−0.053 (0.031)*	−0.121* (0.049)
Constant	4.231*** (0.1435)	4.083*** (0.066)	3.985*** (0.078)
Observations	230	230	230
R^2	0.556	0.370	0.372
Adjusted R^2	0.550	0.362	0.363
Residual Std. Error (df = 226)	0.380	0.188	0.298
F Statistic (df = 3; 226)	94.262***	44.299***	44.587***

(White Standard Errors)
OLS: ordinary least squares.
Note: *p<0.1; ***p<0.01.

Looking at the results of model 1 (25 km), we see that all main covariates are significant and show the signs we expected. Regions with a higher settlement density have lower DI values. Furthermore, regions that have a denser urban form are ceteris paribus less dispersed. Importantly, the dummy variable for CEE regions is not significant, which leads to the conclusion that, based on this spatial lens, it is not a specific political or historical difference that explains the difference between the dispersion values but rather the different stages of urbanization.

For the case of the 50-km model (model 2), the results are very similar for settlement density and urban density. Strikingly, in this wider scale, the dummy variable becomes weakly significant (p value of 0.08). Therefore, on this level, there is a weak structural difference (of around 5%) between DI values that cannot be explained only by structural indicators. The same applies to the model based on an accessibility delineated regionalization (model 3). In this case, however, the settlement density becomes non-significant, while the dummy variable is again gaining significance. To further test the validity of our results, we ran regressions that excluded the big city regions from the sample, i.e. London, Paris and Moscow. Importantly, this did not alter the results significantly.

Taken together, our cross-sectional analysis indicates that patterns of urbanization still exhibit a significant East–West divide. East European city regions are less urbanized, less populated and more dispersed than their Western counterparts. We assume that the suburbanization process of the last 25 years has considerably narrowed the gap between the East and the West.

5.2. Longitudinal analysis of urbanization patterns

The comparison of temporal trajectories of urbanization patterns also reveals differences between the categories of East and West. Figure 3 illustrates growth patterns since 1975 for five selected cities per geographical category. Here, we disregard the spatial entities generated via accessibility isochrones, as road networks are not available for past time steps. Two differences appear to be particularly relevant: (1) The urban growth dynamics, and the expansive growth in areas that were characterized as rural or suburban in the 1970s in particularly, are very pronounced in Eastern European city regions, and (2) the spatial change of the LP, i.e. the extensive expansion of the dominating urban patch, is much more pronounced in Eastern European regions.

The calculation of two-dimensional growth rates of urbanized areas confirms these differences (cf. Fig. 4). Eastern regions show a substantially higher rate of urban expansion. In general, we measured, for the core cities (25 km square) in the East, dynamics from 1.6 times as large as in 1975, in Kiev, to 3.1 times as large, in Minsk. These growth rates are considerably higher than in the West: they range from 1.1 times as large as in 1975, in Frankfurt, to 1.9 times as large, in Vienna. For the larger city regions (50 km square), similar differences in dynamics were measured: in Eastern regions, the range is from 2.4 times as large as in 1975, in Kiev, to 5.1 times as large, in Minsk. In comparison, Western regions range from 1.1 times as large as in 1975, in Frankfurt, to 2.7 times as large, in Vienna.

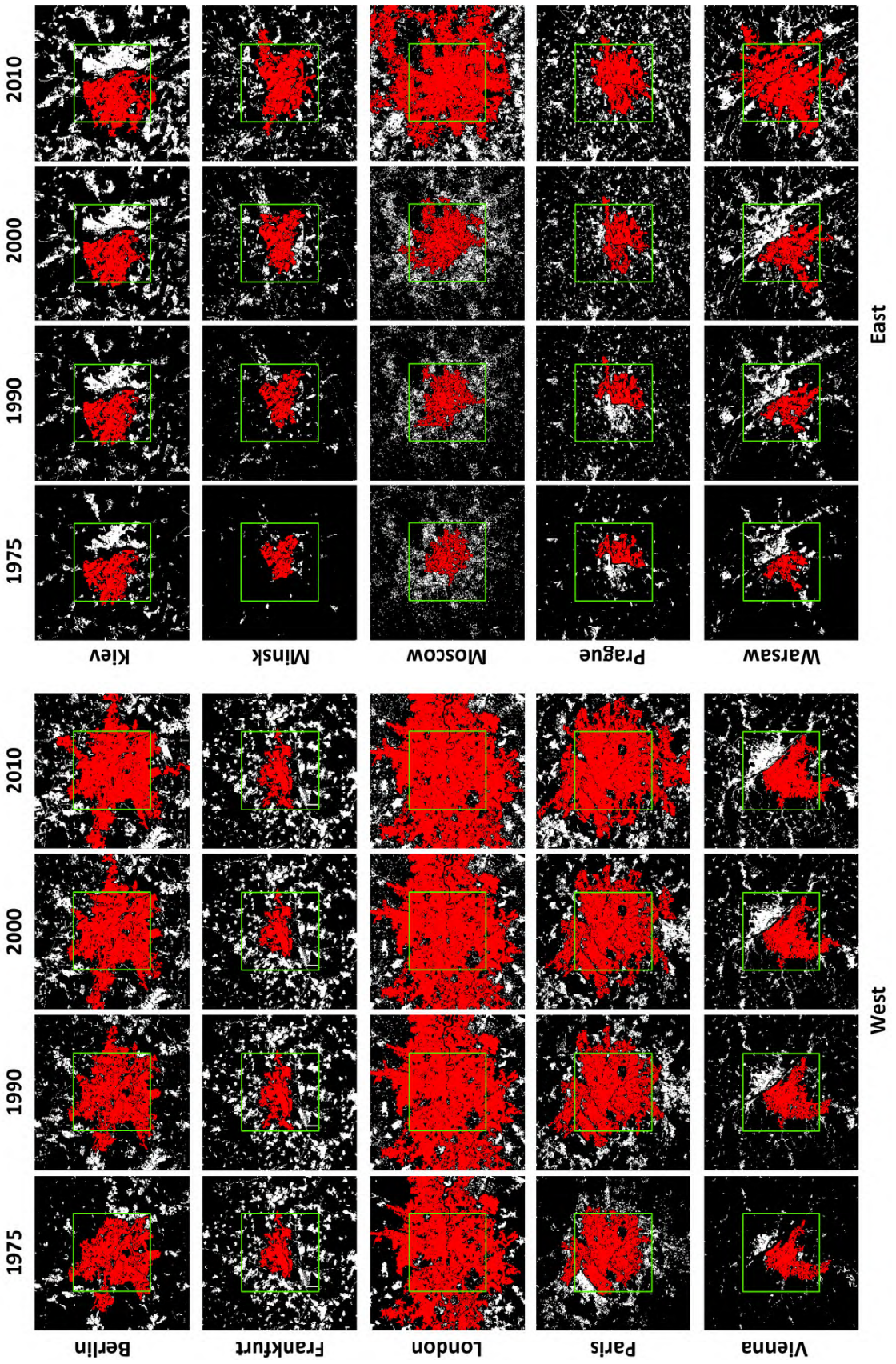


Fig. 3. Settlement pattern evolution from 1975 to 2010 in Western and Eastern European cities. Settlements are displayed in white and red (LP). The spatial extent of each frame is 50 x 50 km; the green square visualises 25 x 25 km; the geometric resolution is 200 metres.

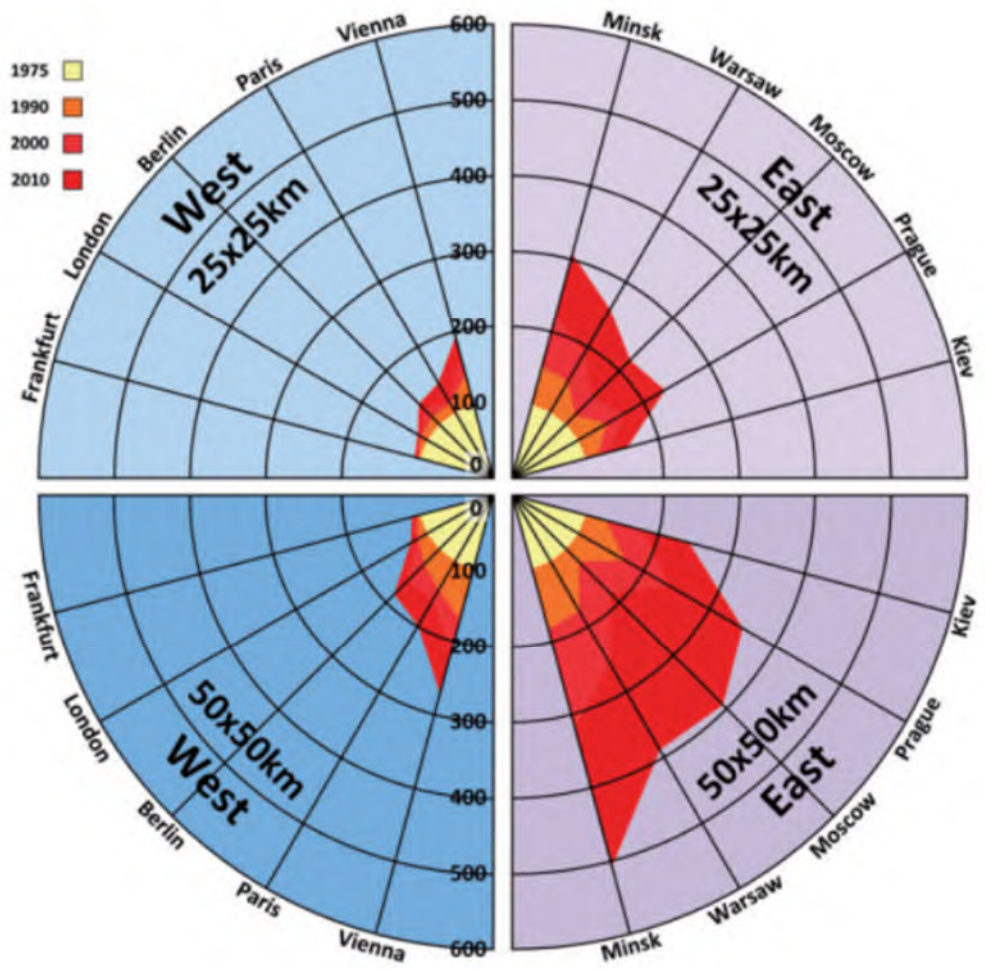


Fig. 4. Relative spatial settlement expansion for the selected cities in Eastern and Western Europe for the 25 km and 50 km units.

Beyond the spatial growth dynamics, we measured the development of the DI across space and time and compared the pattern evolution between East and West (Fig. 5). In general, we found larger changes in urban land use patterns, for the smaller as well as for the larger city regions (25, 50 km square), in Eastern Europe. The trend is towards more compact cityscapes, and the larger changes in Eastern regions can be related to a much more dispersed starting position in 1975. For the core city areas, we measured 7 out of 10 cities, with a trend towards more compact cityscapes (decreasing DI). The highest dynamics were measured for Moscow (DI 1975 of 34.4 to 1.7 in 2010) as a result of densification and merging of urban patches in central areas. Two cities can be considered stable over time, and only one shows a reverse trend towards dispersion (Vienna from DI 13.7 in 1975 to 19.4 in 2010). For the larger spatial entity (50 km), which includes peripheral areas, we measured relatively stable conditions in three Western regions (Berlin, Frankfurt and London). Vienna turns out to be an outlier here, too, as it marks the only Western region with increasing dispersion (from 21.1 to 33.5). In Eastern regions, the settlement patterns reveal high dynamics, with three city regions developing towards dispersion (Minsk from 16.0 to 30.5, Kiev from 31.2 to 36.2 and Prague from 34.0 to 37.0). The comparatively large and dynamic regions of Moscow, Warsaw and Paris are developing towards more compact urban configurations.

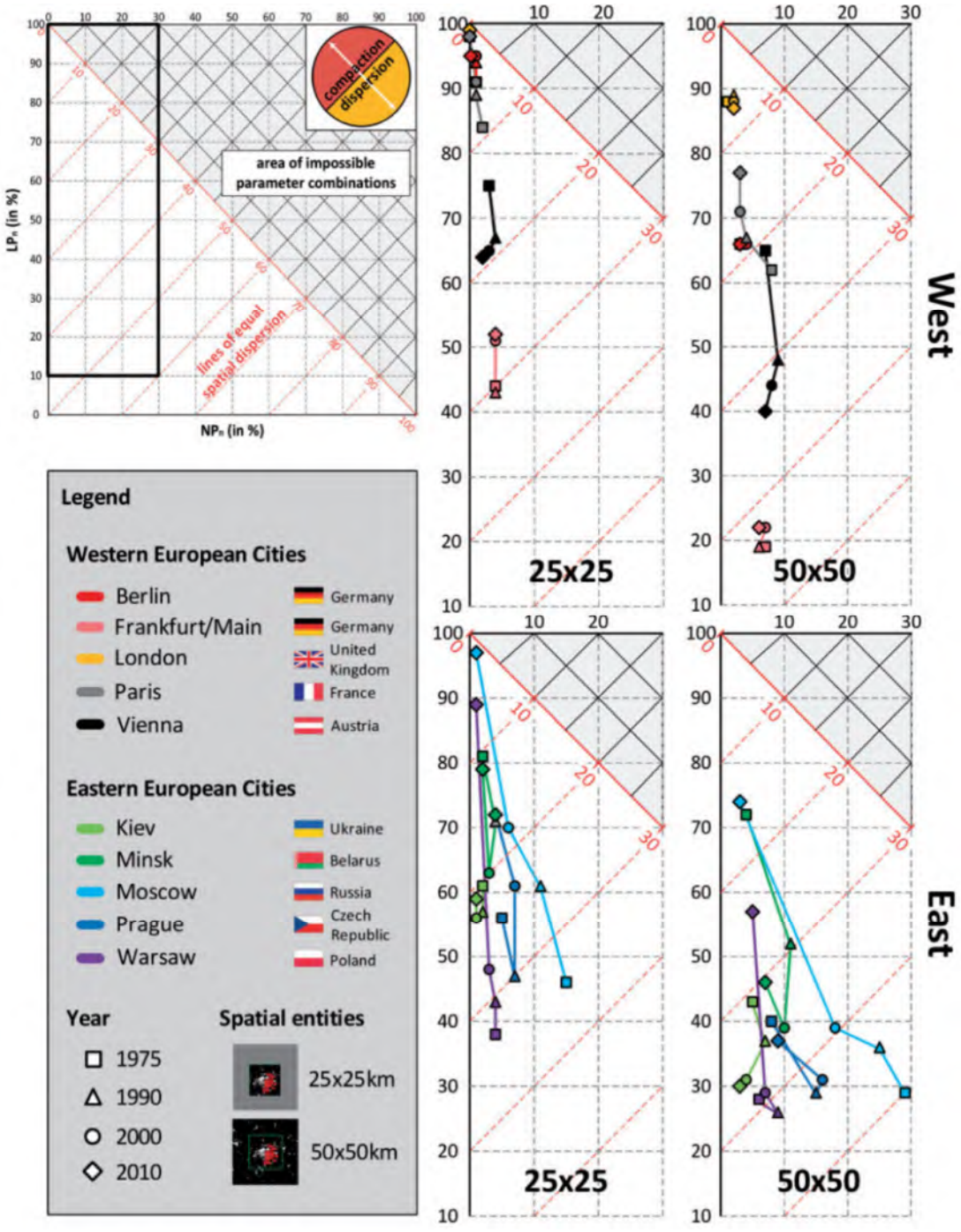


Fig. 5. The evolution of spatial settlement patterns measured by the DI for selected cities in Eastern and Western Europe at the two units of measurement (25 km; 50 km).

6. Discussion

Our investigation offers a wide-ranging overview of the state and process of spatial urbanization in European cities and city regions. The longitudinal analysis over 35 years suggests that a trend towards compaction at higher spatial scales might be a robust scenario of future urbanization in Europe. Following this thought, urban sprawl – often considered as the archetype of urban land use change in the developed world – is simply an intermediate stage of a long-term spatial transition process. Peiser (1989, 2001) has argued that undistorted land markets tend to fill the leapfrogged, scattered territories along the urban

periphery. Rising land prices and the availability of urban services and infrastructure in neighboring developments create incentives for infill development on formerly undeveloped land. What has been portrayed as a 'scattered' or 'discontinuous' form of development will be – according to Peiser – contiguously urbanized and thus 'compact' at a higher spatial scale. Some scholars even think that 'mega-urban' forms of functionally connected and spatially merged networks and clusters of formerly individual cities represent the future of human habitats (Lang & Knox, 2009).

With a view to examining urbanization in CEE countries, our analysis revealed that Eastern European city regions still find themselves in a relatively early phase of suburban expansion. In this historically specific transition phase, a higher degree of dispersion is characteristic. In the rapidly growing capital regions, however, a process of dynamic compaction is evident. Two counteracting processes are, thus, in interplay: on the one hand, new urban development in suburban locations (so-called leapfrog development) increases the NP; on the other hand, new development also results in the merging of spatially isolated settlements so that the LP gains in size, and the NP decreases over time. The urbanization trajectories of regions such as Moscow or Warsaw are characterized by both trends, but the latter was obviously more distinctive. We nevertheless assume that not all city regions follow this trajectory because many CEE regions demonstrate rather weak economic development and lower urban growth pressure. Against this background, the authors believe the East–West '*urbanization gap*' will decrease in the future, but it will likely not – even in a long-term view – diminish.

The indicator utilized, i.e. the DI, is not a one-size-fits-all approach to characterizing patterns of urban land use systems. Within its conceptual structure, the DI allows a quantification of dispersion as the degree of centralization and spatial clustering of urban functions (here addressed in the form of land use patches). We are aware that this metric is incomplete, i.e. other spatial features such as distances between patches remain disregarded. However, the general spatial patterns – as Taubenböck et al. (2019) have shown – can be quantified reliably.

The spatial concept applied must also be subject to careful consideration. The spatial baseline for comparison does influence the DI and thus, the interpreted results. We are aware that our static as well as our city-adjusted spatial entities do not provide results that can be evaluated as general for all possible spatial entities. However, the employment of static spatial entities as well as city-adjusted accessibility zones enables both an equal spatial baseline and a context sensitive spatial approach.

The data sets applied must be subject to careful consideration. When land use patterns are used for interpretation, the initial satellite data source and image analysis provide a product of a certain degree of abstraction. The geometric resolution of satellite images has impact on the measured DI, e.g. lower image resolutions can lead to a smaller number and larger sizes of patches because non-urban areas such as rivers may be suppressed in the data. However, Taubenböck et al. (2019) showed that the relative relation of the DI remained € generally stable across image resolutions. The accuracy of classifications based on remote sensing data in general is subject to errors, which may also alter the DI. In some cases, small errors may transmit to large effects with respect to LP. Although this source of error is difficult to

identify in the results, we assume that the tested high accuracy of the input data (80–90% agreement with reference data; see Klotz et al., 2016) keeps this error comparatively low.

7. Conclusions

Our cross-sectional and longitudinal analysis of urban land use patterns of 230 city regions in 34 European countries indicates a tremendous amount of variation as the outcome of unique topographical conditions, path-dependency of historical settlement systems, varying economic and demographic trajectories, different housing preferences or growth management and land use policies. We, nevertheless, see some structural differences between Eastern and Western city regions. Eastern regions are less populated, less urbanized, more dispersed and denser than in the West. Processes of post-socialist urban restructuring have definitely resulted in a change of land use patterns similar to that of Western Europe since the late 1950s. Suburbanization has led to a strong increase in urbanized land as well as to reductions in urban density. We, nevertheless, do not think that 'catch-up' growth, leading to full convergence with Western-style urbanization, will be a realistic future scenario. Returning to the question in the title of this article, i.e. whether Eastern European urbanization can best be characterized as hybrid: cities and city regions manifest characteristics of convergent adaptation and path-dependency at the same time; they prove typical features of capitalist urbanization, but relics of the socialist past are still omnipresent.

Supplemental material

Supplementary material to this article can be found online at

https://doi.org/10.1177/2399808319846902

Chapter XVIII

Measuring the spatial hierarchical urban system in China in reference to the Central Place Theory

Habitat International, 2020, vol. 105, 102264

Lifeng Shi, Michael Wurm, Xianjin Huang, Taiyang Zhong & Hannes Taubenböck

© 2020 Elsevier Ltd. All rights reserved.

Note: The paper has been utilized in this research report in accordance with the reuse rights granted to the authors as outlined at: https://www.elsevier.com/about/policies-and-standards/copyright.

https://doi.org/10.1016/j.habitatint.2020.102264

Received 17 June 2019; Accepted 23 Sept 2020; Published 3 Oct 2020

Abstract

Dynamic urbanization processes have caused great changes on China's urban spatial pattern and hierarchical structure in the past few decades. In this study, we propose new metrics for measuring the pattern of urban land by considering both, spatial distribution and hierarchical structure of the urban system. The metrics are deduced from Christaller's Central Place Theory and allow evaluating the spatial pattern of urban land based on three

indicators: azimuth heterogeneity, distance heterogeneity and quantity heterogeneity of subordinate urban patches around the corresponding higher-rank urban patch. For monitoring the evolution of the spatial hierarchical urban system, we apply these metrics on China's urban lands for the time period of 1990 until 2015 based on provincial administrative units. The results show that (1) at the national level, we find hierarchical spatial urban systems similar to the ideal pattern suggested by the Central Places Theory in the eastern and central regions; (2) the hierarchical urban systems develop between 1990 until 2015 in most areas in the opposite direction of the ideal pattern suggested by the Central Place Theory; however, this tendency is more evident in the eastern and central regions than in western inland regions. Besides, we find underdeveloped urban systems, i.e. patterns that have some missing high levels in the hierarchical gradation and over developed urban systems, i.e. patterns that have some missing low levels in the hierarchical gradation. We discuss alternative urban systems for explaining the inapplicability of the Central Places Theory in some examples.

Keywords

Spatial pattern; landscape; central place theory; hierarchical urban system; urbanization; China

1. Introduction

The process of urbanization in China has been marked by dramatic growth in recent decades (Chen et al., 2013; Liu et al., 2019). Various aspects of it have been attracting much attention in research, such as rural depopulation (Li et al., 2018), economic transition (You, 2016), urban growth (He et al., 2016), or lifestyle changes (Lu & Dang, 2016). Among them, urban growth is the most direct embodiment of urbanization in space (Taubenböck et al., 2012). Benefiting from the development of remote sensing technology, urban land monitoring and mapping across China has been widely applied (Zhang et al., 2018; Zhou et al., 2016). The scientific literature commonly revealed that spatial urban growth in China featured highest rates during the past few decades (e.g. Huang et al., 2016; Wang et al., 2018; Yang et al., 2017). The dimensions and hierarchies of urban centers and urban size distribution have changed significantly due to these dramatic transformations (Fu et al., 2019; Taubenböck et al., 2019). All of these changes have influenced and will continue to influence the spatial hierarchical urban system in China (Kuang et al., 2016). Against this background, manifold methods have been developed to quantitatively measure the dynamic urban system in China, which can be mainly classified into landscape metrics and empirical models.

Landscape (or spatial) metrics measure landscape characteristics (McGarigal et al., 2009). Based on it, many pattern specificities of the dynamic urban system are described and analyzed, such as urban land dimensions (Shi et al., 2015), spatial distributions (Taubenböck & Wiesner, 2015), or urban patch forms (Xu et al., 2016). These metrics are designed for measuring one particular feature; therefore, commonly different metrics are combined. However, this includes a series of challenges: 1) Many metrics are abstract and thus simplify complex spatial contexts; 2) many metrics feature high collinearity (Riiters et al., 1995); 3) different metrics feature different dimensions; 4) most metrics are prone to the modifiable areal unit problem (MAUP) (Openshaw, 1984). To address these challenges, few new metrics have been proposed for meeting some particular requirements for measuring the spatial urban system. For instance, the dispersion index (DI) developed by Taubenböck et al. (2014) for measuring the structure of urban patch sizes with respect to being clustered or dispersed; metrics to classify urban growth categories (UGC) and urban growth fluctuations (UGF) were developed by Shi et al. (2019) for describing the spatial distribution of urban growth. Overall, these classical and newly developed *landscape metrics* allow to quantitatively measure the spatial patterns of urban systems; however, the known studies did not apply these metrics with respect to the hierarchical structure of the urban system.

Studies on *empirical models* refer to the hierarchical structure of cities within a given spatial entity (e.g. a country). The relationships between city quantities and city sizes (i.e. the hierarchical structures) are considered as very important characteristics of urban systems. An empirical model known as Zipf's law shows that the size distribution of urban can be approximated by the power law (Ioannides & Skouras, 2013; Newman, 2005). For instance, Soo (2005) collected urban populations from 75 countries, and found that the distribution follows the Pareto distribution with a tail index of a little higher than 1. Shi et al. (2016) found that the tail indices for different time steps in China are between 0.8 and 0.9 using the sizes of urban patches, i.e. the quantity of small cities was insufficient. Besides, the well-known geographical model called Central Place Theory (Christaller, 1966) is often

referenced by studies to explain the hierarchies of the urban system. For instance, Hsu (2012) found that city size distribution exhibits a power law under a central place hierarchy structure. Van Meeteren & Poorthuis (2018) found that central functions proposed by the Central Place Theory have typical ranges relating to population spread. Actually, the Central Place Theory as an empirical model still lacks an internationally accepted measurement approach by spatio-quantitative indicators; therefore, it remains challenging to unambiguously answer to what extent the urban system within a given spatial entity corresponds to the theoretical model idea. Contrary to the studies on *landscape metrics*, studies on *empirical models* mostly explain the rationality of urban hierarchical structures of a spatial entity; however, the spatial characteristics and relations among urban entities of different sizes have not yet been studied by these *empirical models*.

The configuration of both, the spatial pattern and the hierarchical structure of an urban system have not yet been systematically and simultaneously investigated and quantitatively analyzed in China. Focusing on these challenges, we propose new quantitative metrics for measuring both, landscape characteristics and hierarchical structures for regional urban systems. These metrics descend from the proposed characteristics (i.e. distribution, quantity and size) of the urban system related to the Central Place Theory. Based on these proposed metrics, we describe and analyze urban land spatial patterns, hierarchical structures and their evolutions in China from 1990 to 2015 aiming to answer the following specific research questions:

1) Is the spatial urban system in China generally organized close to the Central Place Theory?
2) Does the hierarchical urban system in China develop since the 1990 in the direction of the Central Place Theory?

The main workflow of this study is illustrated as follows (Fig. 1): *First*, we deduce these metrics based on the theoretical background of the Central Place Theory (Section 2); *secondly*, we rely on China's urban land classification products (Section 3.1) and develop the conceptual framework for the physical approach using urban patch sizes (Section 3.2); *thirdly,* we exemplify different spatial patterns in detail by two typical regions (Section 4.1). We reveal the development of recent spatial patterns of urban lands across China (Section 4.2) and its spatiotemporal evolutions for a time period of 25 years from 1990 to 2015 (Section 4.3). In Section 5, limitations of our approach with respect to the Central Place Theory, to our data and metrics are discussed and finally, Section 6 concludes the paper.

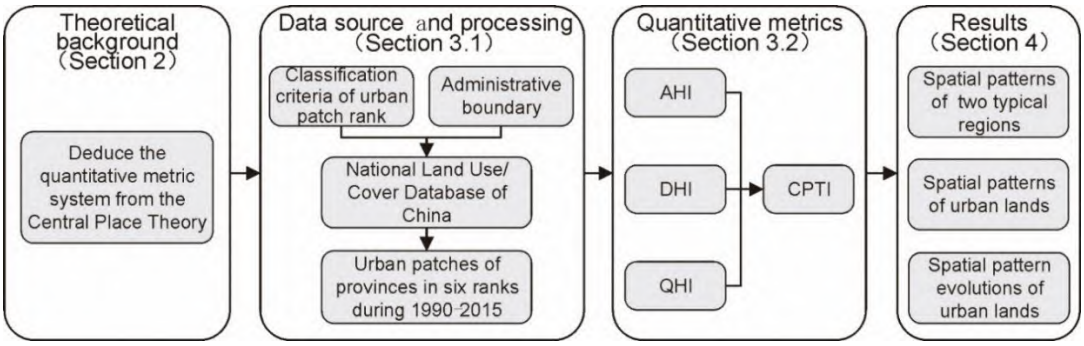


Fig. 1. Conceptual framework and workflow for this study.

2. Theoretical background of urban spatial patterns

Cities do not exist in isolation, but in the form of an urban system (Parr, 2014). An urban system is superficially understood as a group of spatially related cities, which feature complementary functions in dependence of their sizes and types (Anas et al., 1998). All these cities within this urban system are organized by a nested hierarchical struc ture, centered around the core city or cities of high-order functions to smaller cities with low-order functions; they are interlinked and interact with each other (Castells, 1999; Fujita et al., 1999). These characteristics of an urban system were first systematically introduced by German geographer W. Christaller in 1933 through his Central Place Theory by taking the example of the urban system in Southern Germany (Christaller, 1966). In the proposed ideal case, one higher-rank city is surrounded by six subordinate cities, which are spatially distributed around the core city as an equilateral hexagon. The higher-rank city locates in the geometric center of the hexagon, the six subordinate cities locate in the six vertexes of the hexagon. Subsequently, Christaller defined three theoretical models respectively according to marketing principle, transport principle and administrative principle (Fig. 2). Hierarchical cities constitute hierarchical equilateral hexagons, but with different nested structures in the three models illustrated below. The ratios k of quantities of subordinate cities to the higher-rank city are respectively 3, 4 and 7 under the three principles.

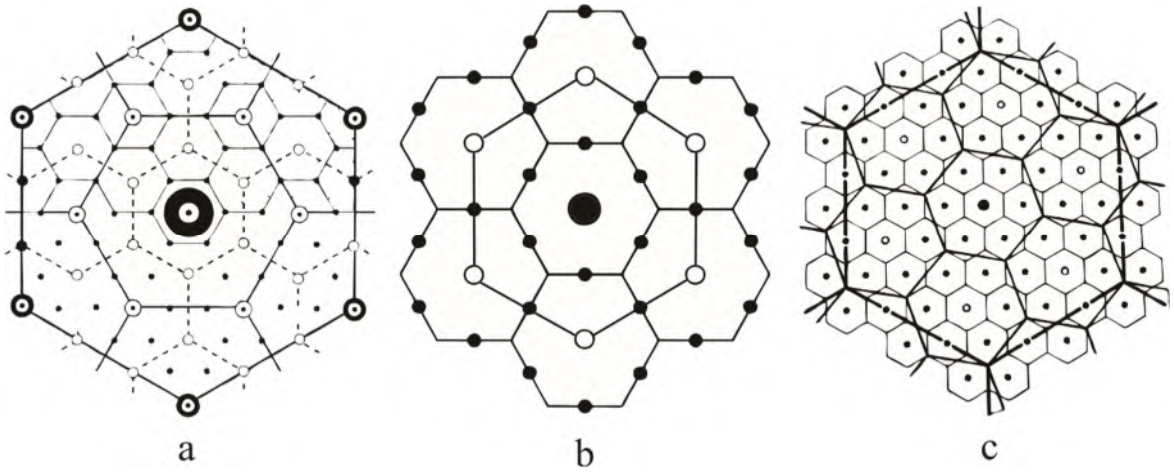


Fig. 2. Christaller's Central Place Theory: (a) marketing principle (k = 3); (b) transport principle (k = 4); (c) administrative principle (k = 7) (Christaller, 1966).

Since the 1950s, urban systems in many regions were matched and explained by this theory by focusing on the relationships among urban ranks, functions and quantities (e.g. Berry, 1964; Berry & Garrison, 1958; King, 1985; White, 1977). With the fast development of regions and the process of globalization, the traditional Central Place Theory is challenged when it is adopted to explain the urban system on a large scale (e.g. on global scale) (Zook, 2001). To account for this, other supplementary theories (or models) to the Central Place Theory have been introduced: examples are the so called '*interlocking network*' which has been adopted to describe the 'non-local' interurban relation (Taylor et al., 2002); further, the Central Flow Theory has been introduced to present interurban processes, i.e. the processes of the networks that link cities together across different regions (Taylor et al., 2010). Actually, the Central Place Theory still plays a crucial role when hierarchical urban systems

are analyzed at a regional scale (Em, 2018), but this theory sometimes needs recalibration due to today's dynamic processes of urbanization (Van Meeteren & Poorthuis, 2018). However, generally agreed upon metrics or methodologies for measuring the spatial pattern and the hierarchical structure of regional urban systems are absent. Thus, to what extents the regional urban systems approach to the Central Place Theory spatially as well as how they evolve over time in a systematic and consistent way remains challenging.

In this paper, we introduce metrics that aim to quantify the urban system in a systematic and consistent spatial way based on spatial distribution, size and quantity of urban patches. Based on the proposed characteristics of an urban system by the Central Place Theory, we deduce three independent metrics, namely *Azimuth Heterogeneity Index* (AHI), *Distance Heterogeneity Index* (DHI) and *Quantity Heterogeneity Index* (QHI) to measure the spatial hierarchical urban system. AHI is designed to measure the azimuthal distributions of subordinate urban patches around the corresponding higher-rank urban patch; DHI is designed to measure the geometric distance differences between sub ordinate urban patches and the corresponding higher-rank urban patch; QHI is designed to measure the quantitative rationalities of subordinate urban patches around the corresponding higher-rank urban patch. Subsequently, we develop the *Central Place Theory Index* (CPTI) by combining the three metrics. We consider for our study the proposed ideal model of the Central Place Theory as our relational starting point. Thus, these metrics aim to measure whether urban spatial hierarchies and their specific spatial distributions are close to or far from the ideal model of the Central Place Theory at a regional scale.

3. Data and methodology

3.1. Data source and data processing

We rely on the '*urban lands*' classification for China from the National Land Use/Cover Database of China (NLUD-C), which has been the main authoritative data source focusing on the status and dynamics of China's land use/cover (Liu & Deng, 2010; Liu et al., 2003, 2005). '*Urban land*' in this data set is defined as the urbanized land of cities and towns. Urbanized land of villages is included in another category, named rural settlement. The latter is not taken into consideration in this study. The mapping of urban land is performed via visual image interpretation based on professional knowledge and experience. The mapping accuracy of urban land is verified higher than 90% (Zhang et al., 2014). The NLUD-C is open access on a resolution at 1 km (http://www.resdc.cn/). We obtain urban lands for entire China at five-year intervals for the time steps 1990, 1995, 2000, 2005, 2010 and 2015 from the NLUD-C.

Cities are classified into seven ranks from villages to metropolis ac cording to their telephone numbers in the process of Central Place Theory by Christaller. In our study, we disregard the villages due to the data resolution and generally classify all urban patches into six ranks from small towns to metropolis. We consider in our study the urban patch size as an indicator for the city size. We empirically find that the radius of higher-rank urban patches in a region is frequently around two times of its subordinate urban patches; this indicates that the area of a higher-rank urban patch is four times the size of its subordinate urban patches. Based

on this empiricism, we adopt the ratio 4 to classify all urban patches into six ranks according to their patch size (Table 1). We find that the ratios of quantities of subordinate urban patches to the higher-rank urban patch (i.e. the k values) mainly locate in the range of 3–7, which are corresponding to the three principles of the Central Place Theory.

As spatial unit we apply provincial administrative units, i.e. the hierarchical urban systems are analyzed within the provincial units. The '*Province*' is the highest administrative unit and the most commonly used "regional scale" in China. It generally contains urban patches of complete hierarchical ranks from the lower to the higher ranks. Thus, we assume planning authorities develop their hierarchical urban system within this spatial entity. Two specific cases are Hong Kong and Macao: they are provincial administrative units; however, both of them feature only one medium or lower-rank urban patch within their very small administrative areas. It indicates that they have no hierarchical urban system inside. Thus, we merge them into the neighboring Guangdong province, which has significantly influenced the development of them in the past few decades. Ultimately, we extract '*urban lands*' for entire China using six ranks in 32 administrative units, with four of them being municipalities (i.e. Beijing, Tianjin, Shanghai and Chongqing).

Table 1. Hierarchical division and quantity statistic of urban patches in China.

Urban patch rank	Urban patch size (km^2)	Urban patch numbers of different ranks during 1990 and 2015					
		1990	1995	2000	2005	2010	2015
6	(1024, 4096)	0	0	0	2	2	7
5	(256, 1024]	2	6	6	11	13	43
4	(64, 256]	43	48	63	79	84	170
3	(16, 64]	185	256	257	350	412	738
2	(4, 16]	923	1035	1127	1134	1156	1319
1	[1, 4]	3233	3663	3095	2992	3030	2035

3.2. Quantitative metrics for urban land spatial pattern measurement

With this classification of ranks, we aim at measuring the urban land pattern with respect to both, hierarchical structure and spatial configuration. In general, we measure the spatial distribution of subordinate urban patches around the corresponding higher-rank urban patch. For instance, we measure the spatial distributions of urban patches in rank 1 around the corresponding urban patches in rank 2, the spatial distributions of urban patches in rank 2 around the corresponding urban patches in rank 3, and so on, until the spatial distributions of urban patches in rank 5 around the corresponding highest rank 6. To do so, we first define the radiation range of the higher-rank urban patch, and then we measure the quantity and distribution of subordinate urban patches around the corresponding higher-rank urban patch. To define which subordinate patch belongs to which higher-rank patch, we postulate that the breakpoint is the midpoint of the connection between two adjacent higher-rank urban patches. We generate the perpendicular bisectors of all connection lines between adjacent higher-rank urban patches for constructing polygons (Fan et al., 2017). We consider these constructed polygons (Thiessen polygons) are the ideal radiation ranges of higher-rank urban patches (Fig. 3). Based on these Thiessen polygons of higher-rank urban patches, the affiliations between subordinate urban patches and higher-rank urban patches are confirmed.

The same steps are repeated for urban patches from rank 2 to rank 6, urban patches in rank 1 without subordinate urban patches.

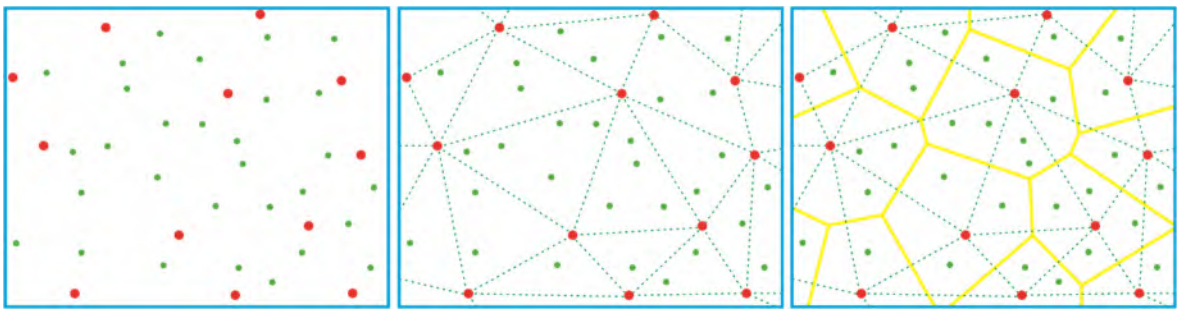


Fig. 3. Radiation range construction of urban patches (red dots) in different ranks.

For the ideal model of the Central Place Theory we assume the following spatial pattern characteristics: *first*, the angles of subordinate urban patches around higher-rank urban patches are constant; *second*, the geometric distances between subordinate and higher-rank urban patches are constant; *third*, the number of subordinate urban patches around a higher-rank urban patch is six. However, the angles, geometric distances and quantities of the subordinate cities naturally vary in reality. We develop three new metrics for measuring the three characteristics, namely, Azimuth Heterogeneity Index (AHI), Distance Heterogeneity Index (DHI) and Quantity Heterogeneity Index (QHI). The three metrics are independent from each other. The equations are as follows:

$$\mathrm{AHI} = \begin{cases} 1 & (n = 0, 1) \\ \dfrac{n\sum_{i=1}^{n}\left|A_i - \frac{360}{n}\right|}{2\times 360\times(n-1)} & (n \geq 2) \end{cases} \tag{1}$$

$$\mathrm{DHI} = \begin{cases} 1 & (n = 0, 1) \\ \dfrac{n\sum_{i=1}^{n}\left|L_i - \frac{\sum_{i=1}^{n} L_i}{n}\right|}{2\times(n-1)\times\sum_{i=1}^{n} L_i} & (n \geq 2) \end{cases} \tag{2}$$

$$\mathrm{QHI} = \begin{cases} \dfrac{6-n}{6} & (n \leq 6) \\ \dfrac{n-6}{n} & (n > 6) \end{cases} \tag{3}$$

where n is the quantity of subordinate urban patches around the corresponding higher-rank urban patch; *Ai* are the angles between subordinate urban patches and the corresponding higher-rank urban patch; *Li* are geometric distances between subordinate urban patches and the corresponding higher-rank urban patch; the index values of AHI, DHI and QHI range from 0 to 1. The closer the values of the metrics are to 0, the closer the spatial pattern is to a perfect distribution of the subordinate urban patches around a higher-rank urban patch with respect to the Central Place Theory. The closer the values of the metrics are to 1, the higher is the deviation of the distribution of the subordinate urban patches from the defined ideal case. We illustrate some examples of different values for AHI, DHI and QHI in Fig. 4. The first example in Fig. 4c exhibits our relational starting point, that we consider in our study as a perfect spatial pattern representing the Central Place Theory. Based on it, we propose the Central Place Theory Index (CPTI):

CPTI = (AHI + DHI +QHI) / 3

The index values of CPTI range from 0 to 1. The closer the CPTI is to 0, the more perfect the spatial pattern of urban patches is with respect to the ideal case. The closer the CPTI is to 1, the higher is the deviation of the spatial pattern. The CPTI value of a certain urban patch in rank *i* indicates the spatial pattern of urban patches in rank *i-1* around the current urban patch in rank *i*.

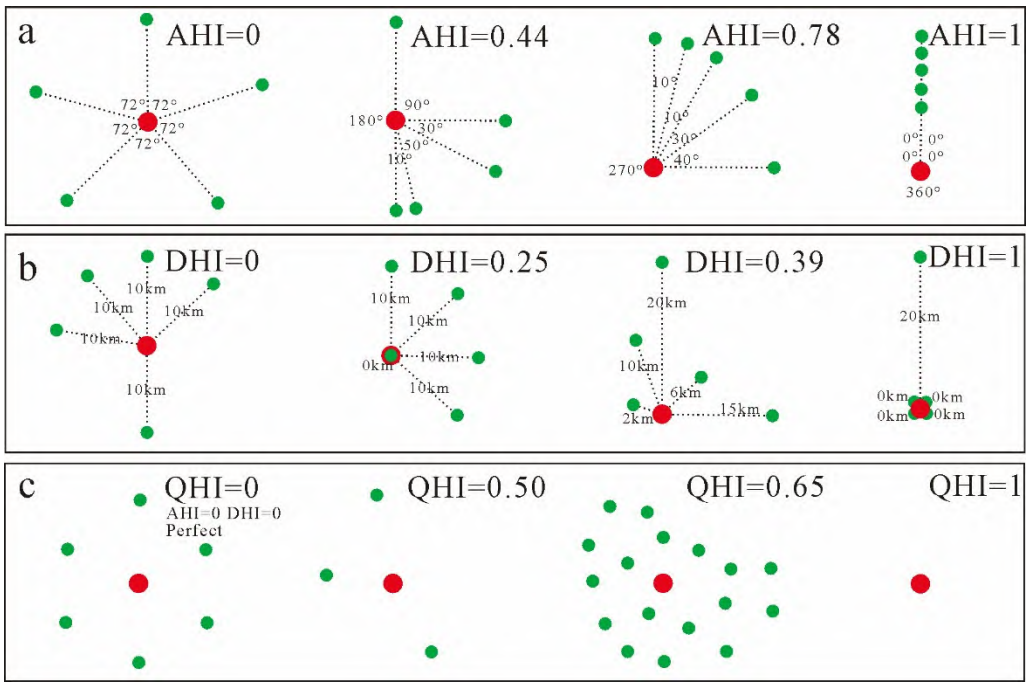


Fig. 4. Illustrations of the three metrics AHI, DHI and QHI: (a) four random patterns exemplifying the AHI; (b) four random patterns exemplifying the DHI; (c) four random patterns exemplifying the QHI.

4. Results

4.1. Spatial pattern evolutions of urban lands in Jiangsu and Sichuan

In a first approach we contrast a measured hierarchical spatial urban system that follows closely the Central Places Theory to a spatial pattern that differs from this spatial configuration. We exemplify appearances of these differing spatial patterns by the provinces of Jiangsu and Sichuan. The urban patches in Jiangsu reveal a spatial pattern close to the ideal Central Places Theory model. Jiangsu also shows only gradual change in the spatial urban pattern over time. On the contrary, Sichuan reveals a spatial pattern with significant deviation from the ideal model and is characterized by fluctuating spatial pattern characteristics.

In Jiangsu, the CPTI values of urban patches were relatively low and generally showed minor changes over time from 1990 to 2010; from 2010 onwards, however, dynamics increase (Table 2). It reveals that the spatial pattern of urban lands start close to the ideal model of the Central Place Theory; however, the evolution gradually evolves away from the ideal pattern, especially after 2010. Urban patches at rank 3 and 5 grow much faster than other ranks and thus disturb the hierarchical urban system as shown by the CPTI. From the three independent metrics, we find that the AHI values of urban patches in different ranks are

mainly the highest. They are followed by the QHI values, while the DHI values are the lowest. It indicates that urban patches in Jiangsu feature a relative constant spatial distance, but a relative inconstant azimuthal distribution of subordinate urban patches around higher-rank urban patches.

Table 2. The AHI, DHI, QHI and CPTI values of urban patches in all ranks of Jiangsu and Sichuan from 1990 to 2015.

Metrics	Ranks	Metric values of Jiangsu from 1990 to 2015						Metric values of Sichuan from 1990 to 2015					
		1990	1995	2000	2005	2010	2015	1990	1995	2000	2005	2010	2015
AHI	Rank 6 to 5	–	–	–	–	–	0.55	–	–	–	–	–	1.00
	Rank 5 to 4	–	–	–	0.60	0.65	0.70	–	–	–	1.00	1.00	–
	Rank 4 to 3	0.71	0.66	0.76	0.75	0.70	0.59	0.04	0.58	0.55	–	–	0.41
	Rank 3 to 2	0.50	0.58	0.63	0.68	0.78	0.84	0.75	0.69	0.52	0.80	0.82	0.64
	Rank 2 to 1	0.60	0.51	0.52	0.51	0.55	0.68	0.61	0.72	0.67	0.72	0.69	0.90
DHI	Rank 6 to 5	–	–	–	–	–	0.22	–	–	–	–	–	1.00
	Rank 5 to 4	–	–	–	0.21	0.18	0.64	–	–	–	1.00	1.00	–
	Rank 4 to 3	0.23	0.20	0.23	0.34	0.44	0.33	0.64	0.44	0.34	–	–	0.26
	Rank 3 to 2	0.28	0.38	0.44	0.52	0.72	0.70	0.60	0.56	0.52	0.55	0.59	0.54
	Rank 2 to 1	0.32	0.28	0.30	0.26	0.23	0.48	0.47	0.51	0.54	0.62	0.60	0.83
QHI	Rank 6 to 5	–	–	–	–	–	0.14	–	–	–	–	–	1.00
	Rank 5 to 4	–	–	–	0.00	0.40	0.71	–	–	–	1.00	1.00	–
	Rank 4 to 3	0.33	0.43	0.33	0.32	0.52	0.45	0.66	0.00	0.16	–	–	0.34
	Rank 3 to 2	0.22	0.43	0.45	0.55	0.75	0.74	0.79	0.58	0.62	0.64	0.63	0.63
	Rank 2 to 1	0.42	0.45	0.48	0.46	0.50	0.58	0.51	0.56	0.62	0.65	0.66	0.84
CPTI	Rank 6 to 5	–	–	–	–	–	0.30	–	–	–	–	–	1.00
	Rank 5 to 4	–	–	–	0.27	0.41	0.68	–	–	–	1.00	1.00	–
	Rank 4 to 3	0.42	0.43	0.44	0.47	0.55	0.46	0.45	0.34	0.35	–	–	0.34
	Rank 3 to 2	0.33	0.46	0.51	0.58	0.75	0.76	0.71	0.61	0.55	0.66	0.68	0.60
	Rank 2 to 1	0.45	0.41	0.43	0.41	0.42	0.58	0.53	0.60	0.61	0.66	0.65	0.86

The hierarchical urban spatial pattern in Jiangsu certainly reflects the dynamic processes of change since 1990. In 1990, three fourth-rank urban patches represented the highest level of the hierarchy; from 2005 onwards, a fifth-rank patch took over this position, and since 2015 there have even been two sixth-rank urban patches. In order to present the hierarchical structure and the spatial evolution of the main urban patches, the subordinate relationships among urban patches of rank 3, 4, 5 and 6 in Jiangsu are illustrated in Fig. 5. In general, we find a clustering of the higher-ranking urban patches in the southern part of Jiangsu. For instance, both of the two urban patches in rank 6 and five of the seven urban patches in rank 5 are located in the southern part of Jiangsu in 2015. This leads to a clustering of urban patches of rank 5 north of urban patches of rank 6 and a less-distribution of urban patches of rank 4 in the southern direction related to rank 5 urban patches. With the fast urban expansion and urban patch annexation, the quantity of lower-rank urban patches reduced continuously since 1990, and especially after 2005. The insufficient distribution of lower-rank urban patches leads to their high and increasing QHI values.

In Sichuan, the CPTI values of urban patches show very significant dynamics after the year 2000 due to the high rank urban patch reclassification and incomplete urban structure formation (Table 2), which are illustrated in Fig. 6. During the 1990s, the urban hierarchy consisted of urban patches from rank 1 to rank 4. The urban patch of Chengdu (the capital of Sichuan) was the only fourth-rank patch. Due to the fast urban growth of Chengdu, the urban patch of Chengdu was reclassified into rank 5 in 2005; however, the third-rank urban patches grew slower and remained in rank 3. Therefore, the fourth-rank urban patches were missing in 2005, and even in 2010. The same process can be observed in 2015, the urban patch of Chengdu was reclassified into rank 6, the third rank urban patches were reclassified into rank 4 at the same time. However, the fifth-rank urban patches were missing. Therefore, the CPTI value of Chengdu remained on 1 (the highest value) since 2005. The CPTI values of urban patches in rank 3 and 4 show a fluctuant decreasing tendency. The spatial pattern

of urban land evolves towards the ideal spatial pattern suggested by the Central Place Theory. The CPTI values of urban patches in rank 2 show a fluctuant increasing tendency. With it the spatial pattern of urban lands evolves away from the ideal spatial pattern suggested by the Central Place Theory. Besides, almost all urban patches in third and higher ranks cluster in the eastern part of Sichuan. Urban patches in the western part of Sichuan are all in rank 1 and rank 2. The missing of urban ranks, together with the unbalanced spatial distribution of urban ranks generate the strongly deviating spatial pattern in Sichuan.

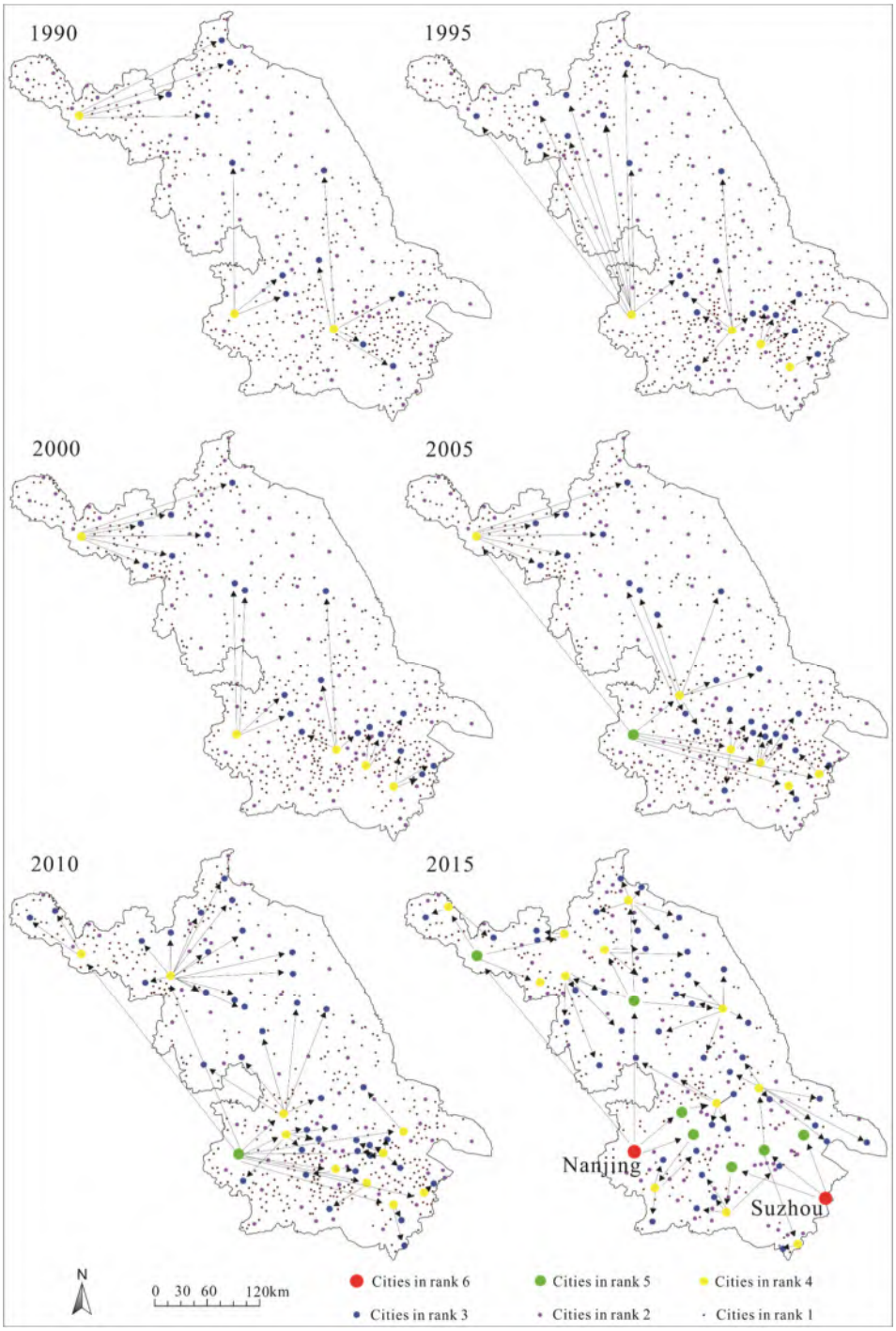


Fig. 5. The spatial distributions of urban patches in all ranks and their temporal evolution from 1990-2015 in Jiangsu.

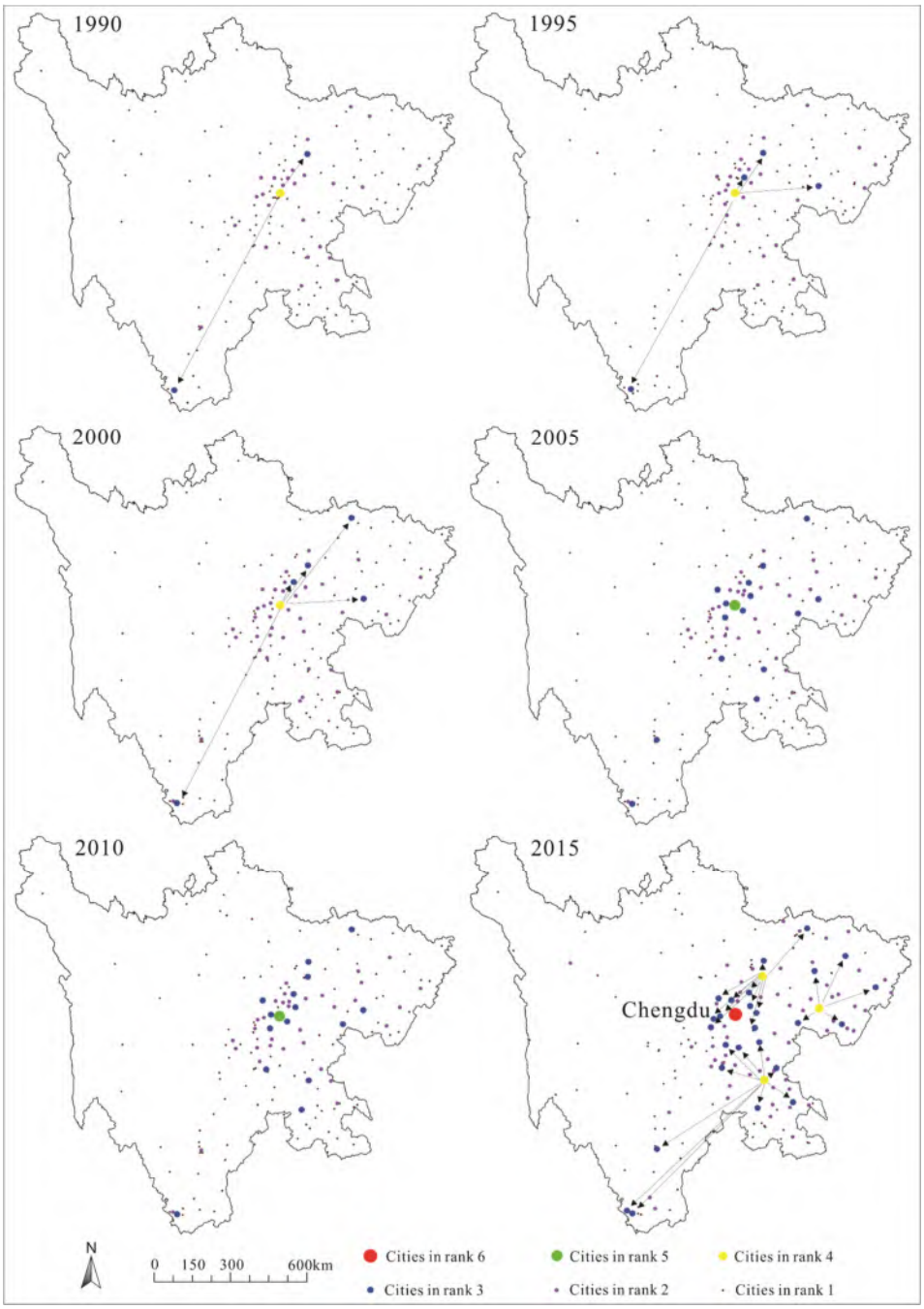


Fig. 6. The spatial distributions of urban patches in all ranks and their temporal evolution from 1990-2015 in Sichuan.

4.2. Spatial patterns of urban lands in China in 2015

Scaling up to all provinces in China allows analyzing whether and where the hierarchical urban system in China is generally close to the Central Place Theory. In the year 2015, we find that the hierarchical spatial patterns vary significantly across China with respect to being close or far from the spatial configuration suggested by the Central Place Theory. Besides, we also find that the averages of AHI, DHI and QHI of every provincial unit are highly correlated (Fig. 7). These characteristics contribute to the distinct divisions among different urban land spatial patterns. Based on it, we classify the spatial patterns of urban lands of the 32 provincial units into *five categories* by using a k-means cluster analysis. The five categories are qualitatively named *very strongly devi ating* (from the spatial features in the

Central Place Theory), *strongly deviating*, *deviating*, *similar* and *very similar* according to their average AHI, DHI and QHI values, and contain 5, 9, 6, 5 and 7 group members, respectively.

When we map these five categories, we find a certain geographical pattern across China (Fig. 8a). Urban lands with very similar spatial patterns to the Central Place Theory are mainly concentrated in central and eastern China (including our example Jiangsu (Fig. 5) as well as Zhejiang, Hubei, Hunan and Jiangxi), where the quantity and distribution of urban patches of higher ranks are relatively proportional (Fig. 8b). Geographical patterns strongly deviating from the spatial pattern of the Central Place Theory are covered by four municipalities and Taiwan. We observe that all of these administrative areas are comparatively small. For China we identify the highest quantity within a group with nine provinces in the category of strongly deviating spatial patterns. The largest concentration of them is located in Southwestern China. An area that features the highest topographic relief (Fig. 8b).

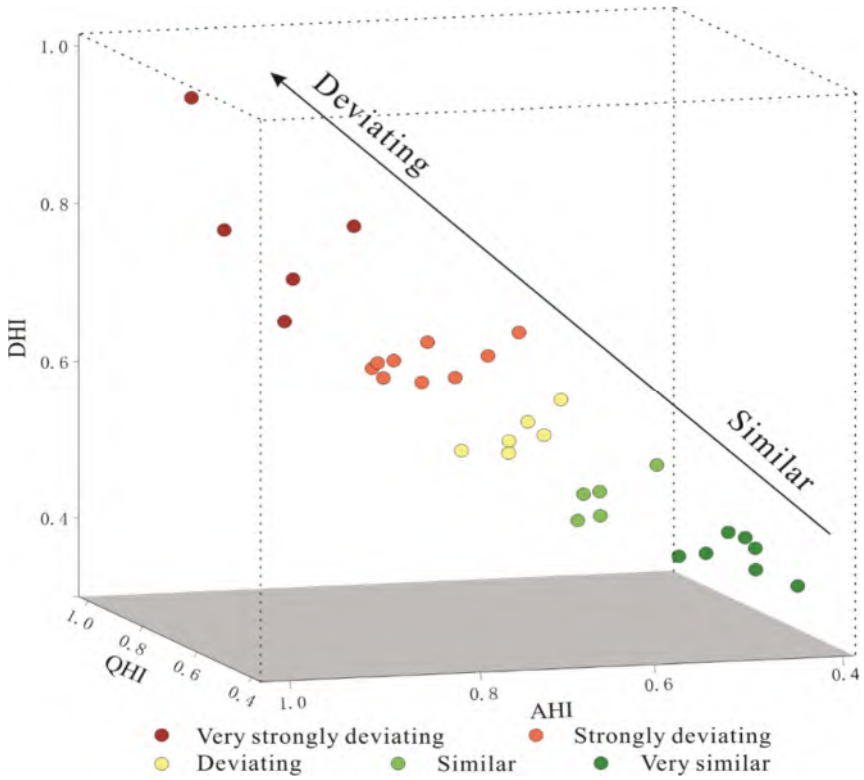


Fig. 7. The average AHI, DHI and QHI values of the 32 provincial units and the five categories classified whether spatially similar or deviating from the Central Place Theory of urban land spatial patterns in 2015.

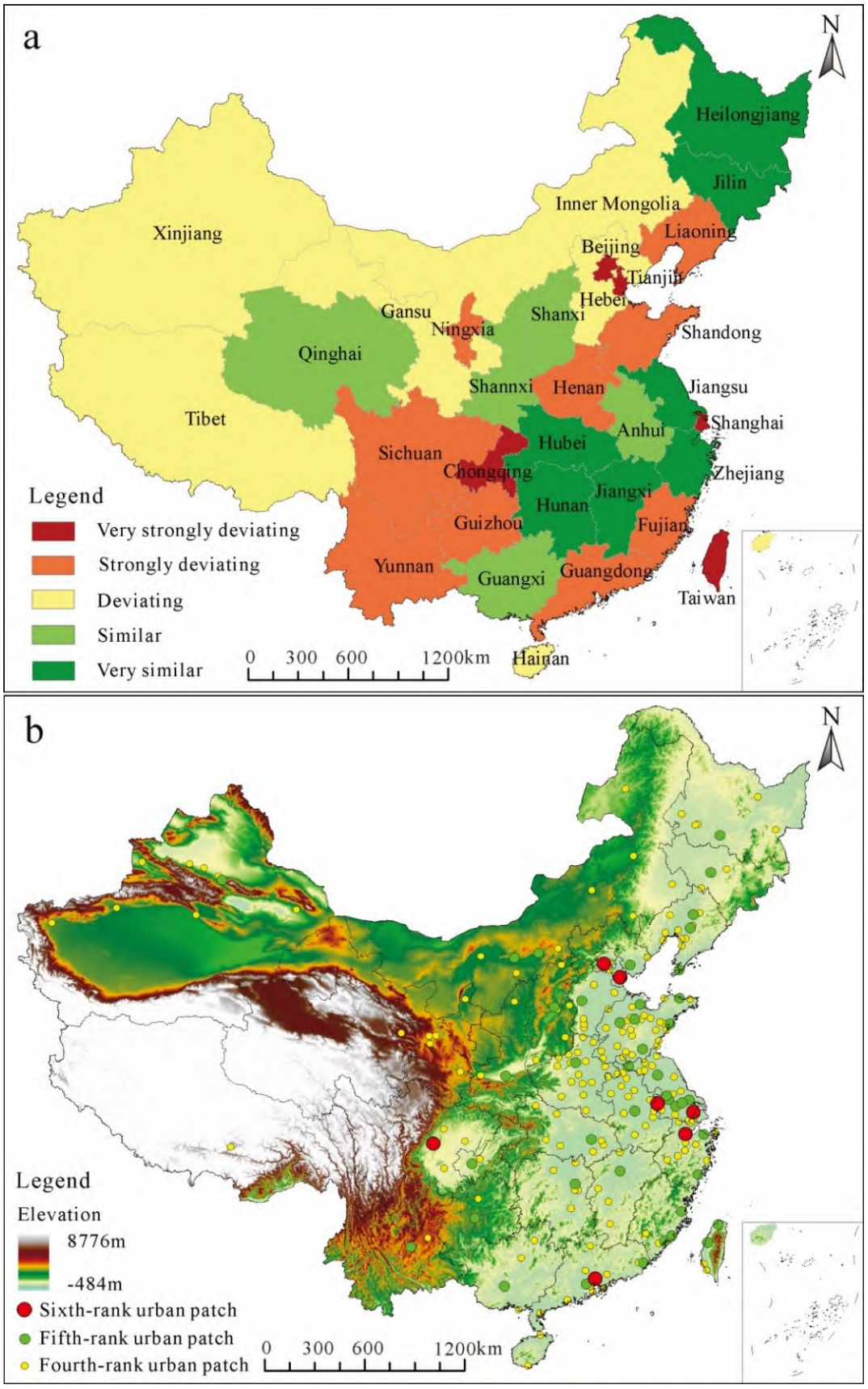


Fig. 8. a): The spatial distribution of urban land spatial patterns classified based on their similarity to the Central Place Theory at provincial level for China in 2015; b) the spatial distribution of higher-rank urban patches in China in 2015 with respect to topography.

4.3. Spatial pattern evolution of urban lands in China from 1990 until 2015

Does the hierarchical spatial urban system in China develop in the direction of the Central Place Theory? We find that the average AHI, DHI and QHI values of most provincial units exhibit an increasing tendency from 1990 until 2015, only few provincial units exhibit almost un changed or decreasing tendency. This indicates that the urban system in most areas develops in the opposite direction of the ideal pattern suggested by the Central Place Theory. We classify the spatial pattern evolutions into five categories based on the variation extents of the three metrics of all provincial units by using the k-means cluster analysis. The five categories are qualitatively named *very strongly deviating tendency* (the average increments of AHI, DHI and QHI values of these provincial units are higher than 60%), *strongly deviating tendency* (ca. 40%), *deviating tendency* (ca. 25%), *slightly deviating tendency* (ca. 5%) and *un changed or similar tendency*, and they contain 6, 5, 6, 8 and 7 group members, respectively (Fig. 9).

When we map the five categories of urban spatial pattern evolutions for China, we find certain geographical tendencies (Fig. 10). The spatial distributions of provincial units with very strongly or strongly deviating tendency are mainly along the eastern coastline in China. As this is among those areas of the highest urban dynamics (cf. Jiangsu as exemplified in Fig. 5), it reveals that these urbanization processes change the hierarchical structure in the opposite direction to that of the Central Place Theory. The spatial distributions of provincial units with deviating tendency or slightly deviating tendency are mainly concentrated in western inland China. The classified provincial units are characterized by vast administrative areas and slow urbanization processes (e.g. Xinjiang). The spatial distributions of provincial units with unchanged or similar tendency exhibit an aggregation around Shanxi, Henan and Shannxi in northern China. This kind of evolution is characterized by size distributions of urban patches in these provincial units closer to the Central Place Theory.

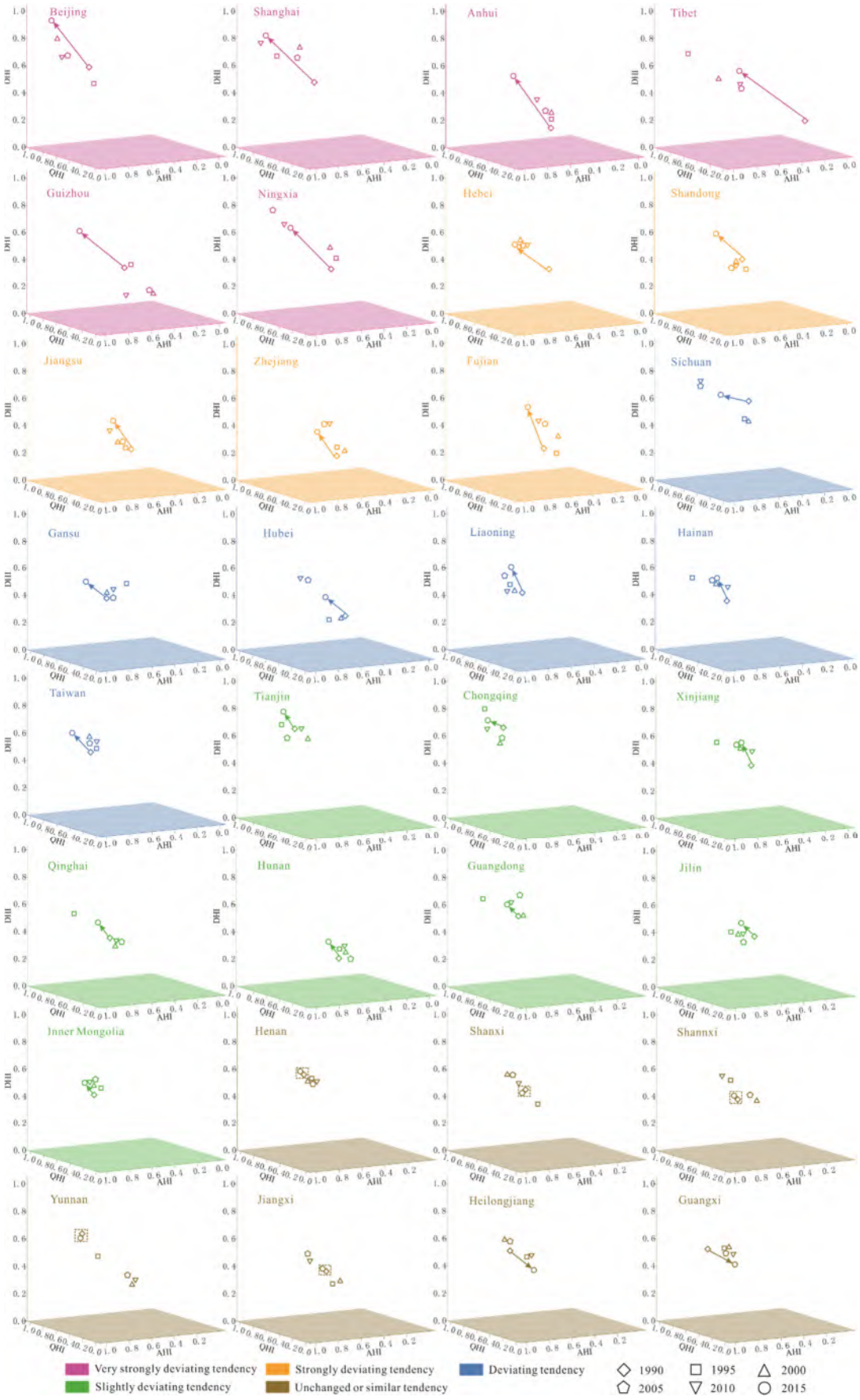


Fig. 9. The temporal evolution from 1990 until 2015 of average AHI, DHI and QHI values of the 32 provincial units classified into five categories of spatio-temporal patterns of urban lands.

Fig. 10. The spatial distribution of the categories of urban land spatial pattern evolutions.

5. Discussion

In this paper, we are measuring whether spatial urban patterns in China relate to the ideal types as conceptualized by the Central Place Theory. We suggest metrics that base on the conceptual foundation of the theory and that allow for a consistent analysis across space.

The proposed metrics are developed based on the capabilities of the spatial input data, i.e. we use the knowledge on spatial distribution, quantity and size of urban patches. They serve as proxies to assess whether patterns are similar or not to the Central Place Theory. Based on these metrics, the hierarchical urban systems are systematically and consistently investigated and quantitatively analyzed at the spatial entity of all provincial administrative units in China. Thus, this study answers the question to what extent regional urban systems approach the Central Place Theory and how they have developed from 1990 to 2015.

However, we are aware that methodological and conceptual challenges remain: first, all spatial patterns of urban lands are measured within the respective provincial units. This conceptual approach aims to respect the administrative planning sovereignty at the regional scale. For it we keep the spatial entities constant over the period of analysis from 1990 to 2015, even if they had been changed. By this conceptual approach, the potential hierarchical relationships with neighboring provincial units are disregarded; second, we consider the size of the main and largest built-up area as the size of a city. We use this size as a proxy to define the hierarchical urban system. However, the size is only a proxy for the importance of the urban node in terms of functionality, economic power, etc.; in other words, if we define the

urban node by different indicators such as trade or economic power, the urban spatial patterns could appear in different orders.

We find on the one hand that the urban systems with complete urban patch hierarchies feature a hierarchy and pattern basically close to the theoretical ideal characteristics of the Central Place Theory (Fig. 5, the case of Jiangsu). It shows that the hierarchical urban system as conceptualized by the Central Place Theory still plays a crucial role in some areas. On the other hand, underdeveloped and over developed urban systems may not be well explained by the Central Place Theory. It indicates that in complex and dynamic urban systems the spatial structures show strong deviations from the Central Place Theory. We reveal underdeveloped urban systems dominated by a single large urban patch (e.g. a mega city) with the next ranks of urban patches missing (the case of Sichuan). Also, we find over developed urban systems lacking small and medium-sized cities due to the coalescence of urban patches. The Guangdong province (especially the Pearl River Delta) is a typical example for this kind of coalescing urban system into a megaregion (Taubenböck et al., 2014). We assume that underdeveloped and over developed urban systems in the different stages of urbanization can be alternative specifications, for whose description the theory of the central location is not suitable.

Actually, both underdeveloped and over developed urban systems reflect a familiar problem during the urbanization processes of China, i.e. the unbalanced regional urban growth. Before the mid-1990s, the Chinese government tried to control the growth of large cities by taking a series of administrative measures (Chan & Wan, 2017). Since the Tenth Five-Year Plan (2001–2005), "*control the growth of large cities*" was replaced by the target of "*a coordinated development of large, medium and small cities (towns)*" (Guan et al., 2018). After entering the Eleventh Five-Year Plan (2006–2010), the megalopolises were intended to cement as leaders of the urban system for exerting their central functions and forming new megalopolises (Han et al., 2018).

The policy restrictions of urban development on big cities (or megalopolises) have been loosening since the 1990s. Small urban patches continuously coalesced with their neighboring big cities (or megalopolises) during these intense urban expansion processes – as our results show. For instance, the ratio of the quantity of urban patches in rank 1 to the quantity of urban patches in rank 2 was 3.50 in 1990 (Table 1), i.e. urban patches in rank 2 were surrounded on average by 3.50 urban patches of rank 1. Until 2015, this ratio was measured to drop to 1.54. These measured developments indicate that the spatial patterns generally deviate away from the ideal pattern suggested by the Central Place Theory. We reveal the unbalanced regional urban growth is intensifying, accompanied with the gradually loosened policy restrictions on big cities (or megalopolises) since the 1990s. On the other hand, the problem of inefficient urban land utilization was monitored to be relatively serious in big cities (or megalopolises) due to their excessive urban growth and coalescence, which has been reflected by abundant vacant residential neighborhoods (Shi et al., 2020) and inefficient industrial areas (Chen et al., 2019). Based on these effects, it can be concluded that - provided the spatial pattern of the Central Place Theory is aimed at - growth control of and efficiency gains for urban areas of large cities (or megalopolises) are the appropriate

approaches. This could help to mitigate the unbalanced regional urban growth and optimize the hierarchical structure of the regional urban system in China.

6. Conclusion

In this study, we focused on spatial urban patterns and spatiotemporal developments in China in the period 1990–2015. Since there is no generally accepted metric or method for measuring urban land patterns from the perspective of hierarchical urban systems, we pro posed four quantitative metrics for measuring both, the spatial distribution and the size structure of urban areas. The ideal model proposed by Central Place Theory is used to assess the degree of homogeneity of urban land patterns. These quantitative metrics represent a stable performance in multi-level, large-scale, and long-term spatial land-pattern measurement.

Based on these quantitative metrics, the hierarchical urban systems and their evolutions for China's provincial units were measured and analyzed. Further, the limitations and applications of the Central Place Theory were discussed when it is adopted to explain China's hierarchical urban system at the regional scale. Besides, we find underdeveloped and over developed urban system, where patterns deviate from the characteristics of the Central Place Theory and potential alternative theories (or models) are necessary to describe them. The results in this study help us to understand urban spatial patterns from a perspective of the hierarchical urban system in China. Nevertheless, an extension with regard to the functionalities of these urban systems is a necessary research direction to capture urban interlinkages in their multidimensional specifications.

Acknowledgments

This work was supported by the National Social Science Foundation of China (Grant No. 17ZDA061); the National Natural Science Founda tion of China (Grant No. 41901230); the China Postdoctoral Science Foundation (Grant No. 2019M651775). Furthermore, the first author would like to thank the German Academic Exchange Service (DAAD) for its scientific research funding.

Part VI

Effects of urban patterns

CHAPTER XIX	The impact of urban configuration types on urban heat islands, air pollution, CO_2 emissions, and mortality in Europe: A data science approach	447
CHAPTER XX	Land consumption in cities: A comparative study across the globe	469
CHAPTER XXI	How we live and what that means – a character study with data from space	495

Chapter XIX

The impact of urban configuration types on urban heat islands, air pollution, CO_2 emissions, and mortality in Europe: a data science approach

Lancet Planet Health, 2024, vol. 8(7), e489-505

Tamara Iungman, Sasha Khomenko, Evelise Pereira Barboza, Marta Cirach, Karen Gonçalves, Paula Petrone, Thilo Erbertseder, Hannes Taubenböck, Tirthankar Chakraborty & Mark Nieuwenhuijsen

© 2024 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

https://doi.org/10.1016/s2542-5196(24)00120-7

Published 3 July 2024

Abstract

The world is becoming increasingly urbanized. As cities around the world continue to grow, it is important for urban planners and policy makers to understand how different urban configuration patterns affect the environment and human health. However, previous studies have provided mixed findings. We aimed to identify European urban configuration types, on the basis of the local climate zones categories and street design variables from Open Street Map, and evaluate their association with motorized traffic flows, surface urban heat island (SUHI) intensities, tropospheric NO_2, CO_2 per person emissions, and age-standardized mortality. We considered 946 European cities from 31 countries for the analysis defined in the 2018 Urban Audit database, of which 919 European cities were analyzed. Data were collected at a 250 m × 250 m grid cell resolution. We divided all cities into five concentric rings based on the Burgess concentric urban planning model and calculated the mean values of all variables for each ring. First, to identify distinct urban configuration types, we applied the Uniform Manifold Approximation and Projection for Dimension Reduction method, followed by the k-means clustering algorithm. Next, statistical differences in exposures (including SUHI) and mortality between the resulting urban configuration types were evaluated using a Kruskal–Wallis test followed by a post-hoc Dunn's test. We identified four distinct urban configuration types characterizing European cities: compact high density (n=246), open low-rise medium density (n=245), open low-rise low density (n=261), and green low density (n=167). Compact high-density cities were a small size, had high population densities, and a low availability of natural areas. In contrast, green low-density cities were a large size, had low population densities, and a high availability of natural areas and cycleways. The open low-rise medium and low-density cities were a small to medium size with medium to low population densities and low to moderate availability of green areas. Motorized traffic flows and NO_2 exposure were significantly higher in compact high density and open low-rise medium density cities when compared with green low density and open low-rise low-density cities. Additionally, green low-density cities had a significantly lower SUHI effect compared with all other urban configuration types. Per person CO_2 emissions were significantly lower in compact high-density cities compared with green low-density cities. Lastly, green low-density cities had significantly lower mortality rates when compared with all other urban configuration types. Our findings indicate that, although the compact city model is more sustainable, European compact cities still face challenges related to poor environmental quality and health. Our results have notable implications for urban and transport planning policies in Europe and contribute to the ongoing discussion on which city models can bring the greatest benefits for the environment, climate, and health.

Keywords

Urban planning; urban development; urban heat islands; air pollution; CO2 emissions; climate change; quality of life; sustainability

1. Introduction

Current and projected rapid urbanization rates highlight the importance of cities and the urgent need to priorities the creation of healthy and sustainable city environments (UN, 2019). Currently, 55% of the global population resides in urban areas, and this proportion is expected to reach 68% by 2050 (UN, 2019). In Europe, 75% of the population already lives in cities, with the urbanization rate predicted to rise to 84% by 2050 (European Commission, 2023). Living in cities offers many advantages, such as improved access to services and infrastructure, higher employment and economic opportunities, and more social connections and networking (Nieuwenhuijsen & Khreis, 2019). However, urban dwellers also face escalating social and economic disparities, a sedentary lifestyle, and increased exposure to environmental stressors that negatively affect human health, including air and noise pollution, rising local temperatures, and the depletion of green spaces (Nieuwenhuijsen & Khreis, 2019). In addition, cities account for approximately 75% of worldwide CO_2 emissions, significantly contributing to the intensification of global warming (Greenhouse Gas Protocol, 2023).

To foster a sustainable and health-centric urban design, the compact city model has emerged as a potential solution. Compact cities are characterized by a greater density and shorter travel distances than sprawling cities, mixed land use, destination accessibility, and shorter distances to public transport (Nieuwenhuijsen, 2020). This layout encourages more walking, cycling, and public transport use, leading to higher physical activity and decreased air and noise pollution and reduced per person carbon emissions (Nieuwenhuijsen, 2020). Additionally, compact cities promote social interactions and a stronger sense of community among residents, contrasting with urban sprawl, which is linked to urban fragmentation, increased infrastructure costs, and socioeconomic inequalities (Nieuwenhuijsen, 2020; OECD, 2018).

At the policy level, there is a growing recognition that the spatial configuration of cities can significantly influence sustainable development and population health through its effect on the natural environment, social context, and human behavior. City leaders across Europe are in the process of implementing sustainability or climate measures, or both, as part of the EU's Green Deal, the Paris Climate Agreement, or voluntary commitments taken by city networks (such as the C40 Mayors' Agenda for a Green and Just Recovery or the Covenant of Mayors) (European Commission, 2008, 2019; UN, 2015; C40 Cities, 2023).

However, despite these efforts, previous studies have shown that cities in Europe still face considerable exposure to adverse environmental exposure and a high premature mortality burden due to high ambient air pollution, road traffic noise, urban heat islands, and insufficient access to green spaces (Khomenko et al., 2021, 2022; Barboza et al., 2021; Iungman et al., 2023). Similarly, one 2021 study indicated that the mitigation targets set for average greenhouse gas emissions across European cities fell short in aligning with the requirements of the Paris Agreement (Salvia et al., 2021). Urgent measures aimed at reducing adverse environmental exposures and greenhouse gas emissions are needed. Therefore, understanding the intricate relationship between urban configuration, environmental exposures, carbon emissions, and health is of crucial importance for urban planners and policy makers.

We built our study upon previous work by Taubenböck et al., (2020), which identified city types on the basis of intraurban morphological configurations, using a sample of 110 cities worldwide. The authors used the local climate zones (LCZs) classification scheme, which provides a harmonized description of urban form, and grouped the majority of European cities within the same city type (Taubenböck et al. 2020). This framework has been widely used as a highly versatile urban form zoning method, which can easily analyze urban characteristics and compare multiple cities under a unified standard (Perera & Emmanuel, 2018).

In this study, we focused on European cities and defined urban configuration, on the basis of available data, as the combination of intraurban morphological configurations, according to the built and natural categories of LCZs, and street design, according to road typologies from the Open Street Map (OSM) database. Our aim was to identify European urban configuration types and provide a descriptive analysis of potential associations with motorized traffic flows, urban heat island (UHI) intensities, and air pollution, based on surface UHI (SUHI) and tropospheric NO_2 proxies and CO_2 per person emissions. In addition, we sought to evaluate the associations between urban configuration types and age-standardized natural-cause mortality rates. Our extensive database of European cities allowed us to disentangle relationships at the European level, shedding light on the associations between urban configuration, levels of environmental exposures, CO_2 emissions, and health, and providing more detailed information on how these urban features are inter-related in Europe.

2. Methods

2.1. European cities and data collection

Our analysis focused on European cities listed in the Urban Audit dataset 2018 (Eurostat, 2018), which uses the city definition as set out by the Organization for Economic Cooperation and Development European Commission, which is based on population density and local administrative boundaries (appendix 1 p 2) (Dijkstra, Poelman, 2012). Overall, we considered 946 European cities (from 31 countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Germany, Estonia, Finland, France, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, and the UK) for the analysis from the Urban Audit database (appendix 2). We collected all data at a 250 m×250 m grid cell resolution based on the Global Human Settlement Layer population dataset for 2015 (Melchiorri et al., 2018), following the same data collection procedure as in our previous studies (appendix 1 p 2) spaces (Khomenko et al., 2021, 2022; Barboza et al., 2021; Iungman et al., 2023).

We evaluated the intraurban morphological configurations of European cities using the LCZs classification, which provides a universal and standardized description of urban form (Stewart & Oke, 2012). The LCZ scheme is based on satellite imagery and classifies urban landscapes considering building density and height, imperviousness, and vegetation variables into 17 LCZs, ten of these representing the built types of urban areas (compact

high-rise, compact mid-rise, compact low-rise, open high-rise, open mid-rise, open low-rise, lightweight low-rise, large low-rise, sparsely built, and heavy industry) and seven representing the non-built or natural types of urban areas (dense trees, scattered trees, bush or scrub, low plants, bare rock or paved, bare soil or sand, and water; appendix 1 p 4) (Stewart & Oke, 2012). For this study, we retrieved the LCZ classification developed by Demuzere et al. (2020), in which every 100 m ×100 m in Europe was assigned an LCZ. We overlaid the LCZ layer with our 250 m grid cell layer and estimated the proportion of area corresponding to each LCZ for each of the grid cells.

To evaluate the street design, we retrieved the density of distinct road typologies from the OSM database (https://www.openstreetmap.org/#map=6/51.33/10.45). We included road types as follows: motorized roads (classified as either motorway or trunk), primary roads, secondary roads, tertiary roads, residential roads (classified as either unclassified, residential, or living streets), pedestrian zones, and cycleways (appendix 1 p 9). For each grid cell, we calculated the length in meters of each road type.

Traffic volume data were retrieved from the OTM database (https://opentransportmap.info/), which details the vehicle per day counts for major roads. To estimate the traffic volume within each grid cell, we aggregated the traffic counts from all roads that intersected each grid cell. Subsequently, to account for variations in road length across grid cells, we calculated a weighted average as follows: grid traffic volume=(grid road length in km ×traffic volume)/road length in km. The resulting variable represented the vehicles per day per km of road within each grid cell.

We assessed the UHI effect on the basis of the SUHI indicator, defined as the higher land surface temperature (LST) in cities compared with surrounding rural areas (Roth, Oke & Emery, 1989). To estimate the LST in cities, we retrieved Landsat-8 images (US Geological Survey, 2023) for 2015 at a 30 m × 30 m resolution and estimated the median summer (from June 1 to August 31) LST for each grid cell. To define the surrounding rural area, we took a buffer zone of 6 km surrounding each city, to ensure sufficient coverage, and defined the rural area on the basis of the Coordination of Information on the Environment Land Cover (CORINE) agricultural, forest, and natural areas categories (appendix 1 p 15). To estimate the average rural LST for each city, we calculated the median summer LST for each of the rural grid cells and afterwards averaged the LST for the whole rural area. The SUHI was estimated as the difference between the average of all median summer rural LST and the median summer LST recorded in each urban grid cell.

For air pollution, we used as our exposure proxy the annual mean tropospheric NO_2 vertical column densities, which constitute a tracer of anthropogenic emissions (Müller, Erbertseder & Taubenböck, 2022). This exposure proxy is suitable to delineate urban pollution islands and it is not influenced by any modelling assumptions on emissions or land use (Müller, Erbertseder & Taubenböck, 2022). Data were retrieved using satellite remote sensing from TROPOMI (European Space Agency, 2023) for 2019, which has daily global coverage at a resolution of 3·5×5·5 km^2. To estimate annual tropospheric NO_2, all daily satellite overpasses were oversampled over a 100 m×100 m grid cell resolution (appendix 1 p 17). We overlaid the NO_2 layer with our 250 m grid cell layer and calculated the mean annual tropospheric NO_2 in µmoL per m^2 for each grid cell. To validate the NO_2 proxy, we used

NO_2 measurements from official monitoring stations, available through the European Air Quality e-Reporting system (European Environment Agency, 2022).

We retrieved CO_2 emissions data from the open-source Data Inventory for Anthropogenic CO_2 (Oda, Maksyutov & Andres, 2018), which provides CO_2 emissions from fossil fuel combustion at 1 km ×1 km resolution for 2019. In this dataset, national emissions are spatialized on the basis of satellite night-time light data and power plant location and profiles.28 To estimate the per person emissions, we overlaid the CO_2 emissions layer with the 250 m grid cell layer and distributed the CO_2 emissions proportionally on the basis of the intersecting area of the 250 m grid cells with the 1 km grid cells (appendix 1 p 20). We divided the emissions between the population in each 250 m grid cell to obtain the CO_2 per person emissions in metric tons. To verify the CO_2 metric, we additionally explored the CO_2 emissions reported in the Copernicus Atmosphere Monitoring Service regional inventory for the residential and transport sectors (Kuenen et al., 2022).

Finally, given that previous studies have indicated associations between urban environment features and health outcomes (Avila-Palencia et al., 2022a; 2022b; Thompson et al., 2020; Olsen et al., 2019), we sought to assess potential differences in adult age-standardized natural-cause mortality rates between the urban configuration types. Natural-cause mortality rates were retrieved at the city level from our previous studies on the health effects of environmental exposures in European cities, estimated on the basis of data from the Eurostat city database for 2015 (appendix 1 p 50) (Khomenko et al., 2021; Barboza et al., 2021). Natural-cause mortality was used due to the absence of data on more specific health outcomes (such as morbidity or traffic injuries) for the large number of European cities in the analysis.

2.2. Data analysis

To analyze the spatial distribution of the variables in a harmonized manner for all European cities, we divided all cities into five rings on the basis of the Burgess concentric urban planning model (Burgess, 1925). For each city, we extracted the urban center coordinates from the OSM Nominatim database (OpenStreetMap contributors, 2015) and calculated the distance from the urban center to the centroid of each of the grid cells. We divided the cities into rings on the basis of the quintile distribution of grid cell distances to the urban centers (appendix 1 p 3). Accordingly, we grouped all grid cells within each city by ring and calculated the mean values of all variables for further analyses.

To identify and characterize distinct urban configuration types, we applied the Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) method, followed by the k-means clustering algorithm. UMAP is a novel non-linear dimension reduction algorithm able to learn the manifold structure of large input datasets and produce a low-dimensional embedding that preserves the basic topological structure of the manifold (McInnes, Healy & Melville, 2018). We used LCZs and OSM variables from each of the five rings for each city as input data for the UMAP algorithm (appendix 1 pp 22–25). Only cities with complete datasets were considered for the analysis, resulting in a sample of 919 cities. The resulting UMAP embedding was evaluated on the basis of the distribution of the

input variables and used as input for the k-means clustering algorithm to identify the urban configuration types. To select the optimal number of clusters we applied the Elbow method and set the number at k=4 (appendix 1 p 25).

Finally, we evaluated statistical differences in traffic volume, SUHI intensity, NO_2 exposure, CO_2 emissions, and natural-cause mortality between the resulting urban configuration types. Given that the SUHI intensity and mortality rates did not comply with the homogeneity of variance assumption, and traffic volume, NO_2, and CO_2 emissions were not normally distributed, we applied the Kruskal–Wallis test, followed by the post-hoc Dunn's test. In all cases, given the variability in exposure and emissions between the city rings, differences in these variables were evaluated for each ring separately. All analyses were conducted using Python version 3.9.6.

2.3. Sensitivity analyses

We conducted sensitivity analyses related to SUHI intensity, given that numerous approaches exist to estimate SUHI, and that it is influenced by other factors in addition to urban configuration, such as geographical location and climatic conditions (Imhoff et al., 2010). For instance, in semi-arid regions, rural surfaces might be warmer than urban areas, especially when vegetation has not been irrigated due to droughts, restricting the evapotranspiration rate (Marando et al., 2022). As a result, Mediterranean cities might present a lower SUHI than expected. To assess this effect and avoid misleading conclusions, first, we compared the SUHI between Mediterranean (n=206; appendix 1 pp 39–40) and non-Mediterranean (n=713) cities, including biome classification (i.e., Alpine, Arctic, Atlantic, Black Sea, Boreal, Continental, Mediterranean, Pannonian, and Steppic), as defined by the European Environment Agency (2002); and, second, we excluded the Mediterranean cities when evaluating differences in SUHI intensity between the urban configuration types. In addition, we also explored an alternative method to estimate the SUHI intensity— namely, the simplified urban extent algorithm developed by Chakraborty and Lee (2019). In this approach, the SUHI is defined as the LST difference between urban and nonurban land uses within the same city boundary, rather than the LST difference with the adjacent rural area.

3. Results

Of the 919 European cities included in this study, the LCZ built types were more prevalent in the central rings of the analyzed cities and their share decreased towards the peripheral rings, whereas the LCZ natural categories increased towards the outskirts of the cities (appendix 1 pp 4–8). Overall, the open low-rise built type was the most common category among European cities (appendix 1 pp 4–8). The density of motorized roads was slightly higher in the outskirts of the cities, whereas all other road types and cycleways had a

decreasing density from the central rings to the outskirts (appendix 1 pp 9–12). We observed higher traffic volume in the city centers compared with the outskirts (appendix 1 pp 13–14).

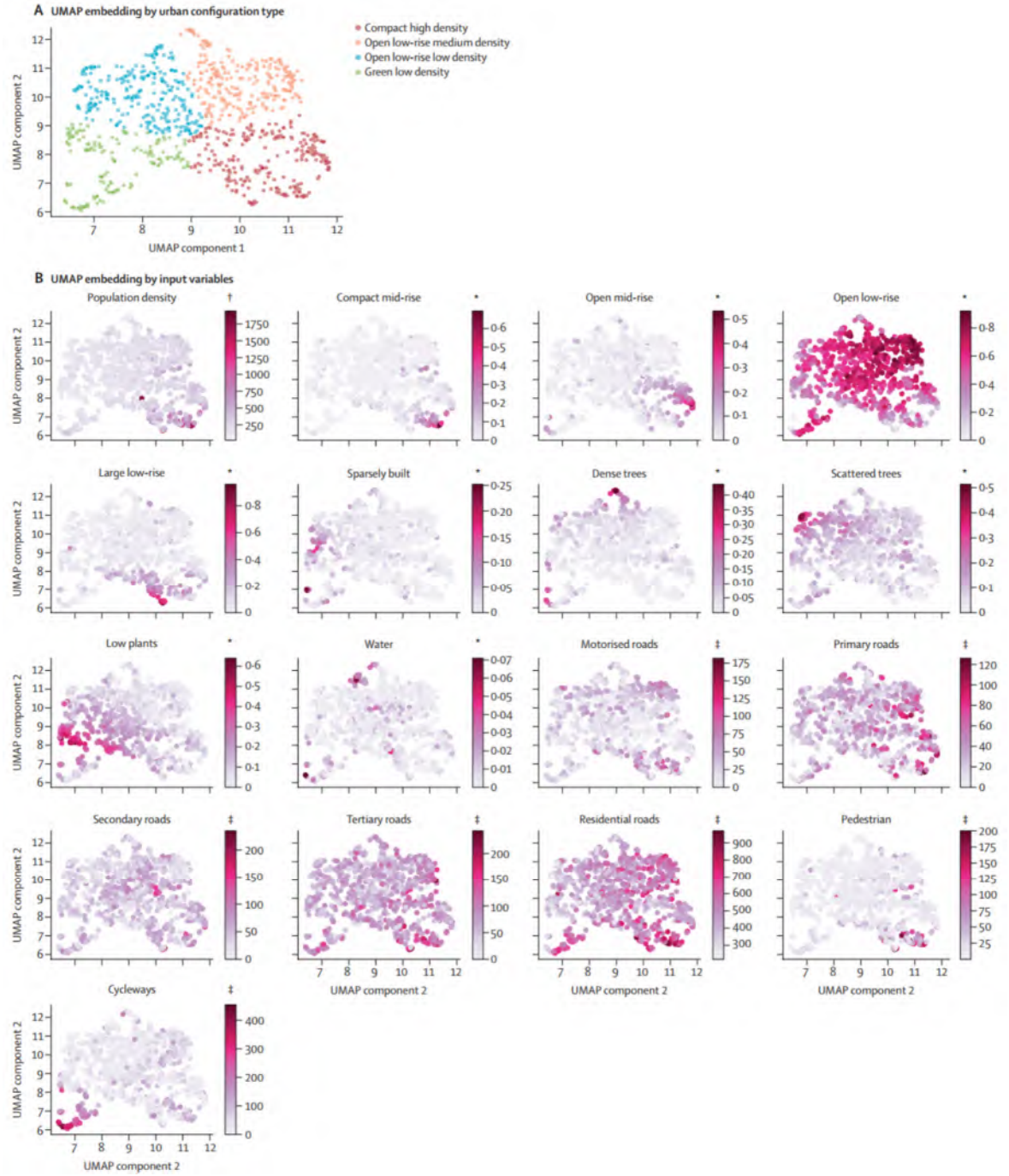


Fig. 1. UMAP embedding coloured by urban configuration type (A) and by each input variable (B) UMAP component 1 and 2 refer to the two relative dimensions to which the UMAP algorithm reduced the data. Each dot in the plot represents a city, and dots closer to each other represent cities that have more similar features in terms of local climate zone and road typologies. UMAP=Uniform Manifold Approximation and Projection for Dimension Reduction. *Percentage. †Mean population per grid cell. ‡Mean meters per grid cell.

Similarly, the SUHI intensity was higher in the city centers, ranging from a mean of 2.1°C to 4.2°C (appendix 1 pp 15–16). Tropospheric NO_2 had slightly higher values in the central rings, ranging from a mean of 53.1 to 56.7 µmol/m² in the periphery of the cities (appendix

1 pp 18–19). Lastly, the CO_2 per person emissions were higher towards the outskirts of the cities, varying between a mean of 0.8 and 1.9 metric tons (appendix 1 pp 20–21). The UMAP embedding showed a continuum without a pronounced segregation of clusters, reflecting similarities in the structural composition of European cities (Fig. 1).

After applying the k-means clustering algorithm, we identified four urban configuration types. The first was compact high-density cities (n=246), which were characterized by small city sizes and high population densities; compact mid-rise, open mid-rise, and large low-rise structures in the central rings; a high density of pedestrian areas; moderate cycleway density; and an overall low availability of natural areas (i.e., dense trees, scattered trees, and low plants; table 1, Fig. 2A).

The second was open low-rise medium density cities (n=245), which were defined by small city sizes and medium population densities, open low-rise configurations prevalent in all rings, higher motorized road density, intermediate pedestrian and cycleway density, and some availability of dense trees and scattered trees towards the city periphery (table 1; Fig. 2B).

The third was open low-rise low-density cities (n=261), which were characterized by medium sizes and low population densities, open low-rise structures prevailing in the city centers, a low density of pedestrian areas and cycleways, and moderate to high availability of natural areas (i.e., scattered trees and low plants) towards the outskirts of the cities (table 1; Fig. 2C).

And fourth was green low-density cities (n=167), defined by large city sizes and low population densities, open low-rise configurations with their share decreasing towards the peripheral rings, moderate pedestrian density, high cycleway density, and high availability of natural areas (i.e., low plants) already starting from the second ring, and increasing in subsequent rings (table 1; Fig. 2D).

Most of the population in all cities resided in compact high density (68 096 496/190 239 499 [35·8%]) and open low-rise medium density (56 108 876 [29·5%]) cities, compared with the open low-rise low density (38 559 619 [20·3%]) and green low density (27 474 508 [14·4%]) city types (table 1).

Traffic volume was significantly higher in compact high-density cities for all rings compared with all other city types. The open low-rise medium density cities had significantly higher traffic volume than open low-rise low density and green low-density cities in all rings (except ring 2), whereas these two types generally had nonsignificant differences between them (table 2; Fig. 3; appendix 1 pp 27–28).

The SUHI intensity was significantly lower among green low-density cities compared with the open low-rise medium density and open low-rise low-density cities in all rings. In contrast, the open low-rise medium density cities had significantly higher SUHI intensity in all rings compared with the compact high-density city type and in rings 2 to 4 compared with the open low-rise low-density city type. Within the compact high-density city type, the proportion of Mediterranean cities was the highest compared with all other city types (89/246

[36·2%]; appendix 1 p 40) and had significantly higher SUHI compared with the green low-density cities in rings 1–3 (table 2; Fig. 3; appendix 1 pp 27–28).

Table 1. Description of urban configuration types. All cities were divided into five rings on the basis of the Burgess concentric urban planning model, with ring 0 being the innermost circle.

	Compact high density (n=246)			Open low-rise medium density (n=245)			Open low-rise low density (n=261)			Green low density (n=167)		
	Mean (SE)	Median	IQR	Mean (SE)	Median	IQR	Mean (SE)	Median	IQR	Mean (SE)	Median	IQR
City area, km²	126 (9·8)	83	41–148	129 (12)	81	43–143	202 (11·6)	148	83–259	381 (41·8)	206	108–352
City population, number	276 815 (28 874)	139 586	73 644–244 439	229 016 (36 089)	135 247	84 426–242 533	147 738 (9330)	106 277	74 024–173 973	164 518 (13 232)	117 287	82 521–176 973
Population density (mean population per grid cell)												
Whole city	421 (19·5)	330	240–450	248 (5·9)	237	193–289	173 (3·8)	172	131–205	179 (6·4)	158	119–221
Ring 0	589 (22·4)	482	354–703	359 (9·7)	324	260–423	280 (6·5)	269	220–312	334 (11·7)	297	236–390
Ring 1	496 (21·3)	391	284–552	275 (7·3)	260	206–322	193 (4·9)	187	144–238	208 (8·4)	188	133–265
Ring 2	419 (20·8)	311	225–469	230 (6·0)	223	178–274	153 (4·4)	152	102–192	143 (7·1)	123	71–195
Ring 3	346 (19·7)	246	172–384	207 (5·9)	203	146–258	129 (4·1)	128	76–169	114 (6·1)	87	52–163
Ring 4	252 (17·5)	174	109–269	170 (5·5)	158	108–221	107 (4·2)	95	59–141	95 (5·5)	72	42–126
Compact mid-rise (%)												
Whole city	8·7 (0·8)	3·7	0·5–11·0	2·1 (0·2)	0·6	0·0–2·7	1 (0·1)	0·1	0·0–1·2	1·9 (0·3)	0·3	0·0–1·7
Ring 0	25·4 (1·7)	15·8	2·5–42·5	8·6 (0·9)	2·5	0·0–11·8	4·3 (0·5)	0·4	0·0–4·8	8·5 (1·2)	1·4	0·0–8·2
Ring 1	9·1 (1·1)	0·4	0·0–9·2	1·3 (0·3)	0·0	0·0–0·1	0·3 (0·1)	0·0	0·0–0·0	0·7 (0·2)	0·0	0·0–0·0
Ring 2	4·3 (0·7)	0·0	0·0–2·1	0·4 (0·1)	0·0	0·0–0·0	0·1 (0·1)	0·0	0·0–0·0	0·2 (0·1)	0·0	0·0–0·0
Ring 3	3·0 (0·6)	0·0	0·0–0·1	0·2 (0·1)	0·0	0·0–0·0	0·1 (0·0)	0·0	0·0–0·0	0·0 (0·0)	0·0	0·0–0·0
Ring 4	1·7 (0·5)	0·0	0·0–0·0	0·1 (0·0)	0·0	0·0–0·0	0·1 (0·0)	0·0	0·0–0·0	0·0 (0·0)	0·0	0·0–0·0
Open mid-rise (%)												
Whole city	10·8 (0·7)	8·9	1·6–17·1	2·7 (0·2)	1·3	0·2–4·0	1·6 (0·1)	0·5	0·1–2·0	1·9 (0·3)	0·3	0·0–1·5
Ring 0	20·8 (1·3)	15·7	0·8–36·5	8·3 (0·7)	3·7	0·3–13·4	5·2 (0·5)	1·5	0·0–6	5·0 (0·7)	0·9	0·0–5·8
Ring 1	15·2 (1·1)	9·1	0·8–25·4	2·5 (0·3)	0·4	0·0–3·2	1·3 (0·2)	0·0	0·0–1·2	2·1 (0·4)	0·0	0·0–0·7
Ring 2	9·8 (0·9)	5·0	0·2–13·2	1·3 (0·2)	0·1	0·0–1·3	0·6 (0·1)	0·0	0·0–0·3	1·1 (0·3)	0·0	0·0–0·1
Ring 3	5·5 (0·6)	1·1	0·0–6·3	0·7 (0·1)	0·0	0·0–0·5	0·4 (0·1)	0·0	0·0–0·0	0·8 (0·3)	0·0	0·0–0·0
Ring 4	2·6 (0·4)	0·0	0·0–2·4	0·4 (0·1)	0·0	0·0–0·1	0·2 (0·1)	0·0	0·0–0·0	0·5 (0·2)	0·0	0·0–0·0
Open low-rise (%)												
Whole city	42·8 (1·2)	46·2	32·6–56·3	69·1 (0·8)	70·6	62·8–72·0	57·2 (0·7)	58·8	51·5–65·3	44·5 (1·4)	46·7	33·3–57·9
Ring 0	25·3 (1·4)	21·3	6·2–39·7	62·3 (1·3)	64·3	50·4–78·2	67·9 (1)	70·8	59·0–80·7	53·7 (2·1)	58·9	30·7–78·8
Ring 1	43·4 (1·6)	44·2	22·0–64·6	75·8 (0·9)	78·0	68·3–86·0	69 (0·9)	71·7	61·2–79·3	56·3 (1·7)	60·2	43·9–74·8
Ring 2	49·9 (1·5)	55·2	34·1–69·3	74·9 (0·9)	77·2	68·0–84·6	59·7 (0·9)	60·9	51·4–70·3	45·7 (1·7)	46·2	30·5–62·9
Ring 3	51·2 (1·4)	57·1	38·9–68	71·3 (0·9)	72·1	64·0–80·9	50·9 (1·1)	52·9	41·8–62·5	37·5 (1·6)	40·5	21·9–51·7
Ring 4	44·3 (1·3)	45·8	30·9–59·8	61·5 (1·2)	63·0	52·6–74·5	38·6 (1·1)	39·7	25·8–50·9	29·2 (1·5)	27·1	13·7–43·2
Large low-rise (%)												
Whole city	17·8 (1·2)	11·0	4·5–26·1	4·7 (0·3)	3·0	1·3–5·5	3·3 (0·3)	2·1	0·8–4·3	6·7 (0·5)	4·1	1·8–10·0
Ring 0	22·9 (1·7)	11·1	3·9–32·4	9·9 (0·8)	5·0	0·9–13·9	8 (0·7)	4·2	1·0–10·8	15·5 (1·5)	7·6	2·9–22·6
Ring 1	21·5 (1·6)	11·9	3·3–31·9	5·3 (0·4)	2·9	0·8–6·3	3·8 (0·4)	2·0	0·3–4·8	8·7 (0·9)	4·6	1·4–11·9
Ring 2	19·3 (1·4)	9·2	3·5–26·9	3·3 (0·3)	1·4	0·3–3·6	2·3 (0·3)	1·0	0·1–2·9	4·4 (0·5)	2·1	0·3–6·0
Ring 3	15·3 (1·3)	7·1	1·5–20·1	2·8 (0·3)	1·1	0·1–2·9	1·4 (0·2)	0·3	0·0–1·6	2·4 (0·3)	1·0	0·0–3·1
Ring 4	9·9 (1·1)	2·5	0·1–11·5	2·1 (0·3)	0·5	0·0–2·1	1·1 (0·2)	0·0	0·0–0·6	2·5 (0·5)	0·3	0·0–2·9
Sparsely built (%)												
Whole city	0·8 (0·1)	0·3	0·0–1·2	1·0 (0·1)	0·4	0·1–1·3	2·8 (0·2)	1·8	0·5–3·9	2·7 (0·3)	1·4	0·4–3·3
Ring 0	0·1 (0·0)	0·0	0·0–0·0	0·3 (0·1)	0·0	0·0–0·1	0·5 (0·1)	0·0	0·0–0·4	0·8 (0·2)	0·0	0·0–0·6
Ring 1	0·3 (0·0)	0·0	0·0–0·1	0·7 (0·1)	0·0	0·0–0·6	1·3 (0·2)	0·4	0·0–1·2	1·9 (0·4)	0·6	0·0–2·0
Ring 2	0·5 (0·1)	0·0	0·0–0·4	0·9 (0·1)	0·2	0·0–0·8	2·3 (0·2)	1·1	0·0–3·5	2·9 (0·4)	0·9	0·0–3·2
Ring 3	0·7 (0·1)	0·0	0·0–0·8	1·3 (0·1)	0·3	0·0–1·5	3·7 (0·3)	1·8	0·4–4·8	3·8 (0·4)	1·6	0·2–5·0
Ring 4	2·5 (0·3)	0·5	0·0–2·7	2·0 (0·2)	0·7	0·1–2·9	6·0 (0·5)	3·6	1·2–8·8	4·0 (0·4)	2·0	0·5–5·7
Dense trees (%)												
Whole city	1·9 (0·2)	0·7	0·0–3·1	5·1 (0·4)	2·7	0·8–6·4	4·6 (0·2)	3·7	1·8–6·7	3·5 (0·4)	1·3	0·1–3·9
Ring 0	0·3 (0·1)	0·0	0·0–0·3	1·9 (0·2)	0·5	0·0–1·9	2·0 (0·2)	0·8	0·1–2·5	1·4 (0·2)	0·1	0·0–1·1
Ring 1	0·9 (0·1)	0·0	0·0–1·2	4·1 (0·5)	1·4	0·2–4·7	3·8 (0·3)	2·4	0·7–5·1	3·0 (0·5)	0·6	0·0–2·9
Ring 2	1·7 (0·2)	0·1	0·0–2·4	5·4 (0·5)	2·1	0·5–6·8	4·8 (0·3)	3·6	1·2–6·9	3·8 (0·5)	0·9	0·0–4·3
Ring 3	2·7 (0·3)	0·5	0·0–4·0	6·1 (0·5)	2·9	1·0–7·9	5·6 (0·3)	4·2	1·8–8·3	4·5 (0·6)	1·6	0·0–4·9
Ring 4	3·8 (0·4)	1·1	0·0–6·1	8·1 (0·7)	4·0	1·0–10·4	7·1 (0·4)	5·3	2·0–10·0	4·6 (0·6)	1·5	0·1–5·2
Scattered trees (%)												
Whole city	5·6 (0·3)	4·6	1·8–8·2	8·1 (0·3)	7·6	5·2–10·3	14·1 (0·5)	12·3	8·2–17·6	7·6 (0·4)	7·4	4·5–10·6
Ring 0	1·8 (0·1)	0·8	0·0–2·7	3·5 (0·2)	3·0	1·1–5·0	5·9 (0·3)	5·0	2·6–8·8	3·1 (0·2)	2·2	0·8–4·5
Ring 1	4·0 (0·3)	2·7	0·6–6·1	6·4 (0·3)	5·8	3·4–8·9	12·1 (0·5)	10·4	6·7–15·1	7·2 (0·4)	6·0	2·9–9·9
Ring 2	5·7 (0·4)	4·4	1·3–9·0	8·5 (0·4)	7·7	4·3–11·1	15·4 (0·7)	12·8	8·5–19·8	8·5 (0·5)	8·2	3·8–11·8
Ring 3	6·8 (0·4)	4·9	1·9–10·2	9·9 (0·4)	9·0	5·6–13·1	17·5 (0·8)	14·7	8·5–23·2	9·1 (0·5)	8·3	3·6–13·2
Ring 4	9·7 (0·6)	6·8	2·9–13·9	12·3 (0·6)	10·2	6·4–16·2	19·7 (0·9)	16·5	9·4–25·9	10·1 (0·6)	8·4	3·8–13·8
Low plants (%)												
Whole city	9·8 (0·4)	9·4	3·8–14·7	5·3 (0·3)	4·7	2·2–8	13·5 (0·5)	13·8	8·1–18·7	27·4 (1·1)	27·8	17·0–36·9
Ring 0	0·8 (0·1)	0·1	0·0–0·9	0·7 (0·1)	0·1	0·0–0·8	2·2 (0·2)	1·0	0·2–2·6	3·9 (0·4)	1·9	0·5–5·4
Ring 1	3·7 (0·3)	2·0	0·4–5·2	2·2 (0·2)	1·0	0·3–3·1	7·2 (0·4)	5·1	2·4–9·9	17·1 (1·2)	13·9	5·3–25·8
Ring 2	7·3 (0·4)	5·5	1·6–11·5	4·4 (0·3)	2·7	1·0–6·9	13·8 (0·6)	12·3	5·7–19·6	31·1 (1·6)	30·5	13·3–46·1
Ring 3	13·2 (0·7)	11·1	3·8–20·1	6·8 (0·4)	5·1	1·7–10·7	19·0 (0·7)	17·8	9·9–26·9	39·3 (1·7)	40·2	21·9–53·9
Ring 4	23·7 (1·1)	22·0	9·1–36	12·5 (0·7)	10·2	3·5–19·4	25·3 (1·0)	24·5	13·2–35·4	45·6 (1·6)	47·3	30·0–61·7
Water (%)												
Whole city	0·5 (0·0)	0·2	0·0–0·8	0·5 (0·0)	0·3	0·1–0·6	0·7 (0·1)	0·3	0·1–0·7	0·7 (0·1)	0·4	0·1–1·0
Ring 0	0·7 (0·1)	0·0	0·0–0·9	0·7 (0·1)	0·2	0·0–0·8	1·0 (0·1)	0·2	0·0–0·8	0·9 (0·1)	0·3	0·0–1·2
Ring 1	0·5 (0·0)	0·1	0·0–0·7	0·5 (0)	0·2	0·0–0·6	0·7 (0·1)	0·2	0·0–0·7	0·6 (0·1)	0·2	0·0–0·7
Ring 2	0·4 (0·0)	0·0	0·0–0·5	0·4 (0·1)	0·1	0·0–0·4	0·5 (0·1)	0·1	0·0–0·5	0·6 (0·1)	0·2	0·0–0·7
Ring 3	0·4 (0·0)	0·1	0·0–0·6	0·4 (0)	0·1	0·0–0·5	0·6 (0·1)	0·1	0·0–0·6	0·7 (0·1)	0·2	0·0–0·9
Ring 4	0·5 (0·1)	0·1	0·0–0·7	0·5 (0·1)	0·1	0·0–0·7	0·9 (0·1)	0·2	0·0–0·7	0·9 (0·1)	0·3	0·0–1·2

(Table 1 continues on next page)

Table 1. *Continued* Description of urban configuration types.

	Compact high density (n=246)			Open low-rise medium density (n=245)			Open low-rise low density (n=261)			Green low density (n=167)		
	Mean (SE)	Median	IQR	Mean (SE)	Median	IQR	Mean (SE)	Median	IQR	Mean (SE)	Median	IQR
(Continued from previous page)												
Motorised roads (mean metres per grid cell)												
Whole city	25 (1·7)	17	3–37	30 (1·4)	29	11–42	22 (1·1)	20	8–33	21 (1·3)	20	9–29
Ring 0	11 (1·6)	0	0–8	30 (2·3)	15	0–53	24 (1·9)	10	0–37	19 (2·1)	3	0–30
Ring 1	22 (2·3)	3	0–31	31 (2·1)	24	1–47	29 (1·8)	26	0–43	27 (2·2)	20	1–40
Ring 2	30 (2·7)	15	0–42	30 (1·7)	26	7–44	24 (1·4)	19	3–38	26 (2·0)	22	7–36
Ring 3	32 (3·1)	16	0–41	30 (1·8)	24	10–44	19 (1·2)	16	3–29	18 (1·4)	15	4–25
Ring 4	29 (3·1)	14	0–38	30 (2·0)	23	6–43	16 (1·2)	12	0–24	18 (1·5)	12	5–25
Primary roads (mean metres per grid cell)												
Whole city	36 (1·7)	33	17–51	34 (1·3)	33	19–47	29 (1·0)	29	16–40	25 (1·2)	25	13–36
Ring 0	50 (3·1)	42	0–77	53 (2·5)	50	22–77	46 (2·1)	49	15–72	41 (2·9)	36	5–67
Ring 1	43 (2·5)	38	12–62	38 (1·9)	31	18–51	30 (1·3)	27	16–42	29 (1·9)	27	7–41
Ring 2	35 (2·0)	28	12–53	31 (1·8)	25	12–47	24 (1·0)	23	11–35	23 (1·6)	21	6–34
Ring 3	31 (1·8)	24	8–48	25 (1·4)	22	8–35	22 (1·1)	18	9–31	18 (1·3)	15	4–27
Ring 4	23 (1·6)	17	2–34	24 (1·5)	21	5–38	23 (1·4)	18	6–31	15 (1·2)	12	2–24
Secondary roads (mean metres per grid cell)												
Whole city	50 (2·1)	48	32–64	48 (1·7)	47	28–63	40 (1·3)	38	25–52	45 (1·6)	44	32–57
Ring 0	76 (3·4)	76	36–105	72 (3·2)	68	35–101	58 (2·4)	52	28–86	77 (2·9)	80	56–101
Ring 1	58 (2·6)	52	29–80	50 (2·2)	46	23–68	40 (1·7)	36	21–53	49 (2·4)	44	25–68
Ring 2	49 (2·8)	39	20–65	43 (1·9)	40	21–58	37 (1·5)	34	19–50	36 (2·0)	33	19–48
Ring 3	38 (2·2)	33	12–51	40 (1·7)	38	21–55	33 (1·4)	29	17–47	32 (1·7)	28	18–45
Ring 4	31 (2·1)	25	8–43	36 (1·6)	34	16–51	33 (1·5)	30	13–46	32 (1·7)	30	15–47
Tertiary roads (mean metres per grid cell)												
Whole city	89 (2·6)	83	64–105	72 (2·0)	70	51–88	59 (1·3)	59	46–72	75 (2·4)	74	52–91
Ring 0	128 (4·0)	119	88–164	90 (2·9)	87	61–114	79 (2·5)	73	53–103	97 (3·5)	96	70–122
Ring 1	101 (3·3)	94	66–127	75 (2·6)	71	48–98	59 (1·7)	57	38–76	76 (2·9)	75	51–97
Ring 2	83 (3·1)	74	52–104	69 (2·5)	66	43–90	53 (1·7)	50	35–70	69 (2·9)	67	45–89
Ring 3	70 (3·3)	64	39–84	67 (2·3)	63	42–88	54 (1·6)	53	37–71	66 (2·8)	64	38–91
Ring 4	62 (2·8)	58	33–82	61 (2·2)	58	37–81	52 (1·6)	51	33–68	65 (2·9)	65	34–90
Residential roads (mean metres per grid cell)												
Whole city	535 (9·6)	498	419–635	520 (6·4)	512	445–589	439 (5·0)	437	381–492	463 (10·0)	442	352–559
Ring 0	681 (14·3)	649	498–829	613 (7·9)	590	525–682	570 (7·9)	551	486–633	669 (12·8)	681	541–792
Ring 1	592 (11·8)	568	452–706	536 (7·6)	523	453–604	476 (7·1)	465	403–542	515 (12·3)	496	393–646
Ring 2	538 (11·0)	499	417–637	504 (7·4)	505	422–583	423 (6·8)	407	348–487	421 (12·5)	386	303–542
Ring 3	474 (11·4)	428	345–557	497 (7·4)	493	415–574	386 (6·0)	383	321–448	379 (12·5)	341	261–483
Ring 4	399 (10·1)	374	284–485	449 (8·6)	436	350–534	346 (6·4)	339	275–405	339 (11·8)	303	242–397
Pedestrian (mean metres per grid cell)												
Whole city	33 (2·7)	16	9–40	14 (1·1)	10	5–15	9 (0·7)	6	4–11	15 (1·1)	10	5–20
Ring 0	73 (4·4)	50	27–98	38 (2·4)	27	16–46	28 (1·7)	20	11–34	41 (2·9)	29	16–56
Ring 1	36 (3·6)	13	4–33	12 (1·3)	5	1–13	7 (1·0)	3	1–7	14 (1·4)	7	2–17
Ring 2	27 (3·1)	8	2–28	9 (0·9)	4	1–10	4 (0·5)	2	0–5	9 (1·0)	3	0–12
Ring 3	20 (2·6)	4	1–16	7 (0·9)	3	1–8	3 (0·3)	1	0–4	6 (0·8)	2	0–7
Ring 4	12 (1·7)	2	0–9	5 (0·7)	2	0–6	3 (0·7)	1	0–2	6 (0·8)	1	0–7
Cycleways (mean metres per grid cell)												
Whole city	49 (2·8)	36	15–75	38 (2·2)	29	12–53	22 (1·2)	17	8–30	118 (8·8)	67	26–194
Ring 0	69 (4·1)	51	19–106	51 (3·1)	37	15–72	39 (2·5)	26	13–50	173 (10·9)	124	52–277
Ring 1	63 (3·9)	45	17–89	41 (2·7)	30	9–58	27 (1·6)	18	7–43	137 (9·5)	92	33–224
Ring 2	50 (3·4)	33	10–70	38 (2·7)	24	8–53	20 (1·4)	12	4–32	111 (9·4)	48	17–205
Ring 3	38 (3·2)	21	4–54	35 (2·5)	25	6–49	15 (1·1)	11	1–21	96 (9·4)	24	7–167
Ring 4	26 (2·4)	13	0–34	26 (1·9)	19	3–39	11 (0·9)	5	1–16	73 (7·9)	15	2–119

For NO_2, we observed that compact high density and open low-rise medium density cities had significantly higher exposure than green low density and open low-rise low-density cities in all rings (table 2; Fig. 3; appendix 1 pp 29–30). The CO_2 per person emissions were significantly lower among cities within the compact high-density type, particularly in rings 0–3. In contrast, the green low-density cities had significantly higher CO_2 per person emissions, specifically in rings 0–3 compared with the compact high-density cities and in rings 1–2 compared with open low-rise medium density and open low-rise low-density cities (table 2; Fig. 3; appendix 1 pp 29–30). The analyses with alternative data sources for the NO_2 exposure and CO_2 emissions validated these trends (appendix 1 pp 31–36).

In the sensitivity analyses, we observed that the non-Mediterranean cities had an overall higher SUHI than Mediterranean cities in all rings ($p<0·001$; appendix 1 pp 39–40). The exclusion of Mediterranean cities accentuated the differences between the compact high density and green low-density cities, in that the compact high-density cities had a significantly higher SUHI than green low-density cities in rings 0–3 (appendix 1 pp 41–43). In the analysis with the simplified urban extent algorithm, we observed overall lower SUHI

estimates compared with the main analysis in all rings (appendix 1 p 45). In this analysis, the green low-density type had significantly lower SUHI in rings 1–3 compared with the compact high-density cities and in all rings (except ring 1) compared with the open low-rise medium density type (appendix 1 pp 46–49).

Finally, concerning the age-standardized mortality rates, we observed a slightly decreasing trend from the compact high-density cities (with an average of 1 124 deaths per 100 000 inhabitants) towards the green low-density cities (with an average of 1 003 deaths per 100 000 inhabitants; table 2; Fig. 4). Green low-density cities had significantly lower mortality rates compared with all other urban configuration types ($p<0{\cdot}001$; appendix 1 p 50).

Table 2. Traffic volume, exposure levels, CO_2 per person emissions, and natural cause mortality by urban configuration type; All cities were divided into five rings on the basis of the Burgess concentric urban planning model, with ring 0 being the innermost circle. SUHI=surface urban heat island.

	Compact high density		Open low-rise medium density		Open low-rise low density		Green low density	
	Mean (SE)	95% CI	Mean (SE)	95% CI	Mean (SE)	95% CI	Mean (SE)	95% CI
Traffic volume (vehicles per day per km)								
City	808 (36)	738–878	495 (18)	460–530	314 (14)	286–342	316 (15)	286–345
Ring 0	930 (82)	768–1093	597 (42)	514–680	397 (37)	324–469	421 (44)	335–508
Ring 1	853 (78)	700–1007	520 (44)	433–607	358 (41)	277–439	345 (37)	273–418
Ring 2	853 (86)	683–1023	454 (33)	388–520	306 (31)	245–366	313 (29)	255–370
Ring 3	748 (84)	582–914	463 (42)	380–545	256 (23)	211–301	255 (28)	199–310
Ring 4	654 (65)	526–782	441 (39)	364–519	253 (25)	203–303	244 (27)	192–296
SUHI (°C)								
City	2·7 (0·1)	2·6–2·9	3·6 (0·1)	3·5–3·7	3·2 (0·1)	3·1–3·3	2·3 (0·1)	2·2–2·5
Ring 0	3·7 (0·2)	3·2–4·1	4·7 (0·2)	4·4–5·0	4·7 (0·2)	4·4–5·0	3·5 (0·2)	3·1–4·0
Ring 1	3·1 (0·2)	2·7–3·5	3·9 (0·1)	3·6–4·2	3·6 (0·1)	3·3–3·9	2·7 (0·2)	2·4–3·1
Ring 2	2·8 (0·2)	2·4–3·1	3·5 (0·1)	3·2–3·8	3·1 (0·1)	2·8–3·3	2·2 (0·2)	1·9–2·5
Ring 3	2·3 (0·2)	2·0–2·7	3·2 (0·1)	3·0–3·5	2·6 (0·1)	2·4–2·8	1·8 (0·1)	1·5–2·1
Ring 4	1·8 (0·1)	1·5–2·1	2·7 (0·1)	2·5–3·0	2·2 (0·1)	2·0–2·4	1·5 (0·1)	1·2–1·8
Tropospheric NO_2 (μmoL/m²)								
City	59·6 (0·9)	57·9–61·3	59·6 (0·7)	58·1–61·0	50·8 (0·6)	49·6–51·9	52·0 (0·9)	50·3–53·7
Ring 0	61·4 (2·0)	57·4–65·4	61·2 (1·6)	58·0–64·4	52·2 (1·4)	49·5–55·0	53·4 (1·9)	49·6–57·2
Ring 1	60·7 (2·0)	56·8–64·6	60·5 (1·6)	57·3–63·7	51·6 (1·4)	48·9–54·3	52·7 (1·9)	48·9–56·6
Ring 2	59·9 (2·0)	56·1–63·8	59·8 (1·6)	56·6–62·9	50·9 (1·4)	48·2–53·6	52·1 (1·9)	48·3–55·9
Ring 3	59 (1·9)	55·1–62·8	58·8 (1·6)	55·7–62·0	50·2 (1·3)	47·6–52·9	51·4 (1·9)	47·6–55·2
Ring 4	57 (1·9)	53·3–60·8	57·5 (1·6)	54·3–60·6	48·8 (1·3)	46·2–51·4	50·2 (1·9)	46·4–54·1
CO_2 per person emissions (metric tons)								
City	1·3 (0·1)	1·2–1·4	1·4 (0·0)	1·3–1·5	1·7 (0·0)	1·6–1·7	1·9 (0·1)	1·8–2·0
Ring 0	0·5 (0·0)	0·4–0·5	0·8 (0·0)	0·7–0·9	0·9 (0·0)	0·8–1·0	0·9 (0·1)	0·8–1·0
Ring 1	0·9 (0·1)	0·8–1·0	1·3 (0·1)	1·1–1·4	1·6 (0·1)	1·5–1·7	1·9 (0·1)	1·7–2·1
Ring 2	1·3 (0·1)	1·1–1·5	1·5 (0·1)	1·3–1·7	1·9 (0·1)	1·8–2·1	2·5 (0·2)	2·2–2·9
Ring 3	1·9 (0·2)	1·5–2·2	1·7 (0·1)	1·5–1·8	2·0 (0·1)	1·8–2·2	2·4 (0·2)	2·0–2·7
Ring 4	1·9 (0·1)	1·7–2·2	1·8 (0·1)	1·6–2·0	1·8 (0·1)	1·6–2·0	1·9 (0·1)	1·7–2·2
City age-standardised natural-cause mortality rate (per 100 000 inhabitants)	1124 (17)	1091–1157	1093 (14)	1065–1121	1091 (15)	1062–1120	1003 (13)	977–1029

All cities were divided into five rings on the basis of the Burgess concentric urban planning model, with ring 0 being the innermost circle. SUHI=surface urban heat island.

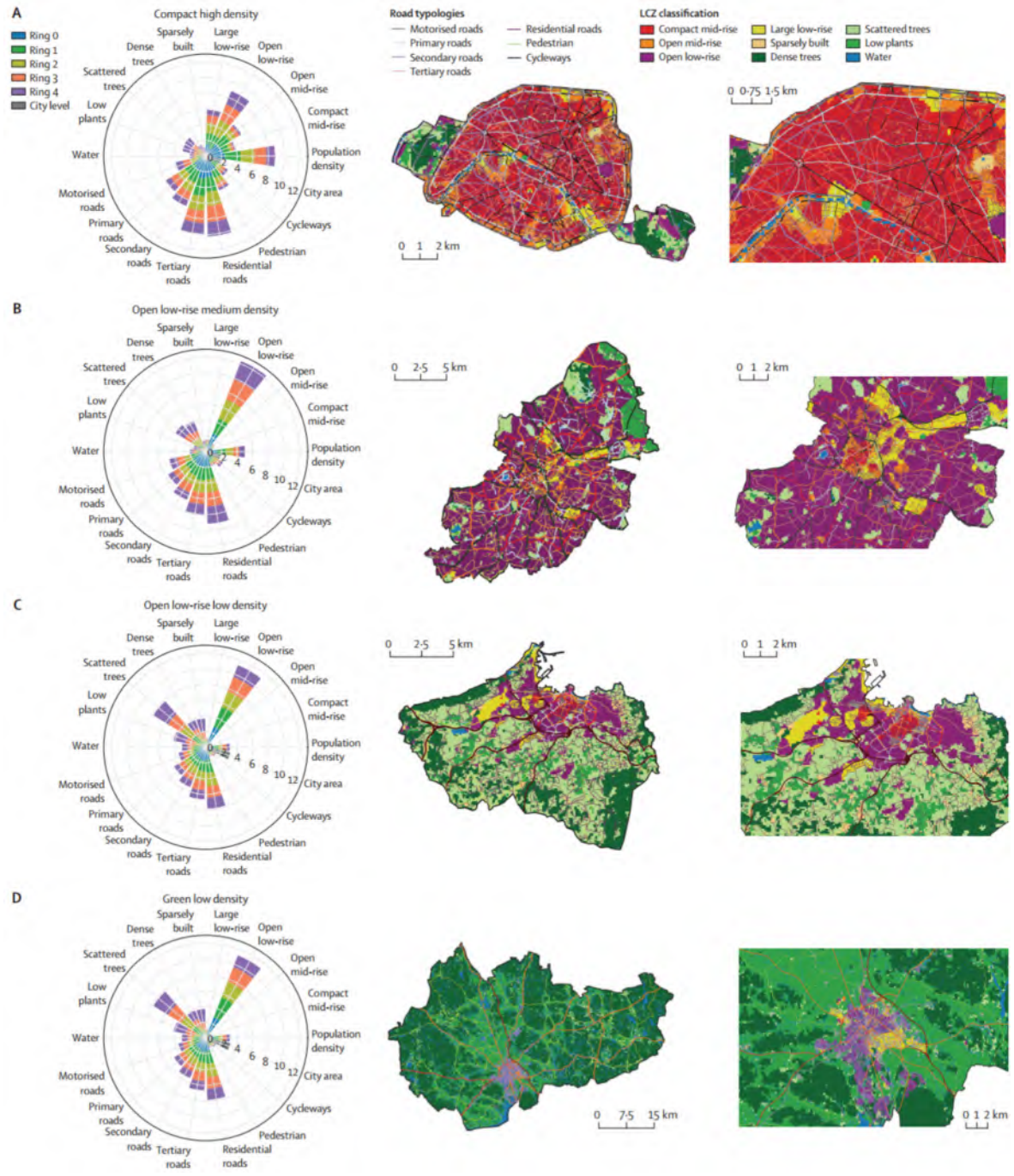


Fig. 2. Overview of urban configuration types (A) Compact high density urban configuration. An example is shown for the city of Paris, France. (B) Open low-rise medium density urban configuration. An example is shown for the city of Birmingham, UK. (C) Open low-rise low-density urban configuration. An example is shown for the city of Gijón, Spain. (D) Green low-density urban configuration. An example is shown for the city of Uppsala, Sweden. LCZ=local climate zones.

4. Discussion

In this study, we integrated information on intraurban morphological configurations and street design and identified four distinct urban configuration types among European cities. Our findings indicated that compact, higher density cities had increased NO_2 and SUHI effects, along with decreased CO_2 per person emissions. Conversely, greener, lower density cities had lower SUHI intensities, NO_2, and mortality rates, but higher CO_2 per person

emissions. The findings of this study hold significant implications for urban and transport planning policies in Europe and contribute to the ongoing discussion regarding which city models can bring the greatest benefits for the environment, climate, and health.

In recent literature (Nieuwenhuijsen, 2020), the compact city model has arisen as the optimal theoretical model to promote healthier and more sustainable cities. The combination of destination accessibility, uniform distribution of employment, management of traffic demand, development of pedestrian and cycling networks, optimum residential density, reduction in distance to public transport, and promotion of active travel modes are all features of compact cities with positive effects for the environment, lifestyle, and social welfare (Nieuwenhuijsen, 2020). A set of previous observational studies have shown the environmental and health benefits of compact cities. A study of 53 metropolitan regions in the USA found that the rate of increase in extreme heat events was lower among more compact cities compared with sprawling cities (Stone, Hess & Frumkin, 2010). In addition, a review on urban form and air pollution and a study with 1 274 global cities indicated that compactness can result in improved air quality (Hankey, Marshall, 2017; Bechle, Millet & Marshall, 2017). Furthermore, a health impact assessment study of six global cities (Melbourne, Boston, London, São Paulo, New Delhi, and Copenhagen) found that there were health gains when cities transitioned to a more compact urban structure, associated with increases in active travel modes (Stevenson et al., 2016).

However, previous research has also outlined the negative effects of compact cities. Some observational studies have indicated that denser and more compact cities have higher air pollution and UHI intensities, and reduced access to green spaces (Rezaei & Millard-Ball, 2023; Zhou, Rybski & Kropp, 2017). In addition, highly densified urban environments can contribute to the formation of pollution hotspots, with increased population exposure and health risks (Hankey & Marshall, 2017; Bechle, Millet, Marshall, 2017).

In this analysis, we found that compact cities in Europe tend to have worse air quality, lower green space availability, and increased SUHI compared with lower density cities (table 3). In addition, we observed a decreasing mortality trend from compact to lower density cities, which could be partly linked to the higher adverse environmental exposure among compact cities (appendix 1 pp 51–54) (Khomenko et al., 2021; Barboza et al., 2021). Our findings align with previous literature that highlight the adverse effects of compact urban configurations. However, it is important to approach our results with care and refrain from drawing direct conclusions, such as avoiding compact cities, because there are various inter-related factors at play that require further examination and understanding.

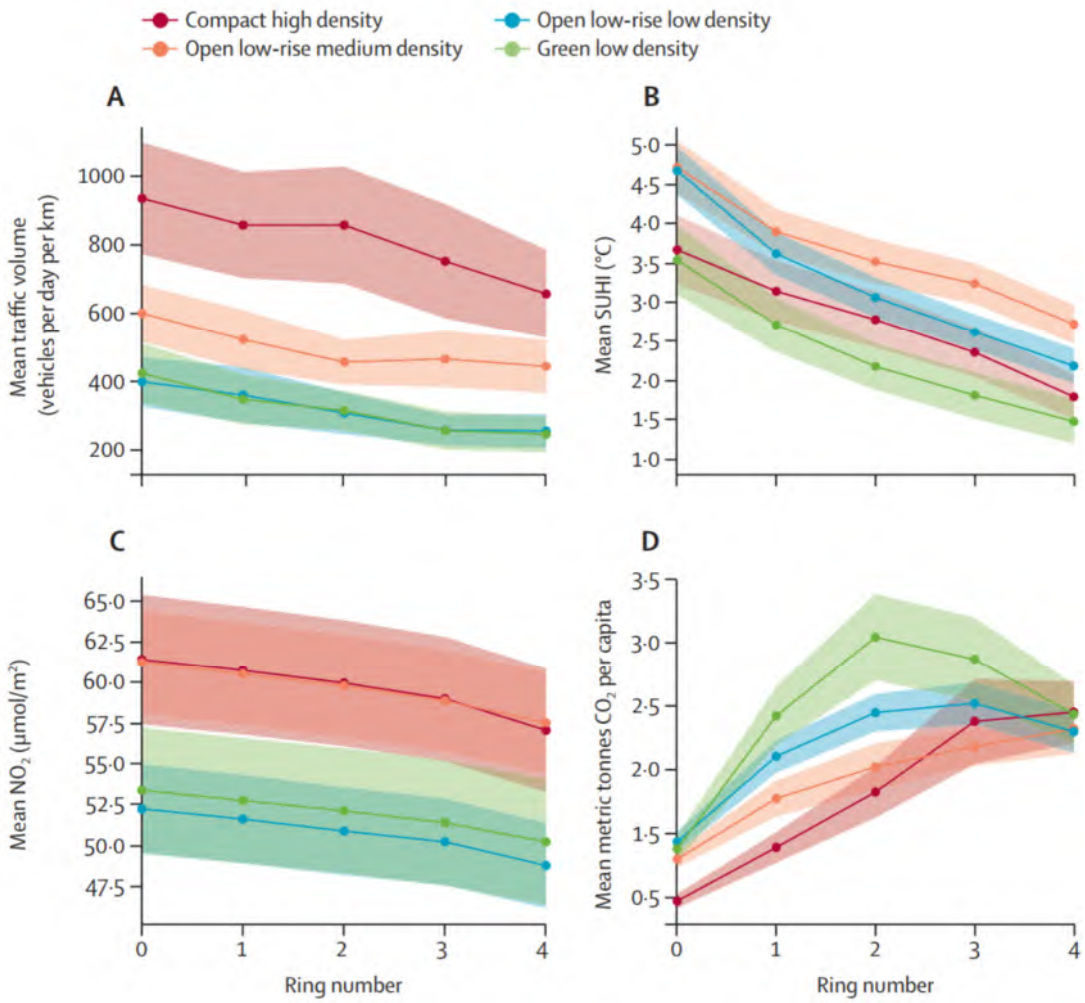


Fig. 3. Motorized traffic flows, SUHI intensity, NO_2 exposure, and CO_2 per person emissions by urban configuration type. The mean and 95% CIs are shown. SUHI=surface urban heat island.

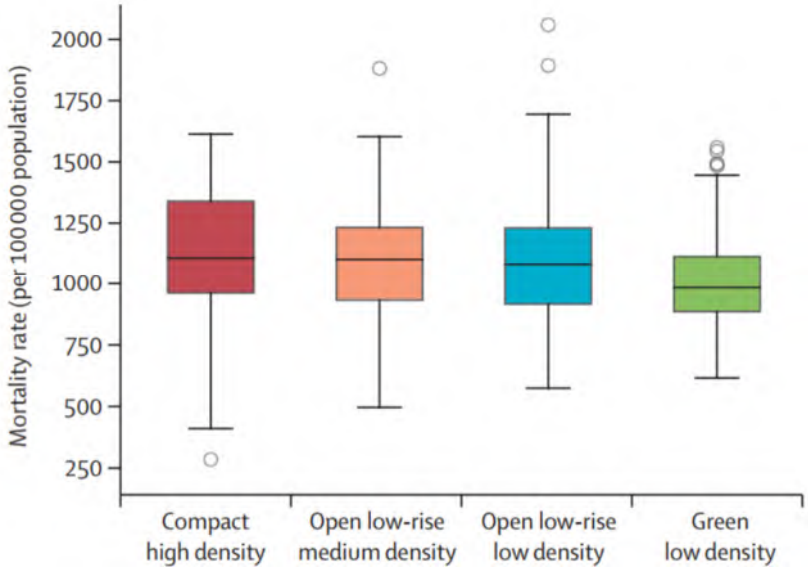


Fig. 4. Age-standardised natural-cause mortality rates by urban configuration type.

Compact cities have many conceptual benefits in terms of reduced car dependency, walkability, proximal access to services, increased social interactions, and an increased sense of community (Nieuwenhuijsen, 2020). Yet, in their current configuration, they are highly densified and probably act as economic hubs that attract large numbers of commuters for education or employment opportunities. In addition, mobility within compact cities is still, to a large extent, reliant on motorized transport modes. As revealed by our analysis, traffic volume is considerably higher within the compact city type compared with all other urban configurations, which is also in line with the higher NO_2 observed among these cities. In this sense, current European compact cities should be understood as being in a transition state towards an optimal form, a state that currently combines some positive features (such as proximity to services) with challenges such as densification, mobility, management of traffic flows, poor air quality, and insufficient green and natural spaces.

Several measures can help to optimize compact cities. Regarding density, a medium to high dwelling density ranging from 45 to 175 dwellings per hectare has been suggested as the optimal solution for providing enough density to create an optimal range of services and avoid social isolation and creating a sense of community (Mueller et al., 2021). However, in cities that are already dense, other measures such as the diversification of local destinations and proximity to public transport might be more relevant (Stevenson et al., 2016). The promotion of active transportation modes is necessary to cover short-distance (generally less than 5 km) mobility needs (Nieuwenhuijsen, 2021). Nonetheless, the development of proximal, frequent, and high-quality public transportation is essential to supply the demand for longer distance trips, which still largely rely on private motorized modes, particularly for commuters from suburban areas and nearby towns (Giles-Corti et al., 2016).

Notably, compact cities can be leveraged for implementing new urban models that promote the creation and better use of public spaces for citizens (Nieuwenhuijsen, 2021). These approaches include the implementation of superblocks in various parts of Barcelona, Spain, the development of low-traffic neighborhoods in London, and the creation of car-free neighborhoods in Hamburg and Freiburg, with potential positive effects for the environment and health (Nieuwenhuijsen, 2021).

Table 3. Summary characteristics of the European urban configuration types. Mean values by city and cluster are shown. SUHI=surface urban heat island. *Sum of percentage of area of dense trees, scattered trees, and low plants. †Based on the results from the sensitivity analysis excluding Mediterranean cities.

	Urban configuration					Environmental effect			CO_2 per person, metric tons	Mortality, deaths per 100 000 inhabitants
	Compactness, percentage compact mid-rise	Density, mean population per grid cell	Pedestrian areas, mean m per grid cell	Cycleways, mean m per grid cell	Green space*	Traffic density, vehicles per day per km	NO_2, µmol/m²	SUHI†, °C		
Compact high density	8·7%	421	33	49	17·3%	808	59·6	3·7	1·3	1124
Open low-rise medium density	2·1%	248	14	38	18·5%	495	59·6	3·8	1·4	1093
Open low-rise low density	1·0%	173	9	22	32·2%	314	50·8	3·5	1·7	1091
Green low density	1·9%	179	15	118	38·5%	316	52·0	3·0	1·9	1003

In addition, there is a need for strategies to effectively integrate green spaces in compact urban areas. Nature-based solutions can take a key role in this regard, not only enhancing

green infrastructure, but also mitigating the effects of climate change. A study by Hansen & Pauleit (2014) proposed the functional and physical integration of green spaces into the planning of grey infrastructure. Green roofs or façade greening can be introduced in highly sealed commercial areas to reduce surface water runoff (Hansen & Pauleit, 2014). Sky gardens can compensate for the absence of ground-level green spaces; street trees mitigate CO_2 emissions and improve the microclimate; and the introduction of small gardens in public, industrial, educational, and institutional spaces can contribute to supporting biodiversity as part of areas of green spaces in the city (Hansen & Pauleit, 2014). Nevertheless, the increased proximity and availability of urban green spaces might lead to higher property values and the displacement of less affluent residents (Anguelovski et al., 2022). Hence, it is crucial for urban planners to monitor the environmental quality, accessibility, and equitable distribution of these spaces among the population to promote environmental and climate justice.

However, our research indicated that compact cities tended to have lower CO_2 per person emissions. This finding is consistent with a previous study involving 53 European cities, which revealed that urban sprawl, characterized by low-density urban fabric and long distances to the city center, was linked to increased greenhouse gas per person emissions (Baur, Förster & Kleinschmit, 2015). High urban complexity, irregularity, and fragmentation were identified as key factors contributing to higher carbon emissions, primarily because of the absence of continuity and connectivity in sprawling cities (Hong, Hui & Lin, 2022). Because compact cities are denser and offer shorter travel distances, these have the potential to be more energy efficient and allow commuting and carrying out daily activities using low emitting transport modes. Accordingly, an analysis of 111 Italian cities found that smaller, more compact, and less monocentric cities have reduced levels of CO_2 emissions per commuter, whereas another study in the greater Dublin region, Ireland, estimated that the compact city scenario can lead to reduced household energy consumption and energy-related CO_2 emissions (Liu & Sweeney, 2012; Cirilli & Veneri, 2013). As such, compact cities have the highest harmful environmental exposures compared with all other city types; however, their compactness plays a pivotal role in facilitating low-carbon urban environments.

Lastly, our results highlighted that large and less dense cities are generally greener and perform better in terms of environmental exposures (table 3). However, as mentioned earlier, sprawled cities also face challenges, particularly related to larger travel distances, higher CO_2 emissions, car dependency, availability of proximal and diverse destinations, and social isolation (Nieuwenhuijsen, 2020). In these cases, measures such as densification, the provision of services (e.g., education, commerce, recreation, and access to jobs), and an adequate mix of buildings with residential and non-residential functions can help in creating more dense and diverse neighborhoods with cultural, social, and employment opportunities. In addition, not all cities require an equal transformation into a compact form. Although the aforementioned factors might be applicable to new urbanizations or large sprawled and growing cities, they might not be so relevant to small cities with low population densities.

In this study, we discuss some general solutions; however, the development and implementation of the compact city strategies depend on the specific physical, geographical,

sociopolitical, economic, and historical aspects of each city. A set of tools and indicators have been developed to ensure more sustainable and healthier urban and transport developments (Mueller et al., 2021; Lowe et al., 2022) and could aid cities to prioritize measures relevant for their local context. Cities need to be understood as complex systems with multiple inter-related dimensions. To address their challenges, comprehensive and integrated solutions are needed that encompass multiple domains including urban and transport planning, housing, education, employment, climate mitigation, and efficient resource use. To achieve this, collaboration among diverse experts in multidisciplinary groups is crucial, as well as the active involvement and commitment of stakeholders through participatory processes.

The main strengths of this study are its innovative analysis, which integrates urban configuration, environmental and climate effects, and health domains; the large number of European cities, high resolution, and the consistency of data; the consideration of spatial variability in all variables; the use of the UMAP algorithm; the incorporation of satellite NO_2 data, which does not depend on any modelling assumptions; and the sensitivity analyses for the definition of SUHI intensity.

Nevertheless, two main limitations need to be addressed. First, we defined European cities on the basis of administrative boundaries, whereas other definitions such as the Morphological Urban Areas definition (Taubenböck et al., 2019) might provide a more accurate view of urban agglomerations, albeit less linked to policy actions. Second, the data availability and definition of detailed indicators were limited for many European cities. For instance, because of inconsistencies in the OSM database, we could not include the connectivity of the general road network or the pedestrian and cycling networks. Similarly, we could not reflect the accessibility to public transportation, natural areas, or services such as food markets or convenience stores. However, these indicators have been proven to be relevant for sustainability and health, and should be considered in future studies (Boeing et al., 2022).

Furthermore, several methodological considerations exist. First, there are multiple methods for estimating the SUHI that, as shown in our analysis, can lead to considerable variability in the results. In this study, we estimated the SUHI as the LST difference between urban and rural areas surrounding each city boundary. Although this is the most accepted method for SUHI estimation, it can also lead to misleading interpretations depending on climatic regions, biome, abundance and diversity of vegetation, irrigation patterns, and droughts, particularly for semi-arid cities (Marando et al., 2022). Given this issue, we conducted a sensitivity analysis with an alternative SUHI definition that takes the non-urban areas within each city boundary as the LST reference. This definition eliminates the influence of extrinsic factors that might affect rural LST. However, given that this definition takes green urban areas as a reference to calculate SUHI, which is artificially originated, it is limited in providing an accurate measure of the excess heat resulting from the anthropogenic modification of natural landscapes and, accordingly, results in lower SUHI estimates. Additionally, differences between UHI and SUHI need to be acknowledged. In this study, the SUHI proxy was used because of limited UHI data availability for many European cities. However, the association between the SUHI and UHI might vary according to other factors

(i.e., elevation and terrain roughness) in a non-linear manner (Leichtle et al., 2023; Mutiibwa, Strachan & Albright, 2015).

Second, we worked with the best available NO_2 dataset derived from satellite observations without any chemical transport modelling. However, it should be noted that satellite-derived NO_2 does not directly represent pollutant concentrations at the ground level. Despite this, it has been shown to constitute a good proxy and a reliable tracer of anthropogenic emissions (Müller, Erbertseder & Taubenböck, 2022), which was further supported by our validation analysis using monitoring stations (appendix 1 pp 35–36).

Finally, it is crucial to acknowledge that the mortality association is based on an ecological analysis and might be influenced by various unaccounted factors, such as socioeconomic conditions, population behaviors, habits, and health-care provision. However, it was the best available indicator for most cities examined in this study. Furthermore, this analysis represents one of the initial broad examinations of the connection between urban configuration and health in Europe. Notably, there is a small body of research in Latin America, with only a few studies globally and in Europe (Avila-Palencia et al., 2022, 2023; Thompson et al., 2020; Olsen et al., 2019). One 2019 study across Europe aligns with our findings, showing a reduction in mortality in cities characterized by natural green spaces (Olsen, et al., 2019). This underscores the need of additional epidemiological studies to deepen the understanding of the intricate relationship between urban configuration and health outcomes.

In conclusion, we identified four distinct European urban configuration types and evaluated their association with motorized traffic flows, SUHI, NO_2 emissions, CO_2 emissions, and natural-cause mortality rates. Our findings suggest that, although the compact city model has arisen as the way forward to healthier and more sustainable cities, present day European compact cities should be understood as being in a transition state that combines positive features, such as access to services and reduced carbon emissions, with challenges such as high traffic volumes and a poor environmental quality. Cities are complex systems and solutions require a holistic approach. A set of policy measures has been proposed to promote more sustainable and healthier urban and transport developments. These measures need to be evaluated at the local level and actions relevant for each city's local context need to be prioritized.

Supplementary

https://www.sciencedirect.com/science/article/pii/S2542519624001207#sec1

Acknowledgments

We acknowledge support from the Spanish Ministry of Science and Innovation and State Research Agency through the grant CEX2018–000806-S funded by MCIN/AEI/10.13039/501100011033, and the Ayudas para la Formación de Profesorado Universitario 2020–24 doctoral funding (grant number FPU19/05210); support from the Generalitat de Catalunya through the Centres de Recerca de Catalunya programme; support from Centro de Investigación Biomédica en red Epidemiología y Salud Pública; and support

from the Urban Burden of Disease Estimation for Policy Making 2023–26 Horizon Europe project (grant number 101094639). TC's contribution was supported by the US Department of Energy, Office of Science, Biological and Environmental Research programme through an Early Career award. Pacific Northwest National Laboratory is operated for the US Department of Energy by Battelle Memorial Institute under contract number DE-AC05–76RL01830.

Funding

Spanish Ministry of Science and Innovation, State Research Agency, Generalitat de Catalunya, Centro de Investigación Biomédica en red Epidemiología y Salud Pública, and Urban Burden of Disease Estimation for Policy Making as a Horizon Europe project.

Chapter XX

Land consumption in cities

A comparative study across the globe

Cities, 2021, vol. 113, 103163

Jingliang Hu, Yuanyuan Wang, Hannes Taubenböck & Xiao Xiang Zhu

© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.1016/j.cities.2021.103163

Received 24 Sept 2020; Accepted 19 Feb 2021; Published 8 Mar 2021

Abstract

Land consumption delineates how effectively we use our living space - whether wastefully in area or efficiently. As the United Nations (UN) projects 5.1 billion population will live in urban areas in 2030, it is crucial to measure, compare, and understand land consumption for our cities across the globe. Currently global approaches for land consumption analysis rely on settlement layers produced by remote sensing data and population data achieved by distributing census data over settlement layers. These layers, however, do not have sufficient spatial details to capture the intra-urban structural variability within a city. In this study, we develop a classification system that consistently produces accurate local climate zone (LCZ) maps at intra-urban scale for 40 cities using Sentinel data. We use the LCZ classes as proxies to disaggregate the global population grids (GHS-POP) to this intra-urban scale. With the refined data, we perform an intra-urban land consumption analysis for 40 cities across the globe. Our measurements show that current per-city land consumption studies severely deviate from the intra-urban variability. We argue urban land consumption at global scale must proceed to an intra-urban resolution. In addition, our measurements also indicate that urban land consumption is differing immensely across the globe.

Keywords

Land consumption; local climate zones (LCZ); population density; Sentinel-1; Sentinel-2; topology enhanced ensemble classification; urban morphology

1. Introduction

1.1. Motivation

The United Nations (UN) reported that in 2018 already 55.3% of the world's population lived in urban areas (UN, 2018). This number is even estimated to increase to 60% by the year 2030. This means that for sustainable urban development, we have to weigh very well the land consumption. With the sustainable development goals (SDGs) set by the UN (UN, 2015b), the 11th goal is to make cities and human settlements inclusive, safe, resilient and sustainable. One specific indicator (SDG 11.3.1) for monitoring is described as "*ratio of land consumption rate to population growth rate*". According to the UN (UN Habitat, 2016), this indicator is particularly designed to monitor urban land consumption and thus, to measure the extent to which we use the space efficiently or wastefully. However, monitoring the intensity of usage or development of urban land at global scale in consistent manner remains a major challenge. Data generation is expensive both in finance and labor. Remote sensing technology, however, provides promising data sources to this challenge because it is affordable and generally available at a global scale. Therefore, in this study we aim at studying urban land consumption of cities across the globe by using remote sensing data.

1.2. Conceptual foundation

The term '*urban land consumption*' is generally defined as modified natural lands to man-made structures due to the living, social, and economic purposes of the urban residences (Boyce, 1963; d'Amour et al., 2017; Gerundo & Grimaldi, 2011). The term '*intensity of land consumption*' is often an important quantitative measurement designed to reflect the status of the land consumption. Since one can study the urban land consumption from various disciplines, the intensity of land consumption has immense implications. From an economic perspective, Kuang et al., (2020) defines it as the amount of urban land consumed per unit of gross domestic product (GDP); (Bimonte & Stabile, 2017) and (Wang et al., 2020) associate the term with income and housing price, respectively. Among urban temperature studies (He et al., 2019; Stokes & Seto, 2016; L. Xu et al., 2019; J. Yang et al., 2017; Yang, Wang, et al., 2020), the intensity of urban consumption is related to the building structure density that influences urban heat flow. From the scope of SDG (UN Habitat, 2016), the indicator focuses on the change of land consumption in time as the SDG indicator suggests by "*ratio of land consumption rate to population growth rate*". From a demographic aspect, the intensity of land consumption relates to building structures which further influence the amount of population per unit of area (Freire et al., 2016; Glaeser, 2011). Moreover, Martinuzzi et al. (2007) found the population density an essential variable in the urban land consumption research with a case study in Puerto Rico. Grekousis & Mountrakis (2015) found a pro-linear relationship between the population and land consumption in the US. In our study, we use the indicator population density (the amount of population per spatial unit) as a measure of the intensity of land consumption. Our indicator reflects the efficiency of urban land usage from the demographic perspective.

Among the urban land consumption studies, many studies focus on a city (Lehnert et al., 2015; Stewart & Oke, 2009; Zhang et al., 2014) or a regional scale (Nicolau et al., 2019;

Salvati et al., 2018; Wang et al., 2020; Yang et al., 2019; Yang et al., 2020). Few studies have approached urban land consumption on the global scale: Melchiorri et al., (2018) analyzed the spatial dynamics of global urbanization from 1990 to 2015 by using the Global Human Settlement Layer (GHSL) (Pesaresi et al., 2013). They reported the planet became substantially more urbanized than what had been previously reported. Another study (Melchiorri et al., 2019) reported the urban land consumption of 10,000 cities by using the GHSL and analyzed the change of land use for these cities. Taubenböck et al. (2019) analyzed 1,692 cities from an urban morphological perspective by using the Global Urban Footprint (GUF) (Esch et al., 2017). They defined the extent of a city by a morphological urban area instead of an administrative boundary. Coupling with a population data, they found the Pearl River Delta is the largest urban agglomeration with the largest population on the planet. By far, thematically and structurally high-resolution analyses are still limited to studies of a city or a regional scale. At the same time, global studies mostly rely on aggregate city so that their basic spatial unit is a city. They derived physical or statistical indicators per city for the analysis. This is because global settlement layers as the GUF (Esch et al., 2017) or the GHSL (Pesaresi et al., 2013) have a spatial resolution at the city scale. Intra-urban spatial information, however, are inexistent on global scale.

Recently, new population layers on global scale have been introduced. Current population data are produced by distributing official census over built-up regions, such as gridded population of the world (GPW) (Doxsey-Whitfield et al., 2015) and global population grids (GHS-POP) (Freire et al., 2016). Additional data might also be used to regulate the spatial distribution. The WorldPop data considers land cover/land use information for distributing the population (Stevens et al., 2015). Census data, building footprint data and open street map data are applied in (Huang et al., 2019) to estimate the population distribution of the United States. In general, the spatial distribution of the global population data relies on the administrative areas of the census data and the spatial details of the settlement layers. Those settlement layers have coarse details, such a binary built-up layer. All these factors result in limiting the spatial resolution of the population data at a city scale. Therefore, settlement layers of improved spatial and thematic resolutions at an intra-urban scale will promise a more accurate population distribution and urban land consumption study.

The available state-of-the-art global urban maps, GUF (Esch et al., 2017) and WSF (Marconcini et al., 2019), delineate the built-up area as a binary mask; and GHSL (Pesaresi et al., 2013), characterizes the urban area as a four-category thematic map considering the density of built-up areas (Melchiorri et al., 2019). Both products follow a classification scheme allowing to describe the spatial patterns of cities. With it, these data sets allow to compare the patterns of cities across space in a consistent manner. However, detailed morphological information within cities, such as the structural types of the settlements, is missing in these classifications. On the other hand, one recent scheme named Local climate zones (LCZ) has often been applied for intra-urban morphological differentiation. The scheme aims at classifying urban areas in a consistent manner into 17 classes (Stewart & Oke, 2012). Although the LCZ classification scheme was originally designed for urban climatic studies (Stewart et al., 2014; Stewart & Oke, 2012), it essentially classifies the urban area by morphological criteria, mainly including the compactness, the structural shape, and

the heights of settlements, as well as the land surface material. Indirectly, those morphological parameters also allow for deriving land use information to a certain degree (Bechtel et al., 2016; Perera & Emmanuel, 2018; Sapena et al., 2021; Zhu et al., 2020). Thus, in comparison to the current state-of-the-art global urban mappings, the LCZ scheme can provide intra-urban morphological information to account for the morphological variance within the settlement landscape. With it and in combination with population data, the analysis of urban land consumption allows for an intra-urban spatial scale.

Based on the conceptual foundation, in this study, we aim at studying and comparing urban land consumption in various cities across the globe at an intra-urban spatial scale to account for the variability within cities. We aim to show the extent to which urban living space is used effectively or ineffectively. In particular, we want to show that indicators that are usually generated at the city scale are susceptible to deviating mean values due to the high variability of land consumption intensity within cities.

1.3. Contribution of this article

In this study we investigate urban land consumption by associating demographic information to intra-urban knowledge on structural variability by the LCZs. The contributions of this article are three-fold: First, in this paper we analyze urban land consumption at global scale, but compared to existing literature we increase the spatial resolution from city- to an intra-urban scale. We argue that monitoring the urban land consumption or urban sustainable development at global scale must proceed to an intra-urban resolution, as we find our estimated city-scale urban land consumption indicators severely deviate from the intraurban variabilities. Second, this paper develops a novel semi-supervised data-topology based classification system which fuses freely-accessible Sentinel-1 and Sentinel-2 data. The developed system is capable of providing urban morphological maps for cities around the globe with intra-urban spatial details in consistent and thus comparable manner. Only with this, we believe that a cross-city comparison is permissible. In this study we apply it for 40 cities around the globe. Third, compared to existing population data that distribute demographic census data over settlement layers, this paper improves the granularity of the settlement layers to 11 types of urban morphological classes for the population distribution. With it, we integrate beyond morphological details also the functionalities of these classes. The refined population grid has a reasonable distribution at the intra-urban scale and with unprecedented 10 m spatial resolution.

The remainder of this paper is as follows. Section 2 introduces the study areas, the data, and the methodology. In section 3, we present the evaluation of the developed classification system, demonstrate the spatially refined population data, and analyze land consumption. Section 4 discusses the outcomes of our studies. And section 5 concludes this paper.

2. Data and method

2.1. Study area

We base our selection of study areas on five rules: *First*, the selection of cities has a global distribution, i.e. we select cities from all continents. *Second*, the selected cities are from different cultural regions, i.e. even within continents we select cities that are of different cultural regions. *Third*, the cities are of different status in terms of development and economy. *Fourth*, these cities are large cities where urban morphology are very complex and challenging for the algorithm development. As large cities, they feature large population as well. *Fifth*, from a pragmatic point of view, we selected cities where LCZ reference data are available and accessible (Zhu et al., 2020). Based on these general rules, we selected 40 major cities. As shown in Fig. 1, the selected cities cover nine cultural regions, namely Australia, East Asia, South-east Asia, the Indian subcontinent, the Middle East, Europe, Africa, North America, and South America. Moreover, the selection includes economically less-developed cities, such as Islamabad, and very-developed cities, such as London. The selected cities all have comparably large population, which in total are according to UN statistic (UN, 2015a; UN, 2016) home to about 343 million inhabitants in the year of 2015.

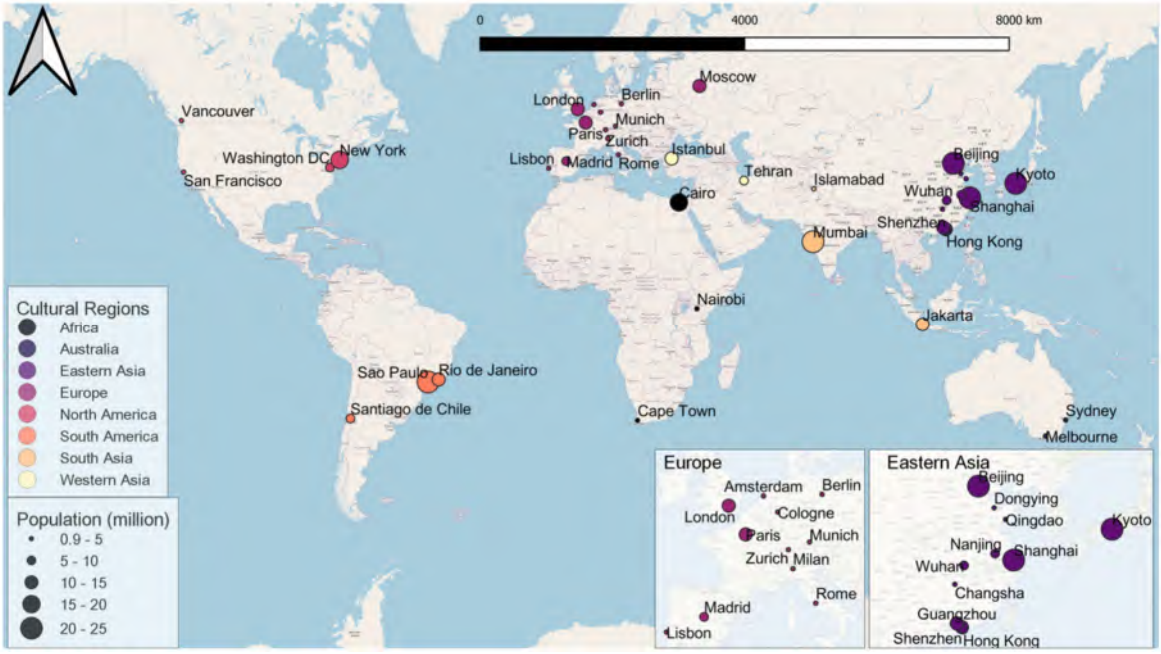


Fig. 1 The geographic locations of the 40 sample cities. The size and color reflect the population and cultural region of the corresponding city.

For a permissible comparative study of cities, comparable spatial units, i.e. spatial extents of cities, need to be defined. The administrative boundary of a city often does not reflect the physical extent of the city. An extent of fixed radius, however, is also not meaningful due to large variety in size and shape of our sampled cities. For example, as shown in Fig. 2, the sizes and shapes of Paris, Istanbul, Zurich, and Cologne have huge differences. Against this background we apply morphological urban areas (MUA) introduced by (Taubenböck et al., 2019). The MUA delineates the extent of a city from the urban landscape along a decreasing

gradient of settlement density. Based on quantitative measure, the MUAs are consistent among cities and thus are permissible spatial extents for meaningful comparative studies. The highlighted areas in Fig. 2 demonstrate the MUA extents of the cities.

2.2. Data sets

We base this study on remotely sensed as well as auxiliary data. For the generation of the LCZ classification we rely on three data sources: 1) *Sentinel-1 data*, 2) *Sentinel-2 data*, and 3) an *LCZ reference data*. For the population information we use 4) the *GHS-POP*.

1) *Sentinel-1 data* used in this paper are level-1 VV-VH dual-Pol single look complex (SLC) products. The data are prepared as an analysis-ready data of 10 m ground sampling distance via a group of pre processing steps using the ESA SNAP toolbox. The detailed procedure of the preprocessing is introduced in Appendix A.

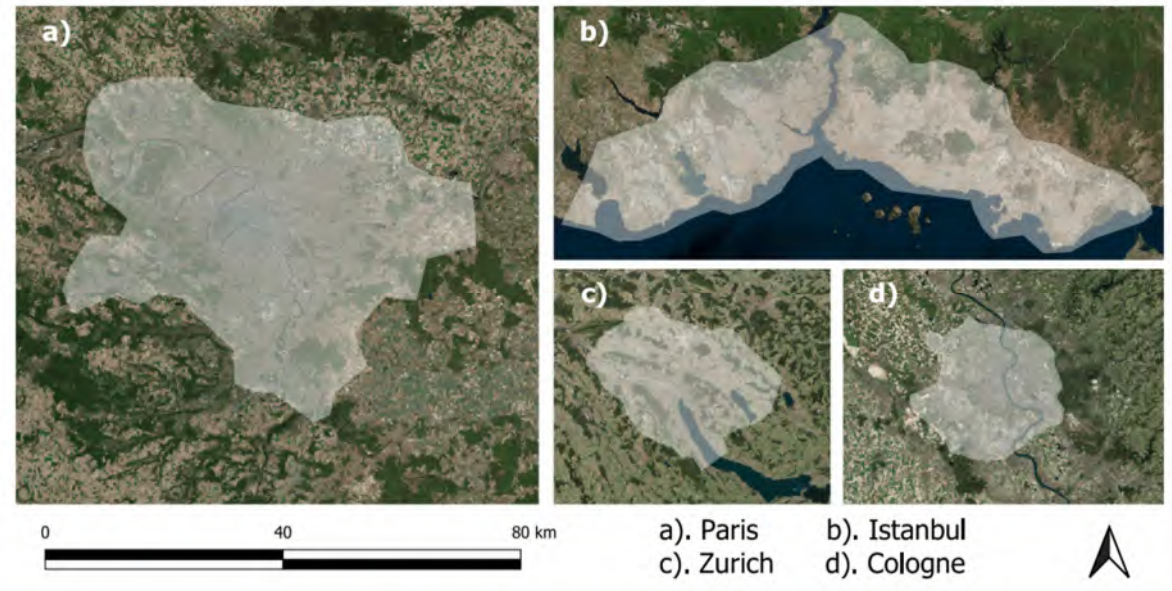


Fig. 2. The cities of Paris, Istanbul, Zurich, and Cologne are shown in the same map scale where the highlighted areas indicate MUA urban extents (Taubenböck et al., 2019).

After the preprocessing, feature extraction is carried out on the analysis-ready data for the classification purpose. It is designed to derive features representing polarimetric information, local neighborhood in formation, and spatial information. The features related to polarimetric information are the intensity of VV, the intensity of VH, the coherence between VV and VH, and the intensity ratio of VV and VH. The intensity features are converted into decibel (dB). The local neighborhood in formation includes the mean and the standard deviation of an 11-by-11- pixel local area for each of the four polarimetric features. The spatial information is derived by the morphological profile (Benediktsson et al., 2003; Fauvel et al., 2008), with radius of 1, 2, and 3 pixels, for each of the four polarimetric features. The feature extraction results in 36 features in total.

2) *Sentinel-2 data* are preprocessed on Google Earth Engine (GEE) for eliminating clouds (Gorelick et al., 2017; Schmitt et al., 2019). Afterwards, feature extraction is carried out on the Sentinel-2 analysis-ready data which also has a ground sampling distance of 10 m for the

classification purpose. The principal components of the data that hold 99% of the total spectrum information are derived to represent the spectral information. The morphological profile is utilized to derive the spatial information on each of the principal component, with radius of 1, 2, and 3 pixels. Thus, the number of extracted S2 features varies from city to city depending on the numbers of principal components.

3) *The LCZ reference data* used in this paper is from the So2Sat LCZ42 (Zhu et al., 2020). These have been annotated by a group of 15 experts who followed a labeling protocol that maximizes its consistence, completeness, and correctness, meanwhile minimizes inevitable human error (Zhu et al., 2020). For each of the 40 cities, the reference data is formatted as a raster image with a ground sampling distance of 100 m. It is important to mention that, instead of the original 17 classes of LCZ, this paper aims to classify a modified version of LCZs with 11 classes. The main reason of modifying the LCZ scheme is that there is no access to global scale data that provides reliable, accurate and complete building height information for urban areas. The 17 classes of the original LCZs describe four properties of urban land: compactness, structural shape, height of settlements and the land surface. Since the satellite data used in our study have no information on height, we modify the LCZ classification scheme by merging the LCZs of '*compact high rise*', '*compact middle rise*', and '*compact low rise*' as *compact structures*, and the '*open high rise*', '*open mid-rise*', and '*open low rise*' as *open structures*. By cross check with satellite images, the large low-rise class very often consists of warehouses or factories located in industrial areas so that we merge the classes of 'large low rise' with 'heavy industry' to an *industrial area*. The sparsely built-up is ignored because it frequently locates in scattered tree area or low plant area. Therefore, we come up with a modified version of LCZ with 11 classes.

4) *The population data* used in this paper is a world-wide grid of the population for 2015, namely GHS-POP (Schiavina et al., 2019). In the grid, each cell has a ground sampling distance of 250 m and provides the number of people. The data is achieved by disaggregating administrative census data over a global built-up layer, the Global Human Settlement Layer (GHSL) (Pesaresi et al., 2013).

2.3. Method

The goal of this study is to analyze land consumption at the intraurban scale. To meet the goal, we *first* develop a classification system that fuses the Sentinel-1 and Sentinel-2 data and classifies the fused data into LCZ maps. Two main modules of the classification system are a semi-supervised data fusion and a data topology-based ensemble classification. Details of the two modules are introduced in Section 2.3.1 and 2.3.2, respectively. To reflect the accuracy of the produced LCZ maps, a sophisticated strategy is designed for a systematic evaluation, which is introduced in Section 2.3.3. *Secondly*, we use the LCZ classification maps to refine population data. We do so by redistributing the population into possible residential areas. Details are introduced in Section 2.3.4. At last, an unsupervised clustering module is developed to identify types of urban morphological formation which enables analyzing the correlation between land consumption and morphological types of cities. This approach is detailed in Section 2.3.5.

2.3.1. Semi-supervised data fusion

The developed classification system fuses Sentinel-1 and Sentinel-2 data, i.e. optical and radar data, via a semi-supervised strategy. In previous studies, researches have investigated optical data (Bechtel, 2011; Danylo et al., 2016; Xu et al., 2017), digital surface models (DSM) (Bechtel & Daneke, 2012; Xu et al., 2018), geographical information system (GIS) data (Gal et al., 2015; Geletic & Lehnert, 2016; Lelovics et al., 2014), Interferometric Synthetic Aperture Radar (InSAR) (Bechtel & Daneke, 2012), and polarimetric synthetic aperture radar (PolSAR) data (Bechtel et al., 2016; Hu et al., 2018) for the classification task. Commonly, optical data have been used for their easy access and good performance. However, these data are limited by weather conditions. Besides being independent from weather conditions, PolSAR data are beneficial by providing spatial information and back scatterings of built-ups. To join the merits of both data sources, we apply Sentinel-1 and Sentinel-2 data for mapping LCZ. The reasons are: (1) the two data sources, the multispectral information and the radar back scattering, provide complementary physical information; (2) both data sources have global coverage and are freely accessible; (3) optical data have a good performance on LCZ classification (Bechtel, 2011; Qiu et al., 2019; Xu et al., 2017), but data availability depends on weather conditions, while Sentinel-1 data is available independent to weather conditions; (4) it has been proven that spatial information and polarimetric information of SAR data benefit LCZ classification (Bechtel et al., 2016; Hu et al., 2019).

Data annotations are of great importance for machine learning algorithms (Reichstein et al., 2019; Zhu et al., 2017). Yet annotating land cover/use information in remote sensing is very expensive in cost and labor. Thus, an algorithm that takes maximum advantage of in formation in data annotations is highly welcomed. For the combined usage of Sentinel-1 and Sentinel-2 data, we follow a semi-supervised data fusion strategy (Hu et al., 2019; Tuia et al., 2014; Wang & Mahadevan, 2009). This strategy propagates the information of annotated data into un-annotated data via the similarity measure between them. It is able to amplify the information in annotations which is beneficial in remote sensing. Therefore, the semi-supervised strategy is utilized in our data fusion module.

The semi-supervised fusion algorithm aims to map Sentinel-1 and Sentinel-2 data into a latent space where (I) the data of the same class locate close to each other, and (II) the topological structure of individual data source is preserved. As shown in the semi-supervised data fusion of Fig. 3, the class-wise relation, the property (I) of the latent space, is derived from reference data and represented as a graph. With a topo logical data analysis tool (Singh et al., 2007; Li et al., 2015), the data structures of the two data sources are presented as two graphs, which stands for the second property of the latent space. Those three graphs can be organized and formulated by using a joint similarity matrix. With the joint similarity matrix, a general eigenvalue problem is formulated and a mapping function for each of the data sources can be optimized. Therefore, the two mapping functions act as the bridges that link the two data sources to the latent space, respectively. More mathematical background of the semi-supervised data fusion algorithm is introduced in Appendix B.

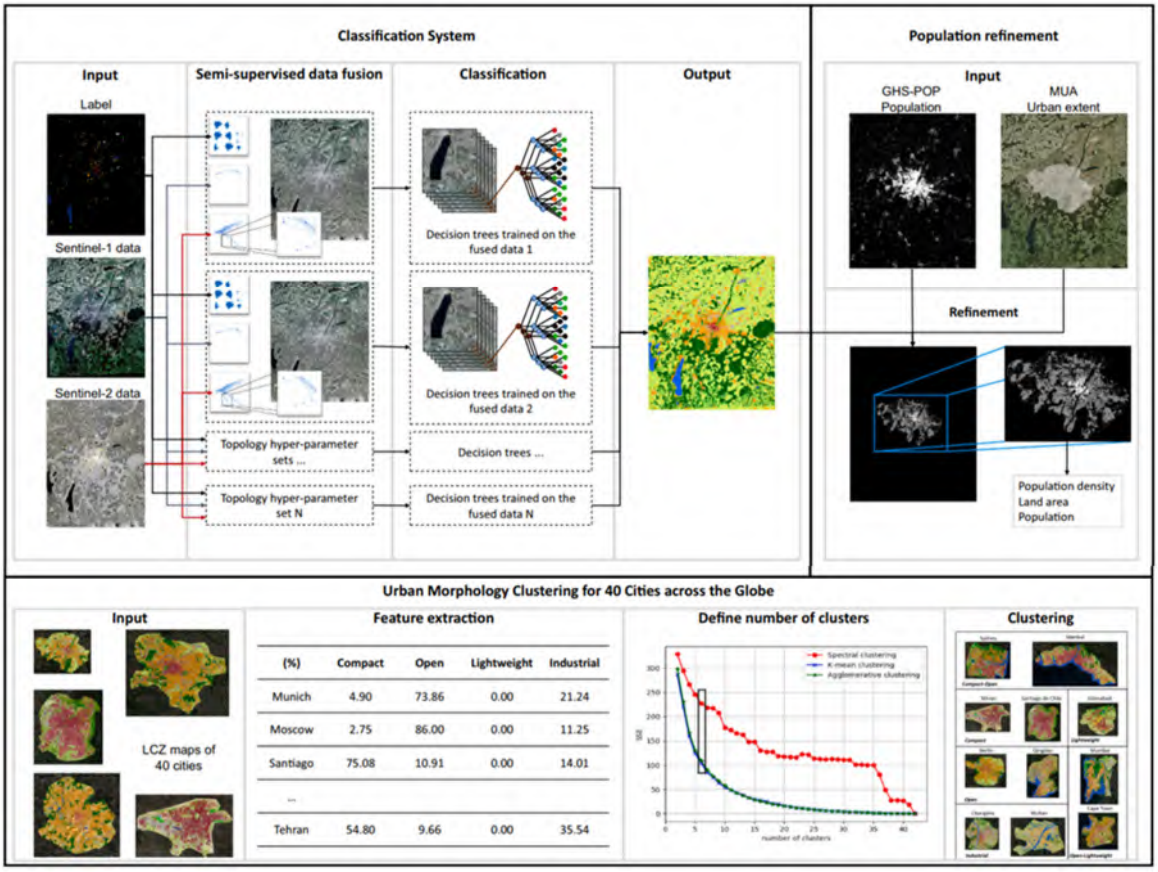


Fig. 3. The methodological workflow of classification, population refinement, and urban clustering. The classification takes Sentinel-1 data, Sentinel-2 data, and LCZ reference data as inputs and produces LCZ maps. The population refinement takes the produced LCZ map, GHS-POP data, and the MUA urban extent to improve the spatial details of the population data. The urban clustering takes 40 LCZ maps as inputs and identifies city types in terms of morphological formations.

2.3.2. Topology enhanced ensemble classification

It has been shown that a group of classification algorithms (Danylo et al., 2016; Geletic & Lehnert, 2016; Lelovics et al., 2014; Xu et al., 2018) is beneficial for producing LCZ maps for a city by a classifier trained with data of that city. Another group of algorithms attempt to train a classifier and produce LCZ maps for cities where reference data are not available (Demuzere et al., 2019; Kaloustian et al., 2017; Tong et al., 2020; Yokoya et al., 2017; Zheng et al., 2020). Due to local specific algorithms of the first group result in better classification accuracy compared to the second group. With this background, we followed the strategy of the first group and produced LCZ maps for selected 40 cities where LCZ labels are available.

Random forest is chosen as classifier for our classification system, since it is a state-of-the-art, non-parametric, and time efficient classifier that is widely used in global scale classification tasks (Li et al., 2017) and local climate zone classification tasks (Bechtel et al., 2015; Demuzere et al., 2019; Yokoya et al., 2017). The random forest trains its decision trees with random subset of features and random subset of training samples. It results in different decision trees predicting different labels for an instance. This configuration creates a diversity that empowers random forest the most. The diversity makes random forest statistically following the *Law of Large Numbers* so that it has a robust performance, avoids

overfitting, and meanwhile preserves the computational efficiency (Breiman, 2001; Elghazel et al., 2011; Rainforth & Wood, 2015).

In the proposed classification system, there is a practical issue of hyper parameter optimization while deriving topological structures. No matter which method is chosen to obtain topological structures of Sentinel-1 and Sentinel-2 data, one would inevitably be confronted with the optimization of hyper-parameter. To handle this practical issue, a study (Tuia et al., 2014) set the hyper-parameter manually based on their empirical experiences. Another study (Hu et al., 2019) reported that no statistically significant optimal solution can be found. In this paper, the authors find that varying values of the hyperparameter leads to topological structures of regional differences. Inspired by the *Law of Large Numbers* concept in random forest classifier, this paper proposes a topology enhanced ensemble classification scheme. It fuses the two data sets with a set of hyperparameter. For each fused data set, a group of decision trees are trained. Therefore, in the prediction phase, the ensemble of decision trees includes diversities of three aspects: different topological structures, random subset of features, and random subset of training samples.

2.3.3. Classification evaluation strategy

To evaluate the performance of the proposed classification system, we design a sophisticated evaluation strategy that is able to avoid overestimation of classification accuracy caused by overfitting and that provides an uncertain range of accuracy in a statistically robust manner.

For evaluating an algorithm, the spatially random selection of training data from the ground reference that often appears in literature will make the training and the testing samples strictly follow the same distribution. Such selection of training data results in the trained algorithm overfitting the test data. In consequence the resulting accuracy would overestimate the performance of the trained algorithm (Hänsch et al., 2017). To avoid the overestimation, we separate the ground reference into blocks for training and testing.

The fractions of reference data used for training and testing would also have an impact on the accuracy assessment. In remote sensing ap plications, the training data normally cover an extremely small fraction of the whole study area. As an example, in subplot a) of Fig. 4, this fraction can be observed by comparing the coverage of reference data to the extent of New York. However, in machine learning evaluation, training data often occupy 90% of reference data and the remaining 10% are used as testing data. The amount of training data is nine times larger than the testing data, which is not realistic in remote sensing. Thus, in our evaluation, instead of only taking 90% of reference data as training data, we take 10%, 50%, and 90% of the reference data as training data for three sets of evaluations, respectively. To achieve this setting, reference data are equally separated into ten block-wise subsets, as shown in Fig. 4. In our configuration, instead of one accuracy, the three settings will provide a range of accuracies indicating possible performance of an algorithm with respect to the amount of training samples used. For each of the three settings, a 10-fold cross validation is per formed to achieve a mean classification accuracy which is a statistical robust indicator.

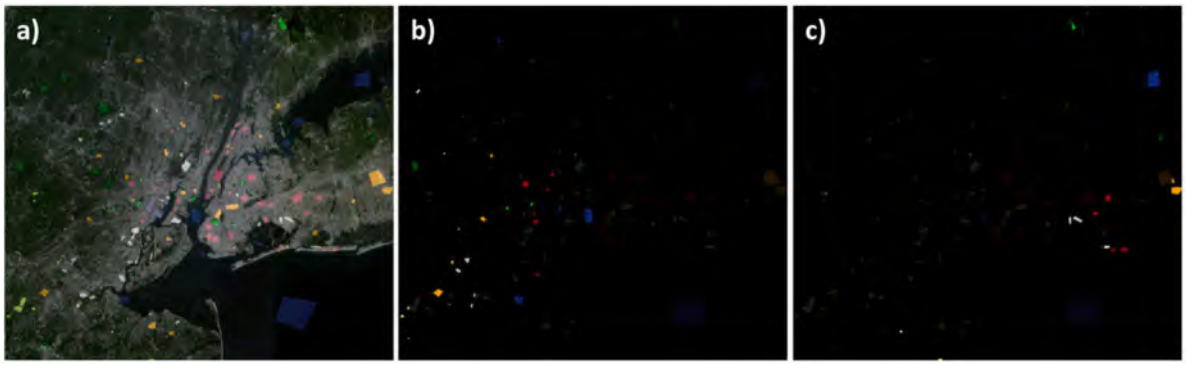


Fig. 4. The reference data are equally separated into ten non-overlapping subsets from west to east for each city. This figure shows the procedure by the example of New York. In this figure, the three images are: a) the reference data overlapped with the RGB imagery; b) the reference data with the first subset highlighted; c) the reference data with the tenth subset highlighted.

2.3.4. Population refinement and population density

Current high-resolution population data are generally achieved by spatially distributing official census data at administrative levels over settlement areas detected from remote sensing data. However, this is in large area applications predominantly done regardless of the functionality of the settlement structures. Therefore, the disaggregation procedure might lead to people assigned to industrial sites and thus might lead to a skewed inaccurate calculation of the population density at intra-urban scale. As we associate population to urban land for analyzing urban land consumption in this paper, the population density is applied as an important quantitative indicator for the analysis. To avoid the miscalculation of population density, we refine the population data using the morphologic and functional knowledge of our classified LCZ maps. Although the local climate zones were originally designed for the study of land surface temperatures, the classification scheme essentially describes the structures of local urban neighborhoods in terms of the compactness, the shape, the height, and the covering materials of the settlements. And these characteristics of local urban neighborhoods reflect the population distribution within a city. We rely on these data to redistribute the given population data from GHS-POP. For the refinement, we assume, (1) people do not live in industrial structures, and (2) the population density of compact, open, and lightweight structures differ, due to the varying available living spaces within these morpho logical structures.

To accomplish the refinement, the population over industrial structures in the GHS-POP data need to be redistributed to the other three structures. The ratios of population that reside in the other three structures are calculated based on the produced LCZ map and the population data. In detail, an LCZ map is presented as a grid L. L_j^i presents a cell in the grid where $i \in \{C,O,L,I\}$ stands for compact, open, light weight, and industrial structures, and $j \in \{1,2,\ldots,n\}$ stands for the j^{th} cell. The population data is presented as a grid P and P^i_j geographically corresponds to L_j^i. Therefore, the ratios of population that reside in compact, open, and lightweight structures are calculated according to (1).

$$R_i = \frac{\sum_j \mathscr{P}_j^i}{\sum_i \sum_j \mathscr{P}_j^i}, i \in \{C, O, L\} \quad (1)$$

Then, the calculated ratios guide the redistribution of the current population to achieve the refined population grid $\tilde{P}$.

$$\tilde{\mathscr{P}}_j^i = \begin{cases} \frac{\sum_{i,j} \mathscr{P}_j^i \times R_i}{n^i} & i \in \{C, O, L\}, n^i : \text{the cell number of the } i \text{ structure} \\ 0 & i = I \end{cases} \quad (2)$$

For our assessment of land consumption, we rely on the population density that is calculated based on our refined population data. We apply our analysis on land consumption at two spatial scales - on city scale and on intra-urban scale. On the city scale we calculate the 'land area' as the sum of all settlement areas from the compact, open and lightweight structures, i.e. from areas where we assume people live. The population is calculated as the total population of a city which refers to the input population data and the applied MUA. The population density equals the total population divided by the possible living areas. Based on the fusion of both products, we provide these three indicators to reflect land consumption. On the intra-urban scale, these three indicators are calculated for compact, open, and lightweight urban blocks. This scale allows for an intra-urban analysis, that would not be possible by using only original population data.

2.3.5. Clustering of cities based on their morphological formations

For our urban geographic analysis, we want to identify morphological types that are similar in their built configuration. For the clustering procedure, we use features from the LCZ maps that represent the morphological formation of a city. We calculate the spatial proportions of the four urban morphological structures, compact, open, light, and industrial structure in the built-up areas. With this feature representation, it becomes possible that cities with large differences in size and shape can be recognized as the same type based on similar morpho logical formations.

We apply clustering to find urban morphological groups. Two critical issues are: (1) choosing a clustering algorithm and (2) defining the number of clusters. For choosing an algorithm, there exist two groups of algorithms that are divided by strategies of deciding the number of clusters. The first group of algorithms has the number of clusters as an input parameter which users must know in advance, such as k-means clustering (Arthur & Vassilvitskii, 2007), spectral clustering (Ng et al., 2002; von Luxburg, 2007), and agglomerating clustering (Frigui & Krishnapuram, 1997). The second group of algorithms detects the number of clusters by physical or statistical variables in an algorithm whose values have to be given in advance. Instead of explicitly giving the number of clusters, these thresholds having physical or statistical meanings indirectly decide the number of clusters. Such algorithms include e.g. DBSCAN (Ester et al., 1996), mean shift (Comaniciu & Meer, 2002), modified spectral clustering (Zelnik-Manor & Perona, 2005), or topological

persistence-based clustering (Chazal et al., 2013). For the sake of simplicity and robustness, we choose K-means. For the second issue, defining the number of clusters, the error sum of squares (SSE) is utilized as a quantitative indicator. SSE is the sum of squared error be tween each data record and the mean cluster center.

$$\mathrm{SSE} = \sum_{i=1}^{n} \left(\mathrm{C}_i - \tilde{\mathrm{C}}_i \right)^2 \qquad (3)$$

where n is the number of data records, C_i is a feature vector representing the i^{th} data record, and $\tilde{C}_i$ is the center of the cluster that the i^{th} record belongs. With SSE, the rule of elbow is a frequent practice of deciding the number of clusters (Jain et al., 1999). It chooses the value whose increase brings a small reduction of the SSE and those non-relevant groups. To decide the number of clusters, we apply k-means clustering, spectral clustering, and agglomerating clustering to cluster sample cities with the number of clusters ranging from 2 to the maximal possible value. Therefore, an SSE curve in terms of the number of clusters can be plotted. By considering the rule of elbow for all three algorithms, the number of clusters can be chosen.

3. Results and analysis

In this section, first, we evaluate the performance of the proposed classification procedure and population refinement. Then, we present and compare the land consumption of the sampled 40 cities at city scale and the intra-urban scale. At last, we summarize our analysis.

3.1. Evaluation

3.1.1. Evaluation of the LCZ classification

Following the strategy introduced in Section 2.3.3, we present the evaluation of the proposed classification system for three morphologically different exemplary cities: Munich, New York and Mumbai.

We applied three other classification systems for comparisons. They are: (1) Sentinel-1 data with random forest (S1); (2) Sentinel-2 data with random forest (S2); (3) the concatenation of Sentinel-1 and Sentinel-2 data with random forest (CON). The three classification systems are chosen for comparison because such systems have appeared on large scale classification tasks in literature (Ban et al., 2015; Demuzere et al., 2019; Esch et al., 2017; Li et al., 2017; Lisini et al., 2017; Yokoya et al., 2017; Zhu et al., 2012). The proposed classification system is indicated by topology enhanced ensemble classification (TEEC) in the evaluation 5.

Figure 5 demonstrates the ranges of the achieved mean overall accuracy by the 10-fold validation. The range shows accuracies achieved by using different amounts of reference data as training data, namely 10%, 50%, and 90% of the reference data. According to Fig. 5, the proposed classification system outperforms the other systems in terms of the mean overall accuracy, under all scenarios of different training sets and cities.

For the city of Munich, the proposed classification system reaches a mean overall accuracy from 83.8% to 89.1%. For the city of Mumbai which has a complex urban environment, our

system has an accuracy from 59.4% to 72.2%. The accuracy for the city of New York ranges from 62.8% to 79.7%. With the sophisticated evaluation strategy, we conclude the proposed classification system allows to produce with respect to commonly used algorithms highly accurate and consistent LCZ classifications for our 40 cities.

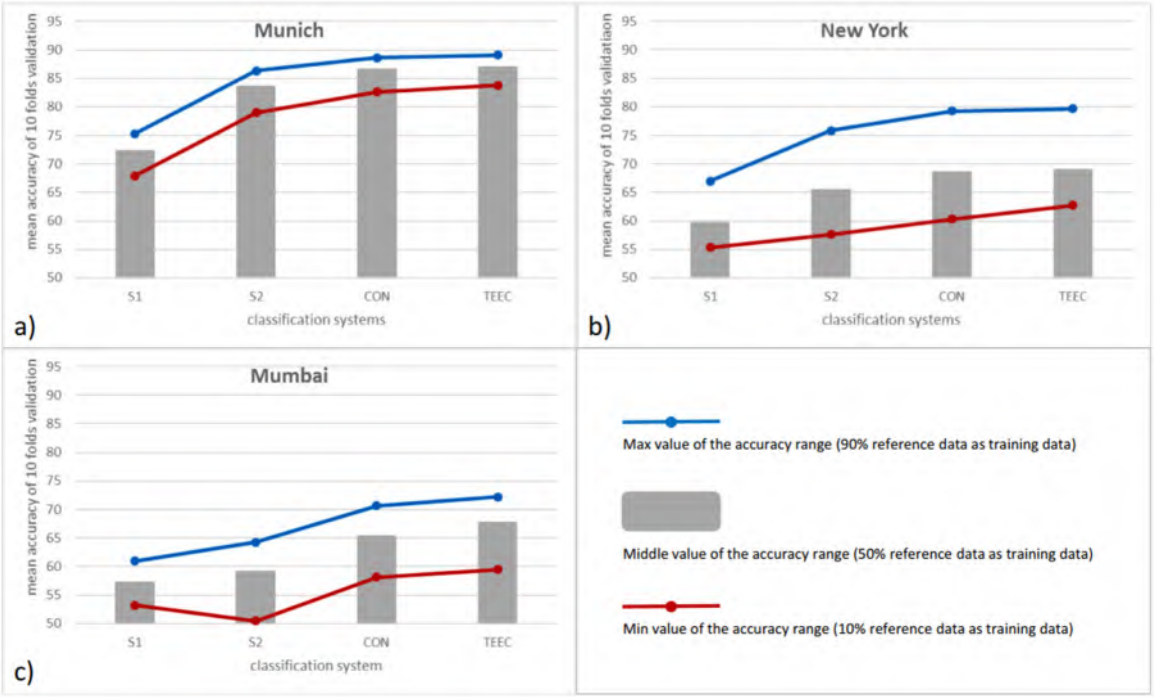


Fig. 5. Evaluation of four classification systems, S1, S2, CON, and TEEC. The x-axis indicates the respective classification system. The y-axis depicts the mean overall accuracy of the 10-folds cross-validation evaluation. The three subplots a), b) and c) represent the evaluation over Munich, New York, and Mumbai. The blue curve shows the accuracies that are achieved by using 90% reference data as training data, which can be regarded as the upper bound of the accuracy. The red curve shows the accuracy produced by using 10% reference data as training data, which can be regarded as the lower bound of each method. And the grey bar shows the accuracy produced by using 50% reference data as training data as a median value between boundaries.

3.1.2. Evaluation of the population refinement

For the evaluation of the spatial refinement of population data, we present the results by the example of Mumbai. As high-resolution ground truth population data are nonexistent, we aim to provide a qualitative analysis for our refinement of the population. As shown in Fig. 6, it seems very plausible that our refinement of the population improves the global data due to the spatial detail in our LCZ classification. For example, the blue, red, and green boxes outline an open area, an airport, and a compact area, respectively. In GHS-POP data, the population density is the same for the three areas. In WorldPop data, the population that reside in the airport and open area is larger than the population in lightweight area. Both cases seem to be very unrealistic. In our refined data, the re-distribution of population appears more reasonable as in the airport we do not have population disaggregated.

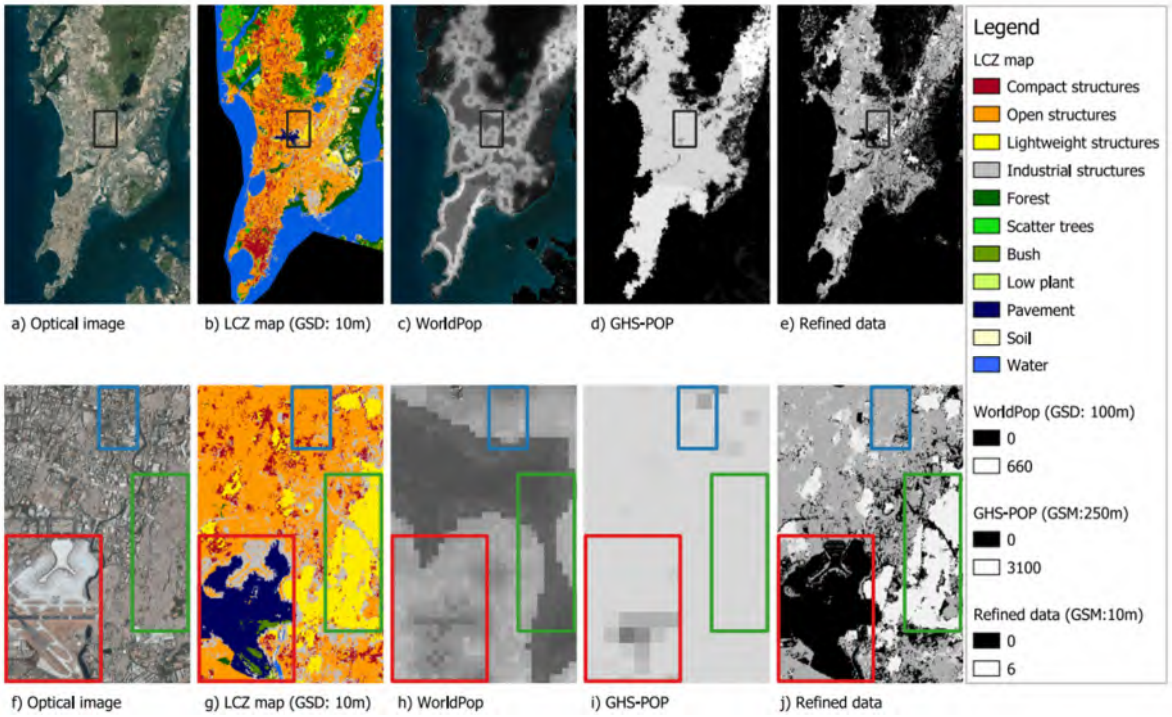


Fig. 6. This figure demonstrates the LCZ map, the WorldPop population data (Tatem, 2017), the GHS-POP population data (Schiavina et al., 2019), and our refined population data for the city of Mumbai. The images in the first-row show: a) the optical image, b) the LCZ map, c) the WorldPop data, d) the GHS-POP data, and e) our refined data of Mumbai. In the second row, we detail the airport district in Mumbai to demonstrate the spatial refinement based on our LCZ classification against the state-of-the-art population data. In the zoomed images, the blue, red, and green rectangles indicate an open area, an airport, and a compact area, respectively. GSD stands for ground sampling distance.

3.2. Land consumption of cities across the globe

Based on the results presented above, we now analyze and compare land consumption in and across our selected 40 cities. The term land consumption describing the lands that have been modified from its natural morphology by human activity are here operationalized by our indicator population density in quantitative manner. We perform the analysis at two scales, the city and the intra-urban scale.

3.2.1. City scale land consumption

At the city scale, we evaluate and compare population densities as proxy for land consumption. In Fig. 7, we compare the population densities of cities that are calculated from GHS-POP data with our refined population data, respectively. As GHS-POP data is produced based on a built-up layer that do not differ industrial and non-industrial areas, we see the effect that population densities are systematically underestimated by compared to our population refinement. According to the estimation from GHS-POP data, the city Hong Kong has the highest population density among the sample 40 cities. However, after the redistribution of population based on our LCZ classification on intraurban scale, we find the city Mumbai taking the first place with 36,741 people/km^2 and Hong Kong slides into the second place with 35,881 people/km^2.

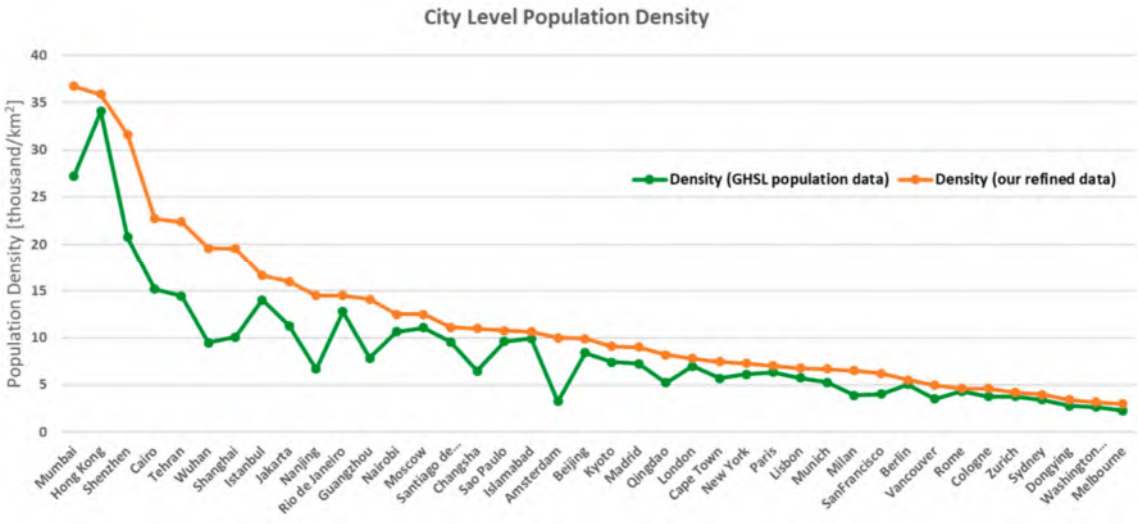


Fig. 7. The per-city population density chart. The densities in green are calculated from the original GHSL population data. The densities in orange are calculated from the refined population data whose spatial distribution at the intra-urban scale is improved.

By our analysis, we generally find the urban land consumption is differing immensely across the globe. At the city scale shown in Fig. 7, Mumbai has a population density of 36,741 people/km^2 which is twelve times denser than Melbourne (2981 people/km^2). Moreover, regarding the urban land area, we found Hong Kong the second most dense city consuming at the same time the least urban land. In contrary, the city of Melbourne is the least dense city, yet, its urban land area is comparably large, ranking at seventh place. Rankings of urban land areas and population are shown in Fig. 11 and Fig. 12 in Appendix C, respectively.

Table 1. The distribution of cities in the Global North and the Global South in terms of land area, population, and population density.

Number of cities		First half in ranking	Second half in ranking
Land area	Global North	9	10
	Global South	11	10
Population	Global North	5	14
	Global South	15	6
Population density	Global North	2	17
	Global South	18	3

It is interesting to relate the land consumption data to geographic regions. We conduct a coarse analysis on the correlation between economy and land consumption by dividing the 40 sample cities into two groups: Global North and Global South. Understanding the Global South as geographic regions with a less developed economy (Therien, 1999), we find 21 cities among our samples belonging to this group and 19 cities to the Global North. We split the rankings of land use, population, and population density into top and bottom halves. For both city groups, the numbers of cities that fall into either half of the three rankings are counted. The results in Table 1 show a specificity for certain geographical regions: 1) cities

among our samples in the Global South have larger population than cities in the Global North, 2) cities in the Global South have significantly higher population densities as 18 out of 20 cities in the first half of the population density ranking are Global South cities, and 3) Land use, however, shows no difference across these geographic regions. It can be deduced that better economies show higher per capita land use.

3.2.2. Intra-urban scale land consumption

At the city scale analysis, we found the densest city, Mumbai, has a population density of 36,741 people/km^2 (Fig. 7). However, by analyzing on an intra-urban scale, it becomes clear that these mean values on a citywide scale, cannot represent the local maximum or minimum. As we demonstrate in Fig. 8, we found even higher population densities at intraurban scale among the sample cities: the lightweight areas of Nairobi and Mumbai, as well as the compact areas of Hong Kong and Mumbai all have population densities larger than 40,000 people/km^2, significantly denser than the mean values on city scale. Among all areas in the 40 cities, the lightweight class, equivalent to complex, high-density slum structures, of Nairobi are the densest areas with a population density of 48,241 people/km^2 (20.73 m^2 land per person). In this particular example, Nairobi has at the city scale, a density of a comparatively low number of 12,409 people/km^2 (80.06 m^2 land per person) and ranks only the 13th place. The difference between the extreme value and the average in this case is fourfold. We observe these differences for many cities among our sample. This shows impressively that generalized land consumption values at the city level cannot do justice to the complex and spatially highly variable land consumption. We therefore propagate that a spatial differentiation to intra-urban scale is absolutely necessary to use this indicator with high significance.

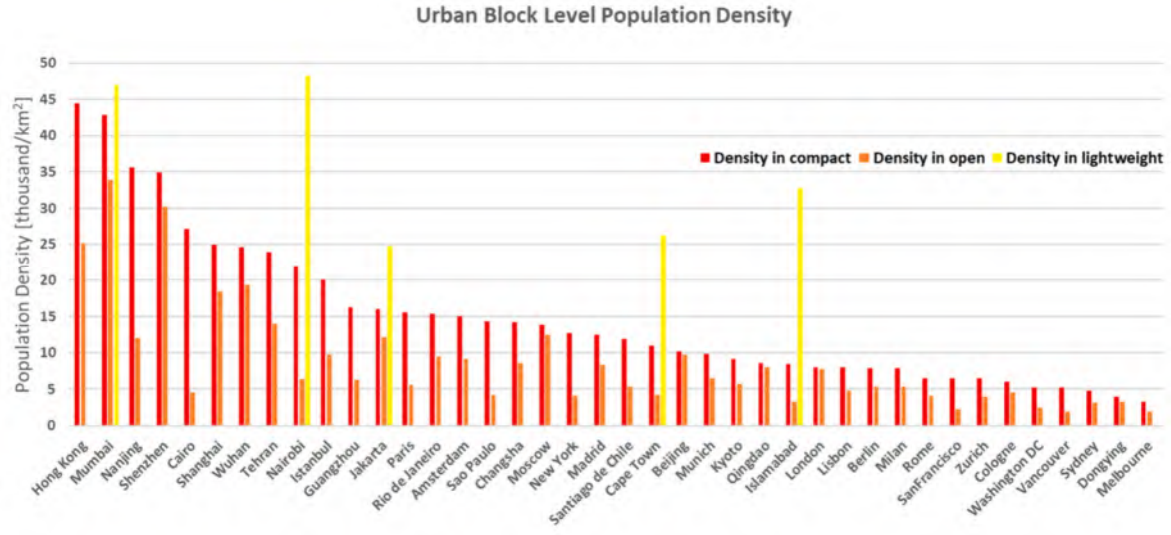


Fig. 8. The population density of intra-urban scale morphological blocks for the 40 cities. These blocks are the compact, the open, and the lightweight blocks.

In addition to the example of Nairobi, in Fig. 8, we also find 17 urban blocks from 12 cities have population densities larger than 20,000 people/km^2 and up to 48,241 people/km^2, while 18 urban blocks from 15 cities have population densities smaller than 5000 people/km^2. The

differences are at least four times. These quantitative numbers suggest that urban land consumption is differing significantly across the globe.

We reveal the uneven distribution of land consumption within cities by differentiating the structural LCZs. To do so, we calculate the standard deviations of population densities of the intra-urban morphological blocks as an indicator. As shown in Fig. 9, we reveal Nairobi having the most uneven population distribution among our sample. In contrary, London reveals the highest homogeneity among our samples. Fig. 9 clearly shows that many cities have very uneven population distributions. This supports our argument that an intra-urban resolution is required for a reasonable assessment of urban land consumption.

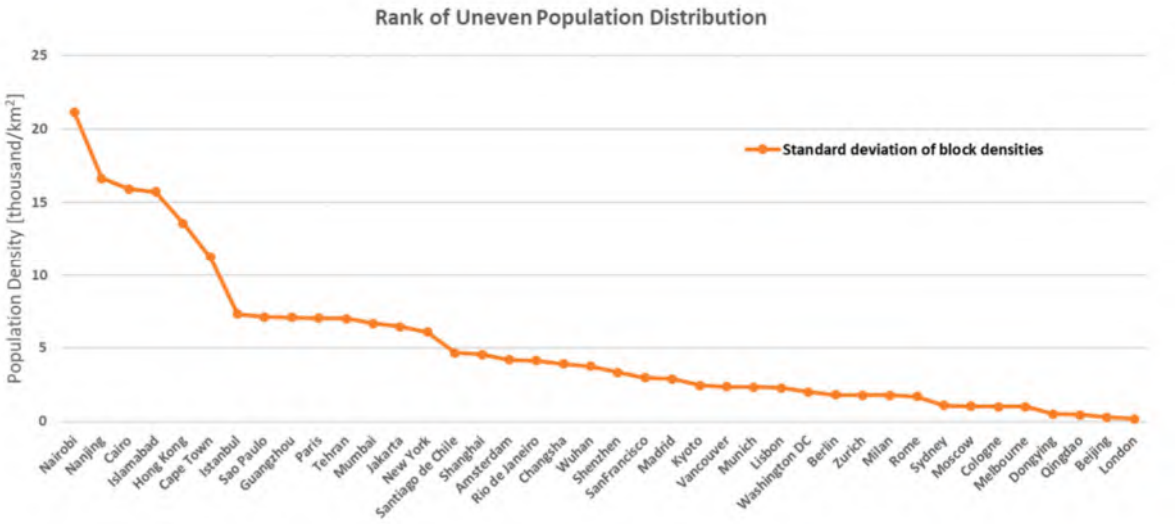


Fig. 9. Unevenness of population distribution within cities indicated by the standard deviation of population densities of intra-urban blocks.

3.3. Morphological types of cities and population density

3.3.1. Morphological types of cities

Based on our proposed clustering approach, we reveal six morphological types of cites among the 40 sample cities. We identified prominent characteristics for each of the six types, by comparing the cluster centers in Table 2 and inspecting their classification maps. The classification maps of representative cities of each type are illustrated in Fig. 13 in Appendix D. We understand the term "structure" here on an intra-urban block scale, and the term '*type*' describes the category of urban morphological configurations. For example, a city consists of 64% compact structures, 9% open structures, and 27% industrial structures; The clustering approach classifies this configuration to a compact type as in the description of urban morphological formation.

Table 2. This table shows feature values of the six cluster centers. The unit of the numbers is percentage (%).

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
Dominating LCZ	Compact	Open	Open-lightweight	Industrial	Lightweight	Compact-open
Share of compact area	63.7	9.3	19.2	13.7	56.1	45.9
Share of open area	8.7	76.9	54.3	39.6	23.6	41.1
Share of lightweight area	0	0	5.0	0	13.3	0
Share of industrial area	27.6	13.8	21.5	46.7	6.9	12.9

Among our six city types, the *first type* features a group of cities which have predominately compact structures. An average 63.7% of their land have been identified as compact structures. 27.6% of their land are industrial structures. The *second type* groups 11 cities where open structures have a dominant proportion of 76.9%. Nine of the eleven cities are European cities. The *third type* includes the cities of Mumbai, Nairobi, and Cape Town. Their prominent characteristics are 54.3% of open structure and 5% of lightweight structures. The *fourth type* includes eight cities and has an average of 46.7% industrial structures. Six out of the eight cities in this type are from China. The *fifth type* includes only one city, Islamabad. Its specifics are 13.3% are lightweight structures, which is significantly larger than in any other cities among the 40 cities. The *sixth type* is characterized by 45.9% of compact structures and 41.1% of open structures. It is an equal combination of the two structures. To sum up, we have identified six types of urban morphological configurations which are: compact type, open type, open-lightweight type, industrial type, lightweight type, and compact-open type.

3.3.2. Morphological types of cities and population density

In the following, we investigate whether different types of cities, which therefore also have different morphological compositions, feature different types of land consumption. To do so, we analyze the correlation between the clustered morphological types of cities and population densities. We reorganize population densities in terms of the six city types as shown in Fig. 10. Since Islamabad is the only city categorized in the lightweight type, it is analyzed with cities of the open-lightweight type which is the most similar type to it.

For cross type analysis, the mean population density of each urban type is calculated and shown in the subplot a) of Fig. 10. This plot demonstrates a decreasing order of population density from types of open-lightweight, industrial, compact-open, compact, to open. It is interesting in this sequences that the compact type locates between types of open and compact-open. By observing the densities in the compact-open group in subplot d) of Fig. 10, we found that, there is an exceptional sample: Hong Kong. By comparing to Istanbul, the second most dense city in the same type, Hong Kong has a density value twice as large as Istanbul. After excluding Hong Kong, the decreasing order be comes open-lightweight, industrial, compact, compact-open, and open. The order of compact, compact-open, and open basically confirms that urban structural compactness is positively related to population density. However, the positive correlation is only valid from the statistical perspective of the mean population densities across types but not for individual cities. For example, Melbourne belongs to the compact type as in subplot c) of Fig. 10, but it is the least dense city among the 40 cities.

Within the cities of compact type, it appears that the cities of low population density are western cities (Melbourne, Vancouver, and San Fran cisco). The same situation can also be observed in the compact-open type, as in subplot d) of Fig. 10, including Sydney, Lisbon, and New York. For the open type in subplot e) of Fig. 10, all cities have a low population density, ten of the eleven cities are western cities (nine from Europe and one from North America). The industrial type rather presents the industrial functionality of cities with an average of 46.7% land as industrial structures. Six of the eight cities are from China. There

is no significant evidence for correlation with population density. As shown in subplot **b)** of Fig. 10, the open-lightweight and lightweight types have an outstanding finger print which features a very high population density in the lightweight blocks.

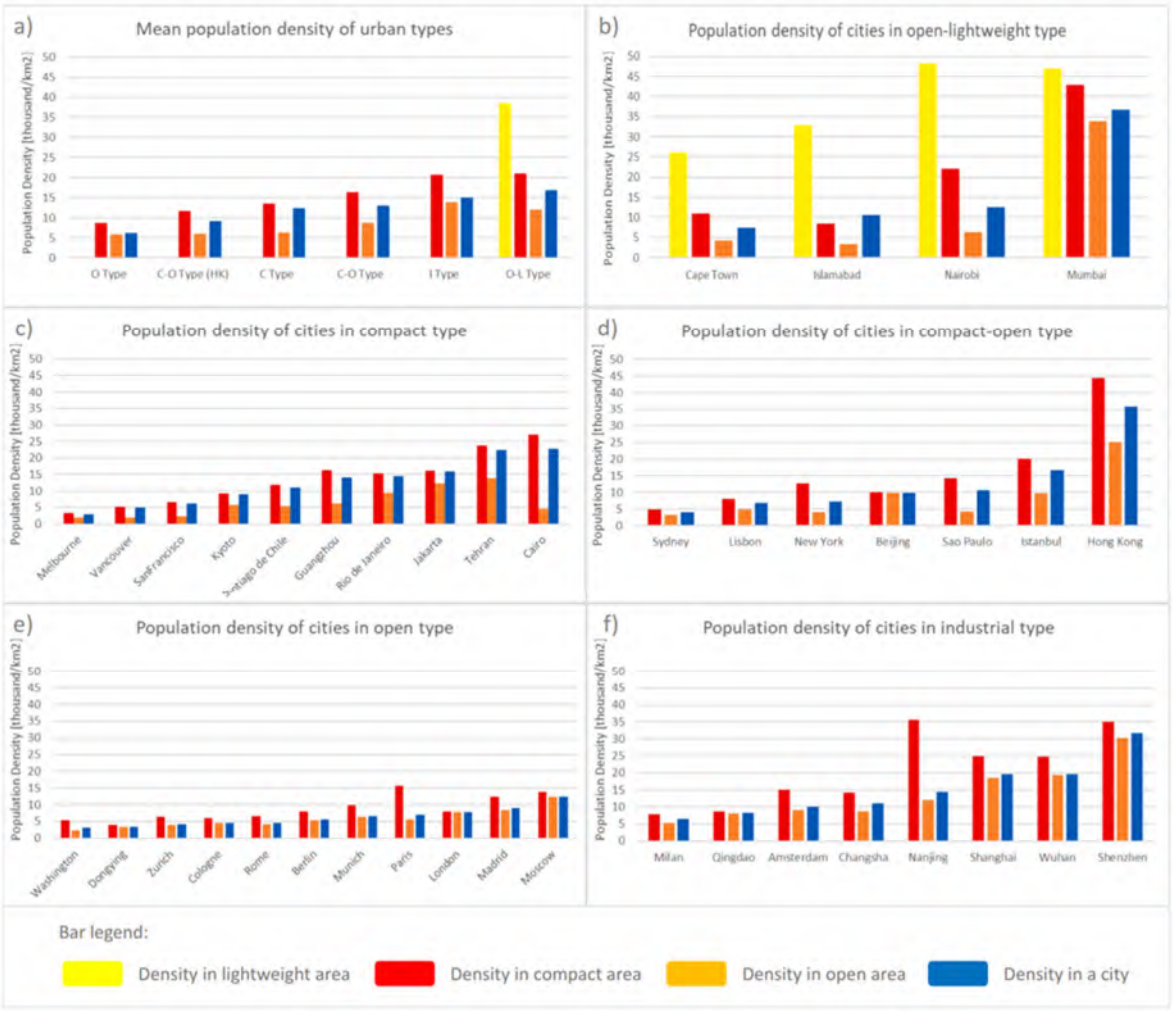


Fig. 10. Population densities of the sample cities organized by the six morphological types. The lightweight type includes only Islamabad so that its density is shown with cities of the open-lightweight type. In a) mean density of city types, O, C–O, C, I, and O-L stand for open, compact-open, compact, industrial, and open lightweight respectively. C–O type (HK) indicates the mean population density of the compact-open type excluding Hong Kong. The density unit is thousand people per square kilometer. The cities in each subplot are sorted by an increasing order of city scale population density.

4. Discussion

In this study, we take advantage of the remote sensing data to study urban land consumption at the global scale with an intra-urban spatial resolution. This approach allows the study of land consumption based on a consistent data at a high spatial resolution which allows to picture the spatial variability often lost in aggregated mean values.

Based on our study, we argue that monitoring the global urban land consumption or global urban sustainable development should proceed to an intra-urban resolution. Because our findings refer that per-city wise land consumption indicators have estimated values that severely deviate from the real situation. We also found that urban land consumption is

differing significantly across the globe. For example, a citizen living in slum areas of Nairobi has 21 m^2 lands for living while one could have 530 m^2 in Melbourne. It is also immensely differing within a city. For example, the population density in lightweight areas of Nairobi is four times denser than the average density of the city. Moreover, we found six types of urban morphological formations among the 40 cities. From a cross type perspective, we reveal a positive correlation between the compactness of urban morphology and population density. We find western cities generally have a lower population density.

In the following we want to critically discuss these main outcomes of our study from various perspectives – conceptual, methodological and geographical:

From a *conceptual aspect*, urban land consumption generally refers to the utilization of land resources due to human activities in cities. Existing studies treat land consumption as modification of natural land to human settlements at city scale (Melchiorri et al., 2018, 2019; Taubenböck et al., 2019). Those studies focus on the aspect of urban land usages. Differ from this we associate land consumption to demo graphical information. And by using the city as well as the intra-urban scale, we unveil a spatially more detailed variability of land consumption, and thus situations of urban inhabitants. We argue the intra-urban scale using morphological descriptions allows to break spatial limitation of administrative boundaries. The aggregated population density analyses inherited from official population data are thus improved. How ever, there still exist limitations in this domain: Due to the functional information in LCZs, urban inhabitants were not distributed to industrial areas, the redistribution of population is still limited due to the definition of our classification scheme to three aggregated urban morpho logical residential areas: compact, open, and lightweight structures. It is in our argumentation indeed a more realistic picture of the city, but neither absolutely correct nor spatially or thematically detailed enough. There are still buildings in these three structures that are not constructed for the residential purpose. Accurate mapping of residential buildings / areas is still a major challenge for remote sensing technology. Extra data sources from other fields are complimentary to remote sensing data (Häberle et al., 2019; Hoffmann et al., 2019; Huang et al., 2018).

which may allow improved geoinformation in the future. The second limitation is the estimation of living spaces. Due to the missing height information in our baseline remote sensing data the available living spaces remain only indirectly assessed by our LCZ classification scheme. Such detailed information requires an advanced accurate three-dimensional classification. Currently, three-dimensional classification of cities across the globe is restricted by limited access to accurate and consistent global data of height information. Even with these limitations, we believe this study still reveals new knowledge about the land consumption status of our sampled 40 cities at the intra-urban scale.

From a *methodological aspect*, of course, the classification inaccuracies involve limitations. As shown in Fig. 5, the classification accuracy of Mumbai ranges from 59.4% to 72.2%. This is not perfect for the purpose of our applications and the effects for the land consumption analysis remain unknown. It needs to be pointed out that this accuracy is an outcome of a strict evaluation strategy revealing an authentic performance of an algorithm. Although its accuracy is not perfect, currently, its performance is still the state-of-the-art for large area or global applications. With respect to our application, this means any mis-

classification will lead to a questionable population redistribution. In general, however, we find our reached accuracy sufficient and the plausibility of our results confirm this. As a second limitation we refer to the difficulty of model transfer of the classification module. There are two classification strategies for our classification task. The first is to train a model on the reference data of the 40 cities, and produce a classification map for any city in the world with this model. The advantage is that one can produce maps for many cities, but the drawback is that the accuracy of those maps is low. The second strategy is to train models on the reference data and only produce maps for these 40 cities. The advantage is the accuracy of these 40 maps is high, but the number of cities is limited. Although some studies (Demuzere et al., 2019; Tong et al., 2020; Yokoya et al., 2017; Zheng et al., 2020) attempt to improve the accuracy of the first strategy, the second strategy is still significantly more accurate than the first strategy. In the context of considering the trade-off in the two strategies, we choose the accurate solution. Thus, our global sample of 40 cities allows a more accurate but possibly not fully comprehensive view on land consumption.

From a geographical point of view, we find a huge difference in urban land consumption across the globe. From 45,000 people/km^2 in some lightweight areas of cities in the Global South to only 1800 people/ km^2 in some cities of the Global North. The limitation is that we cannot provide quantitative accuracy measures for these numbers as reference data on this spatial level are inexistent. However, these numbers seem to be very plausible. In any case, we see that the land consumption is highly variable across cities and geographic regions. Therefore, this analysis allows to reflect on urban land consumption. It might support the analysis of the efficiency of land use with respect to other domains.

5. Conclusion and outlook

In this work, we have developed a classification system that fuses Sentinel-1 and Sentinel-2 data using a state-of-the-art semi-supervised technique and produces LCZ classification maps by a topological enhanced ensemble classification technique. With the classification system, LCZ maps of 40 cities around the world are produced. The urban morphological information of LCZ maps allowed us to improve the spatial details of a global population data (GHS-POP) for the 40 cities to an intra-urban scale. With the refined data, a new ranking of the population density of the 40 cities is provided which mitigates the under estimation in GHS-POP caused by distributing population over industrial regions. The cross-city population density analysis reveals that the urban land consumption differs immensely across the globe. Beyond, an intra-urban scale land consumption analysis is conducted, which reveals information that is missing in a city scale analysis. We found that, for a same city, an urban block scale population density can be four times larger than the city scale population density. The unevenness of population distribution within cities is also analyzed. Nairobi is reported to be the most unevenly distributed city. Six urban morphological types are identified among the 40 cities. It is found that the compactness of urban morphology has a positive correlation with population density from a cross type perspective. These findings are only possible as we worked on an intra-urban spatial resolution. Therefore, we argue that monitoring the urban land consumption or urban sustainable development at global scale should proceed to an intra-urban resolution.

Overall, this analysis allows to assess the unevenness of land consumption within and across different cities as well as in different regions across the globe. In a world which is constantly and dynamically urbanizing, this work allows to reveal urban land consumption, to support societal development, and to reflect sustainability.

Acknowledgments

The work is jointly supported by the European Research Council (ERC) under the European Union's Horizon 2020 - Research and Inno vation Framework Programme (grant agreement No. [ERC-2016-StG- 714087], Acronym: So2Sat), by the Helmholtz Association through the Framework of Helmholtz AI - Local Unit "Munich Unit @Aeronautics, Space and Transport (MASTr)" and Helmholtz Excellent Professorship "Data Science in Earth Observation – Big Data Fusion for Urban Research".

Appendix

Supplementary material to this article can be found online at

https://doi.org/10.1016/j.cities.2021.103163

Chapter XXI

How we live and what that means

A character study with data from space

IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021

Hannes Taubenböck

© IEEE

Note: A payment was made to CCC RightsLink to reuse the paper in this report.

https://doi.org/10.1109/IGARSS47720.2021.9553084

Published 21 Oct 2021

Abstract

How we develop and shape the space where we live and do business determines different living conditions. How ecologically, socially or economically reasonable and resilient these structures are, however, is not always easy to answer. Geodata are mandatory for an evaluation and assessment. Remote sensing with its diverse data allows us to record objects on the Earth's surface, the topography and the atmosphere. And with it, it becomes possible to describe different living environments in high spatial resolution and in a quantitative manner. Spatial indicators from these data can make developments tangible. In this paper, the focus is on what it means how we live with respect to land consumption. Based on a land-cover classification using Sentinel-2 data as well as on population data from the national census for Germany, it is evaluated what different spatial structures mean in terms of land consumption. Various scenarios are developed to evaluate the impact of settlement types on land consumption. In this sense the aim is to discuss how reasonable and resilient our way of living is.

Keywords

Remote sensing; urban geography; spatial metrics; spatial planning

1. Introduction

How we organize space and our living environments has immense impact: e.g. in economic (Glaeser, 2010), social (Tonkiss, 2013) or ecologic (Artmann, Inostroza & Peilei, 2019) terms, among others. Across our planet, a wide variety of processes are defining how built space, i.e. residential, commercial, industrially or traffic areas, is developed. Various political ideas, concepts of planning and design (Kostof, 1991) are alongside economic interests (Montgomery, 2013), or informal developments (Taubenböck, Kraff & Wurm, 2018) which shape the land surface.

While the effects of differing processes are manifold, this paper aims to exemplify what different forms of built-up density mean with respect to land consumption. The term '*land consumption*' is generally defined as natural lands modified to man-made structures due to living, social, and economic purposes (Boyce, 1963). Land consumption has been studied using various indicators such as the amount of urban land consumed per unit of gross domestic product (GDP) (Kuang et al., 2020) or as a function of income and housing prices (Wang et al., 2020). From a demographic perspective, the intensity of land consumption is understood as the relation of the built environment to the amount of population (Glaeser, 2010).

Processes of global change, above all highly dynamic processes of urbanization and industrialization (Taubenböck et al., 2012), have intensified the rates of land consumption in recent decades. For measurement and monitoring, large area, accurate and up-to-date geodata are in demand to be able to estimate land consumption at all. Remote sensing with its diverse data has developed in recent years to a valuable data source to record the Earth's surface, the topography and the atmosphere. In the research domain of land consumption, remote sensing data have already been applied for this purpose (e.g. Wang et al., 2020; Hu et al., 2021, among others).

In this study, the higher-ranking goal is to determine land consumption rates with respect to various spatial land patterns. Based on remote sensing and census data, the idea is to determine the real land consumption for the study site of entire Germany as well as for certain types of settlement patterns. However, the evaluation whether land consumption is high or not, is challenging. Here scenarios are developed to quantitatively measure what land consumption in '*urban*' compared to 'rural' lifestyles entail.

2. Data

This study relies on the one hand on an area-wide landcover classification for Germany derived from optical Sentinel-2 data (ESA – European Space Agency, 2020) using a mosaic of scenes from years 2015 to 2017. It consists of seven thematic land cover classes (artificial land, open land, water as well as low and high seasonal, and low and high perennial vegetation) (Fig. 1a). Land Use and Coverage Area frame Survey (LUCAS) in-situ reference data (Eurostat, 2015) were used for training a random forest classifier. The classifier has shown to provide sufficient results in highly dimensional feature spaces. Accuracy has been reported at 93.1%. For methodological details see Weigand et al. (2020).

On the other hand, this study bases on population data of the census (DeStatis, 2020) as well as on administrative boundaries as provided by the German Federal Statistical Office (DeStatis, 2020b) for Germany.

3. Methods

The basic idea in this study is to analyze the effects of different types of settlements on land consumption. The current land consumption is used as a reference.

As a starting point, the area shares of impervious surfaces are quantified for the administrative areas of German municipalities. This is put in relation to the population of the respective areas for per capita impervious surfaces shares. On this basis, it is possible to determine how unevenly land consumption per capita is distributed across space.

In a second step, the total impervious surfaces in Germany are spatially cumulated. The evaluation whether land consumption is high or not, however, is not obvious. Therefore, we first look at which settlement forms have which effects on land consumption. To do this, we divide the total population of Germany, i.e. 82.8 million people, into three equal parts. Then three groups consisting of 27.6 million people each are assigned to the municipalities. For this purpose, the municipalities with the highest number of inhabitants are added together until the 27.6 million are reached. From then on, descending from the remaining municipalities, the population figures for the second group are also added up and finally, following the same pattern, the third. For these groups, again the impervious surfaces are cumulated. This allows to evaluate and compare which effects different population numbers in municipalities have on land consumption.

In a third step, a theoretical experiment is carried out: the average impervious surface proportion for municipalities that have very small populations, i.e. between 500 and 1000 people, is calculated. In the same way, the average proportion of impervious surfaces for inner city areas is calculated. These examples are considered representative of a very urban, dense landscape configuration as well as a very rural, low-density landscape. Based on the empirical values determined, it can then be extrapolated for the theoretical case that the entire population lives either very urban or very rural. Applying determined population densities and the respective degrees of impervious surfaces for the two theoretical cases, i.e. if the entire population would live this way, the extrapolation is done in linear form. With it, it can be assessed how much land would be in theory necessary if Germany would change its settlement patterns. The results are then the basis for a discussion on lifestyle resilience with regard to land consumption.

4. Results

In general, the results reveal how unevenly land consumption, understood as the relation of the built environment to the amount of population, is distributed across space in Germany. Figure 1 clearly illustrates the intensity of use: Metropolises or cities show low figures, i.e. per capita land is used very effectively in this respect. Striking are Hamburg in the north, Berlin in the east, Munich in the south, in the southwest the Stuttgart-Karlsruhe area, as well

as the Rhein-Main area around Frankfurt and the Rhein-Ruhr area stretching from Cologne to Duisburg and to Dortmund.

The second issue that becomes clear in Fig. 1 is that land consumption per capita is very high, especially in eastern Germany. It can be assumed that high out-migration rates cause this ratio.

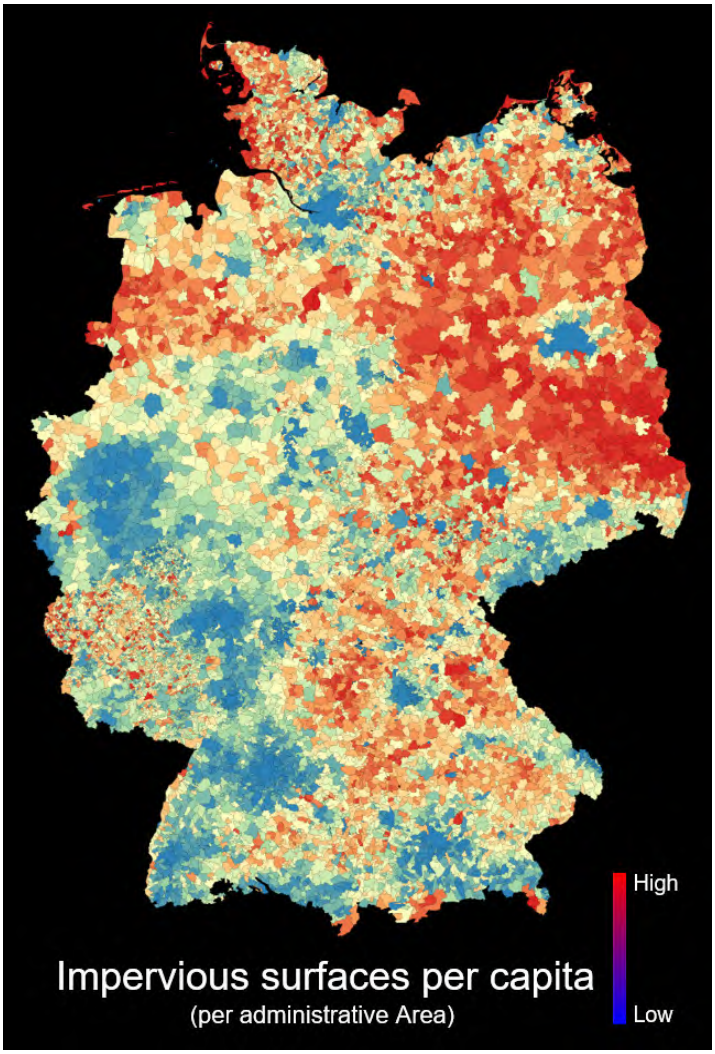


Fig. 1. Impervious surfaces per capita per administrative area, i.e. municipalities, for Germany.

Results also show how fundamentally important it is to be clear about how you build and live as a society. Currently 7.2% of the entire area of Germany are covered by impervious surfaces such as buildings, streets and the like (Fig. 2b). Does this make Germany effective in land consumption or not? This question cannot be answered with pure numbers. In order to be able to evaluate this, a thought experiment may help: As shown in Fig. 1, impervious surfaces are not evenly distributed across Germany. If one extrapolates now linearly the following in this thought experiment, it becomes clear what different types of structural cohabitation mean: if the entire population of Germany, for example, were to live as they do in low dense, small villages, statistically over 42% of the land would be impervious surfaces (Fig. 1a). The opposite effect becomes clear if one extrapolates how much land would be impervious surfaces if everyone lived like in dense inner cities. Then only 2.8% of the area

would be impervious surfaces (Fig. 2a). This shows the effects this has on land consumption, on the loss of natural space, but also in terms of resource consumption.

But even without this theoretical thought experiment, the effects can be made clear: If we distribute the thirds of the German population among the municipalities according to population, we see the following: one third, i.e. 27.6 million people, is living in the 80 largest German cities. These have 100,000 inhabitants or more and are thus classified as major cities in Germany. Their spatial share of impervious surfaces, is 1.3% (Fig. 2c). One third is living in 10,180 municipalities of less than 14,000 inhabitants. The same amount of people needs 3.4%, so considerably more. The third group in between lives in 1,000 cities between 14,000 and 100,000 inhabitants and they need 2.5% of impervious surfaces. This division into thirds shows the effects that city sizes have on land consumption.

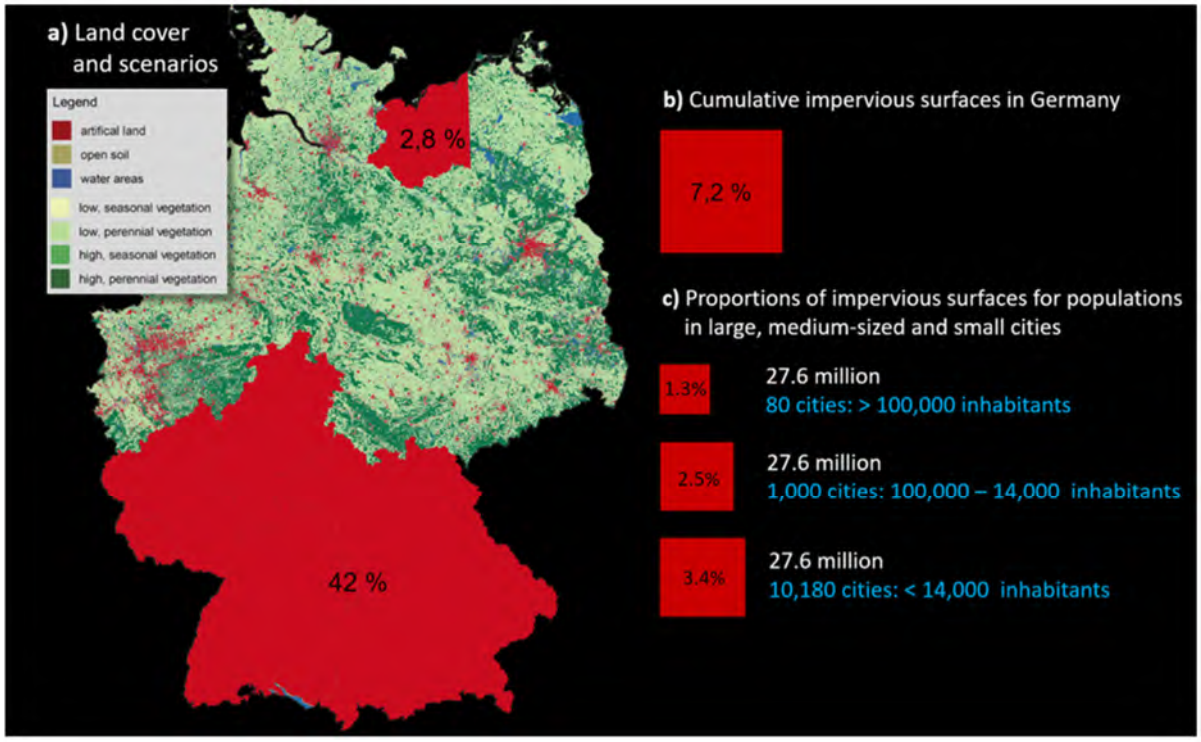


Fig. 2. a) Land cover classification on Germany based on Sentinel-2 data and two theoretical scenarios: if the entire population would live as in typical inner-city areas only 2.8% impervious surfaces would be needed; if the entire population would live as in typical small villages 42% impervious surfaces would be needed; b) Share of cumulative areas of measured impervious surfaces for Germany; c) proportions of impervious surfaces for the population in large, medium-sized and small cities.

5. Discussion

How resilient is our way of life? Although the term ‘resilience’ seems to be not well defined (Meerow, Newell & Stults, 2016), we understand it here as the capability to adapt to stress or risks within a system. In times of climate change, population growth and migration, urbanization and industrialization, among others the ecological system of our planet is coming more and more under stress. Increasing land consumption is one consequence of these change processes.

Land consumption has been the focuses of this study. It becomes clear, especially through the thought experiment carried out, how immense the effects of different types of settlement are. This basically shows that we have it in our hands to lose less natural space if we implement appropriate building structures.

However, it is not easy to say whether land consumption – in our example in Germany – is effective or not. If one relates the land consumption of the German capital Berlin to major cities across the world, as done by the study of Hu et al. (2021), it becomes clear that urban land consumption is relatively high. In other words, many cities across the globe are denser than the largest German city. At the same time, however, if one relates the urban living environments of Berlin to many other areas in Germany, the capital is very effective with respect to land consumption (Fig. 1). These two approaches to evaluation reveal:

Everything is relative and every evaluation is, naturally, only valid in its specific relation. The study, however, makes one issue very clear: living closer together makes from an ecological point of view sense. Land consumption per capita is lower, building materials can be saved, there is less loss of natural space, people live on average in smaller apartments.

And, it has been shown by different studies that dense cohabitation makes also sense from other points of views: higher wages, lower infrastructure costs, higher education (Glaeser, 2010; Bettencourt, 2007). But of course, we have to bear in mind that nothing is black or white. Dense cohabitation also goes along side negative aspects, e.g. infections, AIDS diseases or crime rates scale with a superlinear factor (Bettencourt, 2007). This simple regularity, which goes hand in hand with city sizes and dense cohabitation, is a central building block also in the political discourse to support resilient settlement forms with empirical analyses.

6. Conclusion

In times of highest urbanization dynamics (Kuang et al., 2020), spatial effects such as urban sprawl (Nechyba & Walsh, 2004), informal developments (Taubenböck, Kraff & Wurm, 2018), expansion of urban areas beyond administrative boundaries (Taubenböck et al., 2019), among many other aspects are fundamental drivers and effects of global change. The goal of remote sensing must therefore be to develop not only reliable classification algorithms but also meaningful indicators from its constantly growing, large-scale, up-to-date and consistent data. The spatial data have great potential to make social developments tangible and thus to highlight developments and bring them into the political discourse.

Part VII

Urban adaptation

CHAPTER XXII	Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes	503
CHAPTER XXIII	Cost estimation for the monitoring instrumentation of landslide early warning systems	519
CHAPTER XXIV	Was global urbanization from 1985 to 2015 efficient in terms of land consumption?	545

Chapter XXII

Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes

Nature communications, 2023, vol. 14, 2903

Emanuele Massaro, Rossano Schifanella, Matteo Piccardo, Luca Caporaso,
Hannes Taubenböck, Alessandro Cescatti & Gregory Duveiller

© The Author(s) 2023. This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

https://doi.org/10.1038/s41467-023-38596-1

Received 26 July 2022; Accepted 9 May 2023; Published 22 May 2023

Abstract

The population experiencing high temperatures in cities is rising due to anthropogenic climate change, settlement expansion, and population growth. Yet, efficient tools to evaluate potential intervention strategies to reduce population exposure to Land Surface Temperature (LST) extremes are still lacking. Here, we implement a spatial regression model based on remote sensing data that is able to assess the population exposure to LST extremes in urban environments across 200 cities based on surface properties like vegetation cover and distance to water bodies. We define exposure as the number of days per year where LST exceeds a given threshold multiplied by the total urban population exposed, in person day. Our findings reveal that urban vegetation plays a considerable role in decreasing the exposure of the urban population to LST extremes. We show that targeting high-exposure areas reduces vegetation needed for the same decrease in exposure compared to uniform treatment.

Keywords

Surface temperature; urban pattern; greening; climate change; risk mitigation

1. Introduction

Over half of the world's population lives in cities (United Nations, 2018), and thus in localized hotspots covering less than 3% of the Earth's land surface (Liu et al., 2014). Moreover, the number of urban dwellers is predicted to grow by an additional 2.5 billion by 2050 (Talukder et al., 2015). The large and increasing share of the urban population had stimulated renewed attention to studies that investigate the livability of urban environments (Binder, Wyss & Massaro, 2020). Urban areas are places where humans have altered their local environment and local climate in the most radical way, replacing the natural land cover and vegetation with impervious materials that have lower albedo and heat capacity (Kalnay & Cai, 2003), substantially reducing the cooling effect of water evaporation (Jacobs et al., 2020). Urban ecosystems can be described by having fundamentally three different types of elements: (i) impermeable surfaces, (ii) vegetated surfaces, and (iii) water bodies. The spatial integration of these three elements determines the physical properties of the urban landscape surface, thereby affecting both the microclimate of cities and the frequency of extreme heat (Chatzidimitriou & Yannas, 2015; Krayenhoff et al. 2018; Broadbent, Krayenhoff& Georgescu, 2020).

The combination of global warming and the urbanization process is ratcheting up the number of humans exposed to health endangering heat. This exposure has tripled in recent decades, a faster rise than previous research suggested (Tuholske et al., 2021). As a consequence, heat waves in urban climates have a profound impact on humankind as the climate warms up (Geo et al., 2017) and this will increase in the future (Tong et al., 2021). Regional studies confirm that the population is foreseen to be increasingly exposed to weather-related hazards and, in particular, to heat waves (Forzieri et al., 2017) leading to an ever-higher risk of fatalities (Asseng et al., 2021; Mora et al., 2017). Important aspects of psycho-physical well-being and quality of life crucially depend on the climatic conditions of the environment in which people live (Raymond, Matthews & Horton, 2020). Changes in the global temperature can generate a wide range of consequences for human health. Direct consequences are linked to the bioclimatic well-being of people, whereas indirect consequences are linked to the complex interactions between environmental conditions and the spread of infectious (Patz, 2002) and allergic diseases (Shea et al., 2008). The positive effects of climate-related human well-being are many (Iungman et al., 2023): from the reduction of the risk of numerous chronic diseases in adulthood (such as diabetes and cardiovascular conditions, obesity, and asthma), to the acceleration of recovery after surgery, to the reduction of hospital admissions and premature mortality, up to better pregnancy outcomes and improved mental health (Kingsley & Ontario, 2019). To achieve an intelligent and sustainable development of the urban fabric, there is the necessity to move away from the traditional approach of territorial planning that ignores social and environmental factors (Krueger & Gibbs, 2007). Instead, there is the need to work together with a common goal, using shared methodologies and regulations to handle the impacts of climate change on city growth and development (Ziter et al., 2019). To manage the rapid increase of these direct and indirect climate risks in urban environments, there is a need for knowledge-based policies that foster the design, application, and monitoring of adaptation plans. For this scope, new research efforts are needed to (i) increase the understanding of the specific causes of extreme heat, and to (ii)

offer monitoring and modeling systems of the urban thermal environment at the required spatial and temporal scales. To cope with the increasing need to foster climate mitigation and adaptation, it is therefore important to have spatially and temporally detailed information, possibly at a sub-daily temporal frequency. So far, the lack of adequate data has actually prevented the development of numerical methods to derive appropriate metrics and diagnostics to fully describe, or profile, these factors. The increasing availability of time series of thermal remote sensing information from satellite platforms is now supporting effective quantification of urban heat consistently across the globe and with high spatial detail (Voogt & Oke, 2003; Weng, 2019; Quattrochi & Luvall, 2004). Most previous studies on the subject focused on the Urban Heat Island (UHI) phenomena, or more specifically Surface Urban Heat Island (SUHI) when studied within the remote sensing field (Voogt & Oke, 2003). SUHI is defined as the difference between the urban and the rural land surface temperature (LST), and it occurs because the highly dense artificial materials of city cores heat up considerably stronger than their rural surroundings (see Fig. 1). Depending on the type of surface surrounding the city, the value of SUHI can be either positive or negative (which in the latter case is referred to as an urban cool island or UCI). When cities are surrounded by vegetated areas (see Fig. 1B), the magnitude of the SUHI can be seen as a proxy of the rural vegetation density instead of a measure of the thermal comfort of an urban environment. To make a step-change in the approach to the problem, the design and evaluation of urban adaptation plans should be city-specific and based on the detailed assessment of the city heat characteristics (Marando et al., 2022).

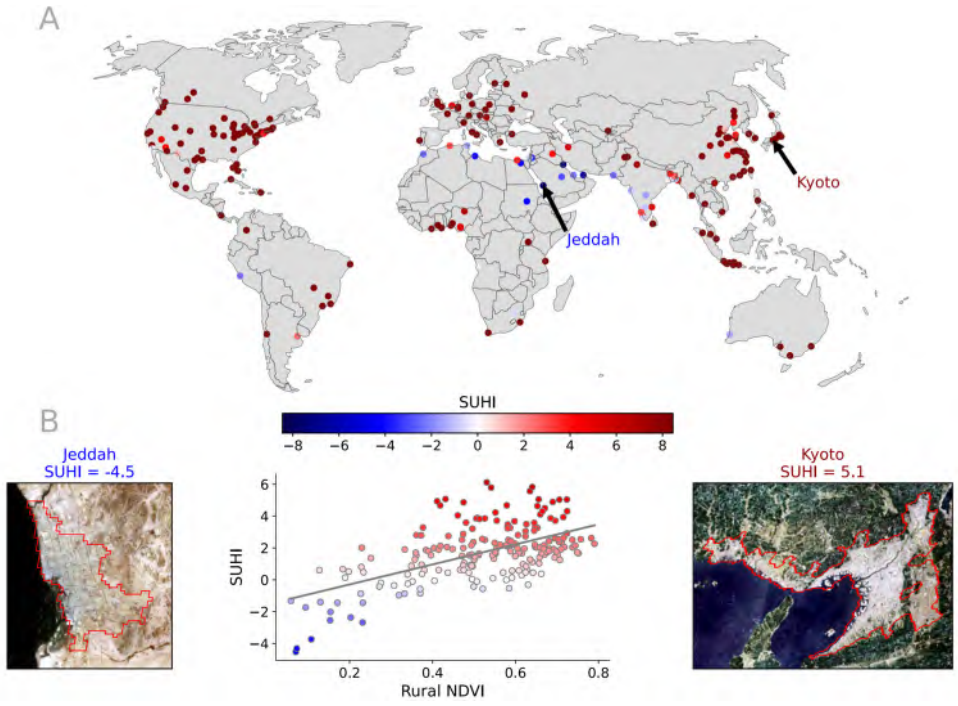


Fig. 1. Surface Urban Heat Island (SUHI). An Averaged value of SUHI between 2010 and 2020 for the 200 cities analyzed in this research. B As an example, we report the opposite cases of Kyoto and Jeddah that have positive and negative values of the SUHI, respectively. The satellite images show the difference in the vegetation of the adjacent rural areas. On the right we show the correlation between SUHI and the average value of the rural greenness expressed with the Normalized Difference Vegetation Index (NDVI) as defined in the Methods section for all the cities. See data availability and code availability for publicly available dataset and codes for generating the figures.

To assess the latter requirements, in this study, we set a spatial model to predict the absolute value of the exposure of the population to LST extremes. We present an approach based on a spatial regression model to predict the exposure of urban populations to LST extremes. Despite its simplicity, the model is able to assess the exposure of the urban population to LST extremes using only remote sensing observations. We measure exposure as the number of days and nights when LST exceeds a certain threshold, multiplied by the number of people affected: this is a surrogate version of population exposure to extreme heat using air temperature (Tuholske et al., 2021; Coffel, Horton & De Sherbinin, 2017). The spatial regression model uses greenness and proximity to water bodies as predictors and has been tested in 200 cities across various climates and has consistently proven its accuracy. Finally, we utilize our findings to assess the impact of urban greening initiatives in reducing the exposure of urban populations to LST extremes. Our research presents a globally validated model and we believe that the same methodology can also be utilized for specific, practical, and localized solutions.

2. Results

We focus the analysis on 200 cities worldwide that are characterized by a wide range of climatic conditions. These are divided equally per major climate zone: Arid, Continental, Tropical, and Temperate (see Fig. S1 in the Supplementary Information (SI)). The urban environment is described by a 1 km spatial resolution multiband raster, with several variables, where for each pixel and for each year, we have aggregated various types of information, namely: daytime and nighttime values per summer where LST exceeds a threshold, average values of the Normalized Difference Vegetation Index (NDVI), distance to water surfaces, and population (see Materials and Methods section and Fig. S2 in the SI). We apply a spatial regression model to predict the exposure of the urban population to LST extremes. In line with the recent literature, we define the urban population exposure TE as the number of days per year where LST exceeds a threshold, called exposure threshold, multiplied by the number of people, expressed in persons · days (Coffel, Horton & De Sherbinin, 2017), where days correspond to the average between daytime and nighttime values over thresholds. For each urban environment, we define the daytime/ nighttime exposure threshold as the 90th percentile of the temporal distribution of the daytime/nighttime LST over 20 years (from 2000 to 2020), as shown in Figs. S3 and S4 in the SI. By this definition it is possible to compare the results obtained in geographic areas with a different climate and in different seasons of the year (Field et al., 2012), and the 90th percentile ensures an effective definition of LST extremes, where the outliers are not taken into account as suggested by the Intergovernmental Panel on Climate Change (IPCC) (Field et al., 2012) and the World Meteorological Organization (WMO) (Klein Tank, Zwiers & Zhang, 2009). For each pixel in the images of the cities, we calculate the number of daytime and nighttime values that exceed exposure thresholds for the three warmest months of the year, as illustrated in Fig. S5 in the SI for Paris. To account for the daily cycle, we define the total exposure TE_T as the average of the daytime exposure TE_D and the nighttime exposure TE_N, i.e., $TE_T = (TE_D + TE_N)/2$, as demonstrated in Fig. S6 for Paris. In Fig. 2 we show the values of TET across space and time: in Fig. 2A we show the average values for each city in the different climate zones. In Fig. 2B we show the trends of the value of the

exposures over time. We can see a clear increment of the urban population exposure to LST extremes from 2010 to 2020 that is mostly caused by the urbanization process as observed in Fig. 2C where we show that the average exposure divided by the population is almost constant over the years for the different climate zones.

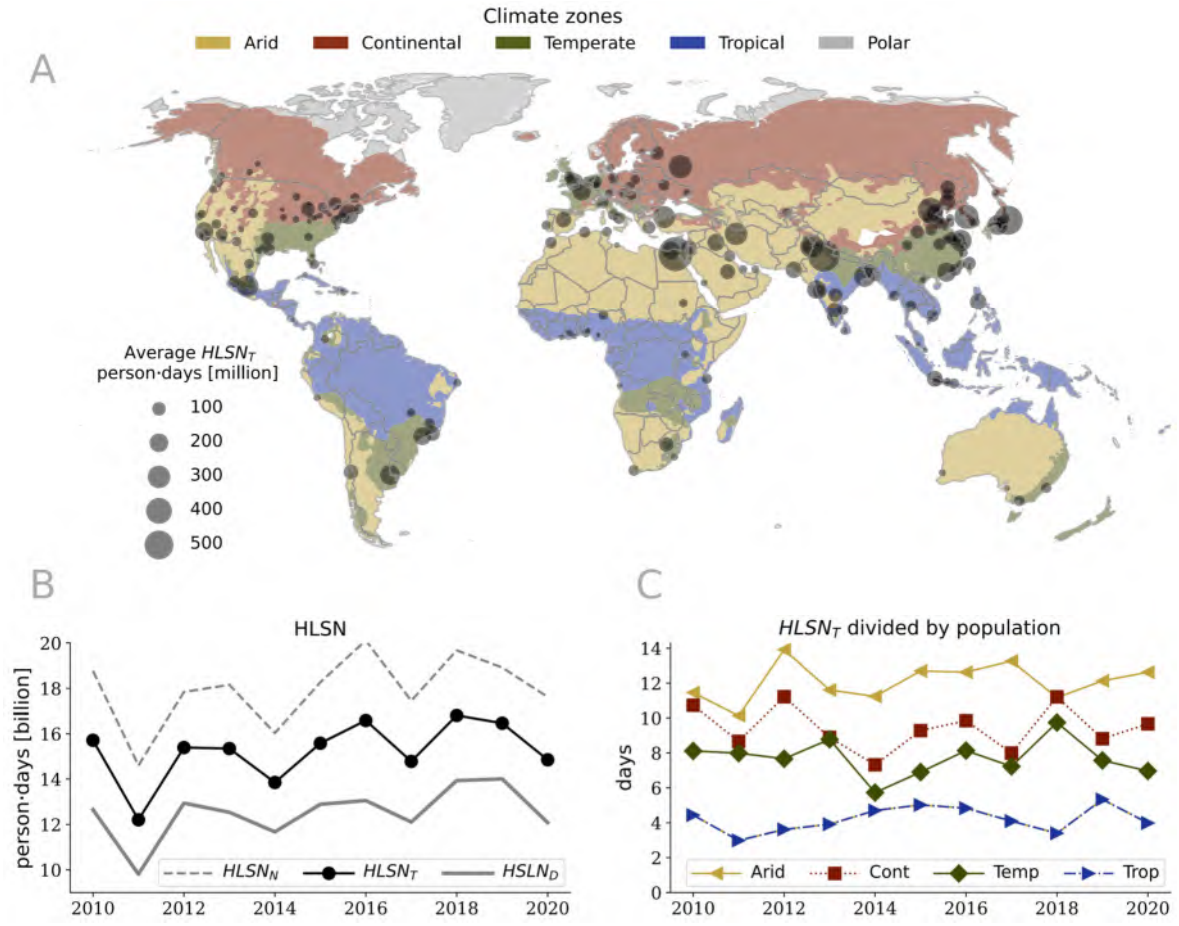


Fig. 2. Population exposure to Land Surface Temperature (LST) extremes. A Population exposure in the cities within the five climate zones defined by the Köppen-Geiger climate classification system (Beck et al., 2018) in mean number of person-days per year. The size of the dots corresponds to the population exposure averaged over the 10 years of the observations. B Values of the exposure over the years. C Average value of the exposure divided by the population that corresponds to the average number of days and nights over the thresholds for each climate zone (in the legend, Cont stays for Continental, Temp for Temperate and Trop for Tropical). See data availability and code availability for publicly available dataset and codes for generating the figures.

Urban greening to reduce the exposure of the urban population to LST extremes Before presenting the findings, it is necessary to elaborate on urban greening, including its benefits and limitations. Vegetation, including trees and other plants, has a significant cooling effect on urban environments, making it a simple and effective way to reduce the exposure of urban populations to extreme heat (Iungman et al., 2023; Marando et al., 2022). Vegetation can lower surface and air temperatures by providing shade and increasing evaporation (Meili et al., 2021). Shaded surfaces can be up to 20-45 °F (11-25 °C) cooler than unshaded surfaces (Akbari et al., 1997). Evapotranspiration from vegetation, either alone or in conjunction with shading, can help mitigate peak summer temperatures (Schwaab et al., 2021; Bonan, 2008), whereas urban surfaces covered in concrete and asphalt contribute to the opposite effect.

Due to these biophysical effects, urban vegetation is an important tool in climate-smart urban planning. Urban vegetation design is a crucial adaptation strategy that reduces heat stress by modifying the surface properties of cities. However, there is still a lack of dedicated modeling tools to aid in the design of city-specific plans based on the level of intervention required to achieve climate targets. For instance, microscale models that simulate the local impacts of urban design and land use on the urban microclimate and the potential for microclimate interventions to mitigate urban heat exposure are needed. While proximity to urban green spaces has advantages for health and cooling (Iungman et al., 2023; Vardoulakis & Kinney, 2019), cities also face competition for space and resources (Melchiorri et al., 2019; Pataki et al., 2021). In this study, we focus solely on measuring urban greening to reduce the exposure of urban populations to LST extremes. However, this is not the only solution to urban heat exposure, and other solutions, such as increasing the albedo of surfaces (Akbari, Pomerantz & Taha, 2001), are known to have positive effects. In practical applications, urban greening can effectively reduce land surface temperatures (Masoudi & Tan, 2019), however, tradeoffs with water resources must be considered (Niinemets & Peñuelas, 2008).

2.1. Limitations and advantages of the proposed approach

In this study, we quantify the exposure of the urban population to LST extremes by measuring the number of days that LST exceeds a certain threshold, multiplied by the number of people exposed: this definition is a surrogate version of exposure to extreme heat using air temperature (Tuholske et al., 2021). We acknowledge that using thermal remote sensing data as a proxy for population exposure to heat in urban environments has its own limits (Martilli, Krayenhoff & Nazarian, 2020). LST may not fully represent the thermal environment experienced by people in urban areas as it measures the temperature of the surface rather than the air (Martilli, Krayenhoff & Nazarian, 2020; Voogt & Oke, 2003). In addition, LST measurements can be affected by the urban canyon effect (Martilli, Krayenhoff & Nazarian, 2020; Voogt & Oke, 2003), and errors can arise due to the complex dependence of emissivity on urban materials and vegetation (Stewart et al, 2021). Recent literature has highlighted the need to assess thermal exposure at a hyper-local level, considering factors such as exposure to the sun, wind, and relative humidity (Turner et al., 2022). For these reasons, in this research, we do not focus on the population exposure to heat but on urban areas with high values of LST. Despite these limitations, the study of LST provides advantages for research purposes, including a more detailed and global representation of temperature patterns (Manoli et al., 2020) that allows for inter-city comparisons at a fine temporal (daily) and spatial (1 km resolution) resolution that would not be feasible with air temperature data. Our study focuses on global patterns and both seasonal and longer-term trends, averaging values of day and night LST at a spatial resolution of 1 km over the 3 warmest months of the year for a period of ten years (Martilli, Krayenhoff & Nazarian, 2020; Manoli et al., 2020). We do not consider hourly or daily conditions at specific locations, where the differences between air and surface temperatures might be considerable due to rapid changes in weather-scale processes (Martilli, Krayenhoff & Nazarian, 2020; Manoli et al., 2020). Our methodology can be adopted and reproduced for any city in the world due to the availability of consistent satellite information captured

globally with the same resolution, methodology, and acquisition time every day. In this study, we only consider the effect of urban vegetation as a mitigation strategy for population exposure to LST extremes, while other strategies such as promoting energy efficiency and regulating building codes (Li, Bou-Zeid & Oppenheimer, 2014) or increasing the albedo of urban surfaces (Zhao et al., 2014) have not been considered.

2.2. The spatial linear regression model

We present a simple and global model to predict the number of days (Y_d) and nights (Y_n) over thresholds for each city where Y_d/Y_n correspond to the average number of days/nights over the threshold observed from 2010 to 2020. The model uses as predictors two local land surface properties: X_1 the fraction of green area within a pixel (as represented by the Normalized Difference Vegetation Index or NDVI); X_2 distance to water bodies d_w as described in the Methods section. Given the presence of a clear spatial autocorrelation of the variables (see Table S1 in the SI), we adopted a spatial linear regression model called spatial lag model (SLM) (Anselin & Rey, 2012). We studied two models for both TE_D and TE_N: we use the Google Earth Engine (GEE) platform (Gorelick et al., 2017) to compute the days where LST is greater than the thresholds (day and night). To assess the performance of the spatial model, we compute R^2 (see Fig. 3 and Fig. S8 in the SI) and mean absolute error (MAE) within a k-fold cross-validation setting (see Methods and Fig. S7 in the SI for details). In Fig. 3(A) and in Fig. S8 in the SI we show that, despite its simplicity, the model is able to predict TED and TEN averaged over the three warmest months of the years in the 200 cities with high accuracy. The result shows that an $R^2 \geq 0.8$ in the test-set within a k-fold cross-validation setting is obtained. Our analysis shows that the SLM performs significantly better than a Ordinary Least Squares (OLS) regression model as shown in Fig. S9 reporting negative low values of R^2 in the training phase and negative values during the validation phase, underlying the importance of accounting for the spatial correlation. Finally, model inferences suggest that the magnitude of the exposure during the warmest months of the year are largely controlled by urban vegetation with average regression coefficients ($\beta_{NDVI;}$ β_{dw}), calculated as the mean value of the best coefficients estimated during the training/validation steps of the k-fold cross validation (see Fig. S7 in the SI), are (-4.7,-1.2) for TE_D, and (-2.7,-0.6) for TE_N. It is noteworthy that the coefficients related to vegetation β_{NDVI} and distance to water bodies β_{dw} are negative, indicating that an increase of the surface covered by green areas and water bodies within cities would result in a reduction of the exposure to LST extremes of the areas that are affected by such an intervention.

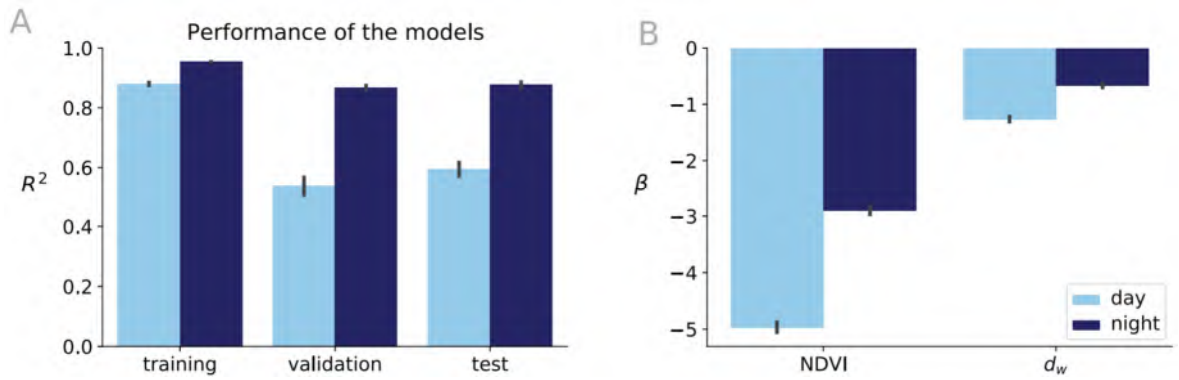


Fig. 3. Spatial lag models for the estimation of the daytime (TED) and the nighttime (TEN) exposures. A Performance of the models in terms of R^2 and (B) the coefficients of the models for the Normalized Difference Vegetation Index (NDVI) and distance to water bodies (d_w). See data availability and code availability for publicly available dataset and codes for generating the figures.

2.3. The mitigation strategy through vegetation increment

We use the SLM with the values of the coefficients computed during the cross-validation setting to estimate the vegetation increment required for reducing the exposure of the urban population to LST extremes as a function of the NDVI (see Methods section), while keeping the values of the distance to the water bodies fixed to the observed values. It should be noted that the vegetation increase is here applied without taking into account any urban or social considerations (e.g., space availability, financial means, energy efficiency, etc.), as these are out of the scope of what we want to do in this particular work. In practice the considerations with respect to mitigation strategies for a specific city need to be analyzed in detail, as there are, for example, limits in how much a pixel can be greened up without compromising its capacity to contain living and functional places. In Fig. 4 we show the results for the city of Paris as an example. There we simulate TE_T when NDVI is increased by 0.3 with respect to the observed values, to the maximum value of 0.85. We apply this increment for all the pixels that have an observed valued of the NDVI < 0.85 of the city and we quantify the exposure reduction as $\Delta TE_T(\%) = (TE_T^1 - TE_T^0)/TE_T^0$, where TE_T^0 and TE_T^1 are respectively the values of the total exposure before and after the NDVI increment. Figure 4(A) shows the entire city as well as these pixels where 80% of the entire population live respectively where the black boundary corresponds to the urban boundary as defined by the Global Human Settlement Layer (GHSL) dataset (Pesaresi et al., 2016) and the blue boundary is the urban environment as defined in the Methods section and in Fig. S2 in the SI. Figure 4(B) and (C) show the value of the NDVI and the total exposure divided by population for three different scenarios. The values of the NDVI and the total exposure for the different scenarios are reported in Table S2 and in Fig. S10 in the SI. We show the observed values of the NDVI that lead to a total exposure $TE_T^0 \sim 164$ million person · days. Scenario 1 corresponds to the value of the NDVI increased homogeneously at all pixels that leads to a total exposure $TE_T^1 \sim 144$ million person · days. We show that by increasing the NDVI homogeneously by 0.3, it is possible to reduce the exposure of urban population to extreme heat by ~ 12%. We also show that by increasing the NDVI homogeneously, we count an increment of 44% of the entire NDVI of the city (see Table S2 and Fig. S10 in the SI). In Scenario 2 we estimate the

NDVI increase in just the most populated areas which would lead to the same exposure reduction of -12%. This reduction could be achieved with a 0.38 increase of the NDVI in the corresponding pixels, shown in Fig. 4(B), and it corresponds to a 14% NDVI increment over the whole image (see Table S2 and Fig. S10 in the SI). With this simulation we show that it is possible to optimize the vegetation increment by targeting specific areas of the city.

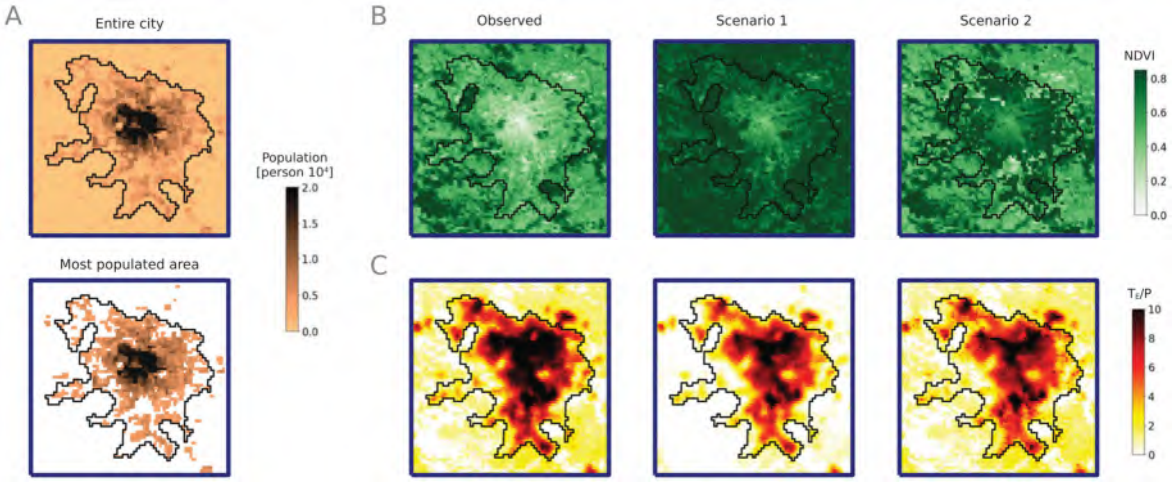


Fig. 4. Exposure reduction for the city of Paris. A We show the number of inhabitants at each pixel of the entire city and the most populated areas. B We present the values of the Normalized Difference Vegetation Index (NDVI) for each observed pixel and for three different scenarios. The Observed scenario corresponds to the observed values of the NDVI and the corresponding measured exposure. In Scenario 1, the value of the NDVI is increased homogeneously at all pixels by 0.3. In Scenario 2, the value of NDVI is increased only at the most populated areas of the city to achieve the same exposure reduction as in Scenario 1. C We show the value of the total exposure divided by the population for each pixel in the different scenarios. See data availability and code availability for publicly available dataset and codes for generating the figures.

In a subsequent analysis, we make simulations in which we increase NDVI for all 200 cities considered in this study as reported in Fig. 5. We show that, on average, increasing the NDVI by 0.3 results in exposure reductions of 12%, 13%, 16%, and 32% for cities located in temperate, arid, continental, and tropical regions, respectively (Fig. 5B). The required amount of NDVI increase per pixel was computed for the most densely populated areas of each city to achieve the same exposure reduction. Our findings demonstrate that targeting specific areas with higher exposure led to significant differences in NDVI increments. On average, 90%, 20%, 26%, and 33% of the NDVI could be saved for arid, continental, temperate, and tropical regions, respectively (Fig. 5C). The average local NDVI increment for the most populated areas is approximately 0.45 to optimize the NDVI increment in the urban environment (Fig. 5C). We also considered different NDVI increments ranging from 0.1 to 0.5. Figure 5(E) illustrates the exposure reduction for various NDVI increments. Results indicate that exposure linearly decreased with NDVI increments, but stable patterns of relative NDVI savings were observed as shown in Fig. 5(F). The relative NDVI savings were calculated using the equation NDVI(%) = 100(NDVI80 - NDVIall)/NDVIall, where NDVIall and NDVI80 represent the sum of NDVI values after interventions on the entire city and on the most populated areas, respectively. On average, it was possible to save up to 70% of greening by targeting specific areas to achieve the same exposure reduction as random homogeneous interventions as shown in Fig. 5F. In Fig. 5G we show the difference

between local and global NDVI increments and we observe that an average NDVI increment of ~0.15 could achieve the same exposure reduction for different global increments.

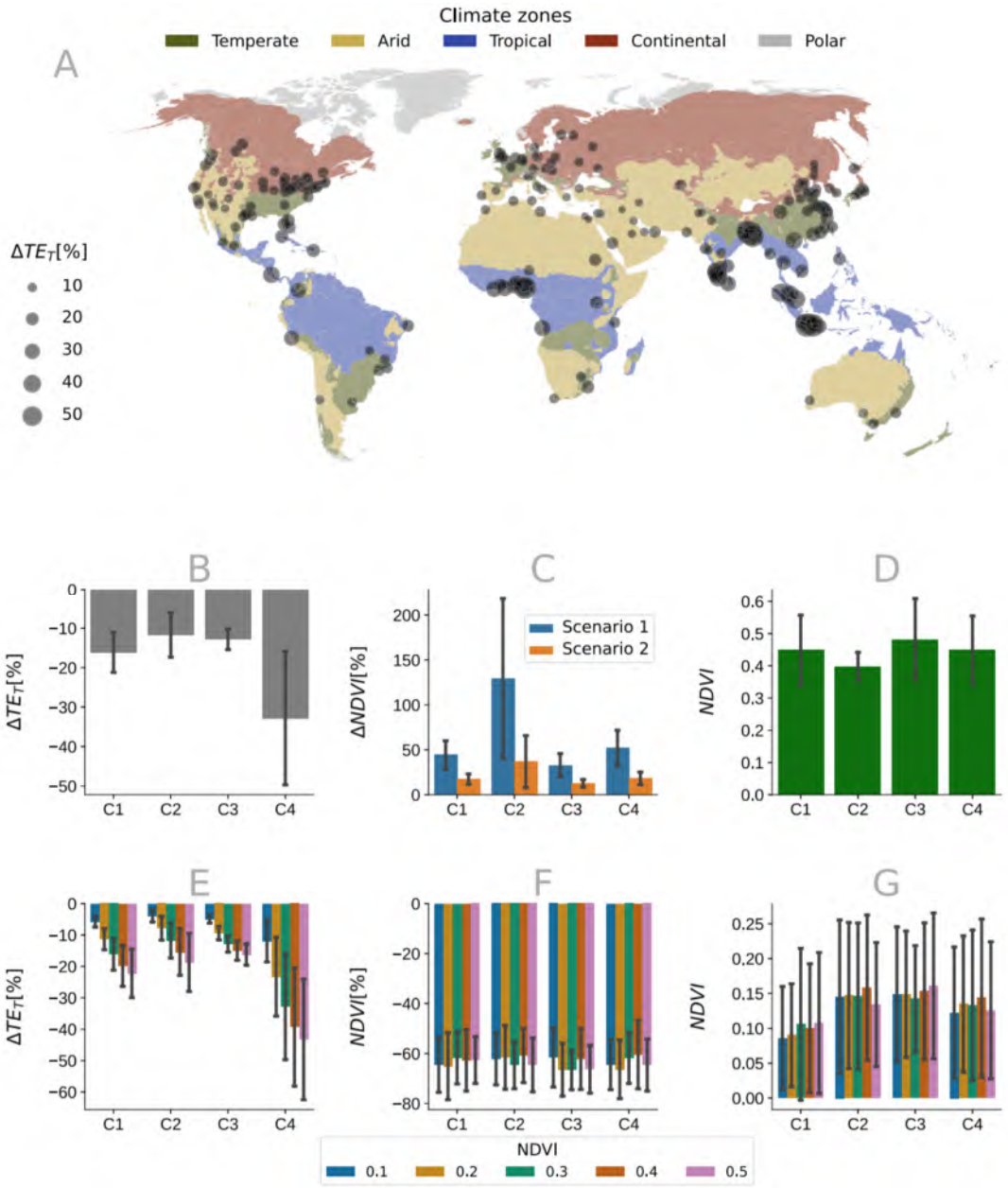


Fig. 5. Exposure reduction for all cities across the different climate zones. The x-axis displays the four climate zones: Temperate, Arid, Continental, and Tropical (C1, C2, C3, and C4, respectively). A Show the values of the total exposure reduction ($\Delta TE_T(\%)$) for all cities with an increment of the Normalized Difference Vegetation Index (NDVI) of 0.3. B We show the aggregated values of $\Delta TE_T(\%)$ for the different climate zones. C We compare the global NDVI increment required to achieve exposure reduction by targeting the entire city versus the most populated pixels. Scenario 1 shows a global NDVI increment to achieve exposure reduction by targeting the entire city, while in Scenario 2, NDVI is increased only in the most populated pixels as described in Fig. 4. D We present the average and standard deviation of the local NDVI increment for Scenario 1 to achieve the same exposure reduction with a homogeneous increment of 0.3. E The plot displays the average and standard deviation of the exposure reduction for different values of NDVI increment for all cities in the different climate zones. F We show the relative NDVI savings achieved by targeting the most populated areas for different values of NDVI increment. G We present the difference between the local and global NDVI increments for different values of global NDVI increment. See data availability and code availability for publicly available dataset and codes for generating the figures.

3. Discussion

In this study, we set a spatial lag model (SLM) to reproduce the exposure of the urban population to LST extremes for 200 cities worldwide: we defined exposure as the number of days per year where LST exceeds a heat exposure threshold multiplied by the total urban population exposed, in person · days: this definition can be seen as a surrogate version of population exposure to extreme heat using air temperature (Tuholske et al., 2021; Coffel, Horton & De Sherbinin, 2017). Our approach, which is exclusively based on remote sensing data, is able to predict the population exposure to LST extremes of urban environments. We have analyzed the impact of the vegetation change in reducing the exposure of urban population to LST extremes and computed this value for each pixel of each city. This provides theoretical options for already highly dense urban areas, often characterized by a lack of space and housing. Lastly, it should be mentioned again that our approach is large-scale and in some ways generalizes patterns across broad geographic and climatic regions (Manoli et al., 2020). Even if the SLM has been tested in cities across various climates and has consistently proven its accuracy, climate adaptation strategies need to be implemented locally, e.g., how additional green spaces are distributed, what types of greenery are planted, details that, for space constraints, are not part of this study.

We also make clear that is not simply a matter of increasing the total amount of urban greening in a city, but where the greening is targeted. While in this paper we focused on the most populated areas of the city, in the next steps we suggest to focus on targeting other relevant parts of the urban environments. It has been shown that the effects of urban greening can vary depending on the location, with irrigated landscapes in arid regions lowering nighttime temperatures only in the least vegetated neighborhoods (Gober et al., 2009). In addition, there are multiple studies that have highlighted the relationship between land surface temperature (LST) and historically marginalized areas within cities (Locke et al., 2021). It is important to also target interventions in more vulnerable areas (Hoffmann, Shandas & Pendleton, 2020). Although the marginalized areas of cities may not be part of the current study, the inequity of where vegetation is currently located should be addressed in the introduction and interpretation of the results to provide a more comprehensive understanding of the issue (Dialesandro et al., 2021). Moreover, exposure to heat is not limited to the outdoors and that people may also be exposed to heat within buildings. The availability of indoor cooling, as well as the gradient in air conditioning between wealthy and poor neighborhoods, is a crucial factor to consider when assessing exposure to extreme heat (Harlan et al., 2006). These considerations emphasize the requirement for a deeper analysis of a multi-dimensional approach when thinking about exposure to extreme heat, and the significance of considering the built environment and socio-economic factors in assessing vulnerability: this is a complex topic that deserves further investigation.

In summary, this research used a spatial lag model to assess and predict with high accuracy the exposure of urban population to extreme heat and it reveals quantitative results that give decisionmakers concrete guidance and general options to develop climate adaptation strategies in planning the urban landscape. In this study, we utilized Land Surface Temperature (LST) to assess exposure to extreme heat, considering both daytime and nighttime observations to account for the diurnal temperature exposure. While there may be

some limitations in using LST for estimating population exposure at specific spatial and temporal scales (Martilli, Krayenhoff & Nazarian, 2020), previous studies have shown that LST and air temperature are closely correlated over extended periods of time (Manoli et al., 2020), such as the 3-month period over a ten-year span we analyzed in this research. In addition, this method can easily be replicated by anyone for any city worldwide due to the availability of consistent satellite data that is captured globally at the same resolution, using the same techniques, and obtained at almost the same time every day and every night (Savtchenko et al., 2004). Our research presented a globally validated model, and we believe that the same methodology can also be applied to specific, practical, and localized solutions. Finally, we employed our findings to assess the efficacy of urban greening initiatives in reducing extreme heat exposure for urban populations. Our results demonstrated that targeting areas of the city where the population is more exposed to extreme heat leads to a substantial reduction in the amount of vegetation coverage needed to achieve the same decrease in exposure compared to a uniform treatment. As a next step, in order to achieve even greater improvement, it would be beneficial to use urban greening in combination with other mitigation strategies such as increasing surface albedo (Zhao et al., 2014) and water bodies (Jacobs et al., 2020).

4. Methods

4.1. Data and key variables

This study makes use of the Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI) products from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on board of the NASA's Terra and Aqua platforms: all data have been collected from the Google Earth Engine (GEE) platform (Gorelick et al., 2017). For each sample city and its relative urban boundary, the average data for the three warmest months of the year from 2010 to 2019 were collected and extracted from GEE as data points in a table. The urban boundary, the population layer and land use (i.e., presence of water bodies) was taken from the EU Global Human Settlement Layer (GHSL) (Pesaresi et al., 2013). To define the three warmest months for each year and city we rely on the ERA5 Copernicus data https://cds.climate.copernicus.eu/ . All the data have been aggregated at spatial resolution of 1km in the Mollweide projection.

Land surface temperature and heat thresholds. LST is from the MOD11A1 V6 product that provides daily LST and emissivity at 1 km nominal spatial resolution. We used the Google Earth Engine (GEE) platform to compute the values of the number of days and nights over thresholds. For each city, the day and night thresholds have been computed as the 90th percentile of the distribution of the LST over 20 years of observation (from 2002 to 2022) of all pixels of the city as shown also in Fig. S3 and Fig. S4 in the SI. This product provides daily per pixel Land The acquisition time of the MOD11A1 Version 6 is around 10:30am and 10:30pm respectively for day and night.

Normalized difference vegetation index. NDVI is from MOD13Q1 V6 product at 250 m nominal resolution and defined as:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

where NIR and Red are the atmospherically corrected bi-directional surface reflectances (masked for water, clouds, heavy aerosols, and cloud shadows) measured at 250m nominal spatial resolution in the near-infrared and red wavebands respectively (see MODIS product description (Savtchenko et al., 2004)).

Distance to water bodies. The third predictive variable of the model is an indicator of each point to the closest water body. Information on water presence was retrieved from the Global Human Settlement Data Layer produced by the European Commission Joint Research Center for the period 1984–2015 (Pekel et al., 2016). The water occurrence is a measurement of the water presence frequency (expressed as a percentage of the available observations over time actually identified as water). The provided occurrence accommodates for variations in data acquisition over time in order to provide a consistent characterization of the water dynamic over time. The variable used in our model is defined as:

$$d_w = \frac{1}{D_w^2} \quad (2)$$

where D_w is the distance of each pixel to the water surface. In the model, we used the normalized value of d_w between 0 and 1 in order to make this value in the same range as the other predictors. We used the Guidos toolbox (Vogt et al., 2022) to estimate the distance to all pixels to water bodies worldwide.

Definition of cities and urban boundaries. The urban boundaries are set to the areas provided by EU GHSL. In particular, we used the Degree of Urbanization, a new global definition of cities, urban and rural areas dataset available online (https://ghsl.jrc.ec.europa.eu/degurba.php). The city boundaries are defined taking a rectangular buffer of 5 km around the GHSL urban boundaries, as shown in Fig. S2 in the SI as an example for the city of Guangzhou.

Data aggregation. The workflow to generate an aggregated dataset with all the necessary information starts with the definition of the urban boundary as defined above. Once the urban boundary has been defined, we divided the urban environment into a 1 km × 1 km grid as shown in Fig. S2 in the SI. In each cell of the grid, we then aggregated the information we collected. Thereafter, all the information necessary for running the model and performing the analysis has been spatially merged and measured over the grid. For the population, we counted the total population living in each cell. For the LST we assigned each data point to its relative cell, for NDVI we assigned several points to a single cell given the higher spatial resolution and we computed the average, while for the distance to the water bodies (D_w) we computed the distance from the epicenter of the grid to the closest water body point. We used the Guidos toolbox (Vogt et al., 2022) to compute the distance of each pixel to the closest water body as defined by the GHLS (Pesaresi et al., 2013).

4.2. Climate aggregation

Cities are classified based on their climate zone following the Köppen climate classification which is one of the most widely used climate classification systems. The Köppen climate classification divides climates into five main climate zones, with each zone being divided based on seasonal precipitation and temperature patterns. The five major climate zones are then further divided into subzones. The second letter indicates the seasonal precipitation type, while the third letter indicates the level of heat. We define five main clusters as shown in Fig. S1.

4.3. The spatial model

In this research we make use of the so-called Spatial Lag Model (SLM) that primarily addresses spatial autocorrelation48 (see also Table S1 in the SI). Spatial autocorrelation refers to the case when the dependent variable exhibits a non-random pattern over our spatial units after controlling for other covariates. Positive autocorrelation reflects value's similarity in space, and negative autocorrelation reflects value's dissimilarity in space. In a matrix notation the SLM reads as (see SI for a detailed description of the SLM):

$$Y = \beta X + \rho W_y Y + \varepsilon \qquad (3)$$

where $Y = \beta X + \varepsilon$ corresponds to a standard linear regression model, while W is the spatial weighting matrix applied to the observed variable, Y, together with a spatial autoregression parameter, ρ that reflects the spatial dependence inherent in our sample data, measuring the average influence on observations by their neighboring observations (Anselin & Bera, 1998). In a spatial lag model the dependent variable among our neighbors influences our dependent variable. The spatial weights matrix, W is standardized such that its rows sum to 1, hence it is effectively including a weighted average of neighboring values into the regression equation. The neighbors can be identified by different methods, such as distance-based methods (k-nearest neighbors, distance specified) and contiguity methods (queen, rook)48. In this research we use a k-nearest neighbors approach with $k = 8$ where all the neighbors are equally weighted.

Acknowledgements

The study was funded by the Exploratory Project Thermopolis of the European Commission, Joint Research Centre. R.S. acknowledges partial support from the European Union's Horizon 2020 innovation action project Go Green Routes under grant agreement No. 869764.

Supplementary information

Supplementary material to this article can be found online at

https://doi.org/10.1038/s41467-023-38596-1

Chapter XXIII

Cost estimation for the monitoring instrumentation of landslide early warning systems

Nat. Hazards Earth Syst. Sci., 2023, vol. 12

Marta Sapena, Moritz Gamperl, Marlene Kühnl, Carolina Garcia-Londoño, John Singer & Hannes Taubenböck

© Author(s) 2023. This work is distributed under the Creative Commons Attribution 4.0 License For more information, see https://creativecommons.org/licenses/by/4.0/

https://doi.org/10.5194/nhess-23-3913-2023

Received 14 Mar 2023; Accepted 30 Oct 2023; Published 24 Dec 2024

Abstract

Landslides are socio-natural hazards. In Colombia, for example, these are the most frequent hazards. The interplay of climate change and the mostly informal growth of cities in landslide-prone areas increases the associated risks. Landslide early warning systems (LEWSs) are essential for disaster risk reduction, but the monitoring component is often based on expensive sensor systems. This study presents a data-driven approach to localize landslide-prone areas suitable for low-cost and easy-to-use LEWS instrumentation, as well as to estimate the associated costs. The approach is exemplified in the landslide-prone city of Medellín, Colombia. A workflow that enables decision-makers to balance financial costs and the potential to protect exposed populations is introduced. To achieve this, city-level landslide susceptibility is mapped using data on hazard levels, landslide inventories, geological and topographic factors, and a random forest model. Then, the landslide susceptibility map is combined with a population density map to identify highly exposed areas. Subsequently, a cost function is defined to estimate the cost of LEWS monitoring sensors at the selected sites, using lessons learned from a pilot LEWS in Bello Oriente, a neighbourhood in Medellín. This study estimates that LEWS monitoring sensors could be installed in several landslide-prone areas with a budget ranging from EUR 5 to EUR 41 per person (roughly COP 23 000 to 209 000), improving the resilience of over 190 000 exposed individuals, 81 % of whom are located in precarious neighbourhoods; thus, the systems would particularly reduce the risks of a social group of very high vulnerability. The synopsis of all information allows us to provide recommendations for stakeholders on where to proceed with LEWS instrumentation. These are based on five different cost-effectiveness scenarios. This approach enables decision-makers to prioritize LEWS deployment to protect exposed populations while balancing the financial costs, particularly for those in precarious neighbourhoods. Finally, the limitations, challenges, and opportunities for the successful implementation of a LEWS are discussed.

Keywords

LEWS cost estimation; landslide susceptibility; data-driven analysis; Medellín

1. Introduction

The lives and livelihoods of billions of people around the world are disrupted by human-induced hazards, exacerbated by climate change. The Intergovernmental Panel on Climate Change (IPCC) of the United Nations has recently reported that climate change is causing more frequent and severe storms, floods, droughts, wildfires, and other extreme weather events (IPCC, 2022). The climate crisis has global implications and poses several challenges to governments, societies, and science (Marchezini et al., 2018), with some geographic regions being more affected than others. For example, Colombia is one of the most landslide-prone countries in the world. The majority of its population lives in areas that are prone to landslide hazards (Ruiz Peña et al., 2017). In Colombia, the increasing frequency of intense and persistent rainfall, coupled with unplanned urban growth in areas prone to landslides driven by limited land availability, significantly increases the likelihood of disasters, particularly impacting the most vulnerable populations (World Bank, 2012).

Due to the occurrence of a disaster, or to mitigate the effects of an imminent natural hazard, people are occasionally compelled to evacuate their places of residence. Displacement severely disrupts people's lives, raising unemployment, interrupting education, and hindering access to essential services, ultimately leading to increased vulnerability and impoverishment. Implementing preparedness measures is crucial in mitigating risks associated with displacement. These measures enhance risk awareness among individuals at risk of displacement, empowering them to make informed decisions and comply with the warnings (UNDRR, 2021). In fact, one of the seven global targets of the Sendai Framework for Disaster Risk Reduction 2015–2030 is to "Substantially increase the availability of and access to multi-hazard early warning systems and disaster risk information and assessments to people by 2030" (UN, 2015).

Landslide early warning systems (LEWSs) play a crucial role in reducing the risk associated with landslides. LEWSs provide timely information of slowly changing slope stability or acutely dangerous situations, enabling proactive measures, enhancing public awareness and education, and facilitating better planning and decision-making. Additionally, LEWSs collect valuable data that can be utilized for scientific research, monitoring, and analysing landslide behaviour. Thus, in turn, the understanding of landslides, their triggers, and their impacts can be understood better, leading to the development of improved predictive models and more effective LEWSs (Guzzetti et al., 2020; WMO, 2018; Segoni et al., 2023). LEWSs have the potential to minimize the loss of lives and mitigate the economic and social impacts of disasters. They provide significant economic benefits by reducing damage and loss, facilitating cost-effective planning and response, preserving economic activities, and saving costs in emergency response operations (Rogers and Tsirkunov, 2011; Grasso, 2014). Therefore, they offer a viable alternative to relocating exposed populations, especially since relocation is in most regions economically unviable and relocation often faces strong opposition from residents (Werthmann and Echeverri, 2013). These economic advantages make investing in LEWSs a prudent choice. However, it is important to note that the cost of implementing LEWSs can vary significantly depending on factors such as the size and complexity of the monitored area, the technology and infrastructure employed, and the level of system sophistication. Furthermore, beyond the technical implementation and

maintenance, the effectiveness of a LEWS relies on actively involving at-risk individuals, improving education and awareness of risks, efficiently disseminating messages and warnings, and ensuring preparedness (WMO, 2018).

In general, however, LEWSs offer promising potential when well-integrated and properly managed. However, these systems also face several challenges, shortcomings, and untapped potential: the monitoring components of LEWSs often rely on expensive high-end sensor systems, such as multiphase GNSS (global navigation satellite system), GB-SAR (ground-based synthetic aperture radar), and tacheometry. These systems require highly trained personnel for operation and are specifically tailored to the local situation, making their transfer to other regions or countries difficult. Nevertheless, there are also low-cost and easy-to-use sensor systems available, such as MEMS (micro-electromechanical system) tilt inclinometers, acceleration sensors, continuous shear monitors, and piezometers. By installing these geosensors in a local network, valuable information about the surface and subsurface processes on landslide-prone slopes can be obtained (Thuro et al., 2014; Uchimura et al., 2015; Singer et al., 2021). When combined with data analysis and numerical landslide process models (Huggel et al., 2010; Thiebes et al., 2014), these sensor systems have the potential to enhance the quality, reliability, and usability of hazard warnings, as well as to reduce the need for extensive manual work. An example is the low-cost subsurface monitoring system implemented in the Alps, specifically at Sudelfeld in Bayrischzell, Germany, which operated from 2008–2014. This system included cost-efficient ground deformation measures and groundwater level and precipitation monitoring (Singer and Thuro, 2006; Thuro et al., 2010). In recent decades, advancements in technology and affordability of slope monitoring have improved, enabling more widespread applications, including in low-income countries. However, continuous monitoring is still limited to only a few slopes worldwide, with some site-specific LEWSs in Latin America (Guzzetti et al., 2020). The dissemination of experiences, challenges, and limitations associated with LEWSs is not a priority and is therefore rarely done, particularly in English (e.g. Reinoso Jerez, 2013; Departamento del Quindío, 2018; Castro Bonilla, 2021).

In general, there are estimates available regarding the economic benefits of early warning systems (EWSs), particularly in European countries, the United States, and Japan (Hallegatte, 2012). For instance, in Europe, the annual benefits of EWSs are estimated to range between EUR 470 million and EUR 2.8 billion. Similarly, it has been projected that low-income countries could experience potential benefits with a cost–benefit ratio ranging from 4 to 35, along with associated co-benefits, provided similar EWSs were available (Hallegatte, 2012). However, estimating the costs and benefits of implementing local LEWSs remains challenging due to various factors influencing the overall expenses. The implementation of a comprehensive LEWS requires substantial investments, including but not limited to equipment costs, infrastructure development, ongoing maintenance, personnel expenses, and social integration. Recognizing this knowledge gap, this study aims to contribute to the literature by providing cost estimations for the instrumentation of lowcost, local, and site-specific LEWSs. To the best of the authors' knowledge, the estimated cost of the monitoring instrumentation of local and site-specific LEWSs is largely unknown or

unpublished despite being highly relevant for policy-makers involved in disaster risk reduction.

The purpose of a LEWS is to reduce risk and improve preparedness for hazards in specific locations. Thus, it is imperative to identify landslide-prone areas, the location of people and assets in exposed locations, and their vulnerabilities. In this context, Earth observation (EO) plays an important role in early warning, mapping, and monitoring natural hazards (Grasso, 2014; Casagli et al., 2017). EO can be used, for example, to identify populations exposed and vulnerable to different natural hazards (Taubenböck et al., 2008; Geiß et al., 2017; Müller et al., 2020; Kühnl et al., 2023), as well as to estimate landslide susceptibility at different regions and scales (Cantarino et al., 2019; Palacio Cordoba et al., 2020; Ospina-Gutiérrez and Aristizábal, 2021).

Landslide susceptibility modelling has witnessed an increase in popularity due to the advancements in remote sensing and machine learning models. Traditional knowledgedriven methods, such as the multicriteria analytical hierarchy process (AHP) developed by Saaty (1980), rely on weights assigned to several landslide-influencing factors. Thus, the result depends on the experience of the user and the potential to identify factors that are important for a special case (Günther et al., 2014; Skilodimou et al., 2019). In contrast, data-driven methods rely on reference data (e.g. landslide inventories) and conditioning factors (i.e. factors influencing landslide risks), which are used to identify their interconnected relationships and predict landslide susceptibility based on statistical models. Data-driven approaches have demonstrated significant potential in effectively mapping areas prone to landslides, particularly in situations where the availability of comprehensive geotechnical data required for physically based methods is lacking. Some of the most common data-driven methods include random forest (Taalab et al., 2018; Calderón-Guevara et al., 2022; Ado et al., 2022; Abu El-Magd et al., 2021), logistic regression (Ado et al., 2022; Azarafza et al., 2021), convolutional and artificial neuronal networks (Nikoobakht et al., 2022; CalderónGuevara et al., 2022; Ado et al., 2022; Azarafza et al., 2021), boosted regression trees (Calderón-Guevara et al., 2022; Pourghasemi et al., 2021), weight of evidence (Calderón-Guevara et al., 2022), support vector machine (Nikoobakht et al., 2022; Ado et al., 2022; Azarafza et al., 2021), k-nearest neighbour (Nikoobakht et al., 2022; Abu El-Magd et al., 2021), naïve Bayes (Abu El-Magd et al., 2021; Azarafza et al., 2021), and linear discriminant analysis (Eiras et al., 2021; Pourghasemi et al., 2021). Previous studies have compared the performance of AHP and statistical methods, and the latter was found to perform better (Erener et al., 2016; Ali et al., 2021; Vojtek et al., 2021). Nevertheless, currently, there is no definitive data-driven method established as the optimal choice for empirical landslide susceptibility modelling. In the recent literature, various methods have been employed and compared, and their accuracies and suitability have shown regional variations. In this study, the random forest method is implemented due to its demonstrated high accuracy in Colombia (Calderón-Guevara et al., 2022). Additionally, this method offers the advantage of being nonparametric, allowing for the inclusion of influencing factors that are not normally distributed (Breiman, 2001).

In this context, the landslide-prone city of Medellín in Colombia is notable for its comprehensive landslide inventories (which are crucial resources for collecting information

related to landslide occurrences, for deriving empirical knowledge, and for creating landslide susceptibility maps) and its implementation of both a city-wide EWS and a local scale LEWS. On the one hand, the Aburrá Valley and the city of Medellín implemented an EWS (Sistema de Alerta Temprana de Medellín y el Valle de Aburrá, SIATA) that monitors real-time hydrological, meteorological, seismic, and geotechnical variables to forecast natural and anthropic phenomena and to strengthen risk management in the territory (SIATA, 2023). On the other hand, a unique local, sitespecific, and low-cost participatory LEWS was implemented in a local community known as Bello Oriente by a research project called Inform@Risk (Strengthening the Resilience of Informal Settlements against Slope Movements; Werthmann et al., 2023). Bello Oriente is a settlement that was originally built informally on one of the eastern slopes of the city in an area exposed to landslide hazard. The LEWS includes a wireless network of sensors that is based on Internet of Things (IoT) technologies. It monitors movements in the subsurface and their effects on the built infrastructure (e.g. tilting, opening of cracks), groundwater levels, and other parameters that, in combination with weather variables and forecasts, are used to inform people about the level of risk. With this, the system aims to provide exposed people enough time to react in the case of an event and to improve their risk awareness and resilience.

For the above-mentioned reasons, Medellín is selected as the study area. Using lessons learnt from the LEWS prototype installation in Bello Oriente in the frame of the Inform@Risk project and assuming that sufficient financial resources are not available on an ad hoc basis for citywide instrumentation with LEWSs, a cost function is developed which allows us to weigh the cost, the location(s), and how many people can possibly benefit from a LEWS in landslide-prone areas. The ultimate goal is to assist decision-makers in their prioritization efforts by providing a comprehensive assessment that facilitates the strategic implementation of LEWSs, ensuring maximum impact and cost efficiency by offering valuable recommendations based on different cost effectiveness scenarios. Therefore, the specific objectives of this study are (1) to identify highly exposed landslide-prone areas that are suitable for the implementation of LEWSs with automatic monitoring sensors, (2) to determine a cost function based on topographic and socio-economic parameters for the implementation of a LEWS, and (3) to provide suggestions for decision-makers on where to start with the implementation of new LEWSs based on cost effectiveness and prioritized areas.

2. Material and methods

In this section, the developed workflow for a city-wide cost effectiveness analysis for LEWSs is introduced. First, the study area and datasets are presented. Subsequently, the process of deriving a landslide susceptibly map is explained, along with the preselection of exposed sites where the installation of a LEWS is preferable. Then, the calculation of physical, social, demographic, and infrastructure parameters is conducted for the preselected sites within the pool of possible LEWS installations. This analysis supports the identification of suitable sites for the LEWS installation. Lastly, the cost function is presented, which aims to prioritize the installation of LEWS based on cost effectiveness (Fig. 1).

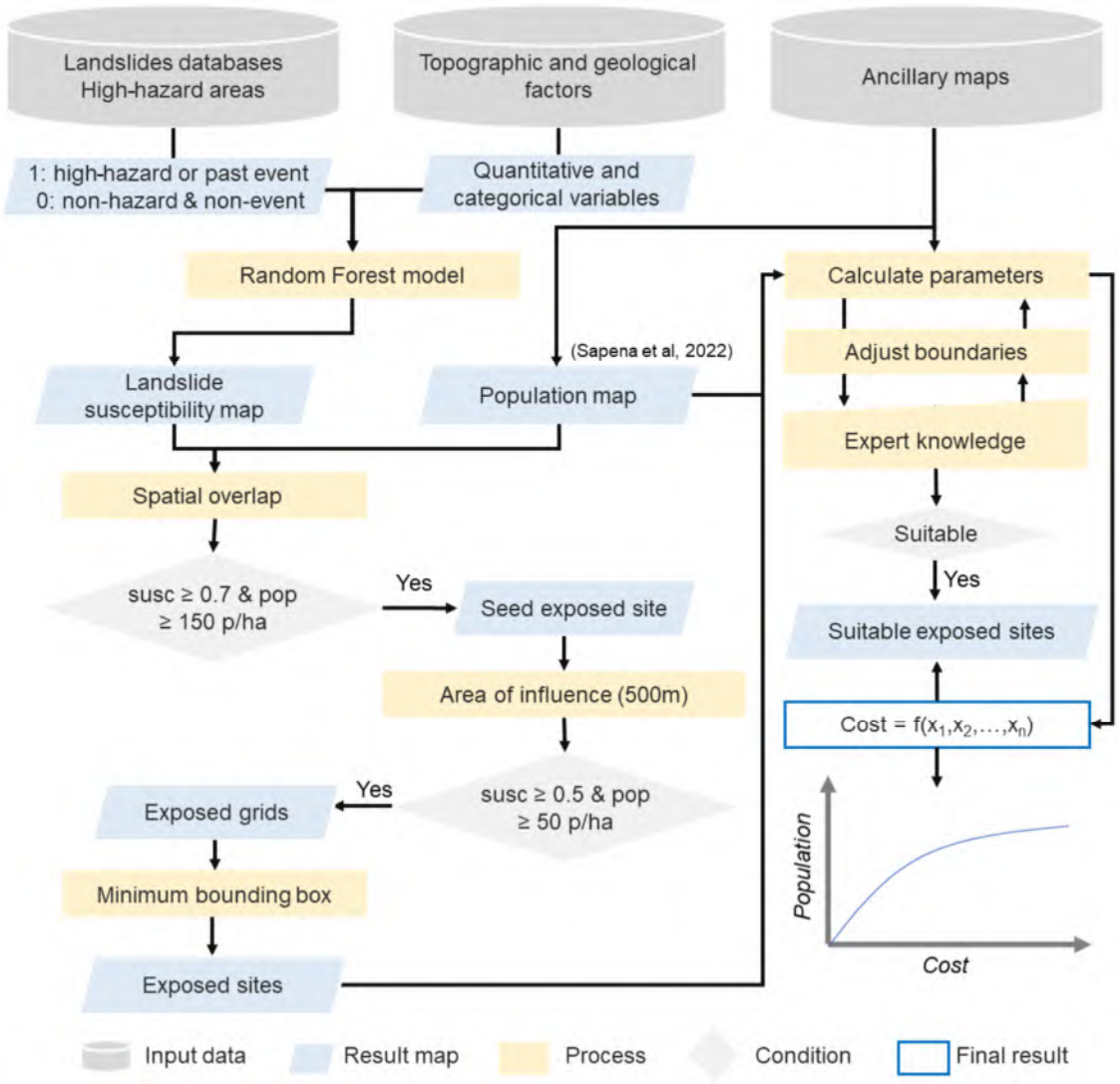


Fig. 1. Workflow of the study.

2.1. Study area

This study takes place in Medellín, Colombia, a highly populated city situated in the Aburrá Valley. The city is characterized by a steep topography, which poses a significant landslide risk, particularly on the eastern slopes. These slopes are predominantly composed of heavily fractured dunite rock, which undergoes deep and intense weathering in the region's tropical conditions due to its high iron content (Thuro et al., 2020). The dunite is overlaid by colluvium material, previously displaced by landslide processes, forming a block-in matrix structure with varying thicknesses ranging from 5 m to over 30 m (Breuninger et al., 2021). This colluvium reveals that substantial portions of the slope have been affected by landslides in the past, and the steep topography of the hills makes them susceptible to future landslides.

In Medellín an increasing portion of the population is exposed to landslide risks (Kühnl et al., 2023). The city showcases a diverse range of urban structural configurations and socio-economic levels, encompassing high-rise buildings in the business district and wealthier neighbourhoods, as well as light-weight and low-rise buildings in precarious neighbourhoods. The latter are primarily located in landslide-prone regions (Kühnl et al.,

2021). These factors contribute to significant inequalities in risk exposure for various social groups.

2.2. Landslide inventories and hazard map

The extensive history of landslides in Medellín has led to the establishment of multiple databases that document landslide events, causalities, and damage to buildings and infrastructure. Specifically, there are three main databases where individuals report landslides: SIMMA, DesInventar, and DAGRD.

SIMMA (Sistema de Información de Movimientos en Masa), managed by the Colombian Geological Survey, is a national landslide inventory (SIMMA, 2022). For Medellín, it provides precise coordinates for 13 landslide events between 1985 and 2013.

DesInventar is the international Disaster Information Management System, which helps in analysing disaster trends and impacts (DesInventar, 2022). The "Medellín Área Metropolitana" database, managed by the Universidad Nacional de Colombia, contains information from 1880 to 2022, including 21 landslides with precise coordinates between 2018 to 2022.

DAGRD (Departamento Administrativo de Gestión del Riesgo de Desastres) registers past or foreseen landslides based on emergency calls from citizens, which are subsequently evaluated by DAGRD technicians. The department provided data from 2004 to 2019, encompassing over 30 200 reports of potential mass movements with their coordinates (DAGRD, 2018). It is important to note that the vast majority of these reports are concentrated in the urban areas, with limited coverage of events in rural regions.

In addition to landslide inventories, remote sensing data and techniques have proven effective for detecting and mapping landslides (Guzzetti et al., 2012), especially when inventories lack precise spatial accuracy. In this paper, the information from the three inventories was combined with landslide locations identified through remote sensing in the urban–rural border and rural areas. This involves utilizing multi-temporal imagery from the Landsat-7, Landsat-8, and Sentinel-2 satellites and applying change detection methods to vegetation indices in areas where landslides are indicated by news articles. It is assumed that significant changes in these vegetation indices in areas of steep slopes are indicative of variations in the soil surface or vegetation removal. These areas can then be attributed to landslides (Mondini et al., 2011). In total, eight landslides were identified between 2008 and 2019.

Additionally, the latest "master plan" of the city of Medellín from 2014 (POT 2014 – Plan de Ordenamiento Territorial) incorporates a zoning system for landslide hazards (Alcaldía de Medellín, 2014a). The hazard map results from a comprehensive analysis that incorporates various sources of information. These include the hazard map from the POT 2006, a probabilistic hazard map developed by the Universidad Nacional, and the DAGRD landslide inventory until 2014, as well as heuristic processes, fieldwork, and the expertise of technicians from DAGRD and the Administrative Department of Planning (Departamento Administrativo de Planeación). To model landslide susceptibility, the mass movements inventories and the hazard map from the POT 2014 are utilized as reference data. Given the

absence of specific details concerning the type of mass movements and considering that the majority of mass movements in Medellín are landslides, it is assumed that all available reports correspond to landslides. These datasets form the foundation for training and evaluating the model for the identification of possible areas prone to landslides, as elaborated further in Sect. 2.4.

2.3. Factors influencing landslide risk

For modelling landslide susceptibility, the proposed methodology incorporates a range of factors influencing landslide risks including topographic, geological, and precipitation data. In addition, to support the search for suitable locations for the implementation of LEWSs, socio-demographic factors are also considered. The database consists of the official cartography from open-data platforms of the city of Medellín and the metropolitan area of the Aburrá Valley such as "GeoMedellín" and "Datos Abiertos del AMVA" (Alcaldía de Medellín, 2023; AMVA, 2023), precipitation data from SIATA (SIATA, 2023), OpenStreetMap data (downloaded in 2022, openstreetmap.org), a high-resolution population map that provides estimates of the number of people per building and a grid of 100 m from Sapena et al. (2022), and a map of precarious settlements from Kühnl et al. (2021).

Various topographic factors are derived from contour lines to analyse the terrain. Contour lines at different scales (1 V 2000 for urban areas and 1 V 5000 for rural areas) are utilized to generate a triangular irregular network (TIN) surface. By interpolating the altitude data from the TIN, a digital elevation model (DEM) with a spatial resolution of 5 m is generated. This DEM is used for deriving several topographyrelated factors, including slope, aspect (which indicates the downhill direction of the slope), and curvature (which represents the shape or curvature of the slope) (Fig. 2). Additionally, the DEM is used to model water flows, resulting in the extraction of the stream network, stream order, landslide travel paths, and angle of reach of landslides (also known as the "fahrböschung angle" or α/. The angle of reach provides insight into the potential mobility of a landslide (Hungr et al., 2005). Regarding the stream network, there is no universally agreed upon flow accumulation threshold for determining streams due to its dependence on various factors, including desired stream density, data scale, resolution, and landscape attributes. Nonetheless, it is common practice to employ threshold values within a range of approximately 100 to 1000 pixels or 0.05 to 5 km2 drainage area (Reddy et al., 2018; Tarboton et al., 1991). In this study, multiple values within this range are evaluated, and it was determined that a threshold of 500 contributing pixels (equivalent to a minimum stream inflow of 12 500 m2) achieved the closest correspondence with the official drainage system map from the POT of Medellín. Furthermore, a visual examination of the result reveals a satisfactory representation of the majority of streams. The Strahler method (Tarboton et al., 1991) is used to classify the streams into numerical orders, distinguishing between main streams classified as major tributaries (ranging from order 7 to 5) and other streams categorized as outermost tributaries (ranging from order 4 to 1). Two Euclidean distance maps are then generated to calculate the distance from a given pixel to both the main streams and the other streams (Fig. 2). Subsequently, the travel path and angle of reach of a landslide are used to support identifying

exposed sites. This is done by identifying unpopulated areas where landslides are likely to occur and where their runout can extend into the inhabited area.

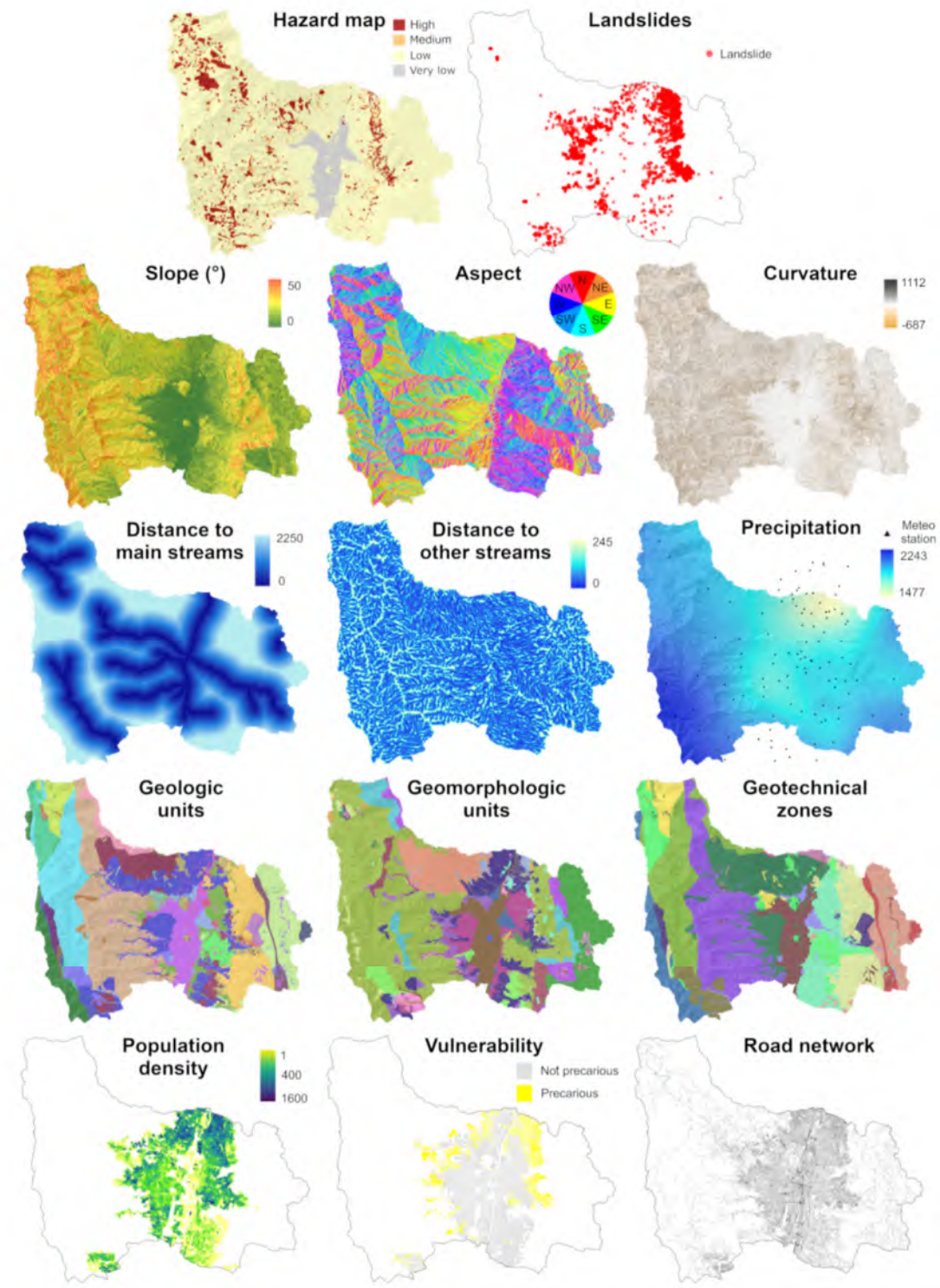


Fig. 2. Dependent variables in the model (hazard map and landslide events), independent variables or conditioning factors (topographic, geological, and precipitation), and ancillary data (socio-demographic and infrastructure).

For the geology, geomorphology, and geotechnics of the study area, data are gathered from "Datos Abiertos del AMVA". Three maps containing categorical information on geological units, geomorphologic units, and geotechnical zoning are used. The surface geology plays a

crucial factor as it provides information about the physical composition of materials, their properties, and mechanical strength characteristics. Geomorphology offers valuable information on the landscape, i.e. the stability, slope, and shape. Additionally, geotechnics provide details about soil types (Universidad Nacional de Colombia, 2009) (Fig. 2).

Precipitation plays a crucial role in triggering landslides, particularly in mountainous areas. In the study area, the average yearly precipitation exhibits significant variations due to local luv and lee effects. Thus, precipitation data obtained from the meteorological stations of SIATA are used. Precipitation measurements from the 215 available stations in the Aburrá Valley for 2021 are acquired. Using these data points, interpolation techniques are employed to generate a continuous map of rainfall accumulation with a spatial resolution of 5 m. To achieve this, the ordinary kriging optimized smoothed (OKOS) method is employed, accompanied by cross-validation. For this process, 70 % of the stations are selected for training and the remaining 30 % for testing. The evaluation provides a root mean square error (RMSE) value of 506 mm yr-1, which represents the discrepancy between the predicted and observed yearly precipitation accumulation. In relation to the mean value of precipitation across all stations, i.e. 1425 mm yr-1, the normalized RMSE, indicating the relative error, is approximately ±35 %. This error is in line with the existing literature, which commonly reports normalized RMSE values ranging from 30 %–35 % (Bostan et al., 2012) and 17 %–29 % (Antal et al., 2021) (Fig. 2).

When it comes to decisive factors for assessing the suitability and prioritization of potential locations for implementing a LEWS, socio-demographic and infrastructure variables are used. To begin with, the population map is used to identify sites with both high population density and a high susceptibility to landslides. These areas allow us to

quantity the exposed population. However, from a socioeconomic perspective the exposed populations at these sites do not have equal vulnerabilities, i.e. the financial resources to recover from a disastrous event. In consequence, the calculated population of these sites is combined with information on the socio-economic precariousness of the residents. The exposed population living in precarious settlements is referred to hereinafter as the highly vulnerable population. This comprehensive analysis enables us to prioritize exposed areas that exhibit a higher vulnerability to disastrous events. In addition, data on the road infrastructure are employed, which is a relevant factor in determining the placement of sensors within the LEWS. The official road network cartography from “GeoMedellín” is supplemented with additional road data from OpenStreetMap (Fig. 2).

2.4. Mapping landslide-prone areas

In this study, baseline data on recent landslide events and an official hazard map are employed to train a random forest statistical model. The aim of this model is to predict the probability of landslide events, also referred to as landslide susceptibility, by considering multiple conditioning factors. Therefore, the landslide hazard map, which incorporates local expert knowledge, includes additional landslide events from the compiled inventories that were not originally included in the hazard map (i.e. all landslides from Sect. 2.2 starting from 2015 onwards). Both datasets are considered as reference data in the model. Consequently,

a set of sampling points is created using these reference data: within high-hazard zones (5000 points); recorded landslide events (2800 points); and non-hazardous areas including medium-, low-, or very low-hazard zones (8000 points). Each point is assigned a numeric attribute representing the dependent variable: high-hazard (1) and non-hazard (0). Additionally, the values of all conditioning factors, explained in Sect. 2.3, are also assigned to each point and are used as independent variables in the model. The sampling set is then divided into training data (70 %) and testing data (30 %). The nonparametric random forest algorithm (Breiman, 2001) is employed, using 100 trees, 3 variables at each split, and 5 terminal nodes. The susceptibility map is generated using the training data, providing the probability of a pixel being classified as high-hazard (1), calculated based on the proportion of votes across the ensemble of trees. The testing data, which were not used for training, are used to evaluate the model. Ultimately, the accuracy statistics are measured using the remaining testing set.

2.5. Selection of exposed areas

A zero susceptibility level is assigned to areas with slopes equal to or lower than 10◦ in the resulting landslide susceptibility map. This is in accordance with the assessment by Alcaldía de Medellín (2014b) as these slopes are considered non-hazardous in Medellín. Next, to identify the exposed areas, the average susceptibility per 100 m grid cell is calculated and combined with the population density per grid. This analysis focuses on the urbanized area, which contains the most exposed elements and where population data at a fine grain level are available.

The identification of exposed sites follows an iterative process, as depicted in Fig. 1. First, seeds of exposed sites are identified based on high susceptibility (≥ 0.7) and high population density (≥ 150 people per hectare). If two or more

seeds are contiguous or within 100 m of each other, the centroid is used as the seed location. Second, these seeds serve as starting points to search within a 500 m area of influence, which is set to match the size of the reference LEWS prototype in Bello Oriente. Pixels with medium to high susceptibility (≥ 0.5) and medium to high population density (≥ 50 people per hectare) within this area are considered. This process identifies all susceptible and populated pixels surrounding the seeds. Third, a minimum bounding box is calculated for each cluster of pixels to automatically define a preselected exposed site. Fourth, the shape of the preselected sites is adjusted based on factors such as the urban structure, topography, travel path, and angle of reach of a landslide. Finally, several parameters for the preselected sites are calculated, including area, mean susceptibility, total population, population density, vulnerable population, built-up and road density, number of buildings, mean slope, main orientation of the slope, and open areas for the preselected sites. These parameters are then used by experts to inspect and select the suitable exposed sites based on the requirements of the LEWS and previous experience in Bello Oriente.

2.6. Developing a cost function

To estimate the costs of monitoring instrumentation for LEWSs in suitable sites, the existing LEWS prototype in Bello Oriente serves as a reference. It provides information on the required workforce (working hours for the sensor system construction and installation), node density, node types, and individual node costs. The Bello Oriente LEWS prototype covers a 39 ha area with a total population of 4600 residents, including 1800 individuals residing in high-hazard landslide zones. The Bello Oriente LEWS prototype utilizes an innovative wireless geosensor network based on IoT technologies, such as LoRa® wireless data communication and MEMS sensors. It incorporates a continuous shear monitor measurement cable and extensometer (CSM-EXT) system spanning 1100 m (Thuro et al., 2010) to monitor subsurface movements and near-surface groundwater levels, forming the foundation for generating warnings. Given the unpredictable nature of future landslides based on geological investigations, extensive coverage with high spatial and temporal observation density is essential. The wireless geosensor network in Bello Oriente consists of 115 nodes, of which 45 monitor subsurface deformation and groundwater levels (subsurface nodes), while 70 nodes detect movement in existing built infrastructure (infrastructure nodes). The spacing of nodes is adjusted according to the level of landslide risk, with high-risk areas having on average eight nodes per hectare and areas with no risk having no nodes. The nodes are opensource and can be replicated using published printed circuit board schematics, 3D printing models, and building instructions. Further details about the measurement system can be found in Gamperl et al. (2021) and Singer et al. (2021), as well as on the Inform@Risk wiki (https://www.informatrisk. com/, last access: 13 July 2023).

It is important to note that the current cost estimation only considers the implementation of the wireless geosensor network and does not encompass other factors. Costs related to risk evaluation, social interventions, dissemination, continuous maintenance, and social work are not included in the analysis. The CSM-EXT measurement system was excluded due to its complexity and high installation costs, particularly in densely populated areas. Consequently, its widespread implementation as an alternative to the wireless sensor network might not be viable or cost effective without conducting a detailed on-site survey. The costs associated with social work are highly site-specific and depend on various factors. These factors include whether the municipality conducts a risk assessment on the site, the community's acceptance of the LEWS installation, and the involvement of NGOs working with the community. Furthermore, the extent and nature of social work can vary significantly based on the risk awareness and social structure of the community. Drawing from the experience of the Inform@Risk project, the initial social implementation costs of the LEWS, which mainly involves social workers accompanying and explaining the installation process, producing and distributing information materials, and conducting training workshops and emergency drills, are expected to be at least comparable to the cost of implementing the technical system.

Next, the variables included in the cost function are explained. The cost of the sensor system elements was determined based on the required working time for production and installation; the 3D printing time; and the material costs such as electronics, sensors, cables, connectors, and accessories. For the working time, an hourly rate of EUR 15 was used, which

corresponds to the approximate hourly cost (inclusive of all insurances and benefits) of a geotechnician in Colombia (as of 2022). The cost of 3D printing time was estimated at EUR 3 per hour, considering the filament cost, power consumption, maintenance, and the overall investment cost for the 3D printer spread over the lifespan of 10 000 operating hours. Material costs encompass all the components needed to construct and install the system. The costs are calculated without including VAT. For a detailed list of the required materials and step-by-step instructions for constructing and installing the sensor system, please refer to the Inform@Risk Wiki.

Regarding the sensor system design, a node density of eight nodes per hectare is established for areas with the highest susceptibility level of 1, resulting in a dense grid of nodes. This density is scaled down based on the susceptibility (i.e. a susceptibility of 0.5 corresponds to four nodes per hectare). The built-up density determines the distribution of node types. In highly urbanized areas, more infrastructure nodes are preferred, while in less urbanized areas, more subsurface nodes are considered. Therefore, the node density is multiplied by the total area to determine the necessary number of nodes per site. Subsequently, this number is multiplied by the built-up density to calculate the number of infrastructure nodes, with the remaining nodes being assigned as subsurface nodes. As for the gateways, it is assumed that at least one gateway per 25 ha is necessary, although having at least two gateways for redundancy is suggested to ensure a backup in case of failure. This is a conservative assumption, as typically not as many gateways will be required. Previous tests conducted in the city of Medellín indicated that using two to four gateways in the city centre could provide sufficient transmission reach to cover the entire eastern slope of the city. The cost for the three different sensor systems is presented in Table 1.

The cost function is calculated following Eq. (1):

$$\begin{aligned}\mathrm{COST} = {} & S \times 8 \times A \times (B_{\mathrm{DENS}} \times \mathrm{EUR}\,250 \\ & + (1 - B_{\mathrm{DENS}}) \times \mathrm{EUR}\,535) + G \times \mathrm{EUR}\,2220, \end{aligned} \quad (1)$$

$$\text{where } G = \begin{cases} A \le 25\,\mathrm{ha} = 1 \\ A > 25\,\mathrm{ha} \,\&\, A \le 50\,\mathrm{ha} = 2 \\ A > 50\,\mathrm{ha} = 3 \end{cases}.$$

COST represents the estimation of monitoring instrument costs for a LEWS in a specific site. The variables considered in the cost function include the landslide susceptibility (S), the area of the site in hectares (A), and the built-up density within the site (B_{DENS}).

The cost function results are used to develop alternative cost-effectiveness scenarios to support decision-making for the implementation of new LEWSs. Five scenarios are evaluated, including prioritizing (1) the total cost of the LEWS, (2) the cost per person, (3) the total population (exposed and vulnerable), (4) the landslide susceptibility, and (5) a combination of the aforementioned scenarios (1–4). In scenarios (1) and (2), priority is given to the lowest costs, while in scenarios (3) and (4), priority is given to the highest population and susceptibility. In the combined scenario (5), the values are normalized using the min–max scale, where 1 represents the highest priority and 0 the lowest. The normalized values are then mapped and plotted on a graph, providing decision making support.

Table 1. Cost for the different monitoring sensor systems.

System	3D printing time and cost (EUR 3 per hour)	Working time and cost (EUR 15 per hour)	Material cost	Total cost
Infrastructure node	4.4 h EUR 13.2	1.5–1.75 h EUR 22.5	EUR 215	EUR 250
Subsurface node	40.7 h EUR 122.1	3.75 h EUR 56.25	EUR 355	EUR 535
Gateway	–	8 h EUR 120	EUR 2100	EUR 2220

3. Results

3.1. Landslide susceptibility map: exposed and suitable sites for LEWSs

The mapping of landslide susceptibility achieved an overall accuracy of 75.26 % (Fig. 3a). The sensitivity and specificity of the generated map were 80 % and 71 %, respectively. Sensitivity represents the percentage of correctly predicted high-hazard (1) class instances, while specificity represents the percentage of correctly predicted non-hazard (0) class instances. These metrics indicate that the model tends to slightly overestimate landslide susceptibility. It is important to note that, although the reference data are discrete, the susceptibility map is a probability map with continuous values ranging from 0 to 1 (a detail can be seen in Fig. 3b). Therefore, the accuracy is measured by considering probabilities equal to or higher than 0.5 as high-hazard (1) and probabilities lower than 0.5 as non-hazard (0), without considering the degree of susceptibility for the validation metrics.

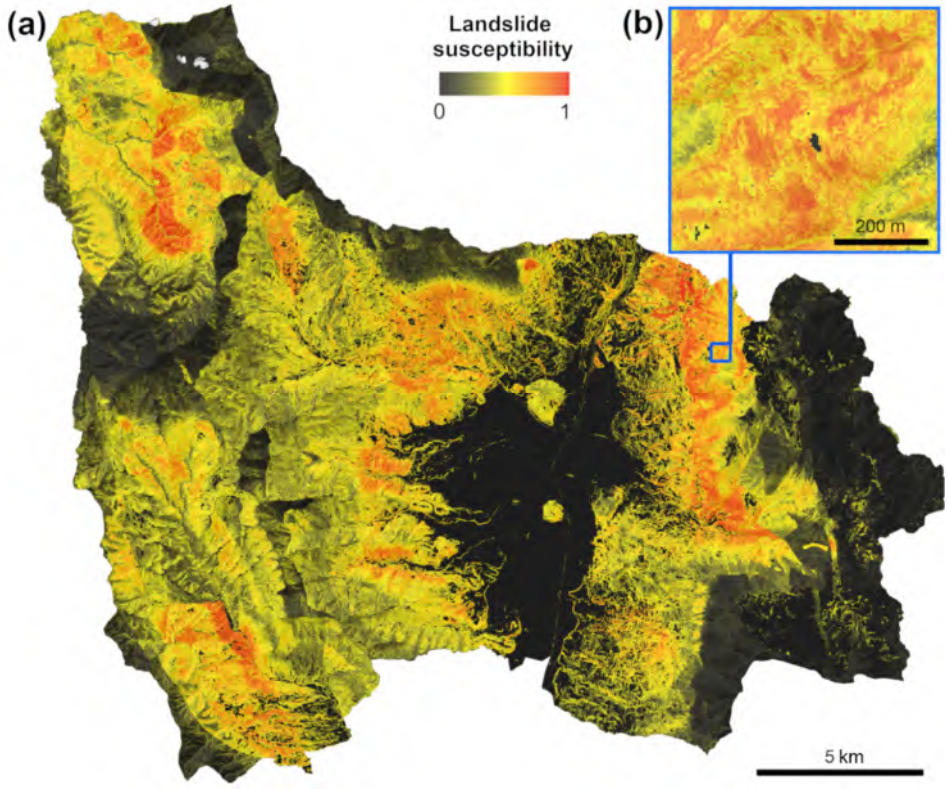


Fig. 3. (a) Landslide susceptibility map (after filtering slopes equal to or lower than 10◦) and (b) detail in Bello Oriente neighbourhood.

To identify exposed sites suitable for potential LEWS implementation considering the risk factors, the combination of landslide susceptibility and vulnerable and exposed elements was performed. A total of 44 seeds were identified (Figs. 4a and b), which served as reference points for identifying susceptible and populated areas surrounding them, thereby delineating the boundaries of the sites. Subsequently, sociodemographic and topographic factors were calculated and used to assess site suitability. Among the identified sites, a total of 16 sites were deemed unsuitable for node-based LEWS implementation based on expert recommendations (Fig. 4c). The primary reason for discarding these sites was their high building density, with limited available open space for installing subsurface nodes and restricted monitoring capabilities of the LEWS to only infrastructure nodes. Additionally, some of the remaining 28 pre-selected sites were subdivided into smaller areas or modified based on considerations such as topography, built-up density, and road network in order to accurately define the boundaries of the suitable exposed sites (Fig. 4d).

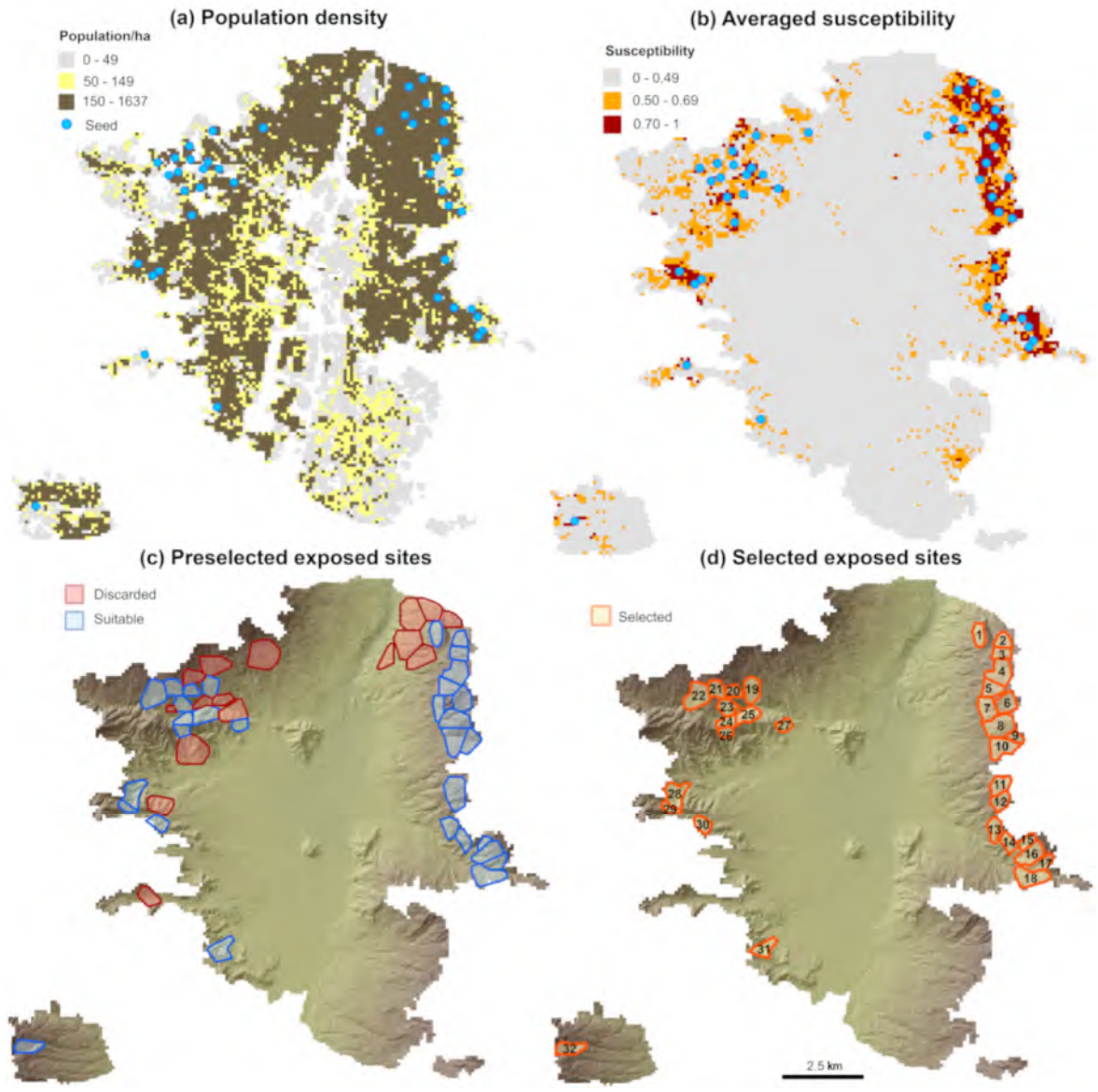


Fig. 4. Identification of seeds based on **(a)** highly populated areas and **(b)** susceptible grids. **(c)** Automatic delineation of exposed sites by means of seeds' areas of influence and preselection of suitable sites. **(d)** Suitable exposed sites after the manual delineation of their boundaries.

The manual delineation process led to the identification of 32 preferred sites suitable for node-based LEWS installation (Fig. 4d). Socio-demographic and topographic factors used in the cost estimation function were calculated for these definitive sites (Table S1 in the Supplement). These sites are primarily located in the north-eastern part of Medellín, with additional sites in the east and west, characterized by high population densities and high landslide hazards (see Figs. 2 and 4). On average, these sites cover an area of 27 ha, with an average built-up density of 20 % and an average population density of 224 people per hectare (Table S1). The most densely populated site, located in Área de expansion Pajarito (site 21), has a population density of 512 people per hectare and a low built-up density. This area corresponds to a city expansion zone with high-rise buildings on the west side of the valley and no identified vulnerable population. On average, the sites are approximately 34 % open land, which is essential for installing subsurface nodes. In terms of slope, the selected sites have an average slope of 24◦, ranging from a minimum of 15◦ in the west to a maximum of 35◦ in the east. As for landslide susceptibility, the average value is 0.68, with a minimum susceptibility of 0.51 and a maximum of 0.85 for the selected sites.

3.2. Cost estimation for the instrumentation of LEWSs

Based on the generated landslide susceptibility map and the factors outlined in Table S1, the estimated costs for installing monitoring sensors in the selected sites for LEWSs are calculated using Eq. (1). In order to facilitate and guide decision making regarding the next LEWS installation in the city, various cost-effectiveness scenarios are evaluated. These scenarios go beyond the monetary efficiency and consider additional priorities such as the number of exposed and vulnerable individuals, as well as the landslide susceptibility of each site. By considering these factors, a more comprehensive understanding of site characteristics is obtained. Consequently, five priority scenarios are examined including prioritizing the total cost of LEWS, cost per person, total population exposed within the site (including vulnerable individuals), landslide susceptibility, and a combination of these factors.

Figure 5a provides information on the population that could potentially benefit from a LEWS in relation to the economic resources required in the combined scenario. It represents the overall costs of the system plotted against the total population per site. The priority level is indicated by a greyscale gradient, with darker shades indicating higher priority. The cost of the systems ranged from EUR 26 000 (≈ COP 132 million with a conversion rate of EUR 1/ COP 5040 at the time of writing) to EUR 157 000 (≈ COP 789 million). In terms of cost per person (p.p.), the estimated range was EUR 5 to EUR 41 p.p. (≈ COP 23 000 to 204 000). Therefore, if the goal is, for instance, to prioritize the most affordable system (i.e. focusing on LEWS cost), the LEWS in El Corazón (site 29) located on the western slopes of the city is the least expensive option (Fig. 5a). However, it covers a smaller population compared to other sites with similar

costs. In this sense, the LEWS in El Pesebre (site 27) is the second least expensive option for instrumentation. It covers more than double the exposed population compared to site 29, although it is important to note that most of the population is not classified as vulnerable.

To demonstrate the potential of the proposed cost function, a case scenario is simulated where the city of Medellín has a budget of COP 2000 million (≈ EUR 397 000) allocated for the implementation of LEWSs. This simulation aims to illustrate the number of LEWSs that could be installed within a specific budget, considering different cost-effectiveness scenarios. Based on these scenarios, several sites are suggested as potential starting points for implementing the systems, depending on different priorities. Figure 5b shows the locations of the sites where LEWSs could be installed using the allocated budget, according to the different scenarios (overall cost, cost p.p., exposed and vulnerable population, landslide susceptibility, and the combined scenarios) using values from Table S1. The colour assigned to each site represents the prioritization based on the corresponding scenario. Sites that have more than one colour are prioritized in multiple scenarios. The accompanying table in Fig. 5b provides information on the total number of LEWSs that could be instrumented under each priority scenario, the total cost, average cost p.p., the total number of exposed and vulnerable people, and the average susceptibility.

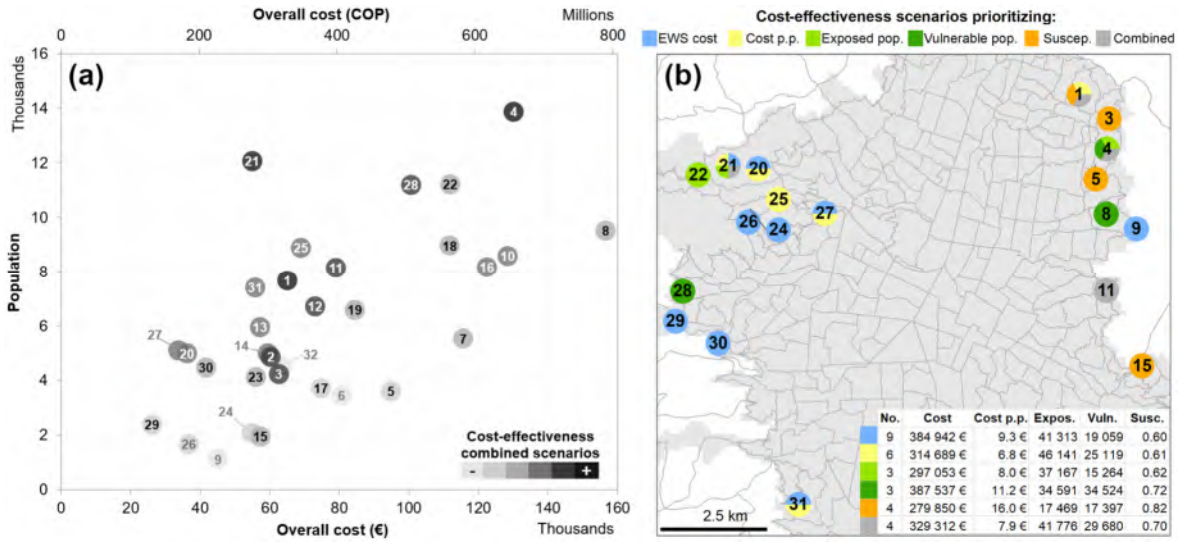


Fig. 5. (a) Overall costs of the monitoring system installation of a site-specific LEWS versus the number of people for each site. The label number represents the site ID, while the grey tone represents the priority per site based on the cost effectiveness of the combined scenario. Cost is given in euros (EUR) and Colombian pesos (COP). (b) Based on the priority, the map shows the sites where LEWSs could be installed with an allocated budget of COP 2000 million. Values are summarized in a table with the number of LEWSs, total cost, cost p.p., people exposed, people vulnerable, and susceptibility.

With the same available budget, it is feasible to instrument nine LEWSs if the objective is to minimize the total cost (shown in blue in Fig. 5b). These LEWSs are predominantly located on the western slopes and cover a population of approximately 41 000 individuals, including 19 000 who are considered highly vulnerable. The average landslide susceptibility in these areas is around 0.6. However, when considering all priorities, four LEWSs can be installed (represented in grey in Fig. 5b) at a lower average cost p.p. These LEWSs are primarily located on the eastern slopes, with only one situated in the western part of the city. In this scenario, the number of exposed individuals covered by the LEWSs remains at 41 000, but the proportion of highly vulnerable individuals increases to nearly 30 000. Additionally, the sites exhibit higher susceptibility levels.

For instance, one site of great significance in several scenarios is located on the north-eastern slope of Medellín, between the south of Carpinelo and north of Maria Canocarambolas (site 4). This site is home to almost 14 000 highly vulnerable individuals residing in precarious settlements, with a relatively high average landslide susceptibility of 0.74. The estimated cost p.p. at this site is EUR 9.4 (COP 47 000), indicating that with an approximate budget of EUR 130 000 (COP 656 million), a substantial proportion of the exposed and highly vulnerable population can be covered. Likewise, the LEWS in Santo Domingo el Savio 1 (site 1), located on the north-eastern slope, is the most effective in terms of cost p.p., susceptibility, and combined priorities. However, if the objective is to cover the maximum number of exposed individuals, this site may not be the most suitable option. In that case, the LEWS in Área de expansion

de Pajarito (site 21), situated on the western slopes, covers a greater number of exposed people, offers a more effective cost p.p., and has a lower overall system cost compared to site 1. However, its lower landslide susceptibility reduces the likelihood of a landslide occurrence. On the other hand, the LEWS in El Corazón (site 29) would be the most affordable option with a low cost p.p., yet it exhibits the lowest susceptibility and covers fewer exposed individuals, which may impact its suitability. Similarly, the most expensive LEWS in La Cruz (site 8), on the north-eastern slope, has a reasonable cost p.p., while the number of exposed individuals and the landslide susceptibility are significantly higher.

Figure S1 in the Supplement illustrates the 32 sites in descending order of priority for the installation of site-specific LEWSs. The figure presents the cumulative value of the five priorities, where the exposed population is divided by vulnerability. By considering all priorities together, decisionmakers can assess the combined value in Fig. S1, enabling them to identify sites that are highly prioritized across multiple cost-effectiveness scenarios simultaneously. This graphical representation assists in locating the optimal site for implementing the next LEWS, taking into account various considerations such as available funds (as depicted in Fig. 5b) or the size of the highly vulnerable population.

Furthermore, Fig. 6 illustrates the cost function developed for prioritizing the installation of LEWSs based on the different cost-effectiveness scenarios. If the city intends to implement LEWSs in all 32 proposed suitable sites, a total budget of EUR 2.4 million (≈ COP 12 100 million) would be required to cover the 200 000 exposed individuals. The trend lines in Fig. 6 depict the population covered by LEWSs based on the available funds and the priority scenario. Prioritizing the cost p.p. yields the most efficient result in terms of budget utilization and population coverage. However, as observed in Fig. S1, this approach overlooks landslide susceptibility and the overall exposed population. In this regard, the combined scenario, which considers all relevant factors, demonstrates a similar trend and is recommended for prioritizing the installation of new LEWSs. As an example, the priority of site 4 is displayed for all cost-effectiveness scenarios. By utilizing these scenarios, policy-makers can make conscious and informed decision regarding the installation of LEWSs, including the purpose and location.

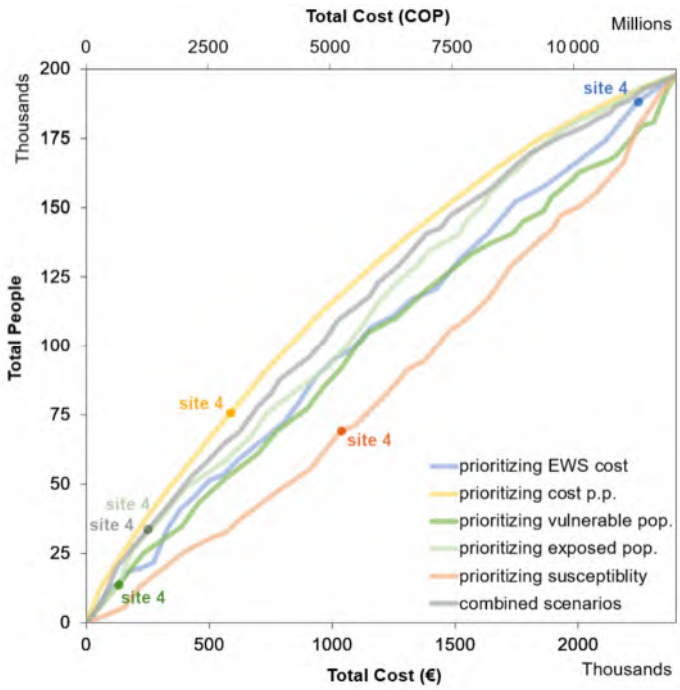


Fig. 6. Cost function based on different cost-effectiveness scenarios and their combination. It shows the number of people that can be covered by the available funds, based on the preferred priority. Site 4 is highlighted to exemplify the particular priority within the cost-effectiveness scenario being considered.

4. Discussion

In recent years, there has been a significant increase in the utilization of data-driven methods and EO-derived data for mapping landslide susceptibility (e.g. Abu El-Magd et al., 2021; Ado et al., 2022; Azarafza et al., 2021; CalderónGuevara et al., 2022; Eiras et al., 2021; Nikoobakht et al., 2022; Taalab et al., 2018) and for generating detailed population distribution maps (Sapena et al., 2022; Metzger et al., 2022). These advancements have enabled the identification of highly exposed areas prone to landslides in various regions worldwide (Garcia et al., 2016; Modugno et al., 2022; Kühnl et al., 2023). This study proposes a comprehensive workflow that can be applied to identify exposed areas prone to landslides that are suitable for the implementation of low-cost and site-specific LEWSs. Furthermore, it adds to the current literature, as a cost estimation function for the instrumentation of the LEWSs has been developed that considers factors such as area, landslide susceptibility, and building density, allowing for the assessment and comparison of estimated costs across multiple sites. This integrated approach facilitates informed decision-making processes by prioritizing actions based on cost effectiveness. One of the key contributions of this study is the provision of an open and transparent cost estimation for LEWSs, serving as a valuable reference for other regions.

Through the application of the proposed workflow, more than thirty critical locations characterized by high exposure, high vulnerability, and high susceptibility to landslides were successfully identified in Medellín. These locations can be assessed by the municipality of Medellín to implement LEWSs based on the available budget. Implementing LEWSs in these areas has the potential to enhance the resilience of thousands of individuals residing in various parts of the city. Moreover, by utilizing the developed cost function the price to

instrument the monitoring component of a LEWS in each location was estimated, and several cost-effectiveness scenarios that align with the financial resources allocated for risk management were suggested. As a result, this study provides valuable decision-making support on where to proceed with LEWS implementation following the successful deployment in Bello Oriente (Werthmann et al., 2023). With this, a conscious, informed, and transparent policy decision can be supported – where to install a LEWS under limited available financial funds. At the same time, however, the developed scenarios show the complexity of planning and political decisions: if one decides for the most cost-effective way or to protect the most people, the most endangered areas are not necessarily instrumented. Moreover, decisions may mean reducing or increasing inequalities, depending on whether precarious settlements are preferred or not. Every decision implies advantages and disadvantages – science can at least support the choice in an information-driven way. The planning or political decision itself, however, must be fought out according to democratic processes.

This study successfully demonstrated the feasibility of accurately mapping landslide susceptibility by using a data-driven approach that integrates remotely sensed data and ancillary datasets. The application of a non-parametric random

forest model allowed for the incorporation of multiple conditioning factors, capturing non-linear relationships within the data. This approach overcame the limitation posed by the non-normal distribution of factors and the requirement for prior knowledge of complex interactions among them. Consequently, an updated landslide susceptibility map was generated, complementing the city's official hazard map from 2014. The new map incorporated new instances of mass movement obtained from landslide inventories and remote sensing.

A novel contribution was done by transparently estimating the cost of instrumenting the monitoring component of LEWSs, which, to the best of the authors' knowledge, has not been previously reported. This new information is valuable for decision-makers involved in disaster risk reduction, where LEWSs play a crucial role (UN, 2015). The proposed automatic monitoring system was designed to be highly modular, scalable, and customizable, aligning with the objectives of community-based LEWSs (Gumiran et al., 2019). This design approach facilitated the transferability study, allowing for area-wide implementation across the entire city based on the experiences gained from installing a LEWS in one neighbourhood. Moreover, the system has the potential to be easily transferred to other regions worldwide with similar characteristics such as densely populated and mountainous areas. Furthermore, being based on the open LoRa® standard, the system can leverage existing infrastructure (i.e. gateways with a cost of EUR 2200 each). This feature streamlines the transferability within a city and reduces costs when scaling up the systems.

The impact of various cost-effectiveness scenarios on the overall cost and the number of exposed and highly vulnerable populations when allocating a specific budget, highlighting the potential for optimization, was demonstrated. The combination of a highly detailed on-site warning system with city-wide information about susceptibility and population provides unique opportunities for researchers and decisionmakers. From an economic standpoint, it

is possible to determine the most effective allocation of resources in terms of the cost-to-exposed-inhabitant ratio. The analysis offered a clear indication of where a LEWS can be the most cost effective. However, it is important to acknowledge that there are several additional factors that contribute significant value to the benefits but are difficult to quantify. These factors include reducing risks by increasing risk awareness, fostering social networks, and facilitating communication between informal dwellers and city planning agencies.

LEWSs have long being recognized for their effectiveness in providing timely warnings to at-risk communities. However, their widespread adoption has been hindered by high implementation costs. This study tackles this challenge by proposing a low-cost LEWS tailored for highly exposed areas, thus lowering the barriers to global adoption. To enhance affordability and sustainability, a wireless network of low cost sensors equipped with solar panels is proposed. This not only improves system maintenance but also optimizes cost effectiveness by strategically targeting areas with the highest vulnerability. By focusing on these areas, the LEWS can maximize its benefits while keeping implementation costs manageable. This approach ensures that resources are allocated where they are most needed, benefiting the local community and enhancing the overall effectiveness of the system. By combining innovative technology and cost optimization strategies, this study aims to promote the wider adoption of LEWS in landslide-prone regions. By doing so, it lowers the barriers associated with expensive implementation, ultimately making LEWS more accessible and beneficial to communities at risk.

However, there are still challenges and limitations to address. In terms of the landslide susceptibility map, the accuracy was affected by evaluating the probability map using a dichotomous variable (high-hazard or non-hazard). The generalization into a dichotomous value influenced the measured overall accuracy, especially considering the frequency of values ranging between 0.4 and 0.6, which were classified as either high hazard or non-hazard. Most of the misclassifications in the accuracy analysis occurred in the medium hazard class, which is here considered non-hazard. Additionally, misclassifications were observed in low-hazard areas with medium–high susceptibility values, as well as highhazard areas with low susceptibility values. These discrepancies may arise from using two different sources as reference data for model training, which can lead to contradictions, particularly when community reports in the DAGRD inventory may not always pertain to landslides. For instance, areas with medium or low hazard levels in the POT 2014 hazard map, but with similar conditioning factors to areas with recent landslides in the inventories, could exhibit high susceptibility levels in the model, resulting in validation errors. It is essential to highlight that data-driven methods rely on the quality and veracity of the reference data, as well as on the conceptual approach. In this study, on the one hand, most of the landslides from inventories are reported by people, leading to a potential bias of higher landslide frequency in urban areas. However, since the focus is on implementing LEWS in urban areas, this bias is not critical to the study. Furthermore, all historic landslides in Medellín have a strong anthropic triggering factor, so it is logical to observe a higher density of landslide events in the urban areas. On the other hand, while the official hazard map is expected to be of high quality due to expert evaluation, a significant number of reported landslides between 2014 and 2021 occurred in – according to the hazard map of 2014 – low-

hazard areas. These areas potentially include mass movements reported by citizens that may not necessarily be landslides. Another factor that may impact the results is the population increase along the urban– rural border since the creation of the official hazard map. Given the significant influence of anthropogenic factors on landslides in Medellín, this may account for the differences

between the generated landslide susceptibility map and the official landslide hazard map.

The selection of suitable exposed locations for the installation of LEWSs followed an iterative process using seeds of high susceptibly and exposure, along with thresholds set manually. These thresholds were established to identify the areas with the highest exposure based on population density and susceptibility while also limiting the number of initial seeds to a manageable quantity. However, it is important to acknowledge that the specific thresholds and the chosen model for mapping landslide susceptibility may have influenced the identification of exposed suitable sites in Medellín. Additionally, the size of the sites was constrained based on the previous experience from Bello Oriente. These factors could impact the size, complexity, and overall accuracy of the identified sites, consequently affecting the cost estimation. Nevertheless, the final selection of more than thirty locations covers a significant portion of highly populated exposed locations in Medellín, encompassing a diverse range of sizes and complexities. This allows for comparability in cost estimation across different sites. It is important to note that the proposed thresholds and sizes are adapted to the specific context of Medellín and should be adjusted accordingly in different regions.

The cost estimation for the proposed system is inherently subject to uncertainties, both in the cost function itself and the underlying data. An important aspect to consider is that the proposed system has only been implemented once in Bello Oriente, which possesses unique characteristics such as being an informal settlement with low-quality building structures, medium building density, steep slopes, and ample open spaces. The cost and density of sensors are based solely on the experiences at this specific site. While the cost of instruments and working hours have been provided for this case, extrapolating these costs to other areas involves certain assumptions and may not account hidden costs that could arise in different conditions. However, this cannot be assessed with the current knowledge. Additionally, local circumstances and the availability of trained personnel also contribute to the overall cost. While the absolute costs might have significant uncertainties, the proposed cost function allows for the evaluation of relative cost differences between multiple sites. Therefore, the proposed approach can be used to identify prioritized areas within a city for the initial installation of new LEWSs based on cost-effectiveness considerations.

The implementation of the LEWS prototype in Bello Oriente was a collaborative effort involving academia, private companies, government agencies, local civil society organizations, and the local community. The active involvement of these stakeholders played a vital role in addressing a range of challenges, including social conflict, insufficient risk awareness, limited political commitment, changes in local government, resources constraints, and inadequate territorial planning (Werthmann, 2023). This means that the

implementation of a LEWS must always be seen as a socio-technical challenge. This involves costs that were not considered in the technical estimates.

Given the complexity of implementing a LEWS, it is important to emphasize that the estimated costs presented in this study solely pertain to the installation of the monitoring instruments. Therefore, expenses related to system maintenance, including significant working hours, additional instruments, and protective measures for the instruments, have not been accounted for. The costs of warning elements, safety signs for emergency routes and meeting points, community engagement, and collaboration with various sectors (government, local civil society organizations, etc.), as well as the operational aspect of the LEWS, have also been excluded from the results. Finally, it is crucial to emphasize that beyond financial aspects, efficient implementation and maintenance, and cooperation among diverse stockholders, the success of the LEWS relies on social work and community engagement. The cultivation of risk awareness and fostering trust in the system define the willingness of individuals at risk to actively participate in the process and respond promptly to warnings. This social aspect is essential for ensuring the long-term sustainability of the LEWS after the monitoring instruments have been installed.

5. Conclusions

The implementation of multi-hazard EWSs is crucial in mitigating disaster risks and safeguarding lives, as emphasized by the Sendai Framework for Disaster Risk Reduction 2015–2030 (UN, 2015). This is particularly significant in countries like Colombia, where a considerable proportion of the population is exposed to landslide hazards and is highly vulnerable.

Building upon the insights gained from the LEWS installed in Medellín through the Inform@Risk project, this study identified 32 highly exposed areas in the city suitable for the installation of a highly modular, scalable, and customizable LEWS. By estimating the required investment for the monitoring component of these systems, the city would need approximately EUR 5 to EUR 41 per protected inhabitant, varying based on site-specific characteristics.

In an approach for prioritizing the selection of exposed sites for a LEWS considering various cost-effectiveness scenarios, budget constraints, landslide susceptibility, total population exposed, and territorial planning agenda, an informed decision-making is supported by science. These findings are intended to provide guidance to decision-makers and support disaster risk reduction efforts not only in Medellín, Colombia, but also in other regions facing similar challenges. By leveraging the lessons learned from this study, policy-makers and stakeholders can make informed decisions to enhance resilience and reduce the impact of landslides, contributing to the broader goals of disaster risk reduction.

Supplement

The supplement related to this article is available online at:

https://doi.org/10.5194/nhess-23-3913-2023-supplement

Acknowledgements

This research is part of the Inform@Risk project (Strengthening the Resilience of Informal Settlements against Slope Movements). We thank Leibniz Universität Hannover, Technische Hochschule Deggendorf, Technische Universität München, Deutsches Zentrum für Luft- und Raumfahrt e.V., AlpGeorisk, Sachverständigenbüro für Luftbildauswertung und Umweltfragen, EAFIT University–URBAM, Alcaldía de Medellín, Área Metropolitana del Valle de Aburrá, Sistema de Alerta Temprana del Valle de Aburrá, Sociedad Colombiana de Geología, Colectivo Tejearañas, Corporación Convivamos, Fundación Palomá, and Red Barrial Bello Oriente for their support.

Financial support

This research was funded by the German Federal Ministry of Education and Research as part of the FONA Client II initiative, grant number 03G0883A-F. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Chapter XXIV

Was global urbanization from 1985 to 2015 efficient in terms of land consumption?

Habitat International, 2024, vol. 160, 103397

Hannes Taubenböck, Johannes Mast, Richard Lemoine Rodríguez, Henri Debray, Michael Wurm & Christian Geiß

© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (https://creativecommons.org/licenses/by-nc/4.0/).

https://doi.org/10.1016/j.habitatint.2025.103397

Received 23 Dec 2024; Accepted 30 Mar 2025; Published 11 Apr 2025

Abstract

Dense cohabitation is associated with higher ecological sustainability and economic benefits than their less dense counterparts. Of course, there are also many challenges associated with high density, but from the perspective of land consumption, the global trend of urbanization is viewed positively. Urbanity and density, however, are and continue to develop very heterogeneously within and between cities. In this study, we analyze morphologic and demographic urban expansion across 1567 major cities around the globe from 1985 to 2015. Thereby, we investigate the changes in the morphological and population densities in the newly built areas in comparison to the status quo in 1985. Furthermore, we use the different densities in cities around the world to design theoretical retroactive scenarios of alternative land consumptions and determine possible carrying capacities. The key findings are: areas built between 1985 and 2015 are less dense in terms of morphology and population than the areas built before 1985. Thus, urban expansion consumed comparatively more land which means that it has become more inefficient with respect to the usage of space. We also find that this inefficiency varies around the globe, with South Asia and East Asia having higher densities in contrast to Europe and North America. The scenarios developed show the unutilized potential: If urban expansion would have been developed more densely, e.g. by ‘compact mid-rise’ structures, our model calculations show that 1.466 billion more people could live in the same newly built urban areas.

Keywords

Urbanization; built-up density; population density; urban morphology; local climate zones; urban expansion; remote sensing; sustainability

1. Introduction

Urbanization is viewed in the literature as a multidimensional process of physical, demographic, social, economic, political and environmental components in which the benefits outweigh the costs (e.g. Angel et al., 2021; Glaeser, 2010; Hollis, 2013). Cities are, for example, more productive than smaller communities (World Bank, 2022), their innovative strength is higher (Bettencourt, 2007), and land consumption is significantly less than in rural landscapes (Schiavina et al., 2022; Taubenböck, 2021), which leads to better preservation of natural habitats on the peripheries of cities (Johnson & Klemens, 2005). At the same time, however, urbanization also brings the burdens of density. The exposure towards lower air quality (e.g. Erbertseder et al., 2024), the effects of the urban heat island (Leichtle et al., 2023; Lemoine-Rodríguez et al., 2022), traffic congestion (Rahman et al., 2021), among many other issues that have been documented to increase with urban density or, in other words, the challenge to manage density.

Considering both, beneficial and negative aspects of density, it is neither obvious nor clear which type of urbanization we should favor. Against the background of global population growth – by 2050 we will be around 10 billion people (IPCC, 2022) – and the unbroken trend of urbanization – with urban areas now home to more than 56 % of the world's population, and projected to increase to 68 % by 2050 (United Nations Department of Economic and Social Affairs Population Division, 2018) – we need to ask ourselves how we want to develop the cities of the future.

In this study, we aim to lay the groundwork to empirically frame this central question based on multitemporal building structural developments in cities across our planet. Although the dynamics of global urbanization in its physical and demographic manifestations have been documented in manifold studies (e.g. Angel et al., 2011; He et al., 2019; Huang & Xu, 2022; Seto et al., 2011; Taubenböck et al., 2012; United Nations Department of Economic and Social Affairs Population Division, 2018) and regional differences in their dynamics have been identified (e. g. Güneralp et al., 2020; Huang, Li, Yang, & et al, 2021; Liu et al., 2020; Melchiori et al., 2018; Taubenböck et al., 2024), there is no systematic, global study of the structural forms at intra-urban level analyzing in which this urban expansion has taken place.

Even if cities are primarily defined by people (Gehl, 2009; Jacobs, 1961) and the complex networks, material flows and exchanges (Batty, 2013; Inostroza & Zepp, 2021; Sassen, 2002), all of this is embedded in the physical built space. The built and natural urban spaces form the basis for actions, everyday rhythms, etc. and thus have a great influence on the livability as well as the sustainability of a city (e.g. Sennett, 2018, p. 368; Tonkiss, 2013).

Buildings, streets, plots, and open spaces can be seen as the basic ingredients (Kostof, 1991; Salat, 2011) of the built urban space or '*technomass*' (Inostroza & Taubenböck, 2024). Combinations of these ingredients, just like DNA, turn into unique configurations with no two cities across the world being identical. Many studies have demonstrated the variability of built structures and patterns in cities across the globe: From theoretical observations (e.g. Burgess, Park, & McKenzie, 1925; Hoyt, 1939; Harris & Ullman, 1945) to various models (e.g. Alonso, 1964; Garreau, 1991; Vance, 1964), to empirical studies measuring and comparing the built space and its urban patterns (e.g. Schneider & Woodcock, 2008;

Frolking et al., 2013; Braun et al., 2023; Huang et al., 2007). With an improved availability of global remote sensing data and new possibilities for evaluation using AI techniques (e.g. Gong et al., 2020; Sun et al., 2022; Zhou & Weng, 2024; Zhu et al., 2022), data sets that make these empirical studies possible are now available. These range from multi-temporal global classifications of settlement evolution (e.g. Huang, Li, Yang, & et al, 2021; Marconcini et al., 2018; Pesaresi et al., 2015; Zhou et al., 2018) to intra-urban classifications of the built landscape (e.g. Bechtel et al., 2015; Demunzere et al., 2022; Li et al., 2022; Zhu et al., 2022). Approaches to documenting the intra-urban configurations of built landscapes are based on parameters such as building density, heights or types (e.g. Taubenböck, Kraff, & Wurm, 2018; Xu et al., 2019; Liu et al., 2024), the fragmentation of urban landscapes (e.g. Angel et al., 2012; Liu et al., 2022) or the spatial arrangements in complex or ordered geometric patterns (e.g. Lemoine-Rodríguez et al., 2020; Debray et al., 2023). The variability of structural city configurations was empirically proven and structural city types were identified (e.g. Thomas et al., 2012; Taubenböck et al., 2020; Iungman et al., 2024). It therefore stands to reason that not only are cities as a whole structurally different, but that expansion in recent decades has also brought about structural differences.

Given the documented variety of structural configurations in space, it is likely that this variation also exists over time due to the shifting of certain paradigms: for example, changes in construction technology have impacted the nature of architecture (Mumford, 1934, 1961). Another example is demographic change, which has been shown to affect the viability of certain infrastructures (Siedentop & Fina, 2010) and thus influences urban development (Schneider & Woodcock, 2008). A further example is the change in planning paradigms, which is proven to have a significant impact on urban structures (Debray et al., 2023). These can be seen, for example, in the implementation of regulations, i. e. development plans, ordinances or planning codes (Kostof, 1991) or the change of ideologies that planning follows (Taubenböck, Murawski, & Wurm, 2018; Kostof, 1991). This temporal dimension of variety in intra-urban structure, however, has not yet been empirically researched at a global level.

In this study, we therefore aim to specifically analyze urban expansion in terms of morphological-structural developments over 30 years for all 1567 cities larger than 300 000 inhabitants. Our approach focuses on four aspects of urban expansion over the period 1985–2015. *First*, we characterize quantitatively the urban fabric, here operationalized by the Local Climate Zones (LCZs) (Stewart & Oke, 2012), of the yearly additions to the built-up area of the cities. *Second*, we quantify the differences between the state of the urban fabric of the cities in 1985 and in 2015. *Third*, we compare the population density in both, the old and the newly built-up areas while controlling for the same LCZ type. And *fourth*, we provide scenarios for comparison to how urbanization occurred over the 30 years of monitoring, i.e. we calculate how many people could have become urban dwellers if the urban structures would have been developed in a denser or less dense way.

In all these analyses, we unravel the morphological variability of urbanization across regions by grouping cities at the global level as well as on the level of continents and geographic regions. For the latter, we apply regions defined by the Intergovernmental Panel on Climate Change (IPCC). The empirical results should then lay the foundation to show how cities

typically and, if necessary, regionally had grown in terms of morphological structure from 1985 to 2015 and what the consequences would be if this trend would continue over the next 30 years.

2. Conceptual foundation

Global studies on the expansion of urban settlements rely mostly on a binary understanding of the '*urban*' as built-up, artificial land cover vs. natural landscapes (Angel et al., 2011; Melchiori et al., 2018; Huang, Li, Yang, & et al, 2021; Taubenböck et al., 2024; among many others). Figure 1a illustrates the settlement expansion for three sample cities of varying dynamics based on this dichotomic classification.

The terms '*city*' or '*urban*', however, conceal very different forms of morphological design. Morphological differences in cities across the globe range, for example, from nearly 90 % building density – if you like, an almost complete usage of space – and population densities estimated at about 500 000 people per square kilometer in slums (Taubenböck, Kraff, & Wurm, 2018) to low-density, low-rise detached, single-family housing developments with about 10 % building density and only 450 people per square kilometer (Ottensmann, 2020). Thus, the terms '*city*' or '*urban*' subsume in these morphologic appearances a large variability of structural compositions of land cover and usage intensities. In previous studies, urban expansion has not been mapped at this intra-urban structural differentiation, but in dichotomous classifications. A long-term study of urban expansion with a higher resolution on intra-urban structural differences is still absent (Zhu et al., 2019). This is mainly because intra-urban structure classifications have only recently become possible with increased resolutions of satellite data and better methods on a global level (Zhu et al., 2022; Demuzere et al., 2022).

Even if such a multi-temporal global study is lacking, there are manifold studies documenting the large variability of morphological settlement structures (e.g. Conzen, 1980; Kostof, 1992; Lynch, 1984; Salat & Bourdic, 2012, pp. 97–107). These works introduce the various logics, processes and theorized ideas shaping the various urban morphologies. Various studies attempt to make the structures empirically tangible and comparable by means of spatial metrics (e.g. Batty, 2005; Boeing, 2019; Fleischmann, Romice, & Porta, 2021). Further studies monitor trends and attempt to interpret and compare the developments in urban morphology. Lemoine-Rodríguez et al. (2020) document a global trend towards homogeneous urban form. Debray et al. (2023) measure characteristic spatial features that point towards the intensity of plannedness of specific structural city appearances. Taubenböck, Kraff, and Wurm (2018) reveal the variety of physical appearances within one target class of '*arrival cities*'. Beyond such studies, there have been conceptual approaches towards the classification of urban structure types that have been operationalized: The European Urban Atlas, for example, provides a uniform mapping scheme for urban structural types applied to all cities across Europe larger than 100 000 inhabitants (EEA, 2024). The National Land Cover Database (NLCD) is another example featuring different intensity classes for developed areas in the United States of America (NLCD, 2025). One widely accepted concept to map the structural composition of cities are Local Climate Zones (LCZs). The LCZs classification frame is a generic, culturally-neutral description of

morphologic-structural realizations of land-use and land-cover (Stewart & Oke, 2012). A variety of studies attempted to map LCZs at large scales (e.g. Bechtel et al., 2015) or even globally (e.g. Zhu et al., 2022). The latter data set, for example, has been the basis to categorize and compare the variability of morphologic city configurations across the globe (Taubenböck et al., 2020). Figure 1c reveals various intra-urban morphological configurations using the three sample cities.

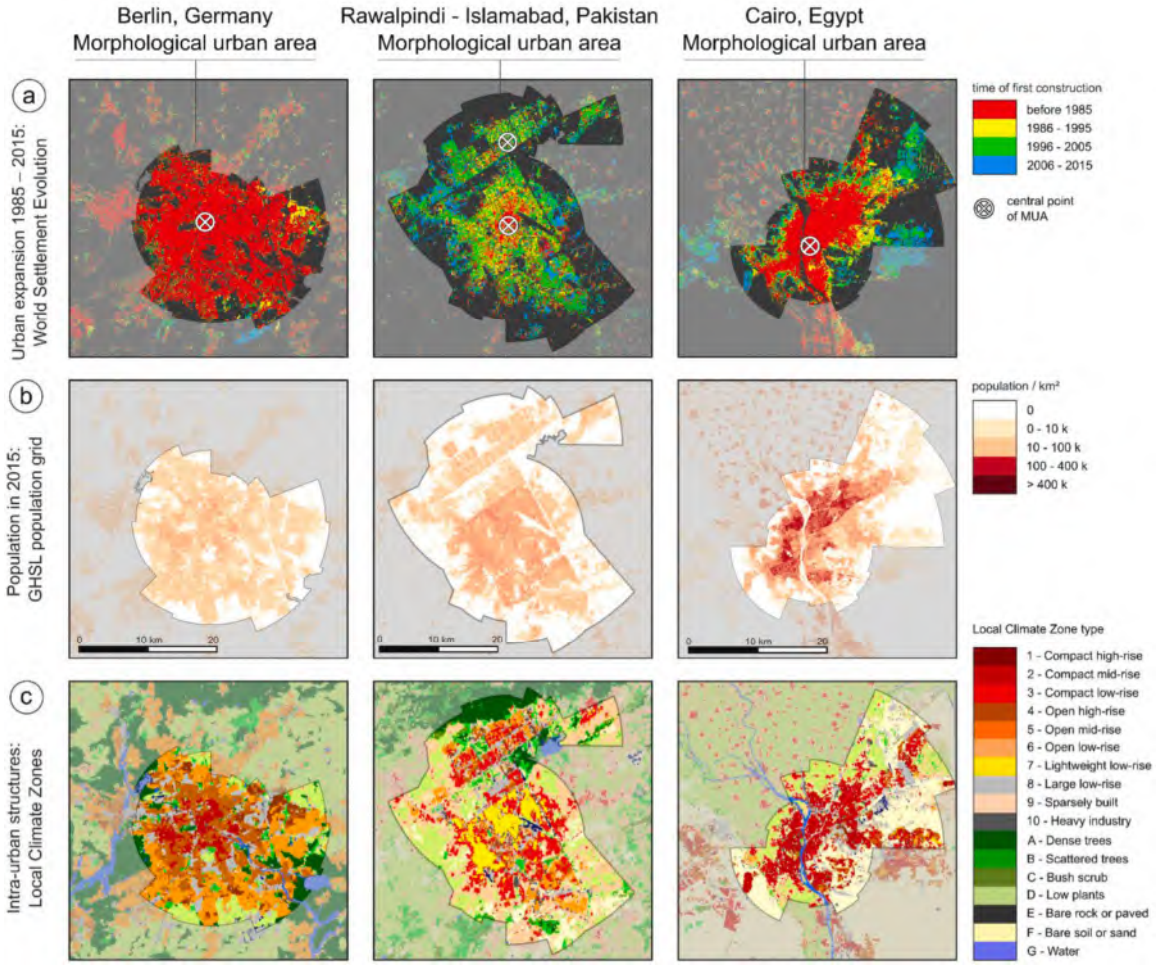


Fig. 1. Data sets used and illustrated for the examples of Berlin, Germany, Rawalpindi-Islamabad, Pakistan and Cairo Egypt. The standardized morphological urban areas (MUAs) (Taubenböck et al., 2019) are used as data-driven consistent delineation of cities. a) Urban expansion 1985–2015 using the World Settlement Evolution layer. b) Population distribution using the Global Human Settlement layer. c) Intra-urban structure using the LCZs classification scheme in 2015.

Against this background and in contrast to previous studies, we focus here on the urban areas that were newly urbanized from 1985 to 2015, i. e. previously undeveloped land that was transformed into urban areas through expansion. We specifically focus on the structural forms in the expansion areas. While the urban fringe where most expansion takes place is commonly associated with low-density, low-rise development, we examine globally whether this association stands up to empirical investigation. The higher-ranking idea in this paper is to document the morphological-structural types of urban expansion between 1985 and 2015. In a thematic perspective, we ask whether urban expansion was sustainable in terms of land consumption.

3. Data

For the analysis, we utilize a set of various data: The *World Settlement Footprint Evolution*, a global *local climate zones* classification, *morphological urban areas*, population data from the *Global Human Settlement Layer* and data on *geographic units* (i.e., continents and geographic regions).

For capturing spatial urban expansion, we rely on the *World Settlement Footprint Evolution* (WSFEvo) (Marconcini et al., 2018) which is based on Landsat data. The mapping product provides the settlement extent at 30m spatial resolution on a yearly basis from 1985 to 2015 at global scale. Figure 1a visualizes the expanding settlement patterns as mapped by the WSFEvo by the examples of the metropolitan regions of Berlin, Germany, the urban agglomeration of two administratively separate cities – Rawalpindi and Islamabad in Pakistan – which have morphologically grown together into one city, and Cairo, Egypt. The three cities illustrate the variety of urbanization dynamics across regions of the world.

For the analysis of intra-urban structural configurations, we rely on the *global local climate zones* classification as provided by Zhu et al. (2022). This data set is based on Sentinel-1 and Sentinel-2 satellite data. The LCZ classification scheme is a generic and neutral description of land-use and land-cover (Stewart & Oke, 2012). The classification scheme relies on universal, standardized and measurable parameters of urban form such as '*density*', '*building size*' and '*building height*' for the built landscape, as well as trees, low vegetation, open spaces and water for the non-built landscape. Figure 2 illustrates this conceptualization of structural types. The arrangement of the LCZs in the figure is organized according to '*high density (compact)*' and '*low-density (open)*' structural types, as this plays a methodological role later on. The numbering of the LCZs is based on the work of Stewart and Oke (2012). Figure 1c exemplifies the intra-urban LCZs structures for Berlin, Germany, Rawalpindi-Islamabad in Pakistan, and Cairo, Egypt. In this study, we assume that the LCZs of today also correspond to the LCZs of built areas in 1985. We also assume that in the newly built areas, the LCZ types that we measure today have not changed over the course of the 30 years. We have to use this conceptual simplification because LCZ classifications are not available for 1985.

For population information, we rely on the *Global Human Settlement Layer* (GHSL POP data) (Schiavina et al., 2023). It provides population information at a 100m resolution. The approach is based on disaggregating population data from censuses to a high granular resolution with the knowledge of settlement locations and respective building volumes. In this study, we rely on data for 1985 and 2015. Figure 1b illustrates the population data for our three sample cities.

As administrative units of cities do not allow for a consistent comparison across cities (e.g. Brenner & Schmid, 2013), we use here a spatial reference unit that delineates cities across the globe in a data-driven way. City boundaries have been delimited for the year 2015 (for details on data and methods, we refer to Taubenböck et al., 2019) using the Global Urban Footprint (GUF) and the GUF Density (GUF-DenS) data sets (Esch et al., 2018). From it, an urban-rural gradient is used to delineate the 'urban' from the 'rural' in a consistent way for

all cities of our sample. These units are called '*morphological urban areas*' (MUAs) and are illustrated in Fig. 1.

Our sample consists of 1567 cities larger than 300 000 inhabitants in 2015 (United Nations& Department of Economic and Social Affairs, 2015). While in the United Nations database, 1692 cities worldwide were recorded with more than 300 000 inhabitants in 2015, our data set is reduced to 1567 cities as the data set provided by Taubenböck et al. (2019) counts cities that spatially coalesced into metropolitan regions as one city. The two administratively separate cities – Rawalpindi and Islamabad in Pakistan illustrated in Fig. 1 have morphologically grown together and are in this conceptual approach counted as one city. These largest cities in the world cover 19.9 % of the entire global settlements in 2015. However, more than 27.3 % of the global population live there. These numbers demonstrate that this comparatively small area can have enormous leverage effects such as the land consumption analyzed in this study.

For the analysis of spatial variabilities of structural urbanization processes across the globe, we relate to the six continents, i.e. Africa, Asia, Europe, North America, Oceania and South America. In addition, we relate our analysis to ten geographical regions as suggested by the IPCC Working Group III in its Fifth Assessment Report (Krey et al., 2014). Here, we use the region categorization 10 (RC10). These correspond to Western Europe, Economies in Transition (Eastern Europe and part of former Soviet Union), North America, Pacific OECD (Japan, Australia, New Zealand), Latin America and Caribbean, Sub-Saharan Africa, Middle East and North Africa, East Asia, South Asia, and South-East Asia and Pacific (IPCC, 2014).

4. Methods

4.1. Indicators for measuring structural urban expansion trajectories

We use two measures characterizing the trajectory according to which urban expansion was physically structured: First, for each year between 1985 and 2015, we retrieve the spatial extent of the newly built-up area of the particular year (cf. Fig. 1a). For each of these incremental newly built-up areas, we retrieve the corresponding LCZs (cf. Fig. 1c). From it, we compute the spatial share of each built-up LCZ class. Comparing this year on year, we show to what extent the incrementally added urban fabric resembles or departs from the original structural status quo of 1985. Second, we compute the absolute changes in the share of each built-up LCZ class for the cities between their 1985 extent and their 2015 extent. With this, we aim to quantitatively evaluate if the cities have changed in their general structural configurations or if they have remained the same.

4.2. Population densities at LCZ level at particular time steps

The LCZs of the built landscape approximate the density in 2D and 3D of the urban structure (cf. Fig. 2). Because all or parts of the floor space is used as residential space, population densities correlate with these structures to a certain degree (e.g. Sapena et al., 2022). However, building usages are often decoupled from building types or volumes or usage efficiency changes due to changing behaviors such as residential space consumption.

Therefore, it cannot be assumed that the correlation is linear and we need to be aware that it also varies over time. With this in mind, we aim to measure the development of population densities in the urban areas considered, while controlling for physical density of the urban fabric.

To do so, we operationalize our analysis of the evolution of the population density at LCZ level over time. We combine the multitemporal population figures of the GHSLpop dataset with each individual LCZ built-up class. First, we calculate this only for the newly built areas between 1985 and 2015. Therefore, we use GHSLpop from 2015 and calculate the population density per LCZ. Second, we use the GHSLpop dataset from 1985 to calculate the densities for the starting year of the analysis. We calculate this for the LCZs in urban areas that existed already in 1985. Third, and finally, we calculate the population densities per LCZs for the entire respective city for 2015.

With these three measures, we aim to 1) compare whether the population densities measured in 2015 are similar or different at LCZ level between the newly developed urban areas and the areas developed prior to 1985. 2) In addition, we aim to see how the population densities in the city have changed from 1985 to 2015, within the 1985 areas, as well as at LCZ level for the entire city.

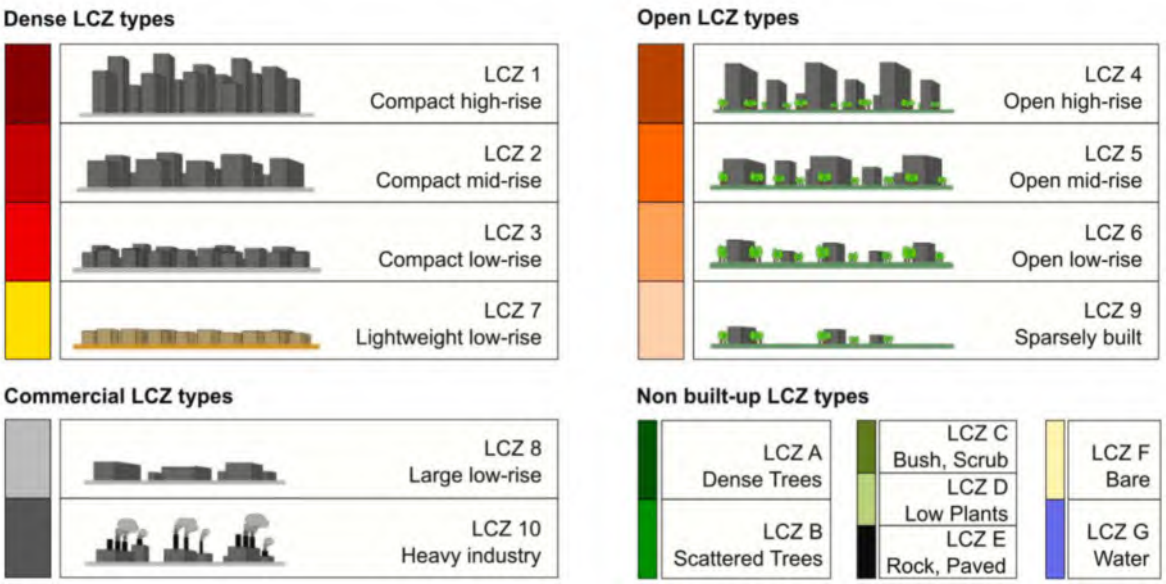


Fig. 2. Local Climate Zones (LCZs) – 17 types of a generic, culturally-neutral description of land-use and land-cover (based on Stewart & Oke, 2012).

4.3. Exploring potential population capacities of urban expansion 1985–2015

It is clear that different structural forms of development have very different carrying capacities in terms of population. While urban living is seen as living in densification, urbanization is seen as more efficient than rural structures in terms of land consumption (e.g. Glaeser, 2010). Nevertheless, population densities are highly variable: Across the globe the same LCZs might have very different population densities due to different economic, socio-demographic or cultural backgrounds in different regions. Accordingly, the question of

whether urban expansion was sustainable in terms of land consumption is not so easy to answer.

Our aim is therefore to estimate the theoretical carrying capacity of the newly built areas between 1985 and 2015. We model what the carrying capacity would be if a structural type with higher or lower population density would have been built.

We basically use four approaches for the scenarios: 1) *Single structural type expansion*; 2) *Structural type specific adaptions*; 3) *Statistical models*; and, 4) *literature-based models*.

1) For the *single structural type expansion*, we assume that the entire urban expansion area between 1985 and 2015 would have been built on with one LCZ, i.e. denser or less dense than in reality, depending on the area. We then derive the theoretical carrying capacity from the average global population density of this particular used LCZ. Subsequently, we use these densities to model the carrying capacities of the de facto expansion areas between 1985 and 2015 with only this one LCZ – denser or less dense – at a time. The result then shows in theoretical manner how many more or fewer people could have lived in the same area if a single type of structure would have been built. Of course, this is a very theoretical thought experiment, as the new settlement area would never completely consist of just one type of structure.

2) With this in mind, we are also developing '*structural type specific adaptation*' scenarios which we consider to be more realistic, as a structural change to slightly denser/less dense or slightly higher/less high built-up structures corresponds to either planning interventions or informal developments that also take place in reality. In four scenarios, we assume in the model calculations that the existing LCZs would have been implemented one structural type denser/less dense or higher/less high (Fig. 3). As example, the LCZ type 'compact high rise (LCZ1)' would be consistently less dense, i.e. 'open high-rise (LCZ4) (Scenario A – Decrease in density). Conversely, we also calculate the inverse scenarios of density increase: e.g. 'open-high rise (LCZ4)' would be consistently built denser as 'compact high-rise (LCZ1) (Scenario B – Increase in density). This is also implemented analogously for the other LCZs in terms of density and for the height of the structures to address the multiple possible scenarios (Fig. 3).

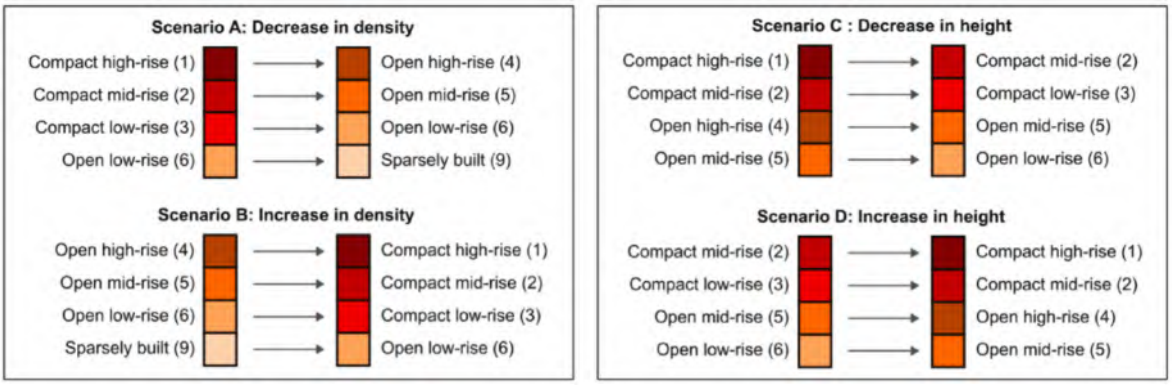


Fig. 3. Four scenarios of '*structural type specific adaptation*': Scenarios A and C depict a decrease in building density and height and Scenarios B and D depict an increase in building density and height.

All these scenarios are calculated using the LCZ-specific population densities at the regional level. For example, the conversion of 'open lowrise' to 'open mid-rise' in East Asia projects the population densities of 'open mid-rise' in East Asia, rather than the global one, to the area of 'open low-rise' in East Asia.

3) In addition, we also use *statistical models* relying on de facto population densities from the complete data set. Based on the GHSL pop, we calculate population densities among our 1567 cities. We calculate quantiles of population density, i.e. the 1 %, 5 %, 10 %, 50 % (median), 90 %, 95 %, and 99 %. We then use these population densities to calculate how many people could live in the expansion areas from 1985 to 2015.

4) And, we apply different population densities as provided by *literature.* Here we rely on densely populated areas such as in slums in Nairobi (more than 48 000 p/km^2) (based on Hu et al., 2021), Mumbai (19 865 p/km^2) (based on United Nations Department of Economic and Social Affairs, 2016) or New York City (10 948 p/km^2) (based on United Nations Department of Economic and Social Affairs, 2016) as well as on very low-density areas such as suburbs (819 p/km^2) (based on Ottensmann, 2020), or a particular low dense case such as in Pittsburgh (459 p/km^2) (based on Ottensmann, 2020).

5. Results

The 1567 cities with more than 300 000 inhabitants in 2015 had a combined built-up area of 140 810 km^2 in 1985. 30 years later, in 2015, the total built area within the MUAs increased by 120 469 km^2 to 261 278 km^2 and thus almost doubled. This expansion of the built-up area considering only these largest cities on our planet corresponds to around one and a half times the area of Austria or three times the area of Switzerland. This shows how dynamically the large cities of our planet have grown.

1.168 billion people lived in these cities in 1985. Until 2015, the total population within the MUAs increased by 844.87 million people, with 729.82 million in the newly built areas. This means that the population in these large cities has increased by almost twice the current population of the European Union, about two and a half times the population of the USA or over 10 times that of Germany.

These numbers show the dynamics and dimensions of urbanization in these large cities. Now, the questions arise as to which built structures this urbanization has taken place in since 1985 (section 5.1.), whether these newly built structures differ from the structures of the cities of 1985 and, if so, how they differ (section 5.2.), how population densities have developed overall and in the respective LCZs (section 5.3.) and what theoretical carrying capacities do the expansion areas have if other densities would have been realized (section 5.4.).

5.1. Urban expansion with increasing shares of low dense built-up structures from 1985 to 2015

The urban fringe where the majority of the urban expansion takes place is generally associated with structures of low-density and low building heights. We can indeed see that

this association stands up to global empirical investigation. Urban expansion has been carried out with increasing shares of low dense built-up structures from 1985 to 2015.

We observe decreasing shares of dense built-up types (LCZs 1, 2, 3) across the globe over the monitoring period in the expansion areas (Fig. 4). Whereas at the beginning of the monitoring period around 1985, almost 25 % of all new-built sites globally were dense LCZs, this had fallen to just over 6 % by 2015. Sparsely built areas (LCZ-9), in turn, have been established in an increasing manner. They had only a share of about 6 % around 1985, but are today the dominating newly added structural type with almost 33 %. It is particularly noteworthy that in the expansion areas, all low dense LCZs (LCZs 4, 5, 6, 9) combined are today the dominating added structural types with about 63 %.

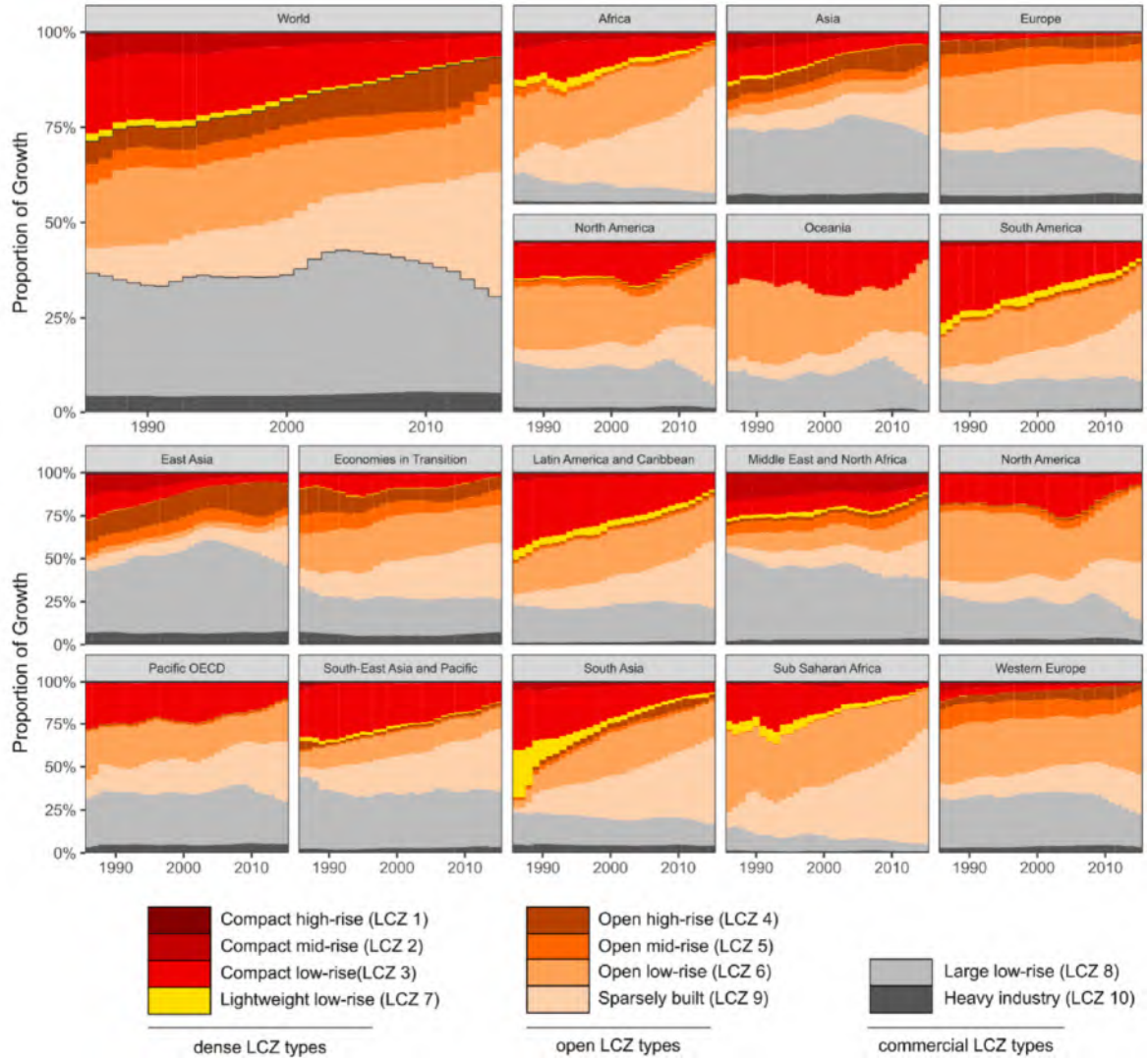


Fig. 4. The share of LCZs in the newly urbanized area in a yearly progression from 1985 until 2015 aggregated to the world, continents and IPCC regions.

Of course, structural developments are not equally distributed around the world. At continental scale, we measure that in particular in Europe 'dense' LCZs (LCZs 1, 2, 3) are almost no longer being built for urban expansion (around 1 % only). Although we are also seeing a decline in this class on all other continents, at least around 5 % of space is still being densely built in this way in Asia, North America and Africa. This type of structure still

accounts for around 10 % of the most recently built areas in Oceania and South America in 2015. It is particularly noteworthy that in South America at the beginning of the monitoring period, 'dense' LCZs were the dominant structural type of urban expansion with over 40 %. With respect to low dense LCZs (LCZs 4, 5, 6, 9), it is striking that the contributions of these LCZs are increasing on every continent. In Africa, for example, the shares of 'sparsely built' (LCZ-9) increased from almost zero in 1985 to more than 62 % in 2015. Combined with the other low dense classes, this type is dominant in urban expansion (about 88 %) in Africa. In North America, Europe, Oceania and South America the shares of the low dense LCZs (LCZs 4, 5, 6, 9) in newly built areas add up for 2015 to around 70 %. Only Asia features among the low dense classes just about over 50 % of shares in 2015. It is worth pointing out that 30–40 % of Asian expansion is related to the 'large low rise' LCZ (LCZ-8) which is substantially more than in any other area.

If we differentiate such large areas as continents according to IPCC regions, certain special features become apparent: In South Asia, Latin America and the Caribbean, and Sub-Saharan Africa the 'lightweight low rise' LCZ (LCZ-7) which is conceptually very close to slums or informal settlement structures reveal a notable share (26 %, 6 % and 5 %, respectively) of newly built areas in 1985. However, the added share of this particular LCZ is decreasing over time. It is also interesting to see that compact urbanization (LCZs 1, 2, 3) played a significant role in Latin America and the Caribbean (about 44 % in 1985), South Asia (ca. 36 %), South-East Asia and Pacific (ca. 31 %), East Asia (ca. 26 %), Pacific OECD (ca. 25 %), the Middle East and North Africa (ca. 25 %) and to a certain degree also South Saharan Africa (ca. 21 %) and in North America (ca. 17 %). In comparison, in Western Europe compact structural developments had only a share of about 11 %. In general, however, dense urban developments have decreased across all IPCC regions. In Western Europe, these developments decreased over the monitoring period from about 11 % to about 1 %, i.e. there is almost no compact development in newly urbanized areas anymore.

5.2. Structurally less dense cities in 2015 than in 1985

The fact that the expansion areas have been built increasingly with low density LCZs (section 5.1.) has also changed the overall structure of cities compared to 1985. The cities in 2015 are on average structurally less dense than in 1985. The proportion of dense, compact LCZs (LCZs 1, 2, 3, 7) has decreased globally from 32 % to 26.1 %, while the proportion of low-dense, open LCZs (LCZs 4, 5, 6, 9) has increased from 37.7 % to 40.5 % (Fig. 5).

We generally see this trend towards fewer area shares of dense, compact LCZs on every continent, but the trend is particularly pronounced in Asia (from 38.6 % to 26.7 %), Africa (from 37.2 % to 31.1 %) and South America (from 56.7 % to 49.5 %). Conversely, the strong increase in the low-dense, open LCZ types is particularly noticeable in Africa where it increased from 44.4 % to 53.3 %, in Asia, where it increased from 26.6 % to 32.1 % and in South America where it changed from 21.9 % to 30 %. 'Sparsely built up' also increased in North America (from 42.4 % to 45.7 %).

This trend of a relative decrease in dense, compact LCZs can also be constantly observed in the IPCC regions. The trend is strongest in East Asia (from 30.8 % to 18.1 %) and South-

East Asia and Pacific (from 48.3 % to 37.8 %) and Latin America and Caribbean from 53.7 % to 46.5 %.

The proportion of commercial LCZs (LCZ 8, 10) remained comparatively stable in most regions, with the notable exceptions of East Asia where it increased from 38.5 % to 48.9 % and Sub Saharan Africa where it decreased from 18.2 % to 14.7 %.

Figure 5 shows an overview of these multi-temporal structural changes structured by the entire world, continents and IPCC regions.

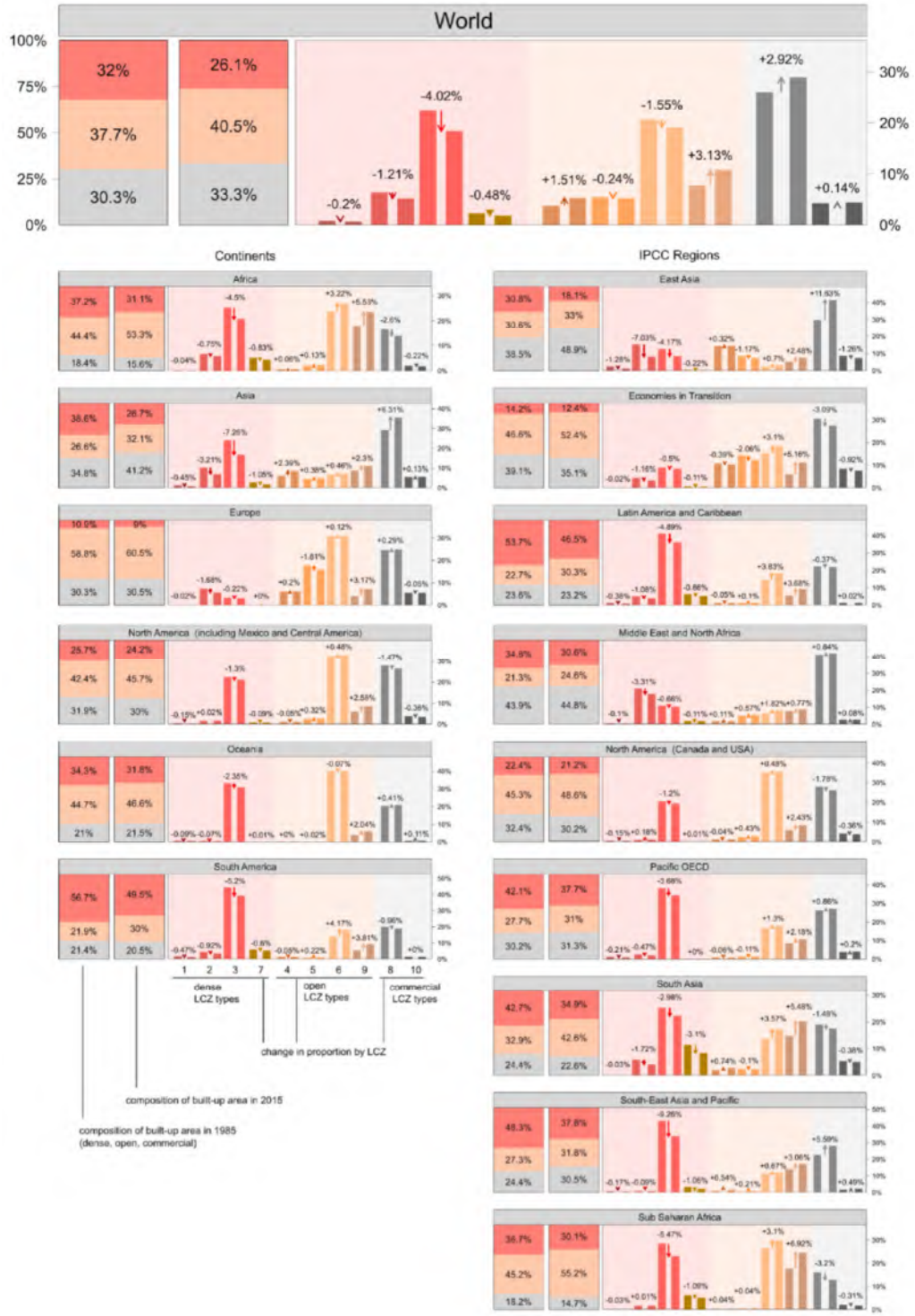


Fig. 5. The shares of built LCZs and their comparison 1985 to 2015 aggregated to the world, continents, and IPCC regions.

5.3. Demographic densification in cities despite low densities in the expansion areas

The analyses have shown that urban expansion has led to lower structural densities in the 1567 largest cities worldwide over the 30 years of monitoring from 1985 to 2015 (cf. section 5.1. and section 5.2.). However, the question arises as to whether this is also accompanied by lower population densities and, in particular, how population densities have developed within certain LCZs.

We find that population densities have slightly increased at the level of the entire cities with 5681 inh./km^2 in 1985 to 5948 inh./km^2 in 2015. However, there are major differences in population densities between the urban area that already existed in 1985 and the expansion areas that were built between 1985 and 2015. The areas that were already built-up in 1985 have increased in population from 970.2 million to 1281 billion and in population density from 5681 to 7500 inh/km^2. The newly built urban areas from 1985 to 2015 add 720.8 million people to the urban population. However, the population density in these expansion areas is with 4349 inh./km^2 at a substantially lower level. Overall, this shows that the urban expansion areas have a lower building density and correspondingly a lower population density than the areas that existed in 1985 already. Nevertheless, it must be stated that for the entire cities the densification of the population in the older urban areas has more than compensated for the low density in the expansion areas.

Distinguishing these developments in population density by LCZ, we measure that more dense, compact built-up LCZs (LCZ 1, 2, 3 and LCZ 7 'lightweight low-rise', the latter being a proxy for slums and informal settlements) have, as assumed, the highest population densities worldwide (Fig. 6).

This relation between the built densities and population densities are also mapped at continental level. However, there are significant differences in population densities between the continents for the same LCZ. Population densities in the 'compact mid-rise' LCZ (LCZ-2) are around 15 000 inh./km^2 in Europe and Asia. In North America or Oceania, however, they are only around 5000 inh/km^2. At the level of the IPCC regions, the population densities in South Asia stand out with around 28 000 inh./km^2 at the 'dense, compact' LCZs. In contrast, North America stands out here with the lowest population densities of around 3000 inh./km^2.

It is particularly interesting that over the course of time, population densities of the respective LCZ class are consistently lower in the areas that were added between 1985 and 2015. For example, in 2015, the LCZ class 'compact mid-rise' (LCZ-2) shows a population density of approx. 20 100 inh/km^2 in the areas that were already built on in 1985. For the same class, we measure a density of only 13 700 inh/km^2 in areas built since 1985. And this trend applies to almost all LCZs (Fig. 6).

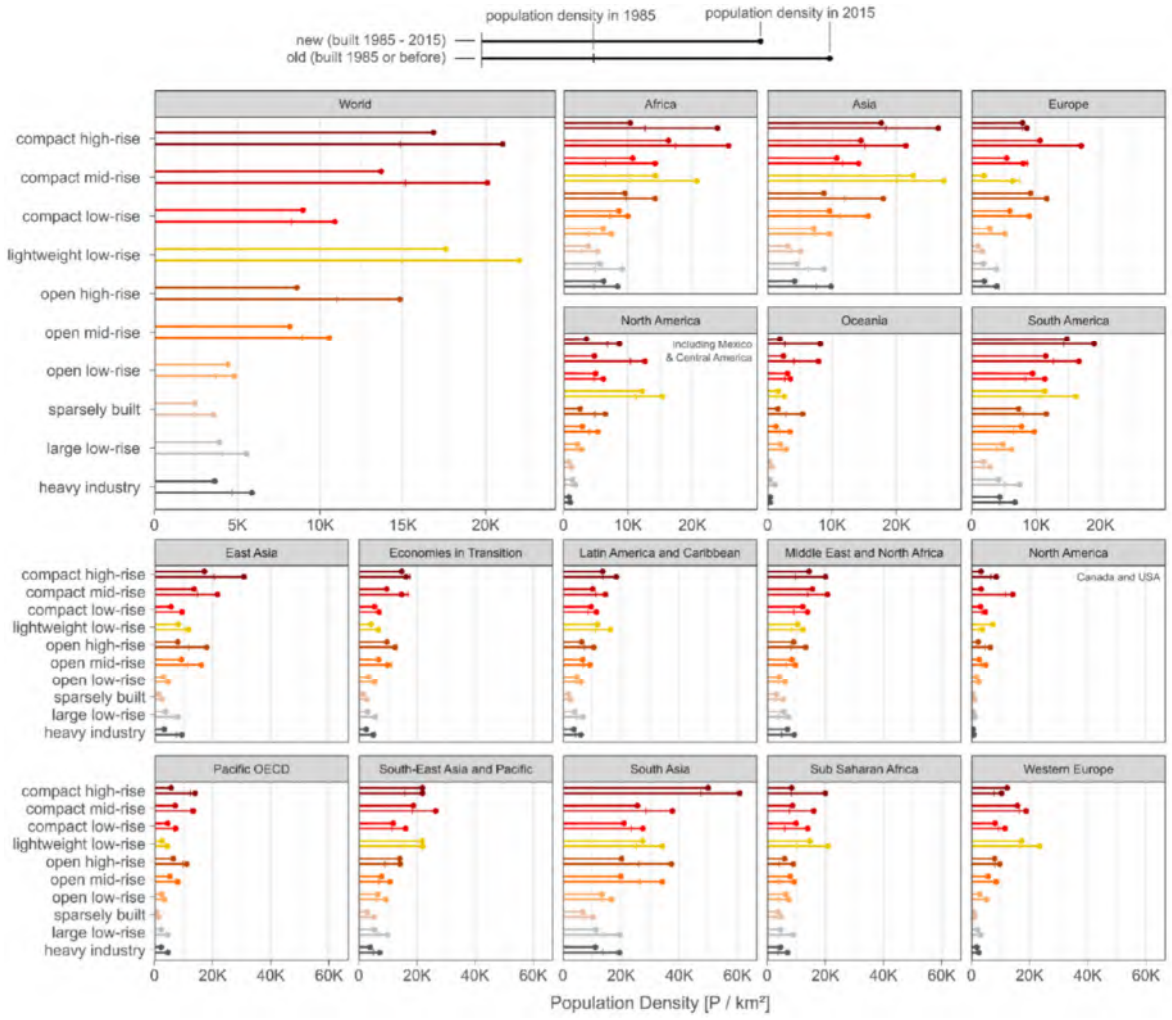


Fig. 6. Population densities for LCZs at global, continental and IPCC region level and over time, 1985 and 2015.

5.4. The great untapped potential: scenarios of land consumption in expansion areas considering theoretical population carrying capacities

Urbanization is seen as the path to less land consumption. Nevertheless, we have seen that very different structures have been realized within the built landscape. And these different structures have different carrying capacities for population. In particular, we have seen that low dense LCZs have dominated the expansion process of cities from 1985 to 2015 (cf. section 5.1.). The 1567 cities larger than 300 000 inhabitants in 2015 carry now an additional urban population within their MUAs of 844.87 million people. Whether this is efficient or not is not easy to answer.

With model calculations, we aim to provide an approximation of this question. The model calculations show that a significantly higher carrying capacity in the urban expansion areas would theoretically be possible, and thus there is a great untapped potential (Fig. 7): With *single structural type expansion* scenarios, we model if the urban expansion areas were not dominated by low-density LCZs (LCZs 4, 5, 6, 9), but were built exclusively with higher densities. For example, ‘compact mid-rise’ LCZs (LCZ-2), having on median a population density of 13 193 inh/km^2 across the globe, would allow 2.187 billion people to live there, i.e. an additional 1.466 billion people.

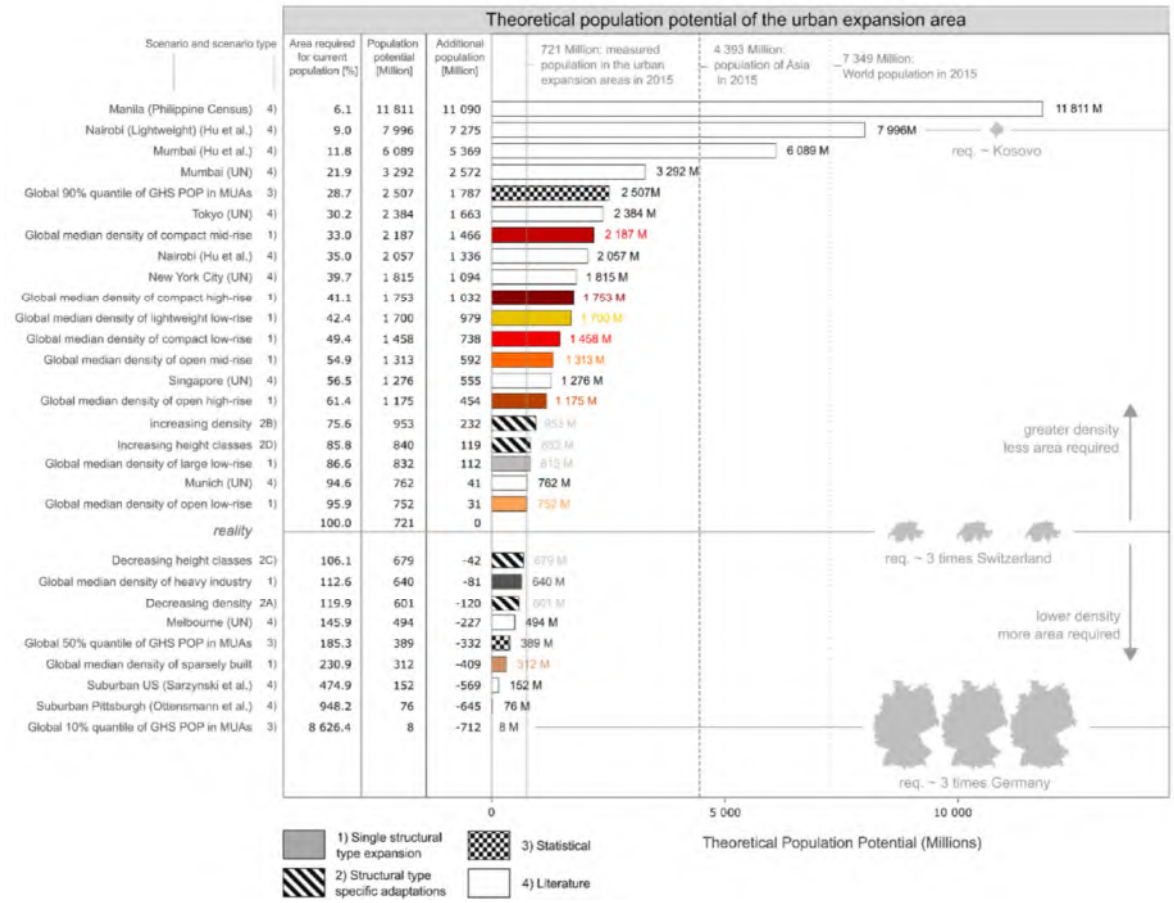


Fig. 7. Scenarios of theoretical carrying capacities of population for the expansion areas between 1985 and 2015. The scenarios refer to single structural type expansion (1), structural type specific adaptation (2 – A, B, C, D), to statistical models (3) and to literature-based models (4) (cf. section 4.3.).

Of course, building the entire urban areas in a homogeneous structural style – as simplistically assumed in the model above – is neither practical nor aesthetic. Therefore, *structural type specific adaptation* are model calculations that come closer to reality, i.e. if each LCZ in the expansion areas were increased by one structural density level. With this single stage increase in density, the same newly built urban area in the observed MUAs can provide a home for an additional 232 million people (cf. Scenario A in Fig. 3). This is around the magnitude of a country the size of Pakistan or Nigeria. Alternatively, 24.6 % of the space could have been saved and used for other purposes or as natural environment. Building higher instead of denser could have had a similar effect, providing housing for 119 million people more (cf. Scenario B in Fig. 3).

If we go one step further and detach the analysis from a structural type and take as *statistical model* the top 10 % of population density within all MUAs, then we could reach 15 128 inh/km^2. This means that 2.507 billion people could live in these expansion areas – and with it far more than double the population could actually be added.

And finally, model calculations using extremes based on *literature* reveal untapped potential. In Nairobi the population density of the 'lightweight low-rise' LCZ (LCZ-7), which is more or less equivalent to slums, is estimated by Hu et al. (2021) to be 48 241 inh/km^2. Extrapolated, this means that 7.996 billion people could live in these expansion areas if they

were similarly densely populated. In other words: In addition to the 721 million people who actually live there, these areas could theoretically accommodate almost the entire current population of the world. Inversely, the urban expansion that took place between 1985 and 2015 that, as mentioned, in reality spans three times the area of Switzerland, could have been reduced to the consumption of only 10 842 km^2, i.e. about the size of only Hawaii or Kosovo, and still absorb the same population growth as it did in reality (Fig. 7).

Of course, it would also be possible to build less densely on urban expansion than was the case in reality. If urban expansion had taken place exclusively with 'sparsely built' LCZs (LCZ-9) in a single structural type expansion scenario, with a median population density of 1883 inh./km^2, then only 312 million people could live there, i.e. 409 million less than actually do. A *structural type specific adaptation* with a decrease in height (cf. Scenario D in Fig. 3), the carrying capacity would be lower by 42 million and with a decrease in density (cf. Scenario C in Fig. 3), the carrying capacity would be lower by 120 million, or required 19.9 % additional area. If it had been built and populated like the *statistical* bottom 10 % quantile across the globe, only 8 million people could live there. It would take, 1 142 282 km^2, i.e. about three times the extent of Germany, rather than three times the extent of Switzerland, to accommodate the measured population growth at this density (Fig. 7). To show one more *literature-based* extreme, if the newly constructed areas had been built and populated like the suburbs of Pittsburgh, with a density of 459 p/km^2 (Ottensmann, 2020), the area could only have housed 76 million people.

This gigantic range of scenarios shows how different built densities and population densities in cities are distributed across the globe and what big effects they have with respect to land consumption. Figure 7 systematizes all these model calculations in relation to reality.

6. Discussion

In the monitoring period from 1985 until 2015, 32.9 % of the global population growth has taken place in the 1567 largest cities on our globe. However, urban land growth in the same period corresponded to only 19.5 % in relation to the global settlement area. This basically shows that urbanization in larger cities is generally space-saving. And although this could be seen as efficient, our analysis shows that the urban development was not ideal in terms of land consumption.

6.1. Urban expansion in structural and demographic low density

The main findings of this study are that in the 1567 largest cities across the globe with more than 300 000 inhabitants, urban expansion has been structurally less dense between 1985 and 2015 than the cities were on average in 1985. Thus, these cities widely have less dense structures in their entirety in 2015 than in 1985. In addition, the population densities in the newly built-up structures between 1985 and 2015 are substantially lower than in the older core areas. Of course, it must be also stated that population densities in these cities increased in general slightly. This testifies to an increase in the intensity of use in the urban core areas, which already existed in 1985, and these areas even compensate for the low population densities in the expansion areas.

It can therefore be said that structural urban expansion in terms of land consumption was unable to maintain the status quo of 1985 and was therefore not sustainable. The idea and call for more compact cities (e.g. by Glaeser, 2010; OECD, 2018; Angel et al., 2021) has thus not been realized, and, although population densities slightly increased overall in these cities, the expansion areas feature substantially lower population densities. It is therefore reasonable to conclude that new built structures need time to achieve population densities similar to those in central areas, even if the morphologic structure is comparable.

Our results are essentially in line with the results found by Angel et al. (2021) for a sample of 200 cities across the globe, by Siedentop and Fina (2012) for European cities and by Schmitt, Siedentop & Fina (2015) for post-socialist cities. Our study extends their findings by the number of cities, following the call of Bell and Jayne (2009) to redirect the distorting focus of research work on the largest cities on the globe to a more balanced analysis of smaller and larger cities. Furthermore, our study extends their findings with the intra-urban structural component as well as the consistent way of delineating cities. One consequence of this highly dynamic global urbanization (cf. Taubenböck et al., 2024) is that humankind is losing large amounts of fertile cropland and forest (van Vliet, 2019). So, if we do not succeed in developing new urban land more effective in terms of land consumption – which was until 2015 not the case – then we lose twice by losing productive agricultural land or forests without adequate compensation by providing housing to a larger amount of people.

Furthermore, we found that population densities in certain LCZs vary greatly around the world. This differs for LCZ type ‘compact mid-rise’ (LCZ-2), for example, between 28 000 inh./km^2 in South Asia to around 3000 inh./km^2 in North America. This shows, of course, that although the global statements in the first paragraph above reflect the basic trends, it is always necessary to differentiate between specific regions. And it shows that structural densification is an essential feature for achieving high population densities, but urban planning policies, cultural context, as well as other reasons can also lead to different densities in these structures.

With this study, we also follow the call of Hu et al. (2021), stating that intra-urban differentiation is necessary in multi-temporal city analyses. Our study here confirms what (Li, Verburg, & van Vliet, 2022) show in their study that population densities especially in Europe and North America and in parts of China are very low, and are therefore in global comparison not sustainable. Various effects play a role here, such as the growing demand for floor space area, especially in wealthier societies (Angel et al., 2021) translated into planning policies (Mangin, 2014; Monkkonen, 2019), the construction boom in China, which leads to so-called ghost towns or areas with a high underload (Shi et al., 2020), among other issues. The systematic analyses of local characteristics of cities against the here presented global trends opens up crucial future research directions.

Overall, however, it must be stated here that humankind is leaving a lot of potential on a global level in terms of structural and demographic urbanization. From the perspective of land consumption, urbanization must be developed more efficiently. And, of course, it is not just land consumption, but many other environmentally relevant variables which potentiate the negative consequences of low dense urban expansion: higher average gasoline usage,

higher electricity usage or higher carbon emissions (Glaeser & Kahn, 2010; Iungman et al., 2024), to name just a few. This study aims to raise awareness of this potential. With this knowledge, we can perhaps answer the question more consciously as to whether we want to and can actively transform our ecologically damaged world, or whether we are being transformed by multidimensional developments.

6.2. The theoretical carrying capacity of urban expansion

From the main findings discussed above, one can conclude, we have a lot of potential to make urbanization more efficient in terms of land consumption. Our scenario calculations show the potential for saving space: If we would have built more densely, then we would have had great potential to accommodate many more people in the same urban expansion areas.

Assuming urban expansion with 'compact mid-rise' structures (LCZ- 2) would allow 2.187 billion people to live there, i.e. an additional 1.466 billion people. In 2015 – the end point of our analysis – 3.86 billion people lived in urban areas worldwide. Using the above scenario, the world's largest 1567 cities alone could have absorbed the entire global urban population growth (to 4.7 billion United NationsDepartment of Economic and Social AffairsPopulation Division, 2018) by 2024 if they had been built more densely. This is not an entirely unrealistic scenario, nor would the population densities be so high that the quality of life would decline by definition. Many studies show that such structures can be developed with a high quality of life. Medium to high-density living has been shown to be acceptable to residents as long as these developments also provide at the same time an increase in quality green spaces, economically vibrant and mixed-use environments (e.g. Glaeser, 2010; Hollis, 2013; Sennett, 2018, p. 368).

Taking it a step further, our more extreme scenarios show the hypothetical potential of these expansion areas: With population densities such as in LCZ-7 proxying slums, almost the entire global population (7.99 billion) could live in these expansion areas. UN forecasts (United Nations Department of Economic and Social Affairs Population Division, 2018) up to 2050, however, only predict an increase of just under 3 billion new urban dwellers between 2015 and 2050. This means that the world's largest cities would not even have to be further built in such an admittedly extreme dense LCZ-7 way as to accommodate all future urban dwellers by 2050. However, this specific analysis in turn reveals the unimaginable population densities in slums (cf. Breuer et al., 2024) which goes along with terrible living conditions. Of course, one can argue that some densities we used for scenario calculations like those in Nairobi's slums are a high price to pay for saving space and that one would sacrifice one goal for another. But that's not the point in these theoretical experiments. These extremes are just to show how inefficient urbanization from the perspective of land consumption has actually been over the 30 years from 1985 to 2015. If we only look at the fact that built structures such as 'compact low-rise' (LCZ-3) or 'compact mid-rise' (LCZ-2) are estimated with the capability to accommodate an additional 738 million or 1466 million people – the latter number doubles the entire population of Europe in 2022 – then we are not talking about structures such as slums, which per se correspond to a low and over-dense type

of low quality. Rather, these gigantic figures show that newly built urban areas have great potential in terms of space and therefore resource savings.

At the same time and in contrast, some densities, such as in American suburbs, show that this particular low-dense type of built-up structure is so land-consuming that if it would be the sole and prime structure type, only 152 million people could live in these new urban expansion areas. This low figure shows that certain forms of housing reflect an intrinsic excessiveness of modern societies embedded in the legal planning system (Mangin, 2014; Monkkonen, 2019).

Nevertheless, we do not want to sound naïve here and present the calculated potential that the same urbanized area would have as a realistic cipher. To be clear: not everything could be developed as efficiently in terms of land consumption as modelled here, nor would it be desirable. We need to be aware of how stable the built environment as well as practices, routines, building codes, planning ciphers or cultural expectations of building and living are and how difficult it is to change routines and attitudes (cf. Nassehi, 2024) even by implementing it into a legal apparatus (Tonkiss, 2013). It has to be said that while the theoretical experiments here show the potential of possible urban redevelopment, it is not possible to react collectively on a global scale. However, if this analysis demonstrates anything, it is that structural urban transformation is physically possible with denser structures that have already proven to be livable, suggesting that the demands of such transformation are not necessarily excessive.

To avoid a naïve perspective, it is important to acknowledge both the potential and the challenges of urban densification. Building denser cities offers a promising solution for accommodating the growing population (United NationsDepartment of Economic and Social AffairsPopulation Division, 2018) and increasing urban migration (Glaeser, 2010) while consuming less land. Defining these goals may seem straightforward, but the methods to achieve them—and their unintended consequences—add complexity. A sustainable strategy must consider multiple factors to prevent conflicting objectives from exacerbating issues such as the urban heat island effect, traffic congestion, and air pollution. Instead of relying on one-dimensional targets, urban density thresholds should be empirically identified to mitigate potential negative impacts.

6.3. Conceptual issues

From a conceptual perspective, there are also some issues that need to be considered: Unfortunately, no multi-temporal intra-urban structural classifications of the built landscape are available since 1985 at a global scale. Although some studies aim at intra-urban multitemporal analysis (e.g. Qi et al., 2024; Wang et al., 2019), they still focus only on a limited number of cities or single countries. The resolution of the satellite data in 1985 make intra-urban structural classification at the level e.g. of the here used LCZs challenging and it needs to remain open whether global multi-temporal data with sufficient accuracy for consistent multi-temporal analysis dating back to the 1980s will become available. Since built structures are comparatively long-lived and fundamental restructurings are generally rather rare or take place very slowly in consolidated urban tissues, we have assumed that the

LCZs of today also correspond to the LCZs of 1985. We also assume that in the newly built urban areas, the LCZ type that we measure today has not changed over the course of the 30 years. It is therefore clear that this data-driven, conceptual simplification represents a certain distortion of the results. Of course, there are changes in the type of structure, which means we are neglecting an important component of urban development. However, we are unfortunately unable to quantify these intra-urban changes over time due to the current data situation. As these structural changes are, however, small in comparison to the dimensions of the data set, we assume that it does not have a significant statistical impact on the analyses.

Nevertheless, it is noticeable that the population densities in the same LCZs that already existed in 1985 are consistently higher than in the expansion areas. As this is systematically the case, we consider this trend to be plausible and assume that newly built areas need a certain amount of time to adjust to the corresponding densities. However, this can also be seen as an indication that structural re-densification has probably also taken place, which we cannot quantify with the current LCZ classifications. The quantity of these structural changes and related processes of densification and infilling or reduction of densities are due to availabilities of high-frequency and high-resolution data now possible and suggested for future studies.

6.4. Data and technical issues

From the *perspective of data*, there are some uncertainties that need to be discussed here: Every data set used in this study – whether the multitemporal classification of urban expansion, the LCZs, the morphological urban areas or the population data – contains inaccuracies or even errors.

We can quantify the accuracies of the individual classifications, even if these are distributed very differently across the world. The WSF 2015 data set reports accuracies of 83–89 % (Kappa 0.55–0.78) (Marconcini et al., 2020), the LCZ classification have an overall accuracy of 87.3 % and a kappa value of 0.65, but with varying accuracies across the various LCZs (Zhu et al., 2022), the morphological urban areas are one consistent way of delineating cities allowing for true geographic comparisons, but they do not represent a universal urban-rural border (Taubenböck et al., 2019), and the GHSL population data set depends on the timeliness and quality of the census data and the accuracy of the building volume which was reported with an RSME of 2.40 m and 3.25 m, i.e. about one floor in general (Pesaresi, Corbane, Ren, & Edward, 2021). In addition, the input data has variable resolutions and projections and the harmonic integration leads to some intrinsic uncertainties (Johnson & Clarke, 2021). These, however, are minor and not quantified here since each data set has already been independently validated. Overall, we have not quantified which amplifying or attenuating combinations of inaccuracies lie in their combination. We suggest this issue to a separate study. It also makes sense to test the reliability of the results obtained from the data sets used here. This means that other current geodata are to be tested for each input data set – for example the GHSL (Pesaresi et al., 2015) instead of the WSF (Marconcini et al., 2018), the LCZ classification from Demuzere et al. (2022) instead of the one from Zhu et al. (2022), the WorldPop data (2025) instead of the GHSLPop (Schiavina et al., 2023) or the degree of urbanization (Dijkstra et al., 2021) instead of the MUAs (Taubenböck et al., 2019) for the

city delineations, among other possible alternatives to our of current data. We suggest this also for a separate study.

In general, however, we can state that our input data have proven high accuracies and the achieved results are plausible. Denser and/or higher built-up LCZs indicate higher population densities and vice versa. The trends of the main findings, although variable across the globe, always point to the same direction which implies consistency and from a logical point of view also plausibility. For some data, we are particularly aware of inaccuracies. For the 'lightweight low-rise' LCZ (LCZ-7), as one example, that proxies slums, we are aware of the accuracy problems with population estimates, where underestimation is consistently documented (Abascal et al., 2024; Breuer et al., 2024; Taubenböck & Wurm, 2015). This makes it clear that although high population densities were measured in the LCZ class 'lightweight low-rise', these are certainly much higher in reality as informal populations are significantly underestimated in the census.

7. Conclusion

This study shows for the first time that urban expansion has made the world's largest cities between 1985 and 2015 less dense in terms of built-up structures. And, from a demographic perspective, the expansion areas feature significantly lower population densities than the older core areas. It can therefore be said that urban expansion was in terms of land consumption unable to maintain the status quo from 1985 and was therefore not particularly sustainable. This shows that urbanization can and perhaps must be made more efficient.

One can therefore argue in two directions: Even if the population densities in the urban expansion areas are low, one could argue that cities have become slightly denser overall and that this fulfils the call for more compact cities. Or, one could also argue much more ambitious that urban expansion has not utilized large parts of the potential for saving space which is generally attributed to it. We do have an incredibly high potential to more or less absorb the expected further urbanization trend with the existing areas through re-densification.

Every scientist, including us, must of course start from exactly the illusion that this new knowledge is a medium for ambitious urban transformation. At the same time, we must realize that we may not even have the control of urban-structural transformation in our hands, as the typically naïve sentence above – that this new knowledge is a medium for ambitious urban transformation – and that we all basically read in any conclusion to this topic should suggest. That means: we are aware that one can only develop what is controllable. This new knowledge on urbanization leads to insight, but not yet to targeted, realistic and realizable actions. Options for adaptations in the urban domain are not trivial. Too often, we natural or data scientists ignore the fact that societal conditions probably play the central role in any adaptation. With this in mind, we conclude this article with the call for inter- and transdisciplinary exchange to pass on this new knowledge to urban planners, politicians, and all other stakeholders taking knowledge into action – to work with ambition towards the goal of denser cities of high quality of life.

Part VIII

Ethical considerations

CHAPTER XXV Earth Observation and Artificial Intelligence: Understanding emerging ethical issues and opportunities 569

Chapter XXV

Earth Observation and Artificial Intelligence

Understanding emerging ethical issues and opportunities

IEEE Geoscience and Remote Sensing Magazine, 2022, vol. 10, issue 4

Mrinalini Kochupillai, Matthias Kahl, Michael Schmitt, Hannes Taubenböck & Xiao Xiang Zhu

© 2022 IEEE

Note: A payment was made to CCC RightsLink to reuse the paper in this report.

https://doi.org/10.1109/MGRS.2022.3208357

Published 17 Nov 2022

Abstract

Ethics is a central and growing concern in all applications utilizing artificial intelligence (AI). Earth observation (EO) and remote sensing (RS) research relies heavily on both big data and AI or machine learning (ML). While this reliance is not new, with increasing image resolutions and the growing number of EO/RS use cases that have a direct impact on governance, policy, and the lives of people, ethical issues are taking center stage. In this article, we provide scientists engaged with AI for EO (AI4EO) research, 1) a practically useful overview of the key ethical issues emerging in this field, with concrete examples from within EO/RS to explain these issues, and 2) a first road map (flowchart) that scientists can use to identify ethical issues in their ongoing research. With this, we aim to sensitize scientists to these issues and create a bridge to facilitate constructive and regular communication among scientists engaged in AI4EO research, on the one hand, and ethics research, on the other hand. The article also provides detailed illustrations from four AI4EO research fields to explain how scientists can redesign research questions to more effectively grab ethical opportunities to address real-world problems that are otherwise akin to ethical dilemmas with no win-win solution in sight. The article concludes by providing recommendations to institutions that want to support ethically mindful AI4EO research and provides suggestions for future research in this field.

Keywords

Ethics; artificial intelligence; guidelines; Earth observing system; big data; machine learning; remote sensing

1. Introduction

Big data lies at the heart of many EO and RS research and development activities. In fact, applications and algorithms emerging from RS and EO research rely heavily on big data collected via satellites, unmanned aerial vehicles, drones, and other state-of-the-art devices. At the same time, RS and EO research is rapidly transforming in the era of AI and ML. AI and ML permit the petabytes of data collected by satellites and other EO/RS devices to be systematically organized and used to train models that can predict a large diversity of events, objects, and circumstances, even in the absence or lack of so-called ground truth. The close relationship between (big) data and AI is reflected in the German Data Ethics Commission's understanding of "artificial intelligence" as "a collective term for technologies and their applications which process potentially very large and heterogeneous data sets using complex methods modeled on human intelligence to arrive at a result which may be used in automated applications. The most important building blocks of AI as part of computer science are subsymbolic pattern recognition, ML, computerized knowledge representation and knowledge processing, which encompasses heuristic search, inference and planning" (BMI, 2018).

In this article, we refer to the expanding field of research that uses AI/ML tools for organizing, understanding, and analyzing EO/RS data as "AI4EO research." At present, AI4EO research can be roughly divided into two categories. On the one hand, 1) algorithmic capabilities are explored to determine whether (geo)information about the land surface and the atmosphere can be derived automatically from the available multimodal EO data with sufficient accuracy. On the other hand, 2) fields of application are explored to identify issues linked to demographics, economics, ecological, social, and other fields from available datasets. Further, where necessary and possible, these latter explorations are increasingly done in combination with other newer sources of data, such as data derived from questionnaires (Marti et al., 2020), censuses (Aubrecht et al., 2009), and in social media platforms (Cervone et al., 2017), among many others.

Ethics is a central and growing concern in the fields of (big) data, AI and ML generally, and each of the preceding two broad categories of AI4EO research. Ethical issues, well beyond the now well-known issues of privacy, explainability, and bias, are increasingly relevant for EO and RS scientists for several reasons: 1) the scale and resolution of EO/RS images are improving rapidly; 2) new use cases are emerging, which go beyond the traditional role of providing information on topics such as climate and environmental changes and have a direct impact on specific human populations and groups as well as governmental policies linked to such groups (Gurung Goodrich, Gurung & Hamal, 2021); and 3) AI/ML tools are increasingly used to manage and analyze the diverse sources of EO/RS data, and several ethical issues are commonly known to arise in several applications that utilize such tools. Beyond these EO/RS-specific reasons, ethical issues can and do arise in the context of AI research as a result of 1) the very fact of data collection, 2) the approaches used for data analysis, and 3) the manner and purpose for which the results from such analysis are used.

While a great deal of work has been done in the European Union (EU) and beyond to compile various guidelines and guiding motifs for data ethics and AI ethics, recent research and

empirical studies suggest that these guidelines may not be practically useful for scientists and academic researchers (Hagendorff, 2020). Various ethics guidelines also note that ethical issues can and should be context and technology specific (European Commission, 2018). In this article, we provide scientists engaged with AI4EO research a practically useful overview of the key ethical issues and opportunities emerging in this field. More specifically, our aim is not only to sensitize scientists to these issues but also to help create a bridge to facilitate constructive and regular communication among scientists engaged in AI4EO research, on the one hand, and ethics research, on the other hand.

We organize these issues and guidelines under six categories, which correspond with the three fundamental values of (business) ethics (namely, honesty, integrity, and fairness) and three centrally relevant emerging issues (namely, privacy, responsibility, and sustainability). Note: A distinction can be made between ethics and business ethics, with ethics being a broader and more general term linked to all/any moral values, including personal values and decisions we make based on our personal conscience. Business ethics, on the other hand, relates to the values and considerations that businesses and institutions, including teams and individuals working in an institutional setup, must bear in mind in the performance of their individual and collective tasks and activities.

The ethical issues under each of these six categories are then explained with the help of illustrations from current AI4EO research. In so doing, we aim to identify ethical issues that are specifically relevant for current and ongoing AI4EO research, given its unique nature and scale. We also hope that with this first-of-its-kind endeavor, AI4EO researchers will be better equipped to identify the ethical issues they need to be most mindful of at every stage of their research. In the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section, using four prominent fields of AI4EO research as examples, namely, 1) artisanal and small-scale mining (ASM), 2) slum mapping, 3) the conservation of biodiversity, and 4) building-level population density estimations, we explain how AI4EO scientists can go beyond merely "avoiding" ethical issues toward actively identifying and embracing ethical opportunities to help resolve some of the most important challenges faced by humankind today. We conclude with some observations and recommendations for various AI4EO stakeholders and for possible future directions in AI4EO research.

2. Understanding Ethics: Issues, Risks, Theories, and Opportunities

2.1. Ethics

The term *ethics* has been defined from numerous perspectives (e.g., philosophical, anthropological, historical, economic, political, and social) in existing academic and scientific literature (Singer, 2022). At the most fundamental level, ethics deals with issues of right and wrong, good and bad. As a discipline, ethics is partially preoccupied with identifying and finding ways of resolving values that come into conflict in various contexts of human life, including scientific and academic research (Dill, 2009).

Ethics is also about emphasizing constructive and "appropriate conduct associated with" basic human "values," such as honesty, responsibility, and so on, and not merely about

"discover(ing) inappropriate conduct" (Ferrell & Fraedrich, 2015; Somers, 2001; Giessner & Van Quaquebeke, 2010). Accordingly, some understand ethics to be about identifying a broad set of desirable duties and (un)desirable consequences (Lussier, Hartmann & Bolander, 2019). Together, these are expected to provide a framework to understand and avoid negative and harmful human actions, interactions, and consequences, on the one hand, and, perhaps more importantly, explore, guide, and/or facilitate positive and constructive human actions and interactions, on the other hand. Within this broad understanding of ethics that is considered by some scholars to be context or even culture specific (so-called "ethical relativism" (Standford Dictionary, 2015), several scholars are of the view that there exists an objective universally applicable set of human values (Solomon, 1998; Shankar, 2007) and responsibilities (United Nations, 1999) that are fundamental to what makes us "human." These values are cherished independently of culture and context, and accordingly, all human actions (and inaction) must be based, or build on, these fundamental values. For example, the most fundamental principle of ethics is the so-called golden rule ("don't do to others what you wouldn't want others to do to you"; Wattles, 1996) and three associated values, namely, honesty, integrity, and fairness (Ferrell & Fraedrich, 2015; Den Hartog & De Hoogh, 1982). Together with other basic human values, such as responsibility and respect for human rights, these principles lie at the heart of all (business) ethics discourses and theories. (Ferrell & Fraedrich, 2015).

2.2. Ethical Risks, Ethical Issues, and Ethical Dilemmas

2.2.1. Ethical Risks

Ethical risks are, simply put, the most likely and predictable negative consequences of any action or inaction. Accordingly, actions and inactions that are highly likely to yield negative and undesirable consequences are labeled as unethical. (Note: Ethical risks must be distinguished from the emerging field of "risk ethics" that debates whether and in what circumstances it may be ethically justifiable to take a risk, e.g., a calculated risk in business. For an overview of risk ethics, see Hansson, 2021; Altham, 1983.)

2.2.2. Ethical Issues

Unlike the term *ethics*, the term *ethical issue* has not been widely defined in academic literature. Nonetheless, this term, which is commonly used in literature dealing with ethics, is arguably self-explanatory. For a fundamental understanding, it is useful to look at the basic definitions of the terms *ethical* and *issue*. According to the Collins English Dictionary, "ethical means relating to questions of right and wrong," and "an issue is an important subject that people are arguing about or discussing." Put together, ethical issues can be understood as issues relating to questions of right and wrong (or good and bad) that are being debated and discussed by people, including scientific and academic researchers.

The previously mentioned idea of "ethical relativism" comes starkly into play when we attempt to identify and define ethical issues. By definition, issues, whether they are linked to ethics, law, economics, natural sciences, society, or politics, are subject to debate and diverse views. This is not least because priorities (and, therefore, considerations) that go into

determining what is right, wrong, good, and bad, can vary from person to person, from context to context, from culture to culture, and depending on the perspective and desired end result from which one views any situation.

Indeed, even within established fields of academic research, ethical issues are not uncommon. For example, in the sphere of medical research, while the collection and analysis of data can help gain a better understanding of a disease and make more accurate diagnoses, the process can also raise ethical issues linked to patient privacy: the ethical issue here is, should we give primary importance to data protection to secure patient privacy (the "ethical concern/risk") or to understanding and curing diseases more efficiently (the "ethical opportunity")? Here, one can see that "ethical opportunities" can be the flip side of "ethical concerns/risks," which together can lead to an "ethical issue."

Ethical issues can also arise at a more fundamental level. For example, while stem cell research has several known and potential (medical) benefits, it is considered by many to be fundamentally unethical Einsiedel & Adamson, 2012; Streiffer, 2007; Devolder & Harris, 2007, as it violates cultural and religious sentiments and, according to a few thinkers, the sanctity of life. It is therefore fair to say that ethical issues arise from a very fundamental and even legitimate space of conflicting values. These conflicting values can take the shape of conflicting understanding of desirable and undesirable goals, approaches, and end results. They can also arise in relation to the benefit that may result (for an individual or for society as a whole) from the pursuit of these goals and approaches, weighed against their impact and consequences on the lives of those who were not involved in the decision to pursue these goals. It is also possible that the same individual or community can, at the same time, stand to gain from an ethical opportunity while losing from the corresponding ethical concern/risk.

When we are faced with an ethical issue that cannot be resolved (e.g., by finding a win-win or middle path) by the application of basic principles of ethics and with the help of recognized theories of ethics, we may be facing what is termed an ethical dilemma. In an *ethical dilemma*, the pursuit and accomplishment of one very important goal/outcome inevitably leads to a substantial or irreversible sacrifice of another, equally important, goal/outcome. The trolley dilemma is a typical ethical dilemma that is hotly debated in the context of decision making by autonomous cars (Kochupillai, Lütge & Poszler, 2020). In the context of illegal ASM that is often surveyed using AI and EO data, there is another ethical dilemma that governments on the ground find hard to tackle: permit such mining to secure the right to a livelihood of poor and indigenous communities or strictly ban such mining to protect the environment and human health (Hentschel, Hruschka & Priester, 2022; Alliance for Responsible Mining, 2013).

2.3. Ethical Opportunities

The term *ethical opportunities* also has not been defined in existing literature. Nevertheless, the term has been used in various contexts. For example, in the context of increasing consumer awareness and demand for labels that are "eco-friendly" and "fair trade," "ethical business" and ethical business models have been described as an "opportunity" to take a

greater share of the market or to carve out a niche market for a business' product or service (Hall, 2007).

Within medical and bioethics, it has been said that "ethical insight into projects may afford opportunities for initiatives that add to the projects' value and the satisfaction of those involved …" and that "ethics is better applied collaboratively, equipping practitioners to maximize the ethical advantages and minimize the ethical costs of their proposals" (Dickens, 2005). Therefore, ethical opportunities can also be seen as means of "maximizing the ethical advantages," while ethical concerns/risks can be seen as "ethical costs" that need to be minimized or avoided altogether.

Further, the Scientific Committee of the AI4People initiative (an Atomium–European Institute for Science, Media, and Democracy initiative) has categorized the chief ethical opportunities offered by AI, in general, under four broad headings corresponding to what it considers to be "four fundamental points in the understanding of human dignity and flourishing" (Floridi et al., 2018). As per "AI4People—An Ethical Framework for a Good AI Society," in each of the following categories, when AI is "used to foster human nature and its potentialities," it creates ethical opportunities:

1) "who we want to become" (opportunity: "enabling human self-realization")
2) "what we can do" (opportunity: "enhancing human agency")
3) "what we can achieve" (opportunity: "increasing individual and societal capabilities")
4) "how we can interact with each other and the world" (opportunity: "cultivate societal cohesion").

The AI4People's framework also warns that in each of the preceding cases, if AI is overused or misused," ethical risks can emerge, as follows (Floridi et al., 2018, pp. 690, 691):

1) "who we want to become" (risk: "devaluing human skills")
2) "what we can do" (risk: "removing human responsibility")
3) "what we can achieve" (risk: "reducing human control")
4) "how we can interact with each other and the world" (risk: "eroding human self-determination").

Much of the work being done by AI4EO scientists arguably falls within the categories of "enhancing human agency" and" increasing societal capabilities." For example, while the satellites serving the Copernicus program are now generating petabytes of information in the form of images of various resolutions, the processing and analysis of these data, if undertaken by human agencies alone, would most likely take too long and might not generate the required information in a timely and usable manner. Here, AI/ML "enhances human agency" and "increases individual and societal capabilities" by providing relevant information, including predictions and patterns, that is presumably as or more accurate and reliable than what would have been possible by human agency alone.

In his famous piece "Tragedy of the Commons," Garrett Hardin opined that certain problems do not have a "technical solution" but need a "fundamental extension in morality" (Hardin, 1968). This is arguably true even today. Nonetheless, technology can aid in overcoming several practical problems as well as the accomplishment of specific societal goals that

humanity has not been able to resolve or accomplish so far through legal, moral, and (other) nontechnical routes. The term *ethical opportunities* can therefore also be understood as including all such instances where attempts are made either by conscious use of specific technologies or by conscious exclusion of specific technologies to conscientiously address existing real-world challenges. In other words, embracing ethical opportunities means acting over and above now commonplace attempts to sidestep and avoid the most well-known ethical concerns/risks. In this article, we discuss such ethical opportunities, with the help of four examples from AI4EO research.

2.4. Ethical Theories and Approaches

Ethical theories and approaches are essentially methods and tools that can help explain, diagnose, and find solutions to ethical issues and dilemmas. In other words, these theories seek to give guidance on appropriate and inappropriate human behavior (MacDonald & Marcoux, 2023). Ethical theories are different from ethical approaches: ethical theories are often rooted in the historical, religious, and philosophical thought of a specific country or region of the world. They are often widely known and accepted or have been subjected to significant academic or popular debate and discussion. Accordingly, ethical theories are more well-developed "systems" for categorizing human behavior and can be used to understand, analyze, and even justify actions/inactions linked to ethical issues and ethical dilemmas.

Ethical approaches, on the other hand, are usually more recent constructs, though they may be rooted in or derived from more established ethical theories. Such approaches may, for example, provide a more systematic or stepwise guide for identifying ethical issues and for determining the correct course of action when faced with an ethical issue or dilemma. In this article, we attempt to explain an approach that AI4EO scientists can take to identify ethical issues in their research. Our approach is rooted in ethical theories that are commonly used in the Western world. Undoubtedly, utilizing ethical theories from other cultures and philosophies can enrich this approach. However, this endeavor will be undertaken in subsequent works.

2.4.1. Ethical Theories Commonly Used in the Western World

Ethical theories commonly used in the Western world are broadly classified under the following headings:

- Ethics of conduct:
  - Consequentialist theories: These theories of ethics focus on consequences and argue that the "right action" is the one that produces the maximum good, either for oneself (ethical egotism) or for the majority of the concerned public/population (utilitarianism). Consequentialist approaches to ethics, therefore, ask questions such as, "Will the consequences of your action/inaction be good?" and "Will the consequences of your action/inaction result in the greatest good for the greatest numbers?"
  - Deontological theories: Here, the theories focus on the "duties" of the people/actors who are in a position to make a decision in relation to an ethical issue or ethical

dilemma. In such theories, the ends do not justify the means, but rather, the emphasis is on ensuring that the means themselves are ethical/good. Deontological approaches to ethics, therefore, ask questions such as, "Are you acting in accordance/compliance with your duties when you act or refrain from acting in a specific way?"

- Ethics of character:
  - Virtue ethics: Virtue ethics places emphasis on the moral character of a person when analyzing ethical issues. In a virtue ethics approach, questions asked can include "Would it be kind/caring/generous/benevolent to act or not act in a specific way?" Scholars believe that several Eastern ethical theories and associated approaches (for example, in Confucianism) are also closely linked to ethics of character and virtue ethics, i.e., giving primary importance to "internal virtues" of an actor/scientist, such as compassion, responsibility, and honesty (Cheng-Tek Tai, 2013).

Consequentialist, deontological, and virtue-based theories are all categorized under the umbrella of "normative" ethics," i.e., the well-established methods and steps one can use to determine whether a course of action or inaction is "good" or "bad." Normative ethics is distinguished from "metaethics" as well as from "applied ethics." Metaethics aims to understand the philosophical and cultural foundations and origins of ethics and moral values that are dominant in any society. It asks fundamental questions, such as "What is morality?" and considers the relationship and impact of convictions and long standing societal (and even academic) expectations and codes that drive one to label specific behaviors as ethical or unethical in the first place. Applied ethics, on the other hand, caters to specific human institutions, organizations, and actions and is aimed at translating ethical principles into practical applications relevant for specific contexts. What we are attempting in this article can therefore be categorized as applied ethics.

2.4.2. Looking Beyond Dominant Western Approaches to Ethics

In AI4EO, the presence of a plethora of uncertainties mandates looking beyond consequences to duties (social responsibilities), intentions, knowledge/human values, and concrete present action: indeed, in several, if not most, instances, the (long-term) consequences of conducting research and translating research results into concrete products and services will remain unknown until a much later date. A focus on consequentialist theories may not provide meaningful guidance, especially at early stages of research.

Being aware of one's broad duties, while very significant, may also not be adequate. Indeed, guidelines for research ethics, which provide, among other things, methodological and procedural guidelines on how and how not to conduct research (e.g., obtaining informed consent from research participants and avoiding plagiarism), enumerate the broad duties of researchers. Fundamental principles of ethics, e.g., honesty, integrity, and fairness, are also basic duties and characteristics of any ethical researcher. However, these duties are abstract and may be inadequate to guide researchers working with emerging technologies, such as AI4EO.

In this situation, approaches that go beyond duties and consequences can be both relevant and useful. In Eastern philosophy, for example, considerable emphasis is placed on the interlinked concepts of karma (action, its causes, and consequences) and dharma (personal

duty, social responsibility, and religious dictates). Considerable emphasis is also placed on "our three powers" (the three levels of shakti): 1) intention/desire/will ("iccha shakti"), 2) action ("kriya shakti"), and 3) knowledge/human values ("gyana shakti"). In Indian mythology, for example, symbols associated with Skanda ("Lord Murugan") vividly describe the link among these three powers: knowledge and human values are meaningless if they are unaccompanied by appropriate intention and will, i.e., a desire to act, and constructive action. Similarly, will and action must be guided by knowledge and human values for long-term benefits for oneself and society (Cheng-Tek Tai, 2013; Shankar, (n.d.); Handelman, 1987; Kochupillai, 2021). Neo-Confucian scholars from China also emphasize a combination of action and knowledge to realize morally appropriate outcomes (Youngmin, 2022).

"Knowledge," in this context, can relate to a broad understanding of fundamental ethical principles, human values, and human rights and, in a more abstract sense, the endeavor to develop a deeper understanding of one's self or true nature. Since this abstract understanding of knowledge is beyond the scope of this article, we focus instead on the concrete understanding of fundamental ethical principles and associated human values and human rights. This foundational "knowledge" and associated understanding of ethics can then pave the way toward the identification of ethical issues and opportunities in a systematic manner and on a case-by-case basis.

3. Categorizing and Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research

Academic and scientific research focusing on ethical issues emerging from AI4EO research is scant. The most prominently identified issues relate to privacy, particularly in light of the increasing resolution of EO and RS images (Coffer, 2020). However, as diverse sources of data as well as AI and ML models are increasingly used with EO/RS data, the number and breadth of ethical issues and concerns are expected to rise. It is therefore not possible to immediately and comprehensively identify all the ethical issues that may arise in the context of AI4EO research in the (near) future.

Accordingly, in this article, following the steps explained in the "Methodology" section, we compile and categorize the most common ethical issues and associated guidelines extracted from the previously mentioned five sources, under six categories. These six categories correspond with the three fundamental principles (or values) of ethics (namely, honesty, integrity, and fairness) and the three most prominent emerging values, namely, privacy, responsibility (including the duty of care and respect for human rights), and sustainability. We are of the view that most, if not all, ethical issues and associated guidelines that may arise in the future can also be categorized under one of these six broad categories. However, the number of ethical issues that can be categorized under one or more of the six categories may grow with expanding technological capabilities in AI4EO research and with expanding human imagination.

After categorizing the ethical issues and associated guidelines, we explain each with examples from AI4EO research. The aim is to assist AI4EO scientists acquire a better and

more practical understanding of the ethical issues/guidelines that are most immediately relevant for their research. This understanding constitutes the first step to identifying ethical issues and opportunities in any ongoing AI4EO research.

3.1. An Overview of Major Artificial Intelligence Ethics Issues and Associated Guidelines

The six fundamental principles, or values, of ethics and the most prominent ethical issues and guidelines that can be categorized thereunder are enumerated in the following and thereafter explained in the context of AI4EO research, with concrete examples. These examples are derived either from a review of the literature or from discussions and deliberations with various AI4EO scientists over a period of 18 months (June 2020–December 2021):

1) Privacy:
   - (non)stigmatization
   - autonomy and freedom
   - self-determination, including national sovereignty
   - data governance, ownership, and licensing
2) Honesty:
   - transparency
   - explainability
   - data veracity
3) Integrity:
   - technical robustness
   - uncertainty
   - safety/security, including national security
4) Fairness:
   - (non)bias (including nonbias in training data)
   - nondiscrimination and diversity
   - sociocultural sensitivities
   - democratic creation of standards
5) Responsibility:
   - human agency and oversight
   - duty of care
   - accountability
   - ensuring social security and cohesion
6) Sustainability:
   - scientific, social, and cultural (including through education and training for the next generation of AI scientists)
   - environmental.

For the purpose of understanding and application, the listed fundamental values/principles that form the six main categories (i.e., honesty, integrity, fairness, responsibility, sustainability, and ensuring privacy where relevant) can be considered the major ethical duties of a researcher. Contextual and appropriate compliance with these major duties will

usually lead to compliance with other ethical duties and with ethical consequences listed as various subcategories. As a corollary, noncompliance with these major ethical principles/duties is likely to lead to the opposite, i.e., to unethical consequences and noncompliance with other ethical duties. In the following, for example, compliance with the duty of privacy will ensure nonstigmatization, autonomy, freedom, and self-determination for those affected by research and for those from whom data have been taken for research. Noncompliance with privacy will lead to the opposite, i.e., stigmatization, a lack of (or taking of) autonomy and freedom, and so on. For a stepwise procedure on how to study and use the information and explanations contained in this section, see the flowchart in Fig. 1.

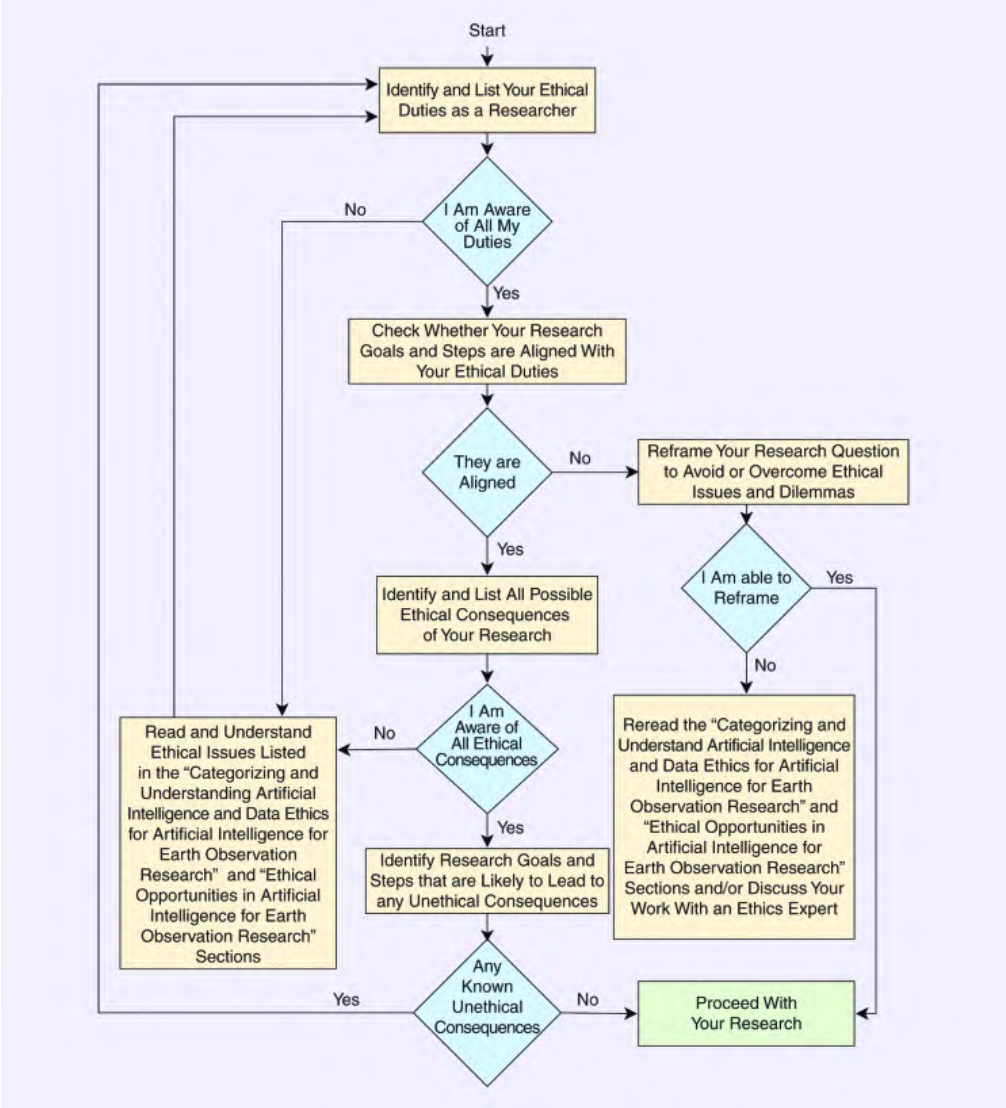


Fig. 1. A workflow suggestion for AI4EO researchers to identify ethical issues in their research.

Before proceeding, it is relevant to note the following. First, needless to say, various expert groups have placed ethical issues and guidelines linked to (big) data and AI under categories that differ from the six we use here. Also, there is a broad understanding that ethical issues (placed under these six categories) can and do differ based on technology and context. Second, it is useful to reiterate that the list of ethical issues under the six broad categories is not a closed set. The ethical issues and associated consequences are potentially ever

expanding, limited only by human imagination. Third, the ethical issues enumerated under the six categories are not independent and exclusive sets. Instead, in our exercise of explaining each of the major ethical issues through examples, it is clear that these issues overlap and can be placed under more than one category. Finally, it is also necessary to note that these issues can arise not only at the real-world application stage of any AI4EO research and development endeavor but also at the conceptual and/or fundamental research stages.

3.2. Understanding Ethical Issues in the Context of Artificial Intelligence for Earth Observation Research

Privacy

Privacy has been a central value guiding the development, implementation, and popular adoption of technologies as old and fundamental as telephony and telegraphy as well as more recent innovations, such as the Internet (Shapiro, 1998). Indeed, technologies and applications that fail to protect privacy may fall into disrepute and lead to declining user acceptance (Rumbold & Pierscionek, 2018) over time. In the context of EO data, privacy concerns had been relatively unheard of (Denis et al., 2017; Sudmanns et al., 2020; Wulder & Coops, 2014). However, with growing image resolutions and the emergence of drone technologies in RS, privacy concerns have become starker and clearer (Coffer, 2020; Chandra & Bhatia, 2020). The problem is aggravated, as images are retrieved at a known point in time (when), at a specific known location (where), and include views of potentially detectable neighborhoods, residences, and even people (who) (Coffer, 2020; Maniadaki et al., 2021). With this information and additional contextual information, it may also be possible to detect specific activities (what) (e.g., gardening, sunbathing, barbecuing, and so on).

In present times, privacy concerns are also commonly flagged in most technologies that rely on big data. Accordingly, in AI4EO research, which uses an ever-increasing variety of complementary (big) data sources, privacy concerns may take on various dimensions based on the type of data used and the purpose for which they are used. Today, very high-resolution EO data are capable of characterizing individual buildings, and sensing techniques may even be capable of assessing building quality (Aravena Pelizari et al., 2021). At these detailed spatial and thematic levels, the identification and analysis of objects (e.g., houses and localities) as well as the subsequent use to which an analysis is put may lead to the violation of privacy. Undoubtedly, the nature of data (whatever their source may be) required in various EO/RS applications is often quite different from data used in, say, the medical field. For example, geoinformation data from social media platforms, such as Twitter and Flickr (Woodcock & Strahler, 1987; Patino et al., 2014; Zhu et al., 2022), mined with the aim of labeling buildings and identifying areas where rescue efforts need to be expedited, is of a different nature than individuals' mental health information mined from the same platforms (Wang, So & Smith, 2015). In this context, recent research refers to Nissenbaum's theory of "privacy as contextual integrity" and suggests that this can be "a useful heuristic to guide ethical decision-making in big data research projects" (Zimmer, 2018).

Based on context, the collection, storing, processing, and transfer of data can be done with much greater attention to legal and ethical requirements by (for example) adopting supporting technologies that give data contributors greater control over their (personal) data, on the one hand, and support the development of the European Data Economy, on the other hand (see the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section for further discussion on this). However, as we see in the context of other ethical concerns and risks, a technical solution is not always possible. What may be necessary in such situations is a fundamental rethinking of the problems on the ground and how research questions can be redesigned to specifically address those concerns. Examples in this regard are provided in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section.

Stigmatization

Closely linked to the issue of privacy is the issue of stigmatization. Take for example, the case of geotagged social media data from platforms such as Twitter and Flickr. Such data are being used with EO/RS data for a wide variety of uses ranging from disaster relief (Bellman et al., 2004) to population density estimates in various regions (Shao et al., 2020; Malyuta et al., 2020). The use of geotagged social media data for disaster relief may be welcomed by the users whose lives are rescued as a consequence of such use (Zhu et al., 2022). However, in case of a negative assessment of housing quality and in the case of slum identification, for example, the classification of certain users (people) as slum dwellers may not only be considered a breach of privacy but may also lead to stigmatization of specific individuals, communities, and regions.

Here, we see that even routine tasks, such as "data labeling," can result in stigmatization. Traditionally, the labeling of large pixelated EO/RS data was done with the support of so-called ground truth data (Malyuta et al., 2020; Kestur et al., 2018), i.e., confirmations and rejections received from the ground as to the accuracy of predictions made from EO images. When these ground truth data are not available or are expensive to collect comprehensively, researchers use visual interpretation based on their subjective perceptions and experiences, or they use social media data, census data, and various ML/AI models to complement and confirm EO data and to make predictions from EO data. While earlier EO/RS work consisted of broad land cover/land use classifications at low resolution (Woodcock & Strahler, 1987), in the present day, these classifications become more and more refined, increasingly supporting the identification of regions that are more prone to crime (Patino et al., 2014), fire (Krylov et al., 2014), and economically deprived living areas, commonly called slums (Kuffer et al., 2018; Kuffer, Pfeffer & Sliuzas, 2016).

Labels, whether they are based on "ground truth" or other sources/algorithms, need to be worded and placed with an ethically mindful eye. Here, ethics requires not only accuracy but also sensitivity. For example, the understanding of what constitutes a "slum" can vary based on country, culture, and geography as well as socioeconomic realities of a region (Dovey et al., 2021) and the experiences of people placing the labels (Kraff, Wurm & Taubenböck, 2020). In this context, as discussed in greater detail in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section,

translational, or applied, ethics may require that stigmatizing labels be replaced with constructive and affirmative ones. Translational ethics calls for the actual application and implementation of theory to practice or, in other words, ensuring that ethical theories and principles can be and are translated to real-world applications and practices. In medical ethics, for example, translational ethics calls for the application of science/scientific findings to clinical practice.

While stigmatization is an ethical issue that will arguably arise only at a very advanced (research) application stage, its avoidance is possible only if researchers are alert to the probable reality of stigmatization (and the various sociocultural and economic harms that can result from it) at a very early stage of research—the stage where individual researchers have the possibility of applying labels responsibly and constructively rather than merely based on convenience and disciplinary custom. This is also discussed in greater detail in the context of bias in the following and in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section.

Autonomy, Freedom, and Self-Determination, Including National Sovereignty

One of the fundamental guiding ideas of data ethics is "enhancing protection for individual freedom, self-determination, and integrity" (Data Ethics Commission of the Federal Government, 2019). When data and associated data analysis limit individual freedom and self-determination, the principles of ethics can be breached while also violating basic human rights. Take again, for example, the issue of slum identification using AI4EO. Under international pressure, country governments may use research findings to clear areas labeled as slums, dislocating thousands from their homes and creating significant social unrest. Or governments may be forced to imprison illegal artisanal miners and deny them their traditional source of livelihood. In the face of such possibilities, AI4EO researchers need to ensure that their research is mindful of "ground truths" beyond the accuracy of labels. Responsible research also needs to go beyond avoiding ethical risks (e.g., by hiding and not disclosing geolocations) to actively changing research approaches to address ethical concerns on the ground. We discuss this in greater detail in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section.

When applied to national contexts, the issue of self-determination and autonomy can assume the form of national sovereignty, which can be (adversely or appropriately impacted) through international sociopolitical pressure. For example, consider the issue of nuclear disarmament: India's preparations to make a first effort at testing its nuclear weapons were discovered by U.S. intelligence via satellite-based monitoring of the test station at Pokhran, Rajasthan. The discovery of these plans led to large-scale international pressure to withdraw the efforts. In a subsequent attempt, great pains were taken by Indian nuclear and defense scientists to prevent the plans from being discovered by U.S. satellites, leading to successful testing without intervention from international pressure. Despite subsequent international condemnation for going against the ideal of nuclear disarmament, India justified its stand based on rising armed attacks and threats from Pakistan and China at its borders (Ganguly, 1999). (The successful nuclear tests were conducted in a desert area of Western India. The

efforts of Indian scientists to "hide" from U.S. satellites to prepare for and conduct the tests were recently dramatized in a popular Hindi movie, Paramanu: The Story of Pokhran.)

Data Governance, Ownership, and Licensing

Data governance has been defined as the "exercise of authority and control over the management of data," with the purpose of increasing "the value of data and minimizing data-related cost and risk" (Abraham, Schneider & Vom Brocke, 2019). Data governance therefore also includes the management of data in accordance with the requirements of the law. In the EU, data are protected under various laws, including the General Data Protection Regulation (GDPR) and the Directive on Copyright in the Digital Single Market (DSM). Accordingly, data governance requires the management of issues of data privacy as well as data ownership.

In the context of data ownership, European Parliament directive 2019/790 on copyright and related rights in the DSM (the so-called DSM Directive) includes two exceptions for "text and data mining" (see articles 3 and 4 and recitals 11, 14, 15, 17, and 18). These exceptions permit text and data mining for scientific research purposes and are also important for AI/ML research that relies on (big) data (Rosati, 2018). Absent these exceptions, the mining of text and data from any source could amount to copyright and database rights infringement. While these exceptions were introduced to support, among other things, R&D, including in the context of AI/ML research (European Commission, 2019), any entity mining and utilizing such data still needs to be mindful of privacy considerations. Whether the subject concerns images of people from high-resolution EO data or individual (image) data mined from social media platforms, the privacy of the people concerned (and of their data) needs to be secured. In addition to being a legal requirement under the GDPR, privacy is a human right and must also be secured as a matter of ethics, even in the absence of legal regulations in most jurisdictions. Arguably, therefore, the ethical standard of privacy may be higher than the legal standard. For example, the argument that data are already public (e.g., since they were posted on Twitter) may be adequate to avoid trouble with law enforcement, but it may not help overcome ethical concerns over privacy (Lomborg & Bechmann, 2014; Zook et al., 2017).

Beyond the GDPR, ethics requires that when data are collected without the consent of people whose lives are closely affected and associated with the data, the analysis and dissemination of such data must not adversely impact the fundamental human rights and welfare of such people. This is centrally relevant, as EO data as well as AI4EO data are being used to make policy recommendations linked to the United Nations (UN) Sustainable Development Goals (SDGs). EO scientists may see this issue come into play when their research helps identify slum regions and illegal mining sites; the research may lead to policy decisions whereby large numbers of people living in identified slums or deriving their livelihood from small-scale artisanal mining lose their homes and means of livelihood. A responsible restructuring of research questions and research designs may be necessary to avoid such dilemmas. Similarly, in the context of agriculture, the EU code of conduct for agricultural data sharing emphasizes that "the farmer remains at the heart of the collection, processing and management of agricultural data" (Copa-Cogeca, 2018). The code recognizes that farmers

who share data have the right to know and participate in the uses to which their data are put. In AI4EO, when data are collected remotely and without the knowledge of farmers, ethical issues become even starker if the data analysis adversely affects farmers' (or farming communities') right to self-determination.

Honesty

Honesty is one of the central and, therefore, most fundamental values of (business) ethics (Ferrell & Fraedrich, 2015). It highlights the importance not only of truthfulness but also of trustworthiness (Mihelic, Lipicnik & Tekavcic, 2010; Schwartz, 2013). Accordingly, hiding a relevant fact and failing to disclose it raises as much of an ethical concern as actively miscommunicating and misrepresenting a fact. In AI4EO, the principle of honesty requires scientists to be transparent about the (possible) shortcomings and limitations of their reasoning as well as the limits to which predictions based on their research may (or may not) be accurate. Therefore, issues of transparency, explicability, and data veracity are closely connected to the broader issue of honesty. Further, as AI4EO research is replete with uncertainties, honesty would require a kind of "laying bare" of these uncertainties so that future researchers, developers/innovators, and policy makers could take necessary precautions while relying on the findings and products/services based on them.

In the context of performance prediction (for an AI4EO model), honesty also comes into play when compiling training datasets and test datasets (i.e., data used to verify the accuracy of the training data). If pixel values corresponding to testing sets are partly or completely seen during the training phase (i.e., as part of the test data), the number of independent testing pixels gets reduced and can lead to overoptimistic accuracy assessments. Good scientific practice requires that test samples remain unseen during training. This is important, as high-capacity models (as with deep neural networks, with their thousands of parameters) are very good at overfitting training data. If the data points in training and test subsets come from the same or an overlapping distribution, good test accuracy/high confidence in test samples does not serve as a good predictor of performance on unseen data.

Honesty requires that AI4EO scientists exercise caution while selecting training and test datasets and be transparent about the scope and overlap in such sets (Molinier & Kilpi, 2019; Audebert, Saux & Lefèvre, 2018). Discussions with AI4EO scientists revealed that overlapping test and training datasets may result not from active dishonesty but because scientists may not be fully aware of good scientific practice in deep learning (or due to lack of experience/inadequate competence, especially in the case of young scientists). Under both Western and Eastern theories of ethics (especially deontological theories), it is the duty of scientists to fully educate themselves on good scientific practices relevant to their field of work. Accordingly, any argument based on inadequate competence and gaps in knowledge of good scientific practices may not release one from ethical duties and from the consequences of "inadvertent" dishonesty in research. At the very least, a researcher will lose credibility, affecting his/her long-term career.

Transparency

In line with the requirement of honesty, transparency calls on scientists engaged with AI4EO research to be transparent about the scope and limitations of their research. This can include, for example, explanations clarifying why specific sources of data are selected instead of or to the exclusion of others, how scientists are extracting such data (e.g., which methods are chosen and which ones are not) (Driscoll & Walker, 2014), what parameters guide the labeling and analysis of a resulting dataset, and what the scope and limitations of the resulting model and analysis are. In this context, the "Ethics Guidelines for Trustworthy AI" (Pekka et al., 2018) state that "humans need to be … informed of the system's capabilities and limitations."

A challenge in the search for scientific truth is that classified and quantifiable results are not always unambiguous because, for example, the target object of interpretation is conceptually ambiguous. Especially in spatial analyses, there is no unambiguous truth. Take the terms monocentric and polycentric as examples. Conceptually, monocentric refers to a center where urban functions are concentrated, and polycentric refers to a reduced dominance of the urban center, giving way to many suburban local centers. While this distinction seems conceptually clear, in spatial science, quantified results are ambiguous in the sense that there can be no absolute truth as to when a dominant center loses its dominance. In such cases, therefore, honesty simply requires that academic researchers declare this ambiguity so that downstream reliance on associated research can take this ground reality into account. As discussed in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section, such ground-level ambiguities can also be addressed by delimiting the focus and intention of corresponding AI4EO research.

Explainability

Transparency, when taken a step further in the direction of decisional transparency, particularly in the context of AI applications using big data (Beerbaum & Puaschunder, 2019) to reach decisions and recommendations, becomes the requirement of explainability, or explicability. Undoubtedly, not all AI and AI4EO applications require explainability. Indeed, explainability may not be necessary in all instances, and it may not serve any additional purpose. In recent times, scholars have also recommended systems that rely on verifiability to complement and substitute complex (and arguably unnecessary) explainability requirements (Patil & Heinrichs, 2022). However, whether explainability is or is not an ethical requirement may depend on the AI4EO use case. For example, if an EO/RS data-based scoring method is used to determine which (rural) areas ought to be prioritized for electrification, the specific types of data and the method adopted to reach that score as well as the parameters that are considered and those that are excluded need to be made transparent and explainable (Engelmeier & Gaihre, 2019). This issue is also closely linked to the issue of democratic decision making (see the following) and inclusive growth, both of which can be either supported or severely compromised by AI4EO research.

Data Veracity

Since AI/ML algorithms are used to train models using data from one source/context to make predictions in various other contexts, the issue of the veracity 1) of the data used to train the models, 2) of the presumptions underlying the models, and 3) of the resulting predictions becomes very relevant. Especially when using data from novel sources, such as texts from social media data, to support building-level population density estimates, the limitations of natural language processing, the absence of accurate geolocation tags (e.g., due to Twitter's decision to remove support for precise geotagging (Hu & Wang, 2020)), the user group of the online platform (which may not be representative of the overall population), and the distortions generated by specific users of the platform need to be borne in mind while relying on such data for training and analysis. The issue of data veracity can also arise as a result of inaccurate and insensitive data labeling (see the "Stigmatization" section). Needless to say, "truth" can have several important and justifiable subjective/local considerations: what is truthfully labeled as a slum from the perspective of a European researcher may be the average standard of living for a South Asian middle-class family living in a congested neighborhood. Contextual data veracity, however, is important to avoid problems, such as the previously mentioned issue of stigma. The importance of context is also seen in archaeological geospatial data (McCoy, 2017).

Integrity

The term integrity has multiple meanings depending on context. In a simple business ethics sense, it can be understood as "walking the talk," or actually doing or implementing what one preaches to be "right" or "correct" (Cording et al., 2014; Koehn, 2005). In the technical sense, the word integrity is often used in the context of "system integrity" (Kim & Spafford, 1993; Khan et al., 2016), which essentially means that the system, algorithm, AI/ML model, and resulting application performs in the manner it is meant to perform and that its outputs/predictions are accurate to the extent the creators and owners of the system claim it to be. System reliability, safety, security, accuracy, and overall technical robustness are therefore subjects that are closely interlinked, have important ethical implications, and can be categorized under the broad heading of "integrity."

In practice, integrity is also closely associated with the previously discussed duty of "honesty" (Russell, 2001). Specifically, if a system does not perform accurately or does not do so in specific circumstances, the ethical bar of "integrity" requires that one is "honest" and transparent about the circumstances and extent to which the system can/does give accurate results (Verhezen, 2008) and/or the extent to which one can rely on the research findings and applications. The implications of meeting or not meeting the "integrity" bar also vary from application to application, depending on the real-world end uses to which the application is put (or intended to be put) and the impact such an application has or is likely to have on actual human, animal, and environmental well-being (Besser-Ones, 2008).

Technical Robustness and Uncertainty

Investigations into "integrity" can be translated into questions such as, "Does the system do what it claims to do?" "What are the limitations?" "What is the possibility and level of error in any result?" In EO and RS data, especially including EO/RS data that are compiled and assessed with the help of AI/ML models, there is a well-known problem of uncertainty. What exactly is the ML model learning (from) and on what basis is the resulting image produced? For example, for many years, it was believed, based on EO data, that the phytoplankton concentration was higher near the poles (Huang et al., 2017). However, it was later discovered that this apparent higher concentration resulted from errors in the atmospheric correction software used in the application (Friedlingstein et al., 2021). To the extent that such measures may lead to major national and international (R&D) investments and changes in governmental policies, integrity requires that the scientific community at least declare the probability and, where possible, the level of error and uncertainty and identify the limitations of its research findings. This will not only permit more responsible research spending but also ensure that policy decisions are not made merely on a kind of "blind faith" in scientific findings (Huang et al., 2017).

Safety/Security, Including National Security

Needless to say, issues of (system) integrity are closely related to safety and security. The failure of any system to perform in the manner in which it is meant to perform can lead to the loss of human/animal lives and to environmental and infrastructural damage. The failure and inadequacy of data governance structures can also lead to security breaches, further leading to compromised privacy and violations of ownership rights. Given the scale of AI4EO research, the scale of security issues that can arise from such research is equally high.

With the possibility of fusing diverse data sources, AI4EO research is marching boldly toward novel use cases that were unheard of just a few decades ago. AI4EO is extensively used in defense and law enforcement to identify enemy/unauthorized vessels (aircraft and ships), track vehicles transporting illegal goods and suspected criminals, remotely identify car license plates, and predict possible target areas during wars and other armed conflicts. These use cases of AI4EO, while ethically questionable, are largely accepted to secure nations and people.

Yet, more novel use cases of AI4EO, such as large-scale projects aimed at predicting building-level population density, can lead to new issues of national security, especially if resulting data and maps are publicly shared. For example, the availability of such data can make specific buildings more prone to terrorist attacks. In this context, therefore, the previously discussed issues of data governance, including decisions as to how, when, and to what extent to make data public, must be given central priority. While open source data are appreciated by the research community, issues of national security may require more responsible disclosure of data and predictions.

Fairness

Fairness is another one of the fundamental building blocks of (business) ethics. Borrowing from Kant's categorical imperatives, fairness can be judged based on whether the person creating a rule or taking an action would find the same rule acceptable/appropriate if it were universally applied, including to him- or herself. Fairness is closely linked with treating equals equally (Kim & Spafford, 1993). At the same time, it also implies, as does the fundamental right to equality, that unequals not be treated equally (Kim & Spafford, 1993). Thus, for example, criminal law does not prescribe the same punishment for juvenile delinquents as it does to adult recidivists, even if the nature of the crime is identical. Similarly, in AI and EO, the principle of fairness requires, again, that appropriate measures be taken, for example, while labeling data and while making recommendations, to ensure that regions that are in similar circumstances are treated similarly. At the same time, the highest standards of ethics would also call for AI and EO research to rethink research designs in ways that can uplift marginalized segments and bring them equal status. The principle of fairness, therefore, can be applied not only at the time of labeling data (e.g., in the context of labeling crime-ridden areas and slums) but also at the time of selecting research focus areas, allocating research funding, and converting research findings into practical applications. Given the scope and reach of EO/RS research in terms of the sheer areas it (potentially) covers, issues of fairness need to be flagged at all stages of research and all stages of data collection and analysis.

Issues of fairness also arise at the macro level when we consider the availability of EO/RS data and the means to meaningfully access, analyze, and use these data. Take the example of climate change: a study of anthropogenic influences driving climate change reveals that it is primarily industrialized nations that contribute to the problem (Friedlingstein et al., 2021); they have significantly higher energy consumption, carbon dioxide (CO2) emissions (see Fig. 2), and greenhouse gas production. Yet, the consequences of climate change are a burden on all nations in the world. An important task of RS is to monitor and collect georeferenced data and datasets on various parameters of anthropological climate change. These datasets are the key information source for ML models and forecast systems of climate change consequences. Therefore, the progress of climate change monitoring depends highly on the availability of such datasets.

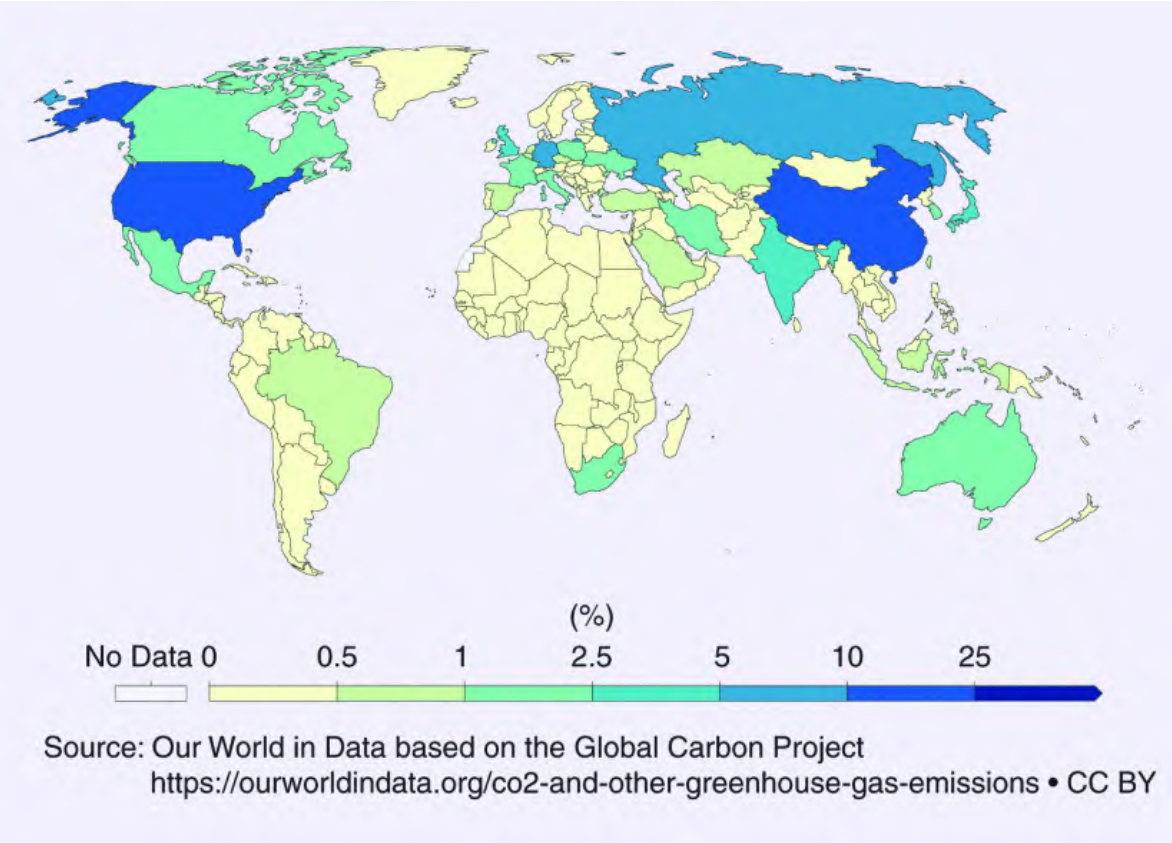


Fig. 2. The cumulative CO2 emissions by country (Friedlingstein et al., 2021). Cumulative emissions are calculated as the sum of annual emissions from 1750 to a given year.

What we observe, however, is a correlation between the number of satellite missions that a country owns and the level of the country's development (see Fig. 3). This trend is also visible in the availability of georeferenced datasets related to that country. In Fig. 4, it is clearly visible that regions of high development (e.g., North America, China, and Europe) are significantly more represented in georeferenced datasets, independent of the degree to which they are affected by climate change. In Figs. 2 and 3, one also observes that countries with a lower Human Development Index (HDI) score have significantly lower CO2 emissions and are also significantly less represented in georeferenced datasets. Yet, these countries are as affected by climate change as the countries with higher emission rates. Fairness requires that more industrialized nations take responsibility to assist poorer nations in combating climate change, for example, by helping them gain access to relevant data from their regions.

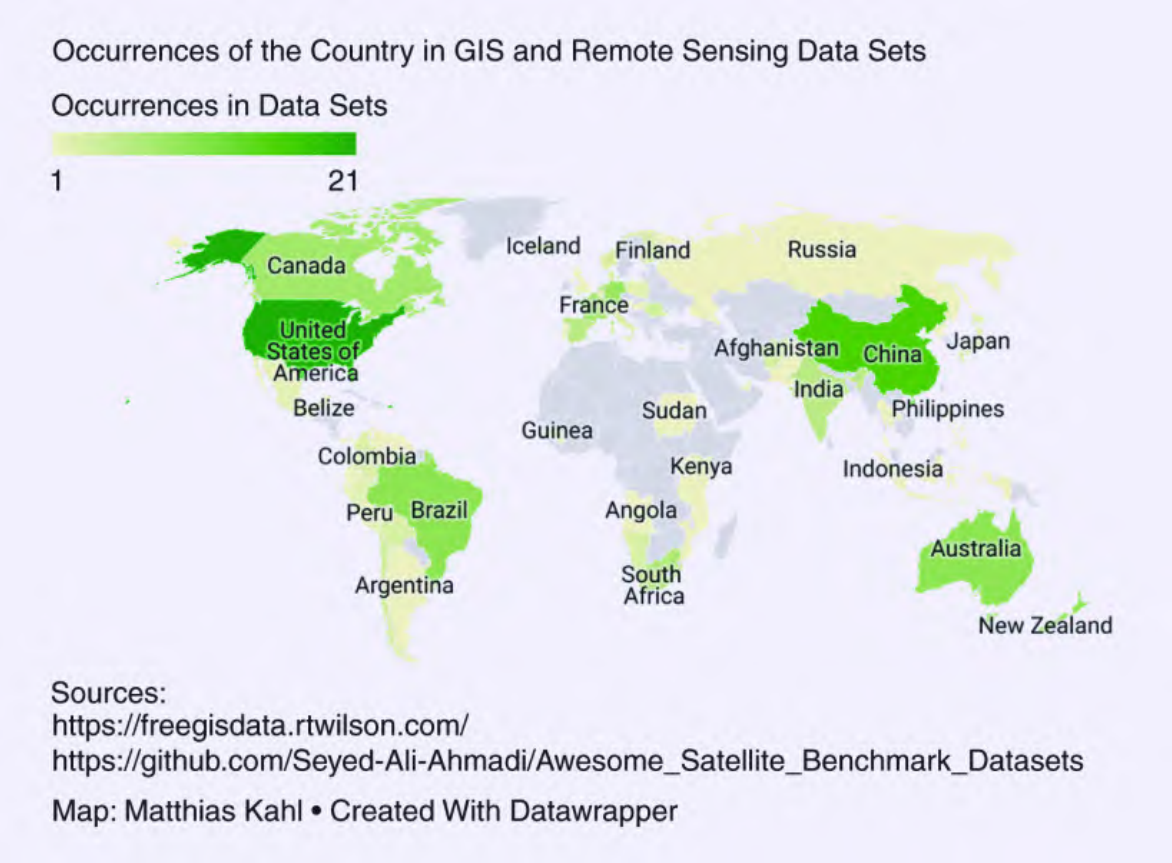


Fig. 3. The number of occurrences and therefore the presence of a country in RS and geographic information systems datasets in (Schmitt, Ahmadi & Hänsch, 2021; Wilson, 2021).

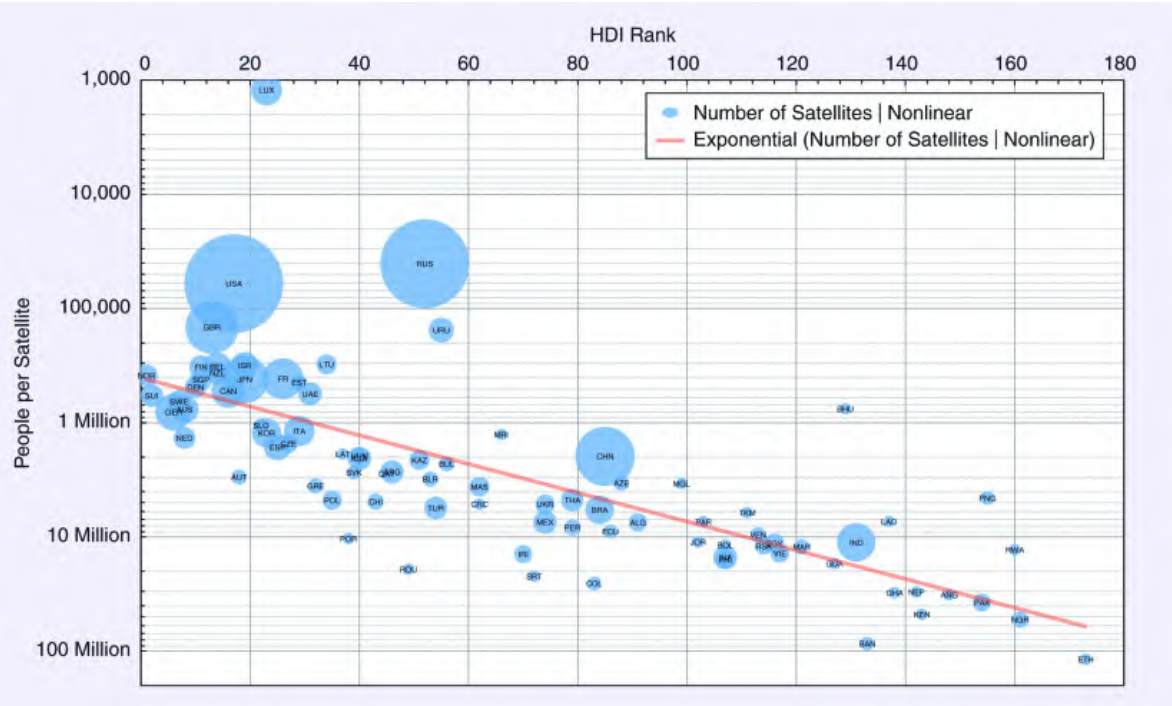


Fig. 4. The correlation between the relative number of satellites in relation to the number of inhabitants of a country and the country's ranking in the Human Development Index (HDI). A correlation between the development of the country and its number of owned satellites can be seen. The chart is based on the analysis of the satellite catalogue in (WMO, 2021), the number of inhabitants for each country in (Wikipedia, 2021), and the ranking of the countries HDI in (Wikipedia, 2023).

SUI: Switzerland; BEL: Belgium: LTU: Lithuania; ESP: Spain; AUT: Austria; KSA: Saudi Arabia; SVK: Slovakia; MAS: Malaysia; CRC: Costa Rica; ROU: Romania: AZE: Azerbaijan; THA: Thailand; PER: Peru; ECU: Ecuador; ALG: Algeria; IRL: Ireland; MGL: Mongolia; PAR: Paraguay; TKM: Turkmenistan; INA: Indonesia; RSA: South Africa; VIE: Vietnam; MAR: Morocco; BHU: Bhutan; USA: United States; GBR: Great Britain; RUS: Russia; NOR: Norway; FIN: Finland; ISR: Israel; NZL: New Zealand; SGP: Singapore; DEN: Denmark; JPN: Japan; CAN: Canada; FR: France; EST: Estonia; UAE: United Arab Emirates; SWE: Sweden; AUS: Australia; GER: Germany; NED: The Netherlands; SLO: Slovakia; KOR: South Korea; ITA: Italy; CZE: Czech Republic; LAT: Latvia; HUN: Hungary; GRE: Greece; POL: Poland; CHI: Chile; POR: Portugal; TUR: Turkey; ARG: Argentina; QAT: Qatar; KAZ: Kazakhstan; BUL: Bulgaria; BLR: Belarus; CHN: China; UKR: Ukraine; BRA: Brazil; MEX: Mexico; COL: Columbia; JOR: Jordan; VEN: Venezuela; BOL: Bolivia; PHI: Philippines; IND: India; GUA: Guatemala; LAO: Laos; PNG: Papua New Guinea; RWA: Rwanda; GHA: Ghana; NEP: Nepal; PAK: Pakistan; ANG: Angola; KEN: Kenya; NGR: Nigeria; BAN: Bangladesh; ETH: Ethiopia; and URU: Uruguay.

It is also necessary to recognize this situation as a lost opportunity to learn from "best practices" for climate change mitigation that may be adopted in developing countries but unknown in Europe and other developed regions of the world. In India, for example, there is rapid farmer adoption of agroecology (a system of farming) under the name of natural farming (Kochupillai & Köninger, 2022). The availability of data from regions that have successfully migrated to these sustainable farming systems can be hugely beneficial for Europe, where policy makers are actively seeking means of optimizing farmer migration to agroecological approaches for sustainable farming.

(Non)Bias, Including Nonbias in Training Data

Bias is an ethical issue that has frequently come into the limelight in relation to AI applications ranging from Amazon's AI-based recruitment tool that showed bias against women (Dastin, 2018) to the PULSE Photo Upsampling via Latent Space Exploration (PULSE) algorithm that upscaled low-resolution images of celebrities, such as Barack Obama and Lucy Liu, into images that were distinctly "White" (The Verge, 2022). Bias in AI applications frequently results from the diversity (or lack of it) in the data used to "train" AI and ML models. This lack of diversity and the resulting bias can arise from using data from any source [including, in our case, visually and automatically generated labels from RS data, from official (cadastral) data databases (Fekete et al., 2015), and from other sources, such as social media platforms] to train (among others) ML models (Beerbaum & Puaschunder, 2019; Gorwa, Binns & Katzenbach, 2020).

Based on the training data used to train AI/ML models, various biases might creep into a system; extending beyond a preference for white/male candidates for jobs, they can lead to unfair categorizations of areas with higher immigrant populations as "slums" and "crime ridden" (O'Brien, Farrell & Welsh, 2019; Perez, 2019). The ethical concern of fairness therefore alerts the AI/ML community to the high possibility (and probability) of bias (including bias in training data, in data labels, and so on). It also actively seeks behavior, labeling, and coding that ensure diversity, are nondiscriminatory, and take into account sociocultural as well as economic and political sensitivities. Closely connected to the concept of fairness, therefore, is the issue of the democratic creation of standards (Wong, 2020) (see the following for further discussion), whether they are standards to be followed for the selection of training data or for labeling data used to train AI/ML models.

Fairness also requires that issues of data ownership and licensing (e.g., of intellectual property) are clearly and fairly addressed. Further, when supplementary sources of data, such as data from social media platforms, are used to complement EO/RS images and increase their accuracy and resolution, it is necessary that the terms under which such data are acquired are in keeping with the licenses of the social media platform, on the one hand, and the spirit of broader regulations, such as the GDPR, on the other.

Nondiscrimination and Diversity

Bias may also arise from personal and cultural prior knowledge. In computer-based pretrained categories, this will simplify, or bias, the world. A sort of "selection bias" can also result from the convenient selection of specific sources of data to the exclusion of others. Indeed, selection bias may also be inevitable due to a lack of data from specific (geographic) locations (see the "Fairness" section with regard to the availability of EO/RS satellites and data. In RS, appropriate datasets for specific tasks may also be neglected due to the limited availability of research funds. Therefore, the selection of any specific source (and even region) for data collection needs to be disclosed, and the analysis resulting therefrom must clarify the scope of applicability (e.g., geographic, cultural, and other limits) and any inherent limitations in the findings and the algorithms used for analysis. This is necessary to avoid the danger of broad policy making based on limited studies. It is also necessary to ensure responsible downstream/applied research.

However, the major challenge in the context of bias in training data results from a lack of open, accessible, representative, and generalizable data (Alsaad, Taamneh & Al-Jedaiah, 2018; Hargittai, 2020; Grace et al., 2019) to permit optimal training and appropriate contextual labeling and semantic segmentation. Further, datasets (e.g., in scientific publications) usually cover only a small range of scenes/cases in which the predictive models and algorithms perform well. As a result, the data used to understand the fundamental context/relations in EO may be inadequate and totally inappropriate to make predictions in other contexts, even where these contexts are somewhat similar. For example, training on crop fields in moderate climate regions cannot be applied in tropical regions. Similarly, data used to train AI/ML models in, say, identifying building types are not reflective and representative of what is seen in diverse real-world contexts. While data scarcity is a reality, it is the authors' responsibility to clearly and honestly state the extents and limitations of their datasets, level of diversity, and overall approach as well as resulting limitations in the scope and applicability of the research models/findings.

It is important here to draw a distinction while acknowledging the relationship between datasets and models. It may be difficult to identify limitations of models (e.g., of feature extractor networks trained in an unsupervised manner), but the scope of the dataset used to train a model is always known. The limitations of the dataset are also likely to impact the output provided by the model trained using that dataset. Therefore, even in models trained in an unsupervised manner, the underlying dataset and its limitations must be identified and disclosed. In the case of AI in health care, for example, it has been discovered that models trained primarily on datasets consisting of male heart patients are unable to accurately predict heart disease in women (Mohammad et al., 2022). Here, the limitations of the dataset (and

by implication, of the model) can be known and predicted if the scientists and developers involved clearly state that most of the training data are from male heart patients so that doctors are more careful when relying on this dataset when faced with a female patient. Similarly, closer to home, i.e., in EO research, if a model for predicting slum areas is trained on data from Brazil, it is necessary to identify the scope and origin of the dataset so that it is not (fully) relied on to predict slum areas in Nigeria and other regions of the world where the population density, local cultures, and average income levels may be very different.

To increase the diversity of available data, methods to incentivize and maximize contributions of more diverse data (in a manner that secures the privacy of contributors) from diverse global locations and circumstances may need to be researched to tap into the full potential (ethical opportunities) offered by AI4EO research. Further, in the context of diversity, the impact of technology on gender and vice versa (i.e., the impact of gender on technology) is an important area of research. The impact of technology on gender is also pertinent in the context of EO/RS research and in the context of geoinformation technology applications. In relation to the impact of gender on EO/RS research, a recent study reveals that in the broad "geospatial" profession, only 19% of the practitioners are women (Gurung Goodrich, Gurung & Hamal, 2021). This gender-skewed landscape can have an impact on the manner in which research is approached and research questions are prioritized as well as the overall effect that the research has on society at large.

Here, it is important to note that a "masculine" or "feminine" research agenda is not exclusively about an emphasis on methods/technical solutions versus emphasis on real-world problems but also about 1) how methods and technical solutions would themselves differ under the scrutiny of female researchers and 2) how the problems for which technical (and other) solutions are sought would differ. For example, in relation to the second point, research questions that are particularly important for the empowerment and security of women "on the ground" may not come to the limelight and be taken seriously in a "masculine" research agenda that emphasizes methods and technical solutions that seek to address more general questions, such as the prediction of biomass content and slum areas.

Sociocultural Sensitivities

In addition to the gender component discussed in the preceding, this important ethical consideration can, once again, be explained in the context of data labeling. In countries of South Asia, for example, for cultural as well as space-related reasons, joint families and denser community living are quite common (Chadda & Deb, 2013). Errors associated with the underestimation of population density and inaccurate labeling of specific buildings and areas as "too" densely populated and as being "slums" can be minimized if researchers are aware of the likelihood of such errors and the local sociocultural and net usage differences that can cause them. This can be done by ensuring diversity in the sources of training data, the culture-sensitive labeling of training data (that also considers inputs from multiple stakeholders), responsible predictions, and the declaration of any residual inaccuracies that may result from inadvertent bias.

Democratic Creation of Standards

Such standards are particularly important to prevent bias resulting, for example, from data labeling. Indeed, every "observer's" label may differ (Caverlee & Niu, 2015). At the same time, this difference can be a source of rich information, and permitting a diversity of labels from researchers of various backgrounds can yield more informed "aggregate" and, hopefully, less biased data labels. In the context of scales and ranges of data labels, multistakeholder and multicultural discussions may be necessary to identify culturally appropriate and democratically created labels and standards that can avoid bias as well as stigmatization.

Responsibility

In the entire domain of "living" creatures, human beings are uniquely equipped with the power to take responsibility. This human ability to take "responsibility" lies at the heart of all laws, regulations, ethical principles, and moral codes. Indeed, absent human ability to take responsibility for actions and inactions, no legal or ethical code would be realistic and implementable. At the interface of law and ethics, the most important responsibility of any human being is to protect human rights, fulfill a "duty of care" (e.g., under tort law), and remain accountable for all his or her actions and inactions. The borders of responsibility and accountability (which can also translate into issues of legal liability), however, become blurry in the face of increasing automation and reliance on AI/ML models. Nevertheless, responsibility is listed as an important and central ethical value in almost all (AI) ethics guidelines. In the following, we discuss some of the most important ethical issues linked to responsibility in AI4EO research.

Human Agency, Oversight, and Accountability

The EU "Ethics Guidelines for Trustworthy AI" highlight the importance of "human agency and oversight" in AI systems. The active involvement of human actors (diverse stakeholders) in every aspect of AI4EO research may not be possible. However, such involvement is closely linked to the previously discussed issue of the "democratic creation of standards," e.g., in the process of labeling. Further, the quantification of uncertainty, a task undertaken by specialized AI4EO scientists, can support human agency and oversight requirements, thereby enhancing trust in AI4EO research. At the same time, it is important to bear in mind that human agency and oversight are not necessary only within the confines of the "system" and its narrow research question. Human agency and oversight are also needed to understand the context in which the designed system is going to be implemented and introduced and the impact that such an introduction may have on real people and their lives, livelihoods, and environment.

The human agency and oversight requirement, therefore, has the potential to evolve considerably in the AI4EO research context. Otherwise, while looking at Earth and its physical elements from afar, intricate human realities, relationships, and expediencies can be overlooked and ignored, leading to significant ethically undesirable consequences. Take, for example, the issue of identifying refugee ships. While a researcher may take great pains

to ensure system integrity, honestly declare system limitations, and ensure that the labeling of pixels is done while taking multiple contexts and viewpoints into account, if the ground reality of how refugee vehicles are treated by a country/government in question are ignored, the sharing of data/algorithms can lead to the loss of several lives. Similarly, in the case of research identifying slum areas, international pressure in the form of HDI scores can lead local governments to tear down areas identified as slums, driving thousands to homelessness.

Current AI4EO research that relies primarily or exclusively on data in the form of images, texts, and numbers (e.g., census information) can benefit greatly from also integrating qualitative real-world information into its analytical framework and research questions. A simple approach could, for example, be conducting a literature review to understand the current sociocultural and political realities of a region in the context of AI4EO research goals. The "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section discusses some concrete examples of how such a literature review may enhance the scope and impact of AI4EO research, on the one hand, while leading to novel methodological approaches and ethical outcomes, on the other hand.

Human Rights and the Duty of Care

EO/RS data and research are rapidly moving toward use cases that have a direct and more concrete impact on governmental policies and real human lives. In such a situation, EO/RS research needs to keep a close eye on the impact it can have on human rights. The sphere within which AI4EO researchers need to exercise "care" is therefore expanding. This is explained in greater detail in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section, with the help of four examples from AI4EO research.

Sustainability

Sustainability is arguably not one of the traditional concerns of (business) ethics (de Bakker, Rasche & Ponte, 2019). However, with the expansion of the concept of "corporate social responsibility," the growing emphasis on the "triple bottom line" (economic, social, and environmental sustainability), and the emphasis placed on accomplishing the UN SDGs, sustainability is now inseparable from ethics (Rendtorff, 2019). A significant and growing body of work in AI4EO purports to support the monitoring and, therefore, the (eventual) accomplishment of various UN SDGs. "Sustainability," therefore, is a major ethical opportunity for the AI4EO (research) community. AI4EO research often aims to monitor various environmental conditions, including climate change (Parselia et al., 2019; Guo, Zhang & Zhu, 2015), sea levels (Guo, Zhang & Zhu, 2015), air and water pollution (Hadjimitsis, Retalis & Clayton, 2002), water quality (Singh et al., 2017), shifts in glaciers and polar ice (Peltier, 2004; Pellikka & Rees, 2009), forests/green cover (Martin et al., 1998), biodiversity (Turner et al., 2003; Turner et al., 2015), drought prediction (Wang et al., 2018), and agricultural/crop type classification (Ndikumana et al., 2018). These diverse uses are also closely linked to the subject matter covered by various UN SDGs, particularly, SDGs 6, 13, 14, and 15. Under SDG 15, Landsat and Sentinel-2 data have been used to predict SDG indicators under SDG 15.1.1: Forest Area as Proportion of Total Land Area, SDG 15.3.1: Proportion of Land That Is Degraded Over Total Land Area, and SDG 2.4.1:

Proportion of Agricultural Area Under Productive and Sustainable Agriculture (Kussul et al., 2019). Together with other datasets, such as Open Street Maps (Kussul et al., 2019), scientists predict that it would be possible to generate training datasets for use at global scales.

In the Tanzanian Morogoro urban municipality, RS data, social media data, and population data have been used to determine urban sprawl from 2011 to 2017. The study demonstrated the usefulness of such an approach to understanding how the city's expansion can impact the whole ecosystem. It has helped to interpret how pristine grasslands and forests, which provide essential ecosystem services, such as carbon sequestration and support for biodiversity, have been replaced by built-up land cover. Combined with Twitter data, this project has provided useful insights for urban planning in the region, where this kind of data was either unavailable or of insufficient quality (Shao et al., 2020). Other work has also demonstrated that EO, RS, and ML can contribute to sustainable development by monitoring land degradation as well as urban expansion, with its impact on the environment (Giuliani et al., 2020; Prakash et al., 2020; Mariathasan & Bezuidenhoudt, 2019; Zhu et al., 2019; Cochran et al., 2020). At the same time, it is essential that AI and ML models relied upon by the community do not result in compromising and sacrificing any one or more prongs of sustainability at the cost of another. For example, a focus merely on environmental sustainability must not ignore and compromise social and cultural diversity (Muhar et al., 2018) and the working/living space of local and indigenous people. The pursuit of "sustainability" can therefore lead to several ethical issues as one attempts to balance out (apparently) divergent goals (Kibert et al., 2011).

In the broader context of sustainability in relation to the impact of research endeavors on the environment, it is important that AI4EO research itself (intrinsically) works toward becoming energy efficient. For example, as accurate, verifiable, and representative (diverse) data are not only difficult to come by but also costly and energy intensive to maintain, the ethical ideal of sustainability also calls for the fundamental research community to work toward building algorithms that support "data minimization." Data minimization applies not only in the GDPR sense of the minimum accumulation and storage of personal data but also in terms of increasing the efficiency of algorithms so that they need fewer data points, emphasizing quality and diversity over sheer volume.

3.3. Ethical Opportunities in Artificial Intelligence for Earth Observation Research

In all fields of AI4EO research, opportunities to do good abound. Even a casual glace through published work provides an overview of the wide array of ethical opportunities presented by technological developments in this field. For example, recent publications suggest that AI4EO research can help save lives during natural disasters (Joyce et al., 2009); track endangered species (Nagendra et al., 2013); monitor and control the spread of harmful algae, such as cyanobacteria, in water bodies (Mishra et al., 2019); track disease progression (Taubenböck, Kraff & Wurm, 2018); support payments for ecosystem services (PES) (Scullion et al., 2011; Oberhauser, 2019); monitor climate change (Yang et al., 2013), sea levels (Mishra et al., 2019; Gens, 2010), refugee settlements (Wang, So & Smith, 2015), and the process of urbanization (Taubenböck et al., 2012); and even detect migrant ships and

vessels to assist humanitarian aid workers and prevent violations of human rights and international laws (Kanjir, 2019).

Yet, as discussed in the previous section, while pursuing these noble intentions, there is a likelihood that ethical risks and concerns can arise (Floridi et al., 2018), justifying the exercise undertaken in the previous section. However, in this article, we aim to step beyond identifying ethical concerns, issues, and dilemmas in AI4EO research. We do so by providing a novel understanding of the concept of "ethical opportunities," with the help of examples related to four dominant fields of ongoing AI4EO research. With these examples, we explain how a simple ground realities-informed restructuring of research questions and research approaches can help identify ethical opportunities and find win-win outcomes in early stages of AI4EO research, without getting entangled in a maze of ethical issues and dilemmas.

With the examples in the following, we also see how restructuring research questions can help accomplish various UN SDGs, particularly, SDG 1: No Poverty, SDG 3: Good Health and Well-Being, SDG 11: Sustainable Cities and Communities, and SDG 15: Life on Land. Finally, as the AI4EO community faces an uncertain future, where not all ethical issues and concerns are immediately clear, we consider this practical approach to grabbing ethical opportunities a means of minimizing ethical concerns.

3.3.1. Ethical Opportunities in Artificial Intelligence for Earth Observation Research: the Classification of Slum Areas

It is estimated that, roughly, there are either 860 million (United Nations Human Settlements Programme, 2013) or 1.5 billion (Moreno et al., 2012) slum dwellers worldwide. These vastly different estimates show that we do not really know the real numbers. AI4EO research attempts to reduce existing knowledge gaps by remotely identifying so-called slum areas and accurately labeling them, for example, to detect extents, morphologies, and changes over time. This, in turn, can help monitor various UN SDGs, such as SDG 1: No Poverty and SDG 11: Sustainable Cities and Communities.

However, the term slum and related labels, as discussed previously, have a clear negative connotation. People living on peripheries of areas labeled as "slums" (and those living within such areas) may object to the area being so labeled. Indeed, there are vast discrepancies in the definition of what constitutes a "slum" in various regions of the world (Nolan, 2015). It has been shown that living areas of the urban poor take on very different morphological forms across the globe (Taubenböck, Kraff & Wurm, 2018). One manifestation of poverty is so-called (morphological) slums, whose physical manifestations are described with very high building densities, small and low buildings, and complex organic structures. It is understandable that this proxy information of building structures for the localization of "slums" is legitimate, especially in many areas in the Global South. It has even been shown that in morphological slums, a large portion of the population is among the urban poor (Taubenböck et al., 2018; Wurm & Taubenböck; 2018).

As legitimate as the approach may sound in data-poor regions, it also involves several ethical risks:

1) The current approach essentially means that the proxy of morphological slum structures is also used to dub people living in such structures as slum dwellers, who may not be poor. In such areas, it is not uncommon that slum dwellers earn higher incomes over time and improve building structures and amenities but still continue to live in the same "slum."
2) Final map products can stigmatize these areas, and residents from these regions may have fewer opportunities, e.g., in jobs, and they may be treated disrespectfully by peers (e.g., other children in schools).
3) What is legitimately proved about specific building structures may be perceived much more positively by residents than by researchers located in countries of the Global North.
4) Therefore, AI4EO research may generate labels that correspond to building structures but do not do sufficient justice to the complexity of the topic of poverty and ground-level realities.
5) Labels are generated by people who often know about slums and slum structures only in theory and have neither lived in nor ever been to one. A bias in labeled data is the consequence, even with a consistent conceptual approach.
6) Algorithms are able to reliably find only very specific housing forms/structures/sizes in image data. Slums, poverty, and housing structures often have local specifics. AI/ML algorithms and models may therefore be inadequate and inappropriate to make global and large-scale predictions.
7) There are also virtually no datasets that clearly spatially locate the target domain (in this case, "slums") so that validation is often accompanied by very subjective criteria.

In the case of slum identification using RS, we are therefore faced with the dilemma of either comprehensively identifying all congested living areas as slums to monitor urbanization and its impacts on safety and standards of city living over time or not identifying (all) congested living areas as "slums" to avoid possible bias and stigmatization resulting from culturally inappropriate labeling of specific areas as "slums."

To avoid or minimize these ethical issues, more constructive labeling can be resorted to, based on the purpose with which the study is conducted; e.g., does the study aim to identify areas that are more prone to fire due to population density? Or does the study aim to identify regions where greater government investment is needed, say, for upgrading sanitation facilities, toilets, and electricity coverage? Perhaps the aim of the study is to support the expansion of green areas in economically poorer regions? More accurate use of labels can lead to affirmative action without stigmatizing populations. To do so, there is a need to rethink research questions as well as research approaches. It may also be necessary to invest funds in acquiring ground-level information and understanding of socioeconomic and cultural realities.

Let us take the example of toilets (especially public toilets) in slum areas and in rural regions. Currently, in EO/RS research, the level of development of an area is often determined by the intensity of night light emissions (Leibrand et al., 2019) and built-up structures (Kraff, Wurm & Taubenböck, 2020). Yet, in the poorest regions and slums, the quality of living conditions can, for example, be better determined based on the availability and quality of public toilets. Slum living often may not permit personal toilets within homes, and in such circumstances,

the availability of clean toilets, with adequate distance between women's and men's toilets, can be a strong indicator not only of the standard of living but also of safety, privacy, and hygiene for women. EO proxy information may not reflect the availability (or not) of toilets. Here, therefore, it is necessary to collect additional ground truth data on the availability (or not) of toilets in specific communities so as to give more concrete direction to policy makers when faced with international pressure to improve living conditions, including the safety of women, in "slum" areas and in economically marginalized areas. This would also help bring a gender-sensitive dimension to AI4EO research that is currently lacking optimal attention.

Further, slum identification research can also (better) help address real-world problems by focusing on the situation inside and outside slum areas, for example, determining whether major landfills and garbage disposal areas are located near slums, increasing the chances of spreading communicable diseases and public health crises. In this context, the story of the father of environmental justice, Prof. Robert Bullard, is both interesting and relevant. Together with his wife, Prof. Bullard sued the state of Texas, the city of Houston, and Harris County, USA, when a private company attempted to locate a landfill in the middle of a predominantly Black middle-class suburban neighborhood in northeast Houston. They conducted a study and created a map identifying the location of all landfills, incinerators, and garbage dumps in Houston. The map showed that "five out of five of the city-owned landfills, six out of the eight city-owned incinerators and three out of the four privately owned landfills in Houston were located in predominantly Black neighborhoods, from the 1930s up to 1978. Even though Blacks only made up 25% of the population, they were getting 82% of the garbage dumped on them" (Watkins, 2021).

The case, Margaret Bean et al. v. Southwestern Waste Management Corp. et al., was the first in the United States to identify the problem of environmental racism. Similarly, AI4EO research can help determine whether factories that pollute air and ground water are located near slums. Such information can, once again, uncover inequality and help improve the quality of life of people living in congested areas, without stigmatizing them and potentially depriving them of living space in case of extreme interventions by local governments.

3.3.2. Ethical Opportunities in Artificial Intelligence for Earth Observation Research: Urban Mapping and Building-Level Population Density

Closely linked to the issue of slum mapping is the issue of rapid urbanization and rural–urban migration. By 2050, around three-quarters of the world's population is expected to live in cities (United Nations, 2014). Since a nearly perfect correlation is verified between urbanization and the economic prosperity of societies (Glaeser, 2010), this development may raise positive expectations. However, current trends in global rural–urban migration pose fundamental challenges, altering the physical dimensions and configurations of cities at all scales. UN population figures, on which our current understanding of global urbanization trends is primarily based, do not provide information on the distribution, pattern, and evolution of the built environment (Zhang & Seto, 2011). With the growing rate of informal settlements, refugee camps, urban corridors, and so on, city profiles are changing rapidly, with local governments and city planners unable to keep track of the changes and their short- and long-term implications.

The immense spatial knowledge gaps we face and will continue to face will impact our societies and way of life by changing risk levels associated with natural hazards and infectious diseases and raising novel governance issues related to infrastructural planning and distributive justice. Despite ongoing efforts, global urban mapping approaches are unable to deliver the necessary spatial data and geoinformation in the geometric, thematic, and temporal resolutions needed to address these issues and adequately prepare for the future.

Data provided by satellite-based RS, especially when fused with data from social media platforms, such as Twitter and Flickr, and with census data, are being explored to assist in mapping global urbanization trends. Related to these efforts are research projects aimed at extracting individual building footprints, building functions, and building-level populations. As previously discussed in the "Categorizing and Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research" section, using AI4EO for urban mapping (including the extraction of building footprints and building-level population densities) can raise several ethical concerns:

1) *Data bias*: Compared to other geospatial applications, urban mapping suffers from a strong data bias: the training data used to train ML models, offered, for example, by Open Street Map, are rich in regions of the Global North, such as Europe and North America, where we already have enough geospatial information, and poor in the Global South, where knowledge and geospatial information gaps need to be filled to address, for example, UN SDG 11: Sustainable Cities and Communities. As a result of this data bias, 1) the quality measures derived from ML models may have representative metrics only for data-rich areas and 2) more severely, urban variables predicted in the Global South may be poor in quality, as ML models may suffer from a distributional shift—a change in the data distribution between an algorithm's training dataset and the dataset when deployed.
2) *Explainability and trustworthiness of the results:* There is no doubt that with the help of cutting-edge AI4EO research, promising results in global urban mapping have been achieved. Yet, to effectively incorporate these results into decision making, for example, the UN SDGs and EU's Green Deal, in addition to ensuring the high accuracy, explainability, and traceability and reproducibility of results, it is indispensable to quantify the uncertainty of these predictions from an AI algorithm. Despite increasing efforts, we are far from identifying appropriate technical solutions.
3) *Data protection:* AI4EO research makes use of various types of data, such as satellite images and social media data. However, improved image resolution and, in particular, the unprecedented geospatial information derived from those data raise questions under the European GDPR.
4) *Security concerns:* Urban variables are produced by AI4EO research for the public good, to support further research, and to aid constructive policy making. Since urban environments accommodate more than half of the world's population on only a small fraction of Earth's land surface, it is particularly important to draw our attention to possible security concerns resulting from making information publicly available, for example, in the context of conflicts. On the bright side, AI applied to satellite imagery

holds promise for the automated detection of war-related building destruction (Mueller et al., 2021). Yet, we cannot ignore the fact that AI4EO also affords opportunities to identify densely built/populated areas and critical infrastructures, such as power plants, hospitals, large industrial facilities, and so on, making them easy targets for terrorist attacks and large-scale targeted destruction by enemy forces during war. How such issues can be resolved remains an open question.

Efforts are ongoing to improve technical solutions to the previously mentioned issues of explainability and privacy (e.g., through privacy-preserving ML). However, the issue of data bias, which is also linked to a lack of availability of diverse datasets, requires fresh thinking. Transparent and trustworthy privacy-preserving mechanisms can facilitate legal and legitimate data sharing by individuals (data generators) as well as by entities that host and "control" large datasets/big data (such as Twitter); this, again, can improve access to diverse datasets. To overcome issues of data scarcity and diversity, it may be useful, once again, to rethink research questions. Instead of merely securing privacy after collecting data, for example, by hiding geolocations and the identities of people contributing data, can systems be designed to permit users to voluntarily contribute data and to incentivize data contribution on a regular basis (in specific formats)?

Evolving research suggests that by using AI/ML models in combination with blockchain and distributed ledger technology (DLT), we can incentivize and encourage the sharing of relevant georeferenced data by individuals and corporate entities. In addition to incentivizing data sharing, these technologies permit data contributors to retain control over their data and ensure the traceability (and transparency) of downstream data access (Kochupillai et al., 2021). Such solutions can empower data contributors (including individual contributors) while also furthering the EU's goal of fostering the development of a "data economy" through data portability. Since these technologies can be made to maintain a record of data access, security concerns can also be better managed.

Technologies such as blockchain could therefore also be researched and deployed, including from an ethical perspective, for fundamental AI4EO research. In this context, the European Space Agency has a declared research focus on the use of EO and blockchain for tracking data flows and adopting privacy-preserving methods for ML (European Space Agency, 2019). Since traceability lies at the heart of blockchain/DLT-based solutions, dedicated research in the field can also support the selective disclosure of data (e.g., data on building functions, building-level populations, and so on) to specific people while maintaining an immutable record of who accessed the data. Such solutions can then also minimize the security concerns listed previously, barring unidentified users from accessing information for illegal purposes. Therefore, we see that once we turn a research question around to specifically accomplish ethical goals and overcome ethical issues, the research approaches, including possible technical solutions to be considered, can also expand, leading to a win-win solution for all concerned.

3.3.3. Ethical Opportunities in Artificial Intelligence for Earth Observation Research: Monitoring (Illegal) Artisanal and Small-Scale Mining and the Right to a Livelihood

The UN recognizes the close relationship between land and peoples' livelihoods. This relationship is particularly obvious in the case of smallholder farming, herding, and ASM. In several developing countries, notably in West Africa countries, such as Ghana, research also reveals the close connection between smallholder farming and ASM. Globally, an estimated 40 million people are employed directly in ASM activities (Hobson, 2019), either because ASM was the way their ancestors traditionally made a living or because limited and unviable income from small-scale farming (Okoh & Hilson, 2011) forces them into ASM for an additional source of income. In current times, communities displaced by large-scale mining (LSM) activities are also seen to resort to ASM as a means of reducing their poverty (Yaaba Baah-Ennumh & Forson, 2017). ASM is also conducted by government-authorized small and medium enterprises.

The African Commission on Human and People's Lives has held that "it doesn't matter whether they (nonindigenous people) had legal titles to the land, the fact that the victims cannot derive their livelihood from what they possessed for generations means they have been deprived of the use of their property under conditions which are not permitted under Article 14." In this decided court case, the African Commission considered the traditional use of land as a more important consideration to bear in mind than the status of being "indigenous" or not (Centre on Housing Rights and Evictions, 2010). Following the rationale of this decided case, communities that have traditionally used land for ASM have the right to continue using land for this purpose unless alternative sources of livelihoods that also secure their traditional and cultural ways of life are offered and accepted. For the remaining categories of ASM activities, the right to a livelihood is significantly, if not equally, implicated if the right to conduct and continue ASM is withdrawn by government regulations and bans without offering meaningful alternatives.

On the other hand, the uncontrolled continuation of ASM can lead to the violation of other rights and obligations. For example, illegal ASM leads to deforestation, the contamination of water bodies (Mwakesi, Wahome & Ichang, 2020), and land degradation, leaving land unfit for agriculture (Yaaba Baah-Ennumh & Forson, 2017). ASM also negatively affects the health of artisans, either due to the use of mercury to extract gold from mines (Perks et al., 2018) or to the lack of equipment to secure their health and safety (Hentschel, Hruschka & Priester, 2022). Currently, although ASM supports subsistence, it does not afford opportunities for the socioeconomic growth and development of artisans. Research reveals that ASM, while contributing significantly to the national economy of various countries, keeps artisans caught in a "poverty trap" (Wilson et al., 2015).

As a result of these realities, there is considerable pressure on governments to formalize and legalize ASM activities and put adequate regulations in place to minimize the environmental and health-related harm resulting from (illegal) ASM. Further, there is an urgent need for governments to create alternative economic and development opportunities for artisans currently caught in a poverty trap, with ASM being their only source of subsistence. This is

also an urgent mandate closely linked to UN SDG 1: Zero Poverty. Undoubtedly, these issues are not unique to ASM but are also observed in the context of LSM activities (Yaaba Baah-Ennumh & Forson, 2017; Mwakesi, Wahome & Ichang, 2020). Nevertheless, the rapid expansion of the area under ASM (including during periods in which national governments imposed a ban on ASM; Nyamekye et al., 2021) and the long-term socioeconomic and environmental concerns linked to ASM (see Table 1) indicate the urgent need to find practical solutions.

Table 1. Costs and benefits of artisanal and small scale mining.

SOCIOECONOMIC, POLITICAL, AND ENVIRONMENTAL ...	
... BENEFITS OF ASM	**... COSTS AND CHALLENGES OF ASM**
It supports livelihoods of four categories of people who either fully or substantially rely on it for subsistence.	It leads to significant environmental degradation, including deforestation, contamination of water bodies, and making land unfit for agriculture.
It supports the national economy of countries where it is widely practiced.	It has a negative impact on the health of artisans.
It can be regulated to minimize environmental and health-related harm.	It keeps artisans caught in a "poverty trap."
	Implementing regulations seeking to legalize ASM under conditions that prevent irreversible environmental damage and mandate revival of mining sites, is challenging. Following the mandates of such regulations is also expensive and may be out of reach for artisans and small companies.
	"Complex issues of land tenure, social stability, mining regulation and taxation, and environmental degradation undermine the viability and sustainability of ASM as a livelihood strategy" [159].

ASM studies that use data and images from the Copernicus/Sentinel-2 mission primarily focus on change detection over time, identifying regions of high and low mine density (Ahmed, Bryant & Edwards, 2020) and seeking to improve the accuracy of images by using ML algorithms that reduce the problems caused by cloud cover and other irregularities in satellite-transmitted images (Nyamekye et al., 2021; Ngom et al., 2020). These studies have been able to monitor the increase in area under ASM with increasing accuracy (Nyamekye et al., 2021) and are aimed at supporting regulatory endeavors that seek to check the unauthorized expansion of illegal ASM.

Yet, these studies have themselves highlighted the inefficacy of governmental bans and the rapid expansion of areas affected by ASM, despite improved remote monitoring and the increasing number of regulatory efforts. This suggests that a major shift is needed in the overall political and regulatory approach to the problem of illegal ASM activities. While RS/EO data have helped track only the expanding scope of the problem, changes in the overall approach taken by AI4EO research can support and optimally direct governmental as well as R&D efforts in the direction of 1) reviving areas that were previously mined and 2) training and employing artisans in alternative, legal, and economically viable activities associated with the revival of ASM sites.

RS/EO and AI4EO research can support the creation of such alternative opportunities if the current approach to tracking and monitoring illegal mining activities is modified. First, as mentioned in the context of human agency and oversight, AI4EO research linked to the monitoring of ASM activities needs to take into account socioeconomic and political "ground realities." A literature review can, for example, provide significant insights (such as the ones discussed in this section) and support the reframing of research questions and approaches vis-à-vis the identification and monitoring of ASM and its expansion.

More specifically, a multidisciplinary and sociopolitical literature review suggests that there is a need to migrate away from the current focus on identifying illegal sites and changes in the extent of illegal mining over time. An alternative approach could, for example, be based on using AI4EO to predict the depth of various mining sites. This information may then be used by governmental entities to calculate the comparative efficiency of various mining sites, on the one hand, and work with environmentalists to find quick means of recuperating mining sites that are not too deep, on the other hand.

Artisanal and small-scale miners are usually able to dig only shallow mines, the environmental impacts of which are, potentially, easier to reverse if acted on immediately after the mining is completed. Indeed, artisans can themselves be trained to mine in ways that minimize environmental damage and to take steps to recoup mined areas before moving on to the next site. Artisanal miners who comply with such guidelines can then be given licenses to mine rather than having their source of livelihood removed. Legalized/licensed ASM sites can then also be converted to tourist attractions, further supporting economic development.

To support this effort, the RS/EO community can also fuse diverse sources of data toward the following long- and short-term ends:

1) *In the short term*: Identify regions that are most likely to suffer long-lasting and significant environmental damage due to ASM and LSM. Such regions should be marked as "highly sensitive" so as to urge local governments to immediately and effectively secure these areas to prevent any and all ASM/LSM activities. Such regions can, for example, be regions that are close to major water bodies, wells, and agricultural lands that are significant sources of food.
2) *In the short term*: Together with ground truth data and spatial depth analysis of EO/RS data on ASM and LSM activities, recommendations can be made to local governments and international support agencies to invest, with priority, in recapturing/reviving lands that are least affected by mining. For example, as discussed, since artisanal mining sites are most likely to be shallower and narrower, it may be easier to revive lands impacted by ASM than by LSM, which uses wider and deeper excavations. Methods to identify the depth and environmental impact of various mining sites may need to be developed with environmental scientists who are familiar with means of calculating land damage due to ASM and the most cost-effective ways of reviving such lands.
3) *In the long term*: Affirmative action initiatives, such as education, as well as concrete monetary incentives need to be given by local governments (supported by international funding agencies) to ensure that artisans have the skills, tools, and economic incentives to support efforts aimed at reviving lands affected by ASM. The environmental impacts of such initiatives and incentives can then be cost-effectively monitored by EO/RS scientists to identify the most successful ones, thereby creating a repository of best practices for the revival of lands affected by ASM as well as means of gainfully employing artisans in jobs that are both sustainable and profitable in the long run.

3.3.4. Ethical Opportunities in Artificial Intelligence for Earth Observation Research: Monitoring and Protecting Biodiversity

One of the areas of ongoing AI4EO work involves monitoring and assessing forest cover, forest density, biodiversity, and changes therein over time. Such efforts can alert governments to the areas where urgent governmental interventions are necessary to revive green cover and protect biodiversity. Ethical concerns and risks, however, can emerge in this sphere of research. In the German state of Bavaria, for example, there has been a long-standing political debate about establishing a third national park under the Bavarian nature conservation policy. The supporters of the idea argue that such a large-scale protected area would "give the diverse animal and plant world a habitat where it can develop and adapt to the climate challenges" (Dunkel, 2021). While recent polls suggest that the vast majority of the German population supports the establishment of such a park, there is a minority voice that has historically disagreed. The reasons, though not widely reported, include the understanding that ideas to establish new parks are merely "an ecological relic from the last millennium." More important from an ethical perspective was the opinion of forest owners that their livelihoods would be threatened.

Similar to the ethical dilemma thrown up by the monitoring and identification of illegal mining sites, therefore, efforts to enrich and revive biodiversity by establishing national parks and, more generally, restricting human activity can result in major human rights violations of minority groups that historically reside in and derive their livelihoods from such areas. The voices of these minorities can also get crowded out (or outnumbered) in large-scale surveys that seek to identify the "dominant," or popular, voice. In such a scenario, once again, the question to be addressed is not (only) how to mitigate the ethical risks but, rather, how to maximize the ethical opportunity, with the aim of finding a win-win outcome. In this case study, we describe two approaches that are emerging from ongoing research.

Creating Naturalness Indices

Various UN SDGs, such as SDG 13: Climate Action and SDG 15: Life on Land, envisage the establishment and designation of certain areas as nature parks and wilderness so as to provide a sanctuary to wild plants and animals (i.e., biodiversity). Such a designation thus essentially entails limiting as well as eliminating human activities within such areas. In his excellent book on wilderness protection in Europe, Bastmeijer (2016) discusses the genesis of the term wilderness as well as the various definitions that exist. While there is no clear mathematical delineation of what a wilderness area is, it is reasonable to use an official governmental definition, such as the one held by the European Commission, which itself builds upon the definition of protected areas used by the International Union for Conservation of Nature (European Commission, 2013): "A wilderness is an area governed by natural processes. It is composed of native habitats and species, and large enough for the effective ecological functioning of natural processes. It is unmodified or only slightly modified and without intrusive or extractive human activity, settlements, infrastructure or visual disturbance."

This definition assigns four central ecological aspects to the concept of wilderness:

1) naturalness

2) undisturbedness
3) undevelopedness
4) scale.

While scale can easily be defined in a mathematical sense, e.g., by defining so-called minimum mapping units, naturalness refers to ecosystems functioning in a natural way and can thus hardly be measured in a technical sense. It is, in fact, subject to judgment by ecology experts based on in situ observations. However, undisturbedness and undevelopedness are concepts that can potentially be well observed from afar, namely, by RS technologies. As per the European Commission, undevelopedness refers to the absence of "habitation, settlements or other human artifacts such as power lines, roads, railways, fences [that] may hinder ecological processes directly or by promoting likelihood of human interference" (European Commission, 2013).

Thus far, Earth's wilderness extent has largely been analyzed by (semi)automatic approaches that rely on the rule-based fusion of existing geospatial data sources. Most of those approaches go back to the work of Sanderson et al. (Sanderson et al., 2002), who proposed to measure the "human footprint" by what they called the Human Influence Index (HII) to map how much human civilization has influenced the land surface. To calculate the HII, they relied on four types of geospatial input data:

1) population density, expressed by population density grids
2) land transformation, expressed by land use maps
3) accessibility, expressed by the proximity to traffic ways
4) electrical power infrastructure, expressed by nighttime lights.

By rescaling all the globally available input data to a resolution of 1 × 1 km and scaling each input source's contribution with a score between zero (for the absence of human influence) and 10 (for high human influence), they created a map of the global human footprint. Some years later, Venter et al. (Venter et al., 2016) used this approach to investigate the change of human influence between 1993 and 2009. In their study, they found that during this period, the human footprint had increased by 9%, with 75% of Earth's land surface experiencing measurable human pressure. In a similar manner, Allan et al. (2017) published temporally comparable maps of terrestrial wildernesses. Since they defined wilderness areas as pressure-free lands with a contiguous area of more than 10,000 km2, they found Europe to be void of any truly wild areas. Ekim et al. (2021) presented the latest adaptation of the Sanderson paradigm to enable a finer-grained calculation of the degree of naturalness at a spatial resolution of 10 m per pixel in the resulting map dataset. They claimed that such maps can support decision making with respect to government-driven conservation efforts on a regional scale.

AI4EO has potential to replace the existing rule-based data fusion approaches by deep learning-based approaches to map wilderness from space, i.e., by using RS satellite images. Given recent trends in explainable AI, this can be extended by adding interpretability and explainability to the deep learning model to gain a better understanding of what makes wildernesses wild. By answering the question "Why would an AI algorithm map certain areas on Earth as wilderness and others not?" AI4EO will not only advance the automation

of RS but also deepen our understanding of what makes nature "wild" and worthy of protection.

However, while there is obviously great opportunity in AI4EO, there are also certain risks involved. The first risk lies, of course, in the technical implementation of deep learning-based AI4EO approaches: supervised deep neural networks are always as good only as the training data they are fed with. If an approach designed to automatically map wilderness areas from satellite imagery is provided only with examples of currently protected landscapes, it will probably be biased toward biomes that can mainly be found in the Western world, e.g., the United States and Europe, as those regions have a larger share of designated conservation areas than other parts in the world (see Protected Planet, 2022). It is thus of crucial importance that extra care is taken to design a model (and its training data) in a way that mitigates the bias risk and allows the identification of natural and wild areas across the whole globe.

At a more human level, as mentioned, there is a major ethical risk that by making the determination of potential conservation areas highly automatic on a purely mathematical basis, there is a chance that local stakeholders will be overlooked. Especially in densely populated countries, banning land exploitation by business, traffic, agriculture, and forestry for the sake of environmental protection does not go without opposition from regional communities that do not want to lose their land use rights. Often, the argument is that there is no real nature left anyway and that prohibiting development is an unjustified expropriation.

Creating Incentives for Biodiversity Conservation

For several decades now, academic research has been preoccupied with the issue of creating appropriate incentives for "correct" behavior. While Hardin's "Tragedy of the Commons" recommended creating property rights over "commons" to prevent their otherwise certain ruin (Hardin, 1968), Ronald Coase recommended a (legal) distribution of these property rights in a manner that would ensure that people use their free bargaining power to reach the most "efficient" conclusion (Coase, 1960). Order ethics, a subdiscipline of ethics, has also evolved to emphasize that moralizing alone is not adequate. Rather, institutional structures and incentive mechanisms need to designed in such a way that ethical (including sustainable) behavior becomes the preferred mode of conduct (Lütge & Muerji, 2016).

There is also a stream of social science and economics research that is looking into the merits of PES (Fripp, 2014). Building on this line of thought, recent work recommends looking at blockchain/DLT to incentivize and reward activities that promote (agro)biodiversity conservation (Kochupillai et al., 2021). Academic research has also demonstrated the possibility of incentivizing biodiversity conservation by using a combination of RS and blockchain-based automated PES in Namibia (Oberhauser, 2019). Here, the habitat integrity of an elephant corridor was assessed by RS algorithms, and a system was demonstrated that could trigger fictitious blockchain smart contract payments to surrounding communities. The surrounding communities could therefore potentially be incentivized to contribute to

biodiversity management and maintenance, without being dislocated from their ancestral homes and even supporting their livelihoods through PES mechanisms.

These research trends are promising in that they point to the possibility of creating a true win-win situation, where human agencies do not need to be eliminated from the equation if conservation goals are to be accomplished. Instead, as seen in the example of mines and PES with blockchain and RS, human agencies can be recruited to repair damage and maintain existing biodiversity.

3.4. Engaging with Ethics in Artificial Intelligence for Earth Observation Research Projects

Once we understand the basic concepts of ethics (the "Understanding Ethics: Issues, Risks, Theories, and Opportunities" section), the major ethical duties and ethical issues related thereto (the "Categorizing and Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research" section), and have an understanding of how reframing research questions can help avoid ethical dilemmas and grab ethical opportunities (the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section), we can use the step-by-step process in Fig. 1 and in the following to identify and address ethical issues in any ongoing AI4EO research:

1) *Step 1:* Identify and list your ethical duties as a researcher. Here, it is useful to study not only the general duties of a scientist (such as avoiding plagiarism, obtaining informed consent from data contributors/other stakeholders, and honestly citating references) but also to take the time to fully familiarize yourself with good scientific practices linked to your research field (e.g., good scientific practices in research using AI/ML and deep learning). The six major categories listed in the "Categorizing and Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research" section also provide an overview of major ethical duties, and a close reading of that section will help you understand the broad scope of these duties in the context of AI4EO research.
2) *Step 2:* Once you have a good understanding of your ethical duties, the next step is to align your research goals and steps with your ethical duties. For example, if you notice that your training dataset and test dataset overlap, steps need to be taken to modify the test dataset. A failure to align your research steps and goals with your ethical duties can lead to unethical consequences and the rejection of your research.
3) *Step 3:* Once your research steps and goals are aligned with your ethical duties, it is important to carefully consider any unethical consequences that can nevertheless result from your research methods and findings in the short or long run. Listing all potential unethical consequences of your research method and findings can help you better understand what changes, if any, need to made to your overall research goal and approach. Again, a close study of the "Understanding Ethics: Issues, Risks, Theories, and Opportunities," "Categorizing and Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research," and "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" sections of this article will help you recognize possible unethical consequences of your research.

4) *Step 4:* If any unethical consequences are found to (potentially) result from your research method and findings, the next step is to reframe your research questions in a manner that either avoids the unethical consequences or, preferably, shifts your research toward the accomplishment of an ethical opportunity. Examples in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section will help you reframe your research questions. It may also be necessary to consult or engage an ethics expert in this exercise and engage in multidisciplinary and empirical research to better understand the ground realities in the region of your study. Once the research questions are appropriately reframed, the research steps need to be aligned and modified to answer/address the new research questions.

It is necessary to note that all the steps in Fig. 1 may need to be revisited from time to time during your research life cycle.

4. Conclusions and Recommendations

This article researched, compiled, and explained the most prominent ethical principles, guidelines, and issues that are relevant for current and ongoing AI4EO research. It also provided a step-by-step approach for identifying ethical issues in ongoing AI4EO research to avoid ethical dilemmas and unethical consequences while embracing ethical opportunities. As it becomes increasingly necessary for (AI4EO) scientists and researchers to familiarize themselves with the most dominant ethical issues and ethical opportunities in their field of research, this article is meant to be an impetus to think about ethics and how it applies to everyday AI4EO research in a more systematic way.

4.1. Summary of Major Points and Their Significance

Based on a comprehensive review of the current literature on AI ethics and from discussions with scientists engaged in AI4EO research over a period of 18 months (June 2020–December 2021), we found the following:

1) Ethics is a major and growing concern for AI4EO research. This is not only because of increasing image resolutions but also because of the fusion of novel data sources and an increasing number of real-world applications (products and services) relying on AI4EO research outputs that affect the sociopolitical lives of real people. Accordingly, AI4EO scientists need to become more aware of ethical principles and guidelines and how these can apply to their everyday work.
2) A set of six broad principles that correspond with the major ethical duties of researchers were identified: honesty, integrity, fairness, responsibility, sustainability, and protecting privacy (where relevant). These principles formed the six categories under which existing ethical guidelines and issues relevant to AI4EO research were categorized.
3) Compliance with these ethical duties is likely to increase the possibility of ethical consequences, and noncompliance is likely to lead to ethical issues, i.e., unethical consequences or a combination of ethical and unethical consequences that correspond with the topics listed under the six broad categories (see the "Categorizing and

Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research" section).

4) Once identified, it is possible to overcome ethical issues and dilemmas in AI4EO research by reframing research questions and approaches.
5) However, AI4EO researchers are often unsure about how to translate, understand, and implement these broad ethical principles and guidelines in the context of their everyday research. Further, no existing work provides concrete examples and explanations of how to translate, understand, and implement existing AI ethics guidelines in the context of AI4EO research.
6) This article makes a major contribution to the literature by being the first to provide concrete examples from EO and RS research to help scientists understand and apply ethics guidelines in their ongoing and planned AI4EO research.
7) This article is also the first to provide a step-by-step iterative process to understand and apply ethics in AI4EO research. This approach can also be used by researchers engaged with research in other emerging technologies.
8) Modifying research questions to overcome ethical issues and optimize the ethical opportunities from AI4EO research can help support the accomplishment of various UN SDGs. With the help of our examples in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section, we aimed to illustrate how researchers engaged in fundamental and engineering research need to understand ethics not just in terms of passively avoiding and sidestepping the most prominent "ethical concerns" but by proactively aspiring to seize "ethical opportunities" in terms of understanding and helping resolve current real-world problems that give rise to ethical concerns in the first place. In this new pragmatic approach to ethical opportunities, we envisage the expansion of AI4EO research approaches and methods to new horizons that go beyond "observation" to support the ground-level implementation of means to accomplish various UN SDGs and solve some of the most urgent problems facing the world today.

4.2. Limitations of the Study

Despite our efforts to provide a comprehensive overview of emerging and current ethical principles, guidelines, and issues relevant to AI4EO research, a limitation of the article results from the nature of ethics, both as a discipline and a human endeavor. As stated in the "Understanding Ethics: Issues, Risks, Theories, and Opportunities" and "Categorizing and Understanding Artificial Intelligence and Data Ethics for Artificial Intelligence for Earth Observation Research" sections, ethical issues and ethical opportunities are difficult to comprehensively define and are therefore difficult to comprehensively list. Ethical issues and opportunities linked to any research are limited only by human imagination, and as technology evolves, ethical issues and opportunities linked to technology will also inevitably evolve. It is therefore important for AI4EO scientists to develop an interest in and intuition for ethics in their own superspecialized subfields of research. Recommendations in this regard are provided in the following section.

Further, we acknowledge that given the cutting-edge nature of AI4EO research, it may not be possible for even the most diligent researcher to be fully aware of his/her ethical duties and possible (negative) ethical consequences of his/her research. Here, we are faced with a limitation of Western approaches to ethics that focus on duty- and/or consequence-based theories and approaches to ethics. In subsequent research, therefore, it may be useful to dive more deeply into multicultural approaches (including, for example, Eastern approaches that look into knowledge, intention, and action as a means of realizing ethics). Such approaches can complement dominant Western approaches to support the delivery of novel and more comprehensive approaches to identifying ethical issues and opportunities in AI4EO research. Some recommendations in this direction are provided in the following section.

4.3. Recommendations and Future Research Directions

In addition to compiling a list of ethical issues and explaining them with examples, during the course of our multidisciplinary interactions over the past 18 months, we collectively made several observations that can help the AI4EO scientific community to conduct more ethically mindful research. We compiled these observations here and hope that they will also propel further research at the interface of ethics and AI4EO:

1) Several ethical issues in the sphere of AI4EO research remain unknown even to highly experienced academic researchers. As a result, general (AI/data) ethics guidelines may not be adequate to ensure ethically mindful research, especially in emerging technologies, such as AI4EO, where a lot of research is either fundamental and "application agnostic" and several ethical issues/concerns are yet to emerge and be adequately understood.
2) Accordingly, existing (AI) ethics guidelines need to be significantly modified in their approach and content so as to become practically usable by AI4EO researchers in every stage of their research. This is also necessary to ensure that ethical issues are identified, addressed, and (at least) flagged at a very early stage of research (rather than at the stage of technology deployment for real-world use, at which time it may be either too late or too expensive to make major technological changes).
3) Ethics guidelines, including AI ethics guidelines, are context specific. Therefore, for these guidelines to become practically useful and "accessible" to AI4EO scientists, there is a need to recategorize and explain the most prominent ethics guidelines, using examples that such scientists can relate to. With this article, we made a first attempt in this direction.
4) However, there is also a need for more joint research endeavors among AI4EO scientists and ethics and social science researchers so that evolving ethical issues can be identified in early stages of research and nipped at the bud before issues assume unmanageable proportions. This is particularly important in light of ambitious projects, such as the Digital Twin Earth, which is not meant to be a mere research tool but a tool to support concrete policy making and governance.
5) To encourage such collaborative and multidisciplinary research endeavors, it is important that various institutions and institutional approaches, such as those adopted by academic journals, funding agencies, and institutions supporting research, give more

importance, credit, and merit to research conducted with careful attention to ethical issues and ethical opportunities than is currently the practice. This is necessary to create a culture of ethically mindful research. For this, institutional practices, such as criteria for accepting papers for publication by high-impact journals, criteria for securing tenure, and criteria for securing funding and extensions of funding, need to be reviewed to include a significant emphasis on methods and approaches adopted for ethically mindful research and flagging and avoiding pitfalls during the course of the research.

6) If ethics, ethical principles, and ethics guidelines appear (to some scientists) to be impractical, this indicates a lack of a common language that can be understood by both ethicists and AI4EO scientists and a lack of concrete effort to build communication channels and bridges. The effort to build such bridges must come from both communities. This means that ethics experts need to take the time to understand the technological goals, aspirations, and limitations of AI4EO research, and AI4EO scientists need to take the time to understand the fundamentals of ethics and rethink research questions and research goals so that they are better aligned with ground realities, human rights, and human values.
7) From a practical standpoint, to minimize ethical concerns/risks while seizing ethical opportunities, a multidisciplinary ground-level understanding of the situation is necessary. Undoubtedly, EO/RS research, by definition, looks at Earth and its various characteristics from afar. In fact, efforts are ongoing to substitute "ground truth" data (that are often difficult and expensive to access) with more easily accessible/available data (e.g., from social media and census data). However, to the extent that applications based on EO/RS are now being used to impact policy and governance, it is important to ensure that EO/RS data, reasoning, analysis, and research approaches are not divorced from ground realities. Efforts to understand "ground truth" not merely in the sense of confirming physical and landscape structures but in the sense of socioeconomic, demographic, political, and cultural realities on the ground need to be reemphasized. As seen in the instance of identifying illegal ASM sites, for example, absent an understanding of ground realities, AI4EO research will merely go around in circles, pointing out the growing number of illegal sites but not making any concrete impact on human and environmental suffering. Such research cannot have a constructive and long-lasting impact on policy and governance.
8) Further, ethics education, especially through the design of comprehensive "continuing ethics education" curriculum, needs to be worked into academic research programs, mandating the engagement of not just bachelor's and master's students but also Ph.D. candidates, postdocs, and professors at all stages of their career.
9) In addition to designing and attending "continuing education" programs, it is necessary for ethics researchers and AI4EO technological researchers to work closely together in all stages of research so that ethical issues can be predicted early enough and avoided with appropriate within-technology and external mechanisms (ethics by design). In this context, it is necessary to reemphasize that as technologies become more complex and unpredictable, research methods, analysis, and results become more uncertain. In the era of uncertainty, what is needed is not just a set of abstract high-level guidelines but a concrete set of open-ended matrices and questionnaires (tools) that AI4EO and ethics

researchers can codevelop and use at various stages of research to conduct ethically mindful research and decision making. Ethical theories and approaches from diverse cultures need to be consulted while creating such tools so that the tools cover diverse culturally appropriate means of reaching ethical ends in early stages of research. This is a subject that needs to be covered by several academic research papers.

If technological advances are pursued without adequate importance given to ethics, human values, and human rights, mankind as a whole will stand to lose. Needless to say, the widespread acceptance (or not) of any technology also depends on whether the technology itself and its impact after implementation is in line with the highest standards of ethics. We are living and working in exiting times, and the (research) directions we choose today will guide several generations of future AI4EO research and innovation.

5. Methodology

The "An Overview of Major Artificial Intelligence Ethics Issues and Associated Guidelines" section, which provides an overview of major ethical guidelines, was compiled using the following methodology. In recent years, several ethics expert groups have been constituted to compile ethics guidelines for AI and/or big data used for AI applications. For our review, we collected the guidelines common to the following: 1) "Ethics Guidelines for Trustworthy AI" (European Commission, 2018), 2) "AI4People—An Ethical Framework for a Good AI Society" (Floridi et al., 2018), and 3) "Recommendations of the Data Ethics Commission for the Federal Government's Strategy on AI" (Bundesministerium der Justiz, 2018). To ensure that relevant ethical issues and guidelines from other experts and expert groups were not skipped, we also looked at two recent publications that attempted to comprehensively consolidate the recommendations of various expert groups from across the globe, namely, Jobin et al. (2019) and Hagendorff (2020). These publications sought to identify a comprehensive list of ethics guidelines, critique the (efficacy of the) guidelines, and determine the prominent or most commonly/globally accepted guidelines (see Table 2).

Jobin et al. reviewed 84 documents enumerating ethics guidelines for AI, issued by various countries and international agencies around the globe. All guidelines contained in these documents were reviewed, and the most dominant/frequently reappearing ones were categorized under 11 major categories, namely, 1) transparency, 2) justice and fairness, 3) nonmaleficence, 4) responsibility, 5) privacy, 6) beneficence, 7) freedom and autonomy, 8) trust, 9) sustainability, 10) dignity, and 11) solidarity. Most of these categories have several ethics issues listed under them as subcategories consisting of similar and closely related issues. Hagendorff selected and compiled a list of ethics guidelines via a literature review and supplemented his list with Algorithm Watch's "AI Ethics Guidelines Global Inventory." His research revealed a list of 22 ethical issues covered by the guidelines issued by various experts and expert groups.

Table 2. The comparison and recategorization of AI and data ethics guidelines.

JOBIN ET AL., "THE GLOBAL LANDSCAPE OF AI ETHICS GUIDELINES," *NATURE MACHINE INTELLIGENCE* [10]	EU "ETHICS GUIDELINES FOR TRUSTWORTHY AI" [7]	FLORIDI ET AL., "AI4PEOPLE—AN ETHICAL FRAMEWORK FOR A GOOD AI SOCIETY," *MINDS AND MACHINES* [8]	"RECOMMENDATION OF THE DATA ETHICS COMMISSION FOR THE FEDERAL GOVERNMENT'S STRATEGY ON AI" [1]	HAGENDORFF, "THE ETHICS OF AI ETHICS: AN EVALUATION OF GUIDELINES," *MINDS AND MACHINES* [6]	NEW CATEGORIZATION FOR AI4EO
Privacy (personal and confidential information)	Privacy	Privacy	Privacy, data ownership and compensation, and controlled data access	Privacy protection	Privacy [including autonomy, (non) stigmatization, freedom and, self-determination, including national sovereignty, data ownership, and data licensing]
Freedom and autonomy (freedom, autonomy, consent, choice, self-determination, liberty, and empowerment)	Respect for human autonomy	Enabling human self-realization, not devaluing human abilities, enhancing human agency, autonomy (the power to decide or not decide), not eroding human self-determination	Strengthening the digital sovereignty of both Germany and Europe, ensuring the human-centered and value-oriented design of technology, self-determination, human-centered design	Human autonomy	
Dignity	Preserving dignity	Human dignity and respect for the rights of parties involved			
Transparency (transparency, explainability, explicability, understandability, interpretability, communication, and disclosure)	Explicability and transparency	Explicability and intelligibility	Transparency (interest oriented), explicability, and comprehensibility, with clear accounting structures	Transparency, openness, explicability and interpretability, and public awareness	Honesty (including transparency, explainability, and data veracity)
Responsibility (responsibility, accountability, liability, acting with integrity)	Technical robustness and safety, communication of capabilities and limitations of AI system to all stakeholders, and adoption of technical and nontechnical methods to ensure adherence to the requirements	Fit-for-purpose data quality	Accountability	Integrity (including technical robustness, uncertainty quantification, safety, security, and national security)	
Nonmaleficence [nonmaleficence, security, safety, harm, protection, precaution, prevention, integrity (bodily and mental), nonsubversion]	Information security (at risk-adequate level), market-related supervisory activities, guarantee of robustness and security, high level of quality and performance, and security	Safety, cybersecurity, human oversight, control, auditing, future of employment/worker rights, dual-use problem, military, and AI arms race			
Trust					
Justice and fairness [justice, fairness, consistency, inclusion, equality, equity, (non) bias, (non)discrimination, diversity, plurality, accessibility, reversibility, remedy, redress, challenge, access, and distribution]	Fairness, data governance, diversity, nondiscrimination; dissemination of research and innovation findings, and accessing inputs from the wider public	Justice	Justice and solidarity, minimizing discrimination, minimizing bias	Fairness, nondiscrimination, justice, inclusion, diversity in field of AI, certification of AI products, and cultural differences in the ethically aligned design of AI systems	Fairness [including (non)bias, nondiscrimination, diversity, sociocultural sensitivities, democratic creation of standards, data governance]

(*Continued*)

JOBIN ET AL., "THE GLOBAL LANDSCAPE OF AI ETHICS GUIDELINES," *NATURE MACHINE INTELLIGENCE* [10]	EU "ETHICS GUIDELINES FOR TRUSTWORTHY AI" [7]	FLORIDI ET AL., "AI4PEOPLE—AN ETHICAL FRAMEWORK FOR A GOOD AI SOCIETY," *MINDS AND MACHINES* [8]	"RECOMMENDATION OF THE DATA ETHICS COMMISSION FOR THE FEDERAL GOVERNMENT'S STRATEGY ON AI" [1]	HAGENDORFF, "THE ETHICS OF AI ETHICS: AN EVALUATION OF GUIDELINES," *MINDS AND MACHINES* [6]	NEW CATEGORIZATION FOR AI4EO
Responsibility (responsibility, accountability, liability, acting with integrity)	Paying attention to situations involving vulnerable groups	Human agency and oversight; accountability; mitigation of risks linked to AI's impact on the human mind; mitigation of risks linked to democracy, rule of law, and distributive justice; and prevention of harm	Accountability	Data use and data sharing for the public good	Responsibility (including human agency/oversight, accountability, protecting human rights, duty of care, and ensuring social security and cohesion)
Enhancing protection for individual freedom, self-determination and integrity, responsibility-guided consideration, and cooperation with applicable legal regimes	Legislative framework, legal status of AI systems, and science–policy link				
Beneficence (benefits, beneficence, well-being, peace, social good, and common good)	Beneficence, promoting well-being, and promoting prosperity	Fostering responsible data utilization that is compatible with the public good, and introducing risk-adapted regulation and effective oversight of algorithmic systems	Foresighted responsibility	Common good and well-being	
Solidarity (solidarity, social security, and cohesion)	Increasing societal capabilities, cultivating societal cohesion, and preserving solidarity	Safeguarding and promoting democracy and social cohesion	Democracy and compatibility with core societal values	Solidarity and social cohesion	
Sustainability [sustainability, environment (nature), and energy, resources (energy)]	Environmental and societal well-being, fostering research and innovation in AI systems in line with requirements, and training new generation of AI ethics experts	Sustaining the planet	Guiding Motifs: Aligning digital strategies with sustainability goals, sustainability, sustainability of design and use of algorithmic systems, and fostering digital skills and critical reflection in the digital world	Sustainability, responsible/ intensified research funds, and education about AI and its risks	Sustainability (socioeconomic and environmental)

We initially used the 11 categories created by Jobin et al. as a starting point to tabulate the common ethics guidelines from each of the selected five sources. Thereafter, the categories created by Jobin et al. were rearranged to bring overlapping issues closer to one another and thereby reduce the number of categories. More specifically, the categories of privacy, freedom and autonomy, and dignity were considered closely associated and clubbed together into one category, privacy. Similarly, the issues of nonmaleficence (including safety and security), trust, and acting with integrity were considered closely linked, especially when we look at the term integrity in the technical context of "system integrity." Accordingly, these were clubbed into one (new/stand-alone) category, integrity. In this context, it is necessary to highlight that we split up Jobin et al.'s responsibility category, which contained "acting with integrity" as a subissue, into two categories, namely, responsibility and integrity. The responsibility category was then clubbed with Jobin et al.'s categories of solidarity and beneficence, as all three are closely linked with the larger human duty to take the welfare of others and society at large into account.

With this method, the ethics issue sets emerging from our five sources were regrouped under six simple categories that we expect AI4EO scientists to relate to and understand more intuitively. The first new category we created is that of privacy, not least because this is the

most commonly discussed ethical issue in AI4EO as image resolutions increase, new data sources are fused, and novel use cases of AI4EO appear on the horizon. Since privacy is closely linked to the issues of autonomy, freedom, and self-determination, these ethical issues were categorized as subissues of privacy. Further, based on discussions with various AI4EO scientists, especially scientists working with slum identification through AI4EO, we identified a new subcategory under privacy, namely, (non)stigmatization. This is an ethical issue that has not been prominently discussed in ethics guidelines. Further, given the scale of AI4EO research and its goal of monitoring compliance with SDGs, we added a new country-level dimension to the issue of self-determination, namely, national sovereignty. Indeed, under international law, the right to self-determination (United Nations, 1966) is a right given not just to individuals but to groups of individuals who view themselves as sharing a common history, culture, values, and heritage (United Nations, 1966). Later, issues linked to data governance, data ownership and licensing were also added under privacy (although they can also be categorized elsewhere, e.g. under "Fairness") because of the close linkage between issues of data privacy and data governance/ownership.

Thereafter, we used the three fundamental values of business ethics, namely, honesty, integrity, and fairness, as the next three categories of the various ethics guidelines distilled from our five sources. Here, as explained previously, we considered it necessary to add a new stand-alone category—integrity—because of its close link to the issue of technical robustness and "system integrity," both of which are relevant for AI4EO applications. Again, because of the scale of AI4EO research, the fairness category included not just (non)bias in relation to individuals but also (non)bias with regard to specific cultures that can be ensured only if AI/ML models and the data used to train them are compiled with an eye toward sociocultural sensitivities.

The final two categories, namely, responsibility and sustainability, were then compiled by combining a few categories of Jobin et al.'s classification (as described previously). The issues linked to the education of new generations of AI scientists, which showed up in some of the guidelines, were placed under the subissue of sociocultural sustainability. The resulting new categorization of ethical issues for AI4EO is summarized in the "An Overview of Major Artificial Intelligence Ethics Issues and Associated Guidelines" section. The examples used to explain the ethical issues in the "Understanding Ethical Issues in the Context of Artificial Intelligence for Earth Observation Research" section derive from numerous discussions with AI4EO researchers between June 2020 and December 2021 (and later, also in April-May 2022). Since significant contributions came from a few scientists who contributed to the compilation of the four case studies in the "Ethical Opportunities in Artificial Intelligence for Earth Observation Research" section, they were invited to coauthor the present article and further enrich the examples in the "Understanding Ethical Issues in the Context of Artificial Intelligence for Earth Observation Research" section.

Acknowledgement

This work is supported by the German Federal Ministry of Education and Research, in the framework of the international future AI lab "Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics, and Beyond" (grant 01DD20001). We would also like to

acknowledge the wonderful research assistance provided by Yasmin Al-Douri, master's student, Technical University of Munich. Mrinalini Kochupillai is the corresponding author of this article.

References

A

Abascal, A., Georganos, S., Kuffer, M., Vanhuysse, S., Thomson, D., Wang, J., Manyasi, L., Otunga, D. M., Ochieng, B., Ochieng, T., Klinnert, J., & Wolff, E. (2024). Making Urban Slum Population Visible: Citizens and Satellites to Reinforce Slum Censuses. In M. Kuffer & S. Georganos (Eds.), Urban Inequalities from Space: Earth Observation Applications in the Majority World (pp. 287–302). Springer International Publishing. https://doi.org/10.1007/978-3-031-49183-2_14

Abed, J., & Kaysi, I. (2003). Identifying urban boundaries: Application of remote sensing and geographic information system technologies. In Canadian Journal of Civil Engineering (Vol. 30, Issue 6, pp. 992–999). https://doi.org/10.1139/l03-051

Abidin, H. Z., Andreas, H., Gumilar, I., Fukuda, Y., Pohan, Y. E., & Deguchi, T. (2011). Land subsidence of Jakarta (Indonesia) and its relation with urban development. Natural Hazards, 59(3), 1753–1771. https://doi.org/10.1007/s11069-011-9866-9

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49, 424–438. https://doi.org/10.1016/j.ijinfomgt.2019.07.008

Abu El-Magd, S. A., Ali, S. A., & Pham, Q. B. (2021). Spatial modeling and susceptibility zonation of landslides using random forest, naïve bayes and K-nearest neighbor in a complicated terrain. Earth Science Informatics, 14(3), 1227–1243. https://doi.org/10.1007/s12145-021-00653-y

Acioly, C., & Davidson, F. (1996). Density in urban development. Building Issues, 8, 3–25.

Acuto, M., Parnell, S., & Seto, K. C. (2018). Building a global urban science. Nature Sustainability, 1(1), 2–4. https://doi.org/10.1038/s41893-017-0013-9

Acuto, M., Pejic, D., & Briggs, J. (2021). Taking City Rankings Seriously: Engaging with Benchmarking Practices in Global Urbanism. International Journal of Urban and Regional Research, 45(2), 363–377. Scopus. https://doi.org/10.1111/1468-2427.12974

Ado, M., Amitab, K., Maji, A. K., Jasińska, E., Gono, R., Leonowicz, Z., & Jasiński, M. (2022). Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey. Remote Sensing, 14(3029), 1–48. https://doi.org/10.3390/rs14133029

Adolphson Marcus. (2009). Estimating a Polycentric Urban Structure. Case Study: Urban Changes in the Stockholm Region 1991–2004. Journal of Urban Planning and Development, 135(1), 19–30. https://doi.org/10.1061/(ASCE)0733-9488(2009)135:1(19)

AdV – Arbeitsgemeinschaft der Vermessungsverwaltungen der Lander ¨ der & Bundesrepublik Deutschland. (2018). Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok). https://www.adv-online.de/GeoInfoDok/Aktuelle-Anwendungsschemata/AAA-Anwendungsschema-7.1.2-Referenz-7.1/binarywriterservlet?imgUid=61e40780-c5f2-bc61-f27f-31c403b36c4c&uBasVariant=11111111-1111-1111-1111-111111111111

African Union Commission. (2021). Africa's development dynamics 2021 digital transformation for quality jobs. OECD Publishing. https://doi.org/10.1787/0a5c9314- en

Ahmed, A. I., Bryant, R. G., & Edwards, D. P. (2020). Where are mines located in sub Saharan Africa and how have they expanded overtime? Land Degradation & Development, 32(1), 112–122. https://doi.org/10.1002/ldr.3706

Akbar, J. (1988). Crisis in the Built Environment: The Case of the Muslim City. Concept Media Pte Ltd.

Akbari, H., Kurn, D. M., Bretz, S. E., & Hanford, J. W. (1997). Peak power and cooling energy savings of shade trees. Energy and Buildings, 25(2), 139–148. https://doi.org/10.1016/S0378-7788(96)01003-1

Akbari, H., Pomerantz, M., & Taha, H. (2001). Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. Solar Energy, 70(3), 295–310. https://doi.org/10.1016/S0038-092X(00)00089-X

Alarasi, H., Martinez, J., & Amer, S. (2016). Children's perception of their city centre: A qualitative GIS methodological investigation in a Dutch city. Children's Geographies, 14(4), 437–452. https://doi.org/10.1080/14733285.2015.1103836

Albert, A., Kaur, J., & Gonzalez, M. C. (2017). Using Convolutional Networks and Satellite Imagery to Identify Patterns in Urban Environments at a Large Scale. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1357–1366. https://doi.org/10.1145/3097983.3098070

Alberti, M., Marzluff, J. M., Shulenberger, E., Bradley, G., Ryan, C., & Zumbrunnen, C. (2003). Integrating Humans into Ecology: Opportunities and Challenges for Studying Urban Ecosystems. BioScience, 53(12), 1169–1179. https://doi.org/10.1641/0006-3568(2003)053[1169:IHIEOA]2.0.CO;2

Albrieu-Llinás, G., Espinosa, M. O., Quaglia, A., Abril, M., & Scavuzzo, C. M. (2018). Urban environmental clustering to assess the spatial dynamics of Aedes aegypti breeding sites Authors Guillermo Albrieu-Llinás. Geospatial Health, 13(1), 135–142. https://doi.org/10.4081/gh.2018.654

Alcaldía de Medellín. (2014a). Plan de ordenamiento territorial, Acuerdo 48 de 2014, Medellín.

Alcaldía de Medellín. (2014b). Revisión y ajuste del Plan de Ordenamiento Territorial de Medellín, Evaluación y Seguimiento – Tomo IIIB, Versión 3-Concertación con Autoridades Ambientales, Medellín.

Alcaldía de Medellín. (2023). GeoMedellín. https://www.medellin.gov.co/geomedellin/

Alexander, C. (1965). A city is not a tree. Architectural Forum, 122(1), 58–62.

Alexander, C. (1979). The Timeless Way of Building. Oxford University Press.

Alexander, C., Ishikawa, S., & Silverstein, M. (1977). A Pattern Language: Towns, Buildings, Construction. Oxford University Press.

Ali, S. A., Parvin, F., Vojteková, J., Costache, R., Linh, N. T. T., Pham, Q. B., Vojtek, M., Gigović, L., Ahmad, A., & Ghorbani, M. A. (2021). GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms. Geoscience Frontiers, 12(2), 857–876. https://doi.org/10.1016/j.gsf.2020.09.004

Allan, J. R., Venter, O., & Watson, J. E. M. (2017). Temporally inter-comparable maps of terrestrial wilderness and the Last of the Wild. Scientific Data, 4(1), 1–8. https://doi.org/10.1038/sdata.2017.187

Alliance for Responsible Mining. (2013). Approaching artisanal and small-scale mining through the lens of human rights: A call for international action. http://www.responsiblemines.org/wp-content/uploads/2018/05/059_Human_Rights_and_ASM_full-version.pdf

Alonso, W. (1964). Location and Land Use: Toward a General Theory of Land Rent. Harvard University Press. https://doi.org/doi:10.4159/harvard.9780674730854

Alonzo, M., Bookhagen, B., & Roberts, D. A. (2014). Urban tree species mapping using hyperspectral and lidar data fusion. Remote Sensing of Environment, 148, 70–83. https://doi.org/10.1016/j.rse.2014.03.018

Alsaad, A., Taamneh, A., & Al-Jedaiah, M. N. (2018). Does social media increase racist behavior? An examination of confirmation bias theory. Technology in Society, 55, 41–46. https://doi.org/10.1016/j.techsoc.2018.06.002

Altham, J. E. J. (1983). Ethics of Risk. Proceedings of the Aristotelian Society, 84(1), 15–30. https://doi.org/10.1093/aristotelian/84.1.15

Altieri, L., & Cocchi, D. (2022). Understanding the expansion of Italian metropolitan areas: A study based on entropy measures. Environment and Planning B: Urban Analytics and City Science, 49(2), 447–463. https://doi.org/10.1177/23998083211012699

Altieri, L., Cocchi, D., & Roli, G. (2021). Spatial entropy for biodiversity and environmental data: The R-package SpatEntropy. Environmental Modelling & Software, 144(105149), 1–12. https://doi.org/10.1016/j.envsoft.2021.105149

Amaral, S., Câmara, G., Monteiro, A. M. V., Quintanilha, J. A., & Elvidge, C. D. (2005). Estimating population and energy consumption in Brazilian Amazonia using DMSP night-time satellite data. Computers, Environment and Urban Systems, 29(2), 179–195. https://doi.org/10.1016/j.compenvurbsys.2003.09.004

Amiri, R., Weng, Q., Alimohammadi, A., & Alavipanah, S. K. (2009). Spatial–temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz urban area, Iran. Remote Sensing of Environment, 113(12), 2606–2617. https://doi.org/10.1016/j.rse.2009.07.021

AMVA. (2023). Datos Abiertos del Área Metropolitana Valle de Aburrá (AMVA) [Dataset]. https://datosabiertos.metropol.gov.co/

Anas, A., Arnott, R., & Small, K. A. (1998). Urban Spatial Structure. Journal of Economic Literature, 36(3), 1426–1464.

Anas, A., & Kim, I. (1996). General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration. Journal of Urban Economics, 40(2), 232–256. https://doi.org/10.1006/juec.1996.0031

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). A land use and land cover classification system for use with remote sensor data (Report No. 964; Professional Paper). USGS Publications Warehouse. https://doi.org/10.3133/pp964

Andersson, E., & Colding, J. (2014). Understanding how built urban form influences biodiversity. Urban Forestry & Urban Greening, 13(2), 221–226. https://doi.org/10.1016/j.ufug.2013.11.002

Andersson, F., Burgess, S., & Lane, J. I. (2007). Cities, matching and the productivity gains of agglomeration. Journal of Urban Economics, 61(1), 112–128. https://doi.org/10.1016/j.jue.2006.06.005

Angel, S., & Blei, A. M. (2016). The spatial structure of American cities: The great majority of workplaces are no longer in CBDs, employment sub-centers, or live-work communities. Cities, 51, 21–35. https://doi.org/10.1016/j.cities.2015.11.031

Angel, S., Lamson-Hall, P., Blei, A., Shingade, S., & Kumar, S. (2021). Densify and Expand: A Global Analysis of Recent Urban Growth. Sustainability, 13(3835), 1–28. https://doi.org/10.3390/su13073835

Angel, S., Lamson-Hall, P., Galarza, N., & Blei, A. (2018). Our Not-So-Urban World | Marron Institute of Urban Management. https://marroninstitute.nyu.edu/papers/our-not-so-urban-world

Angel, S., Parent, J., & Civco, D. L. (2010). Ten compactness properties of circles: Measuring shape in geography. Canadian Geographies / Géographies Canadiennes, 54(4), 441–461. https://doi.org/10.1111/j.1541-0064.2009.00304.x

Angel, S., Parent, J., & Civco, D. L. (2012). The fragmentation of urban landscapes: Global evidence of a key attribute of the spatial structure of cities, 1990–2000. Environment & Urbanization, 24(1), 249–283. https://doi.org/10.1177/0956247811433536

Angel, S., Parent, J., Civco, D. L., & Blei, A. M. (2011). Making room for a planet of cities (No. Code PF027; Policy Focus Report). Lincoln Institute of Land Policy. https://www.lincolninst.edu/app/uploads/legacy-files/pubfiles/making-room-for-a-planet-of-cities-full_0.pdf

Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000-2050. In Progress in Planning (Vol. 75, Issue 2, pp. 53–107). https://doi.org/10.1016/j.progress.2011.04.001

Angel, S., Sheppard, S. C., Civco, D. L., Buckley, R., Chabaeva, A., Gitlin, L., Kraley, A., Parent, J., & Perlin, M. (2005). The Dynamics of Global Urban Expansion. Department of Transport and Urban Development, The World Bank.

Anguelovski, I., Connolly, J. J. T., Cole, H., Garcia-Lamarca, M., Triguero-Mas, M., Baró, F., Martin, N., Conesa, D., Shokry, G., del Pulgar, C. P., Ramos, L. A., Matheney, A., Gallez, E., Oscilowicz, E., Máñez, J. L., Sarzo, B., Beltrán, M. A., & Minaya, J. M. (2022). Green gentrification in European and North American cities. Nature Communications, 13/1(3816), 1–13. https://doi.org/10.1038/s41467-022-31572-1

Anjos, C., Gonçalves Lacerda, M., do Livramento Andrade, L., & Neves Salles, R. (2017). Classification of urban environments using feature extraction and random forest. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1205–1208. https://doi.org/10.1109/IGARSS.2017.8127174

Ankerst, M., Breunig, M. M., Kriegel, H.-P., & Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. SIGMOD Rec., 28(2), 49–60. https://doi.org/10.1145/304181.304187

Anselin, L., & Bera, A. K. (1998). Spatial Dependence in linear Regression Models with an Introduction to Spatial Econometrics. In Handbook of Applied Economic Statistics (Vol. 155, pp. 237–290). CRC Press.

Anselin, L., & Rey, S. J. (2012). Spatial econometrics in an age of CyberGIScience. International Journal of Geographical Information Science, 26(12), 2211–2226. https://doi.org/10.1080/13658816.2012.664276

Antal, A., Guerreiro, P. M. P., & Cheval, S. (2021). Comparison of spatial interpolation methods for estimating the precipitation distribution in Portugal. Theoretical and Applied Climatology, 145(3), 1193–1206. https://doi.org/10.1007/s00704-021-03675-0

Araujo, R. V., Albertini, M. R., Costa-da-Silva, A. L., Suesdek, L., Franceschi, N. C. S., Bastos, N. M., Katz, G., Cardoso, V. A., Castro, B. C., Capurro, M. L., & Allegro, V. L. A. C. (2015). São Paulo urban heat islands have a higher incidence of dengue than other urban areas. The Brazilian Journal of Infectious Diseases, 19(2), 146–155. https://doi.org/10.1016/j.bjid.2014.10.004

Arauzo-Carod, J.-M., & Viladecans-Marsal, E. (2009). Industrial Location at the Intra-Metropolitan Level: The Role of Agglomeration Economies. Regional Studies, 43(4), 545–558. https://doi.org/10.1080/00343400701874172

Aravena Pelizari, P., Geiß, C., Aguirre, P., Santa María, H., Merino Peña, Y., & Taubenböck, H. (2021). Automated building characterization for seismic risk assessment using street-level imagery and deep learning. ISPRS Journal of Photogrammetry and Remote Sensing, 180, 370–386. https://doi.org/10.1016/j.isprsjprs.2021.07.004

Aravena Pelizari, P., Spröhnle, K., Geiß, C., Schoepfer, E., Plank, S., & Taubenböck, H. (2018). Multi-sensor feature fusion for very high spatial resolution built-up area extraction in temporary settlements. Remote Sensing of Environment, 209, 793–807. https://doi.org/10.1016/j.rse.2018.02.025

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2015). Constructing cities, deconstructing scaling laws. Journal of The Royal Society Interface, 12(102), 1–8. https://doi.org/10.1098/rsif.2014.0745

Arellano, B., & Roca, J. (2016). Landscapes impacted by light. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B8, 813–820. https://doi.org/10.5194/isprs-archives-XLI-B8-813-2016

Arino, O., Bicheron, P., Achard, F., Latham, J., Witt, R., & Weber, J.-L. (2008). GlobCover: The most detailed portrait of Earth. In European Space Agency Bulletin (Vol. 2008, Issue 136, pp. 24–31).

Arlt, G., Hennersdorf, J., Lehmann, I., & Nguyen, X. T. (2005). Auswirkungen städtischer Nutzungsstrukturen auf Grünflächen und Grünvolumen (No. 47; IÖR Schriften). Dresden: Leibniz-Institut für ökologische Raumentwicklung e.V. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-396830

Aronson, M. F., Lepczyk, C. A., Evans, K. L., Goddard, M. A., Lerman, S. B., MacIvor, J. S., Nilon, C. H., & Vargo, T. (2017). Biodiversity in the city: Key challenges for urban green space management. Frontiers in Ecology and the Environment, 15(4), 189–196. https://doi.org/10.1002/fee.1480

Arribas-Bel, D., & Sanz-Gracia, F. (2014). The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. Urban Geography, 35(7), 980–997. https://doi.org/10.1080/02723638.2014.940693

Arribas-Bel, D., Ramos, A., & Sanz-Gracia, F. (2015). The Size Distribution of Employment Centers within the US Metropolitan Areas. Environment and Planning B: Planning and Design, 42(1), 23–39. https://doi.org/10.1068/b39038

Arthur, D., & Vassilvitskii, S. (2007). k-means++: The advantages of careful seeding. Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 1027–1035.

Artmann, M., Inostroza, L., & Fan, P. (2019). Urban sprawl, compact urban development and green cities. How much do we know, how much do we agree? From Urban Sprawl to Compact Green Cities – Indicators for Multi-Scale and Multi-Dimensional Analysis, 96, 3–9. https://doi.org/10.1016/j.ecolind.2018.10.059

Aschbacher, J., & Milagro-Pérez, M. P. (2012). The European Earth monitoring (GMES) programme: Status and perspectives. Remote Sensing of Environment, 120, 3–8. https://doi.org/10.1016/j.rse.2011.08.028

Ash, C., Jasny, B. R., Roberts, L., Stone, R., & Sugden, A. M. (2008). Reimagining Cities. Science, 319(5864), 739–739. https://doi.org/10.1126/science.319.5864.739

Asner, G. P., & Vitousek, P. M. (2005). Remote analysis of biological invasion and biogeochemical change. Proceedings of the National Academy of Sciences, 102(12), 4383–4386. https://doi.org/10.1073/pnas.0500823102

Asseng, S., Spänkuch, D., Hernandez-Ochoa, I. M., & Laporta, J. (2021). The upper temperature thresholds of life. The Lancet Planetary Health, 5(6), e378–e385. https://doi.org/10.1016/S2542-5196(21)00079-6

Aubrecht, C., Steinnocher, K., Hollaus, M., & Wagner, W. (2009). Integrating earth observation and GIScience for high resolution spatial and functional modeling of urban land use. Computers, Environment and Urban Systems, 33(1), 15–25. https://doi.org/10.1016/j.compenvurbsys.2008.09.007

Audebert, N., Le Saux, B., & Lefèvre, S. (2018). Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks. Geospatial Computer Vision, 140, 20–32. https://doi.org/10.1016/j.isprsjprs.2017.11.011

Auerbach, F. (1913). Das Gesetz der Bevölkerungskonzentration. Petermanns Geographische Mitteilungen, 59, 74–76.

Avdan, N. (2019). Visas and Walls: Border Security in the Age of Terrorism. University of Pennsylvania Press. https://doi.org/10.2307/j.ctv16t6jpj

Avila-Palencia, I., Sánchez, B. N., Rodríguez, D. A., Perez-Ferrer, C., Miranda, J. J., Gouveia, N., Bilal, U., Useche, A. F., Wilches-Mogollon, M. A., Moore, K., Sarmiento, O. L., & Diez Roux, A. V. (2022). Health and Environmental Co-Benefits of City Urban Form in Latin America: An Ecological Study. Sustainability, 14(22), 1–17. https://doi.org/10.3390/su142214715

Avila-Palencia Ione, Rodríguez Daniel A., Miranda J. Jaime, Moore Kari, Gouveia Nelson, Moran Mika R., Caiaffa Waleska T., & Diez Roux Ana V. (2022). Associations of Urban Environment Features with Hypertension and Blood Pressure across 230 Latin American Cities. Environmental Health Perspectives, 130(2)(027010), 1–10. https://doi.org/10.1289/EHP7870

Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. Scientific Reports, 11(24112), 1–16. https://doi.org/10.1038/s41598-021-03585-1

B

Baccini, P., & Brunner, P. H. (2012). Metabolism of the Anthroposphere: Analysis, Evaluation, Design. The MIT Press. https://doi.org/10.7551/mitpress/8720.001.0001

Bachtrögler-Unger, J., Dolls, M., Schüle, P., Taubenböck, H., & Weigand, M. (2021). Evaluating EU cohesion policy using satellite data (No. 61; Bertelsmann Stiftung). https://doi.org/10.11586/2021047

Baddeley, A., & Turner, R. (2005). Spatstat: An R Package for Analyzing Spatial Point Patterns. Journal of Statistical Software, 12(6), 1–42. https://doi.org/10.18637/jss.v012.i06

Bai, X., Chen, J., & Shi, P. (2012). Landscape Urbanization and Economic Growth in China: Positive Feedbacks and Sustainability Dilemmas. Environmental Science & Technology, 46(1), 132–139. https://doi.org/10.1021/es202329f

Bai, X., McPhearson, T., Cleugh, H., Nagendra, H., Tong, X., Zhu, T., & Zhu, Y.-G. (2017). Linking Urbanization and the Environment: Conceptual and Empirical Advances. In Annual Review of Environment and Resources (Vol. 42, Issue Volume 42, 2017, pp. 215–240). Annual Reviews. https://doi.org/10.1146/annurev-environ-102016-061128

Bakker, F. G. A. de, Rasche, A., & Ponte, S. (2019). Multi-Stakeholder Initiatives on Sustainability. A Cross-Disciplinary Review and Research Agenda for Business Ethics, 29(3), 343–383. https://doi.org/10.1017/beq.2019.10

Baklanov, A., Molina, L. T., & Gauss, M. (2016). Megacities, air quality and climate. Atmospheric Environment, 126, 235–249. https://doi.org/10.1016/j.atmosenv.2015.11.059

Balk, D., Leyk, S., Jones, B., Montgomery, M. R., & Clark, A. (2018). Understanding urbanization: A study of census and satellite-derived urban classes in the United States, 1990-2010. PLOS ONE, 13(12)(e0208487), 1–20. https://doi.org/10.1371/journal.pone.0208487

Ban, Y., Jacob, A., & Gamba, P. (2015). Spaceborne SAR data for global urban mapping at 30m resolution using a robust urban extractor. Global Land Cover Mapping and Monitoring, 103, 28–37. https://doi.org/10.1016/j.isprsjprs.2014.08.004

Banerjee, T. (2012). Megaregions or Megasprawls? Issues of Density, Urban Design and Quality Growth. In C. L. Ross (Ed.), Megaregions. Planning for global competitiveness (pp. 83–106). Island Press.

Banzhaf, E., & De la Barrera, F. (2017). Evaluating public green spaces for the quality of life in cities by integrating RS mapping tools and social science techniques. 2017 Joint Urban Remote Sensing Event (JURSE), 1–4. https://doi.org/10.1109/JURSE.2017.7924559

Barboza Pereira, E., Cirach, M., Khomenko, S., Iungman, T., Mueller, N., Barrera-Gómez, J., Rojas-Rueda, D., Kondo, M., & Nieuwenhuijsen, M. (2021). Green space and mortality in European cities: A health impact assessment study. The Lancet Planetary Health, 5(10), e718–e730. https://doi.org/10.1016/S2542-5196(21)00229-1

Barr, J., & Cohen, J. P. (2014). The floor area ratio gradient: New York City, 1890–2009. Regional Science and Urban Economics, 48, 110–119. https://doi.org/10.1016/j.regsciurbeco.2014.03.004

Barrie, C., & Ho, J. C. (2021). academictwitteR: an R package to access the Twitter Academic Research Product Track v2 API endpoint. Journal of Open Source Software, 6(62)(3272), 1–2. https://doi.org/10.21105/joss.03272

Bartosiewicz, B., & Marcińczak, S. (2020). Investigating polycentric urban regions: Different measures – Different results. Cities, 105(102855), 1–11. https://doi.org/10.1016/j.cities.2020.102855

Bastmeijer, K. (2016). Wilderness Protection in Europe: The role of International, European and National Law. Cambridge University Press. https://doi.org/10.1017/CBO9781107415287

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Batty, M. (1974). Spatial Entropy. Geographical Analysis, 6(1), 1–31. https://doi.org/10.1111/j.1538-4632.1974.tb01014.x

Batty, M. (2005). Cities and Complexity: Understanding Cities with Cellular Automata, AgentBased Models, and Fractals. MIT Press.

Batty, M. (2008). The Size, Scale, and Shape of Cities. Science, 319(5864), 769–771. https://doi.org/10.1126/science.1151419

Batty, M. (2009). Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies. In R. A. Meyers (Ed.), Encyclopedia of Complexity and Systems Science (pp. 1041–1071). Springer New York. https://doi.org/10.1007/978-0-387-30440-3_69

Batty, M. (2013). The New Science of Cities. The MIT Press.

Batty, M. (2016). Evolving a Plan: Design and Planning with Complexity. In J. Portugali & E. Stolk (Eds.), Complexity, Cognition, Urban Planning and Design (pp. 21–42). Springer International Publishing.

Batty, M., Besussi, E., Maat, K., & Harts, J. J. (2004). Representing Multifunctional Cities: Density and Diversity in Space and Time. Built Environment, 30(4), 324–333.

Batty, M., & Marshall, S. (2009). Centenary paper: The evolution of cities: Geddes, Abercrombie and the new physicalism. Town Planning Review, 80(6), 551–574. https://doi.org/10.3828/tpr.2009.12

Batty, M., & Marshall, S. (2012). The Origins of Complexity Theory in Cities and Planning. In J. Portugali, H. Meyer, E. Stolk, & E. Tan (Eds.), Complexity Theories of Cities Have Come of Age: An Overview with Implications to Urban Planning and Design (pp. 21–45). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-24544-2_3

Baur, A. H., Förster, M., & Kleinschmit, B. (2015). The spatial dimension of urban greenhouse gas emissions: Analyzing the influence of spatial structures and LULC patterns in European cities. Landscape Ecology, 30(7), 1195–1205. https://doi.org/10.1007/s10980-015-0169-5

Bechle, M. J., Millet, D. B., & Marshall, J. D. (2017). Does Urban Form Affect Urban NO2? Satellite-Based Evidence for More than 1200 Cities. Environmental Science & Technology, 51(21), 12707–12716. https://doi.org/10.1021/acs.est.7b01194

Bechtel, B. (2011). Multitemporal Landsat data for urban heat island assessment and classification of local climate zones. 2011 Joint Urban Remote Sensing Event, 129–132. https://doi.org/10.1109/JURSE.2011.5764736

Bechtel, B., Alexander, P. J., Beck, C., Böhner, J., Brousse, O., Ching, J., Demuzere, M., Fonte, C., Gál, T., Hidalgo, J., Hoffmann, P., Middel, A., Mills, G., Ren, C., See, L., Sismanidis, P., Verdonck, M.-L., Xu, G., & Xu, Y. (2019). Generating WUDAPT Level 0 data – Current status of production and evaluation. Urban Climate, 27, 24–45. https://doi.org/10.1016/j.uclim.2018.10.001

Bechtel, B., Alexander, P. J., Böhner, J., Ching, J., Conrad, O., Feddema, J., Mills, G., See, L., & Stewart, I. (2015). Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities. ISPRS International Journal of Geo-Information, 4(1), 199–219. https://doi.org/10.3390/ijgi4010199

Bechtel, B., & Daneke, C. (2012). Classification of local climate zones based on multiple earth observation data. In IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Vol. 5, Issue 4, pp. 1191–1202). https://doi.org/10.1109/JSTARS.2012.2189873

Bechtel, B., Demuzere, M., Sismanidis, P., Fenner, D., Brousse, O., Beck, C., Van Coillie, F., Conrad, O., Keramitsoglou, I., Middel, A., Mills, G., Niyogi, D., Otto, M., See, L., & Verdonck, M.-L. (2017). Quality of Crowdsourced Data on Urban Morphology—The Human Influence Experiment (HUMINEX). Urban Science, 1(15), 1–21. https://doi.org/10.3390/urbansci1020015

Bechtel, B., Foley, M., Mills, G., Ching, J., See, L. M., Alexander, P. J., O'Connor, M., Albuquerque, T. T. de A., Andrade, M. de F., Brovelli, M. A., Das, D., Fonte, C. C., Petit, G., Hanif, U., Jiménez, J. S. P., Lackner, S., Liu, W., Perera, N., Rosni, N. A., … Gál, T. (2015). CENSUS of Cities: LCZ Classification of Cities (Level 0) – Workflow and Initial Results from Various Cities. 1–6. https://core.ac.uk/download/pdf/42938825.pdf

Bechtel, B., Pesaresi, M., See, L., Mills, G., Ching, J., Alexander, P. J., Feddema, J. J., Florczyk, A. J., & Stewart, I. (2016). Towards consistent mapping of urban structures—Global human settlement layer and local climate zones. In International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives (Vol. 41, pp. 1371–1378). https://doi.org/10.5194/isprsarchives-XLI-B8-1371-2016

Bechtel, B., See, L., Mills, G., & Foley, M. (2016). Classification of local climate zones using SAR and multispectral data in an arid environment. In IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Vol. 9, Issue 7, pp. 3097–3105). https://doi.org/10.1109/JSTARS.2016.2531420

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific Data, 5(180214), 1–12. https://doi.org/10.1038/sdata.2018.214

Beck, U. (1997). Was ist Globalisierung? Suhrkamp Verlag.

Beerbaum, D., Puaschunder, J. M., & Puaschunder, J. M. (2019). A Behavioral Economics Approach to Digitalization: The Case of a Principles-Based Taxonomy. In Intergenerational Governance and Leadership in the Corporate World: Emerging Research and Opportunities (pp. 107–122). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-8003-4.ch006

Begg, I. (1999). Cities and Competitiveness. Urban Studies, 36(6), 795–809. Scopus. https://doi.org/10.1080/0042098993222

Behrens, K., & Robert-Nicoud, F. (2014). Survival of the Fittest in Cities: Urbanisation and Inequality. The Economic Journal, 124(581), 1371–1400. https://doi.org/10.1111/ecoj.12099

Beirle, S., Platt, U., Wenig, M., & Wagner, T. (2003). Weekly cycle of NO_2 by GOME measurements: A signature of anthropogenic sources. Atmospheric Chemistry and Physics, 3(6), 2225–2232. https://doi.org/10.5194/acp-3-2225-2003

Bell, D., & Jayne, M. (2009). Small Cities? Towards a Research Agenda. International Journal of Urban and Regional Research, 33(3), 683–699. https://doi.org/10.1111/j.1468-2427.2009.00886.x

Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. . (2004). International Differences in Information Privacy Concerns: A Global Survey of Consumers. The Information Society, 20(5), 313–324. https://doi.org/10.1080/01972240490507956

Ben-Dor, E., & Saaroni, H. (1997). Airborne video thermal radiometry as a tool for monitoring microscale structures of the urban heat island. International Journal of Remote Sensing, 18(14), 3039–3053. https://doi.org/10.1080/014311697217198

Benediktsson, J. A., Pesaresi, M., & Arnason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. In IEEE Transactions on Geoscience and Remote Sensing (Vol. 41, Issue 9 PART I, pp. 1940–1949). https://doi.org/10.1109/TGRS.2003.814625

Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. International Journal of Environmental Research and Public Health, 15(2796), 1–9. https://doi.org/10.3390/ijerph15122796

Bennett, M. M., & Faxon, H. O. (2021). Uneven Frontiers: Exposing the Geopolitics of Myanmar's Borderlands with Critical Remote Sensing. Remote Sensing, 13(1158), 1–26. https://doi.org/10.3390/rs13061158

Berghauser Pont, M., Haupt, P., Berg, P., Alstäde, V., & Heyman, A. (2021). Systematic review and comparison of densification effects and planning motivations. Buildings and Cities, 2(1), 378–401. https://doi.org/10.5334/bc.125

Berking, H. (2008). Städte lassen sich am Gang erkennen wie Menschen—Skizzen zur Erforschung der Stadt und der Städte. In H. Berking & M. Löw (Eds.), Die Eigenlogik der Städte (pp. 15–31). Campus.

Berman, P. (2013). Terror and Liberalism. WW Norton & Co.

Berry, B. J. L. (1964). Cities as systems within systems of cities. Papers in Regional Science, 13(1), 147–163. https://doi.org/10.1111/j.1435-5597.1964.tb01283.x

Berry, B. J. L. (1990). Urbanization. In B. L. Turner II, W. C. Clark, R. W. Kates, J. F. Richards, J. T. Mathews, & W. B. Meyer (Eds.), He Earth as Transformed by Human Action: Global and Regional Changes in the Biosphere over the Past 300 Years (pp. 103–120). Cambridge University Press. https://doi.org/10.1177/030913259301700340

Berry, B. J. L., & Garrison, W. L. (1958). Recent developments of central place theory. Papers in Regional Science, 4(1), 107–120. https://doi.org/10.1111/j.1435-5597.1958.tb01625.x

Berry, B. J. L., & Okulicz-Kozaryn, A. (2012). The city size distribution debate: Resolution for US urban regions and megalopolitan areas. Cities, 29, 17–23. https://doi.org/10.1016/j.cities.2011.11.007

Bertaud, A., & Richardson, H. W. (2005). Transit and Density: Atlanta, the United States and Western Europe. In H. W. Richardson & C.-H. C. Bae (Eds.), Urban Sprawl in Western Europe and the United States (pp. 293–310). Routledge.

Bertram, C., & Rehdanz, K. (2015). The role of urban green space for human well-being. Ecological Economics, 120, 139–152. https://doi.org/10.1016/j.ecolecon.2015.10.013

Besser-Jones, L. (2008). Personal Integrity, Morality and Psychological Well-Being: Justifying the Demands of Morality. Journal of Moral Philosophy, 5(3), 361–383. https://doi.org/10.1163/174552408X369718

Besussi, E., & Chin, N. (2003). Identifying and measuring urban sprawl. In P. A. Longley & M. Batty (Eds.), Advanced spatial analysis: The CASA book of GIS (pp. 109–128). ESRI Press.

Bettencourt, L. M. A. (2007). The Origins of Scaling in Cities. Science, 340(6139), 1438–1441. https://doi.org/10.1126/science.1235823

Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. Proceedings of the National Academy of Sciences, 104(17), 7301–7306. https://doi.org/10.1073/pnas.0610172104

Bhanye, J., Shayamunda, R., & Tavirai, R. C. (2021). Social Media in the African Context. In The Palgrave Handbook of Global Social Problems (pp. 1–32). Springer International Publishing. https://doi.org/10.1007/978-3-030-68127-2_366-1

Bharti, N., Tatem, A. J., Ferrari, M. J., Grais, R. F., Djibo, A., & Grenfell, B. T. (2011). Explaining Seasonal Fluctuations of Measles in Niger Using Nighttime Lights Imagery. Science, 334(6061), 1424–1427. https://doi.org/10.1126/science.1210554

Bhaskaran, S., Paramananda, S., & Ramnarayan, M. (2010). Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data. Applied Geography, 30(4), 650–665. https://doi.org/10.1016/j.apgeog.2010.01.009

Biłozor, A., & Cieślak, I. (2021). Review of Experience in Recent Studies on the Dynamics of Land Urbanisation. Land, 10(1117), 1–27. https://doi.org/10.3390/land10111117

Bimonte, S., & Stabile, A. (2017). Land consumption and income in Italy: A case of inverted EKC. Ecological Economics, 131, 36–43. https://doi.org/10.1016/j.ecolecon.2016.08.016

Binder, C. R., Wyss, R., & Massaro, E. (Eds.). (2020). Sustainability Assessment of Urban Systems. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781108574334

Bino, G., Levin, N., Darawshi, S., Hal, N. V. D., Reich-Solomon, A., & Kark, S. (2008). Accurate prediction of bird species richness patterns in an urban environment using Landsat-derived NDVI and spectral unmixing. International Journal of Remote Sensing, 29(13), 3675–3700. https://doi.org/10.1080/01431160701772534

Bissonnette, A., & Vallet, É. (2020). Borders and Border Walls: In-Security, Symbolism, Vulnerabilities (1st ed.). Routledge. https://doi.org/10.4324/9780429352508

Bocquier, P. (2004). World Urbanization Prospects: An alternative to the UN model of projection compatible with the mobility transition theory. Demographic Research, 12(9), 197–236. https://doi.org/10.4054/DemRes.2005.12.9

Boeing, G. (2017). OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Computers, Environment and Urban Systems, 65, 126–139. https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Boeing, G. (2019). Urban spatial order: Street network orientation, configuration, and entropy. Applied Network Science, 4(67), 1–19. https://doi.org/10.1007/s41109-019-0189-1

Boeing, G., Higgs, C., Liu, S., Giles-Corti, B., Sallis, J. F., Cerin, E., Lowe, M., Adlakha, D., Hinckson, E., Moudon, A. V., Salvo, D., Adams, M. A., Barrozo, L. V., Bozovic, T., Delclòs-Alió, X., Dygrýn, J., Ferguson, S., Gebel, K., Ho, T. P., … Arundel, J. (2022). Using open data and open-source software to develop spatial indicators of urban design and transport features for achieving healthy and sustainable cities. The Lancet Global Health, 10(6), e907–e918. https://doi.org/10.1016/S2214-109X(22)00072-9

Bok, R. (2021). Wayfinding in the Long Shadow of City Benchmarking: Or How to Manufacture (an Economy of) Comparability in the Global Urban. International Journal of Urban and Regional Research, 45(2), 381–384. Scopus. https://doi.org/10.1111/1468-2427.12977

Bolund, P., & Hunhammar, S. (1999). Ecosystem services in urban areas. Ecological Economics, 29(2), 293–301. https://doi.org/10.1016/S0921-8009(99)00013-0

Bonan, G. B. (2008). Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. Science, 320(5882), 1444–1449. https://doi.org/10.1126/science.1155121

Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., & Perez, J. R. (2011). GLOBCOVER 2009 products description and validation report (ESA-GlobCover 2005 Project, pp. 1–53). https://epic.awi.de/id/eprint/31014/16/GLOBCOVER2009_Validation_Report_2-2.pdf

Bontje, M., & Burdack, J. (2005). Edge Cities, European-style: Examples from Paris and the Randstad. Cities, 22(4), 317–330. https://doi.org/10.1016/j.cities.2005.01.007

Border Policy Research Institute, & Hodges, H. (2006). Explaining the Decline in Border Crossings Since 1990. Border Policy Research Institute Publications, 66, 1–3. https://cedar.wwu.edu/bpri_publications/66

Borruso, G., & Porceddu, A. (2009). A Tale of Two Cities: Density Analysis of CBD on Two Midsize Urban Areas in Northeastern Italy. In B. Murgante, G. Borruso, & A. Lapucci (Eds.), Geocomputation and Urban Planning (pp. 37–56). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-89930-3_3

Bostan, P. A., Heuvelink, G. B. M., & Akyurek, S. Z. (2012). Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. International Journal of Applied Earth Observation and Geoinformation, 19, 115–126. https://doi.org/10.1016/j.jag.2012.04.010

Boucher, A., Seto, K. C., & Journel, A. G. (2006). A novel method for mapping land cover changes: Incorporating time and space with geostatistics. In IEEE Transactions on Geoscience and Remote Sensing (Vol. 44, Issue 11, pp. 3427–3435). https://doi.org/10.1109/TGRS.2006.879113

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., & Goede, A. P. H. (1999). SCIAMACHY: Mission Objectives and Measurement Modes. Journal of the Atmospheric Sciences, 56(2), 127–150. https://doi.org/10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2

Boy, J. D., & Uitermark, J. (2017). Reassembling the city through Instagram. Transactions of the Institute of British Geographers, 42(4), 612–624. https://doi.org/10.1111/tran.12185

Boyce, R. R. (1963). Changing patterns of urban land consumption. The Professional Geographer, 15(2), 19–24. https://doi.org/10.1111/j.0033-0124.1963.019_r.x

Brakman, S., Garretsen, H., van Marrewijk, C., & Oumer, A. (2012). The Border Population Effects of Eu Integration. Journal of Regional Science, 52(1), 40–59. https://doi.org/10.1111/j.1467-9787.2011.00752.x

Bramley, G., & Power, S. (2009). Urban Form and Social Sustainability: The Role of Density and Housing Type. Environment and Planning B: Planning and Design, 36(1), 30–48. https://doi.org/10.1068/b33129

Brauer, M., Amann, M., Burnett, R. T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S. B., Krzyzanowski, M., Martin, R. V., Van Dingenen, R., van Donkelaar, A., & Thurston, G. D. (2012). Exposure Assessment for Estimation of the Global Burden of Disease Attributable to Outdoor Air Pollution. Environmental Science & Technology, 46(2), 652–660. https://doi.org/10.1021/es2025752

Braun, A., Warth, G., Bachofer, F., Schultz, M., & Hochschild, V. (2023). Mapping Urban Structure Types Based on Remote Sensing Data—A Universal and Adaptable Framework for Spatial Analyses of Cities. Land, 12(1885), 1–41. https://doi.org/10.3390/land12101885

Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brenner, N. (Ed.). (2014). Implosions/Explosions: Towards a Study of Planetary Urbanization. Jovis.

Brenner, N., & Schmid, C. (2015). Towards a new epistemology of the urban? City, 19(2–3), 151–182. https://doi.org/10.1080/13604813.2015.1014712

Brenner, N., & Schmid, C. (2013). The ‘Urban Age’ in Question. International Journal of Urban and Regional Research, 38(3), 731–755. https://doi.org/10.1111/1468-2427.12115

Breuer, J. H. P., Friesen, J., Taubenböck, H., Wurm, M., & Pelz, P. F. (2024). The unseen population: Do we underestimate slum dwellers in cities of the Global South? Habitat International, 148(103056), 1–19. https://doi.org/10.1016/j.habitatint.2024.103056

Breuninger, T., Menschik, B., Demharter, A., Gamperl, M., & Thuro, K. (2021). Investigation of Critical Geotechnical, Petrological and Mineralogical Parameters for Landslides in Deeply Weathered Dunite Rock (Medellín, Colombia). International Journal of Environmental Research and Public Health, 18(11141), 1–26. https://doi.org/10.3390/ijerph182111141

Brezzi, M., & Veneri, P. (2015). Assessing Polycentric Urban Systems in the OECD: Country, Regional and Metropolitan Perspectives. European Planning Studies, 23(6), 1128–1145. https://doi.org/10.1080/09654313.2014.905005

Broadbent, A. M., Krayenhoff, E. S., & Georgescu, M. (2020). The motley drivers of heat and cold exposure in 21st century US cities. Proceedings of the National Academy of Sciences, 117(35), 21108–21117. https://doi.org/10.1073/pnas.2005492117

Brown, W. (2010). Walled States, Waning Sovereignty. Zone Books. https://doi.org/10.2307/j.ctv14gpj55

Bryan, M. L. (1975). Interpretation of an urban scene using multi-channel radar imagery. Remote Sensing of Environment, 4, 49–66. https://doi.org/10.1016/0034-4257(75)90005-X

Budde, R., & Neumann, U. (2018). The size ranking of cities in Germany: Caught by a MAUP? GeoJournal, 84(6), 1447–1464. https://doi.org/10.1007/s10708-018-9930-z

Buettner, T. (2014). Urban Estimates and Projections at the United Nations: The Strengths, Weaknesses, and Underpinnings of the World Urbanization Prospects. Spatial Demography, 3(2), 91–108. https://doi.org/10.1007/s40980-015-0004-2

Bundesamt für Kartographie und Geodäsie. (2021). Dokumentation 3-D Gebäudemodell Deutschland LoD1 (LoD1-DE). https://gdz.bkg.bund.de/index.php/default/3d-gebaudemodelle-lod1-deutschland-lod1-de.html

Bundesamt für Kartographie und Geodasie ¨ (BKG)/GeoBasis-DE. (2020). Verwaltungsgebiete 1:250 000. VG250 und VG250-EW (Version Retrieved, 01.10.2020) [Dataset]. https://gdz.bkg.bund.de /index.php/default/catalog/product/view/id/789/s/verwaltungsgebiete-1-250-000 -ebenen-stand-31-12-vg250-ebenen-31-12/category/8/?___store=default

Bundesinstitut für Bau- Stadt- und Raumforschung (BBSR). (2018). Handlungsziele für Stadtgrün und deren empirische Evidenz. Bundesamt für Bauwesen und Raumordnung. https://www.bbsr.bund.de/BBSR/DE/veroeffentlichungen/sonderveroeffentlichungen/2018/handlungsziele-stadtgruen.html

Bundesministerium der Justiz. (2018). Data Ethics Commission. https://www.bmj.de/SharedDocs/ Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.htm

Bundesministerium für Umwelt, Naturschutz, Bau und Reaktorsicherheit (BMUB). (2017). Weißbuch Stadtgrün: Grün in der Stadt—Für eine lebenswerte Zukunft. Brandenburgische Universitätsdruckerei und Verlagsgesellschaft Potsdam mbH. https://www.bbsr.bund.de/BBSR/DE/veroeffentlichungen/ministerien/bmub/verschiedene-themen/2017/weissbuch-stadtgruen.html

Bundesministerium für Umwelt, Naturschutz und nukleare Sicherheit (BMU). (2019). Masterplan Stadtnatur. https://www.bmuv.de/stadtnatur

Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of Sprawl: A Portrait from Space*. The Quarterly Journal of Economics, 121(2), 587–633. https://doi.org/10.1162/qjec.2006.121.2.587

Burdett, R., & Rode, P. (2018). Shaping Cities in an Urban Age. Phaidon Press.

Burdett, R., & Sudjic, D. (2007). The Endless City. Phaidon Press.

Burdett, R., & Sudjic, D. (2011). Living in the Endless City. Phaidon Press.

Burger, M., & Meijers, E. (2012). Form Follows Function? Linking Morphological and Functional Polycentricity. Urban Studies, 49(5), 1127–1149. https://doi.org/10.1177/0042098011407095

Bürgerbegehren Grünflächen erhalten. (2019). Grünflächen-erhalten.de – Ein Bürgerbegehren für München. https://www.gruenflaechen-erhalten.de/

Burgess, E. W. (1925). The City. The University of Chicago Press.

Burgess, E. W. (1925). The Growth of the City: An Introduction to a Research Project. In J. M. Marzluff, E. Shulenberger, W. Endlicher, M. Alberti, G. Bradley, C. Ryan, U. Simon, & C. ZumBrunnen (Eds.), Urban Ecology (pp. 71–78). New York, NY.

Burrough, P. A., & Frank, A. (1996). Geographic Objects with Indeterminate Boundaries. Taylor & Francis Ltd.

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., DeBeek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M., & Perner, D. (1999). The Global Ozone Monitoring Experiment (GOME): Mission Concept and First Scientific Results. Journal of the Atmospheric Sciences, 56(2), 151–175. https://doi.org/10.1175/1520-0469(1999)056<0151:TGOMEG>2.0.CO;2

Butilă, E. V., & Boboc, R. G. (2022). Urban Traffic Monitoring and Analysis Using Unmanned Aerial Vehicles (UAVs): A Systematic Literature Review. Remote Sensing, 14(620), 1–28. https://doi.org/10.3390/rs14030620

Butt, E. W., Turnock, S. T., Rigby, R., Reddington, C. L., Yoshioka, M., Johnson, J. S., Regayre, L. A., Pringle, K. J., Mann, G. W., & Spracklen, D. V. (2017). Global and regional trends in particulate air pollution and attributable health burden over the past 50 years. Environmental Research Letters, 12(104017), 1–10. https://doi.org/10.1088/1748-9326/aa87be

C

C40 Cities. (2023). Green & Just Recovery Agenda. https://www.c40.org/what-we-do/raising-climate-ambition/green-just-recovery-agenda/

Cai, J., Huang, B., & Song, Y. (2017). Using multi-source geospatial big data to identify the structure of polycentric cities. Big Remotely Sensed Data: Tools, Applications and Experiences, 202, 210–221. https://doi.org/10.1016/j.rse.2017.06.039

Calderón-Guevara, W., Sánchez-Silva, M., Nitescu, B., & Villarraga, D. F. (2022). Comparative review of data-driven landslide susceptibility models: Case study in the Eastern Andes mountain range of Colombia. Natural Hazards, 113(2), 1105–1132. https://doi.org/10.1007/s11069-022-05339-2

Caldwell, J. T., Ford, C. L., Wallace, S. P., Wang, M. C., & Takahashi, L. M. (2016). Intersection of living in a rural versus urban area and race/ethnicity in explaining access to health care in the United States. American Journal of Public Health, 106(8), 1463–1469. Scopus. https://doi.org/10.2105/AJPH.2016.303212

Camps-Valls, G., Tuia, D., Zhu, X. X., & Reichstein, M. (2021). Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences. John Wiley & Sons Ltd.

Cantarino, I., Carrion, M. A., Goerlich, F., & Martinez Ibañez, V. (2019). A ROC analysis-based classification method for landslide susceptibility maps. Landslides, 16(2), 265–282. https://doi.org/10.1007/s10346-018-1063-4

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. Journal of Urban Economics, 61(3), 389–419. https://doi.org/10.1016/j.jue.2006.08.003

Carnahan, W. H., & Larson, R. C. (1990). An analysis of an urban heat sink. Remote Sensing of Environment, 33(1), 65–71. https://doi.org/10.1016/0034-4257(90)90056-R

Carta, M. G., Preti, A., & Akiskal, H. S. (2018). Coping with the New Era: Noise and Light Pollution, Hperactivity and Steroid Hormones. Towards an Evolutionary View of Bipolar Disorders. Clinical Practice and Epidemiology in Mental Health, 14, 33–36. https://doi.org/10.2174/1745017901814010033

Carter, D. B., & Poast, P. (2017). Why Do States Build Walls? Political Economy, Security, and Border Stability. Journal of Conflict Resolution, 61(2), 239–270. https://doi.org/10.1177/0022002715596776

Carter, D. B., & Poast, P. (2020). Barriers to Trade: How Border Walls Affect Trade Relations. International Organization, 74(1), 165–185. Cambridge Core. https://doi.org/doi:10.1017/S0020818319000353

Casagli, N., Frodella, W., Morelli, S., Tofani, V., Ciampalini, A., Intrieri, E., Raspini, F., Rossi, G., Tanteri, L., & Lu, P. (2017). Spaceborne, UAV and ground-based remote sensing techniques for landslide mapping, monitoring and early warning. Geoenvironmental Disasters, 4(9), 1–23. https://doi.org/10.1186/s40677-017-0073-1

Castells, M. (1977). The urban question: A Marxist approach. Hodder & Stoughton Ltd.

Castells, M. (1996). The Rise of the Network Society (1st ed.). Blackwell Publishers.

Castells, M. (1999). Grassrooting the space of flows. Urban Geography, 20(4), 294–302. https://doi.org/10.2747/0272-3638.20.4.294

Castells, M. (2000). End of Millennium, The Information age: Economy, society and culture, vol. 3. Blackwell Publishers.

Castells, M. (2002). Local and Global: Cities in the Network Society. Tijdschrift Voor Economische En Sociale Geografie, 93(5), 548–558. https://doi.org/10.1111/1467-9663.00225

Castells, M. (2006). The Space of Flows. In M. Castells (Ed.), The Rise of the Network Society (pp. 407–459). Wiley. https://doi.org/10.1002/9781444319514.ch6

Castells-Quintana, D., & Royuela, V. (2015). Are Increasing Urbanisation and Inequalities Symptoms of Growth? Applied Spatial Analysis and Policy, 8(3), 291–308. https://doi.org/10.1007/s12061-015-9146-2

Castro Bonilla, J. P. (2021). Sistema de alerta temprana de deslizamiento para el fenómeno de movimiento en masa del sector Altos de la Estancia. Universidad Distritial Francisco José de Caldas, Bogotá, Colombia.

Cataldi, G., Maffei, G. L., & Vaccaro, P. (2002). Saverio Muratori and the Italian school of planning typology. Urban Morphology, 6(1), Article 1. https://doi.org/10.51347/jum.v6i1.3899

Center for International Earth Science Information Network (CIESIN), Columbia University. (2017). Gridded population of the world, version 4 (GPWv4): Population density, revision 10. https://sedac.ciesin.columbia.edu/downloads/docs/gpw-v4/gpw-v4-documentation-rev10.pdf

Centre on Housing Rights and Evictions , COHRE. (2010). Centre on Housing Rights and Evictions (COHRE) v. Sudan, Communication 296/2005. ESCR-Net. https:// www.escr-net.org/caselaw/2010/centre-housing-rights-and -evictions-cohre-v-sudan-communication-2962005

Cervero, R. (2000). Shapeless, spread out, skipped over and scattershot–sprawl sweeps the globe. The World Paper, 5–6.

Cervero, R., & Murakami, J. (2010). Effects of Built Environments on Vehicle Miles Traveled: Evidence from 370 US Urbanized Areas. Environment and Planning A: Economy and Space, 42(2), 400–418. https://doi.org/10.1068/a4236

Cervone, G., Schnebele, E., Waters, N., Moccaldi, M., & Sicignano, R. (2017). Using Social Media and Satellite Data for Damage Assessment in Urban Areas During Emergencies. In P. (Vonu) Thakuriah, N. Tilahun, & M. Zellner (Eds.), Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics (pp. 443–457). Springer International Publishing. https://doi.org/10.1007/978-3-319-40902-3_24

Chadda, R. K., & Deb, K. S. (2013). Indian family systems, collectivistic society and psychotherapy. Indian Journal of Psychiatry, 55(Suppl 2), 299–309. https://doi.org/10.4103/0019-5545.105555

Chai, B., & Seto, K. C. (2019). Conceptualizing and characterizing micro-urbanization: A new perspective applied to Africa. Landscape and Urban Planning, 190(103595), 1–12. https://doi.org/10.1016/j.landurbplan.2019.103595

Chakraborty, T., & Lee, X. (2019). A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability. International Journal of Applied Earth Observation and Geoinformation, 74, 269–280. https://doi.org/10.1016/j.jag.2018.09.015

Champion, T., & Hugo, G. (2004). New Forms of Urbanization: Beyond the Urban-Rural Dichotomy (1st ed.). Routledge. https://doi.org/10.4324/9781315248073

Chan, K. W., & Wan, G. (2017). The size distribution and growth pattern of cities in China, 1982–2010: Analysis and policy implications. Journal of the Asia Pacific Economy, 22(1), 136–155. https://doi.org/10.1080/13547860.2016.1266829

Chandra, G. R., & Bhatia, S. (2020). Issues and Concerns in the Field of Remote Sensing. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 654–658. https://doi.org/10.1109/ICRITO48877.2020.9197901

Chandrashekar, C. M., & Aithal, B. H. (2021). Impact assessment of Corridor Oriented development. International Review for Spatial Planning and Sustainable Development, 9(2), 172–194. https://doi.org/10.14246/irspsdc.9.2_172

Chapman, S., Watson, J. E. M., Salazar, A., Thatcher, M., & McAlpine, C. A. (2017). The impact of urbanization and climate change on urban temperatures: A systematic review. Landscape Ecology, 32(10), 1921–1935. https://doi.org/10.1007/s10980-017-0561-4

Chatzidimitriou, A., & Yannas, S. (2015). Microclimate development in open urban spaces: The influence of form and materials. Energy and Buildings, 108, 156–174. https://doi.org/10.1016/j.enbuild.2015.08.048

Chazal, F., Guibas, L. J., Oudot, S. Y., & Skraba, P. (2013). Persistence-Based clustering in riemannian manifolds. In Journal of the ACM (Vol. 60, Issue 6, p. 41:1-41:38). https://doi.org/10.1145/2535927

Chen, G., Li, X., Liu, X., Chen, Y., Liang, X., Leng, J., Xu, X., Liao, W., Qiu, Y., Wu, Q., & Huang, K. (2020). Global projections of future urban land expansion under shared socioeconomic pathways. Nature Communications, 11(537), 1–12. https://doi.org/10.1038/s41467-020-14386-x

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., Zhang, W., Tong, X., & Mills, J. (2014). Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103, 7–27. https://doi.org/10.1016/j.isprsjprs.2014.09.002

Chen, M., Liu, W., & Tao, X. (2013). Evolution and assessment on China's urbanization 1960–2010: Under-urbanization or over-urbanization? Habitat International, 38, 25–33. https://doi.org/10.1016/j.habitatint.2012.09.007

Chen, W., Chen, W., Ning, S., Liu, E., Zhou, X., Wang, Y., & Zhao, M. (2019). Exploring the industrial land use efficiency of China's resource-based cities. Cities, 93, 215–223. https://doi.org/10.1016/j.cities.2019.05.009

Chen, X., & Nordhaus, W. D. (2019). VIIRS Nighttime Lights in the Estimation of Cross-Sectional and Time-Series GDP. Remote Sensing, 11(9), 1–11. https://doi.org/10.3390/rs11091057

Chen, X.-L., Zhao, H.-M., Li, P.-X., & Yin, Z.-Y. (2006). Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sensing of Environment, 104(2), 133–146. https://doi.org/10.1016/j.rse.2005.11.016

Chen, Y., Liu, X., Li, X., Liu, X., Yao, Y., Hu, G., Xu, X., & Pei, F. (2017). Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method. Landscape and Urban Planning, 160, 48–60. https://doi.org/10.1016/j.landurbplan.2016.12.001

Cheng, X., Luo, M., Chen, K., Sun, J., & Wu, Y. (2024). Intra-annual vegetation changes and spatial variation in China over the past two decades based on remote sensing and time-series clustering. Environmental Monitoring and Assessment, 196(675), 1–17. https://doi.org/10.1007/s10661-024-12816-7

Chmielewski, S., Bochniak, A., Natapov, A., & Wężyk, P. (2020). Introducing GEOBIA to Landscape Imageability Assessment: A Multi-Temporal Case Study of the Nature Reserve "Kózki", Poland. Remote Sensing, 12(17), 1–25. https://doi.org/10.3390/rs12172792

Choe, S. C. (1996). The evolving urban system in North-East Asia. In F.-C. Lo & Y.-M. Yueng (Eds.), Emerging World Cities in Pacific Asia (pp. 498–519). United Nations University Press.

Christaller, W. (1966). Central places in Southern Germany. Prentice Hall.

Christensen, P., & McCord, G. C. (2016). Geographic determinants of China's urbanization. Regional Science and Urban Economics, 59, 90–102. https://doi.org/10.1016/j.regsciurbeco.2016.05.001

Cirilli, A., & Veneri, P. (2013). Spatial Structure and Carbon Dioxide (CO2) Emissions Due to Commuting: An Analysis of Italian Urban Areas. Regional Studies, 48(12), 1993–2005. https://doi.org/10.1080/00343404.2013.827333

Clarke, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. In Environment and Planning B: Planning and Design (Vol. 24, Issue 2, pp. 247–261). https://doi.org/10.1068/b240247

Coase, R. H. (1960). The problem of social cost. The Journal of Law & Economics, 3, 1–44. https://doi.org/10.1086/466560

Cochran, F., Daniel, J., Jackson, L., & Neale, A. (2020). Earth observation-based ecosystem services indicators for national and subnational reporting of the sustainable development goals. Remote Sensing of Environment, 244(111796), 1–14. https://doi.org/10.1016/j.rse.2020.111796

Coffel, E. D., Horton, R. M., & de Sherbinin, A. (2017). Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. Environmental Research Letters, 13(014001), 1–9. https://doi.org/10.1088/1748-9326/aaa00e

Coffer, M. M. (2020). Balancing Privacy Rights and the Production of High-Quality Satellite Imagery. Environmental Science & Technology, 54(11), 6453–6455. https://doi.org/10.1021/acs.est.0c02365

Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. In IEEE Transactions on Pattern Analysis and Machine Intelligence (Vol. 24, Issue 5, pp. 603–619). https://doi.org/10.1109/34.1000236

Conzen, M. (1980). The Morphology of Nineteenth-Century Cities in the United States. Urban History Review / Revue d'histoire Urbaine, 119–141. https://doi.org/10.7202/1020702ar

Conzen, M. G. (1988). Morphogenesis, morphological regions and secular human agency in the historic townscape, as exemplified by Ludlow. In Urban Historical Geography. Recent progress in Britain and Germany (pp. 253–272). Cambridge University Press.

Conzen, M. P. (2001). The study of urban form in the United States. Urban Morphology, 5(1), 3–14. https://doi.org/10.51347/jum.v5i1.3894

Conzen, M. R. G. (1960). Alnwick, Northumberland: A Study in Town-Plan Analysis. Transactions and Papers (Institute of British Geographers), 27, iii–122. https://doi.org/10.2307/621094

Copa-Cogeca. (2018). EU code of conduct on agricultural data sharing by contractual agreement. https://www.cema-agri.org/images/publications/brochures/EU_Code_of_conduct_leaflet.pdf

Corbane, C., Florczyk, A., Pesaresi, M., Politis, P., & Syrris, V. (2018). GHS-BUILT R2018A - GHS built-up grid, derived from Landsat, multitemporal (1975-1990-2000-2014)—OBSOLETE RELEASE [Dataset]. European Commission, Joint Research Centre (JRC); PID: http://data.europa.eu/89h/jrc-ghsl-10007. https://doi.org/10.2905/jrc-ghsl-10007

Corbane, C., Pesaresi, M., Kemper, T., Politis, P., Florczyk, A. J., Syrris, V., Melchiorri, M., Sabo, F., & Soille, P. (2019). Automated global delineation of human settlements from 40 years of Landsat satellite data archives. Big Earth Data, 3(2), 140–169. https://doi.org/10.1080/20964471.2019.1625528

Cording, M., Harrison, J. S., Hoskisson, R. E., & Jonsen, K. (2014). Walking the Talk: A Multistakeholder Exploration of Organizational Authenticity, Employee Productivity, and Post-Merger Performance. Academy of Management Perspectives, 28(1), 38–56. https://doi.org/10.5465/amp.2013.0002

Correlates of War Project. (2002). Direct Contiguity Data, 1816-2016. Version 3.2 [Dataset]. https://correlatesofwar.org/data-sets/direct-contiguity/

Cozzolino, S. (2018). Reconsidering Urban Spontaneity and Flexibility after Jane Jacobs: How do they work under different kinds of planning conditions? COSMOS + TAXIS, 5(3), 14–24.

Cozzolino, S. (2020). The (anti) adaptive neighbourhoods. Embracing complexity and distribution of design control in the ordinary built environment. Environment and Planning B: Urban Analytics and City Science, 47(2), 203–219. https://doi.org/10.1177/2399808319857451

Cozzolino, S. (2022). On the spontaneous beauty of cities: Neither design nor chaos. URBAN DESIGN International, 27(1), 43–52. https://doi.org/10.1057/s41289-021-00170-w

Cozzolino, S., & Moroni, S. (2021). Multiple agents and self-organisation in complex cities: The crucial role of several property. Land Use Policy, 103(105297), 1–7. https://doi.org/10.1016/j.landusepol.2021.105297

Creutzig, F., Agoston, P., Minx, J. C., Canadell, J. G., Andrew, R. M., Quéré, C. L., Peters, G. P., Sharifi, A., Yamagata, Y., & Dhakal, S. (2016). Urban infrastructure choices structure climate solutions. Nature Climate Change, 6(12), 1054–1056. https://doi.org/10.1038/nclimate3169

Creutzig, F., Jochem, P., Edelenbosch, O. Y., Mattauch, L., Vuuren, D. P. van, McCollum, D., & Minx, J. (2015). Transport: A roadblock to climate change mitigation? Science, 350(6263), 911–912. https://doi.org/10.1126/science.aac8033

Criado Perez, C. (2019). Invisible Women: Exposing Data Bias in a World Designed for Men. Abrams Press.

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., Karagiorgou, S., Efentakis, A., & Lamprianidis, G. (2014). Crowdsourcing urban form and function. International Journal of Geographical Information Science, 29(5), 720–741. https://doi.org/10.1080/13658816.2014.977905

Crosetto, M., Castillo, M., & Arbiol, R. (2003). Urban subsidence monitoring using radar interferometry: Algorithms and validation. Photogrammetric Engineering and Remote Sensing, 69(7), 775–783. Scopus. https://doi.org/10.14358/PERS.69.7.775

Curtis, C. (2006). Network City: Retrofitting the Perth Metropolitan Region to Facilitate Sustainable Travel. Urban Policy and Research, 24(2), 159–180. https://doi.org/10.1080/08111140600703691

D

D'Acci, L. (2019). Quality of urban area, distance from city centre, and housing value. Case study on real estate values in Turin. Cities, 91, 71–92. https://doi.org/10.1016/j.cities.2018.11.008

DAGRD. (2018). Base de datos de Visitas Técnicas, Departamento Administrativo de Gestión del Riesgo de Desastres (DAGRD) [Dataset].

D'Amour, C. B., Reitsma, F., Baiocchi, G., Barthel, S., Güneralp, B., Erb, K.-H., Haberl, H., Creutzig, F., & Seto, K. C. (2017). Future urban land expansion and implications for global croplands. In Proceedings of the National Academy of Sciences of the United States of America (Vol. 114, Issue 34, pp. 8939–8944). https://doi.org/10.1073/pnas.1606036114

Danylo, O., See, L., Bechtel, B., Schepaschenko, D., & Fritz, S. (2016). Contributing to WUDAPT: A Local Climate Zone Classification of Two Cities in Ukraine. In IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Vol. 9, Issue 5, pp. 1841–1853). https://doi.org/10.1109/JSTARS.2016.2539977

Dastin, J. (2018). Insight—Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. https://www.reuters.com/article/us-amazon -com-jobs-automation-insight-idUSKCN1MK08G

Data Ethics Commission of the Federal Government. (2019). Opinion of the Data Ethics Commission. https://www.bfdi.bund.de/SharedDocs/Downloads/EN/Datenschutz/Data-Ethics-Commission_Opinion.pdf?__blob=publicationFile&v=1.

Davies, R. J. (1981). The spatial formation of the South African City. GeoJournal, 2(2 Supplement), 59–72. https://doi.org/10.1007/BF00196325

Davis, M. (2007). Planet of Slums. Verso Books.

Davoudi, S. (2007). Polycentricity: Panacea or pipedream? In N. Cattan (Ed.), Cities and Networks in Europe: A Critical Approach of Polycentrism (pp. 65–73). John Libbey Eurotext.

De Castro, M. C., Yamagata, Y., Mtasiwa, D., Tanner, M., Utzinger, J., Keiser, J., & Singer, B. H. (2004). Integrated urban malaria control: A case study in Dar es Salaam, Tanzania. The American Journal of Tropical Medicine and Hygiene Am J Trop Med Hyg, 71(2_suppl), 103–117. https://doi.org/10.4269/ajtmh.2004.71.103

Dear, M., & Dahmann, N. (2008). Urban Politics and the Los Angeles School of Urbanism. Urban Affairs Review, 44(2), 266–279. https://doi.org/10.1177/1078087408320240

Debray, H., Kraff, N. J., Zhu, X. X., & Taubenböck, H. (2023). Planned, unplanned, or in-between? A concept of the intensity of plannedness and its empirical relation to the built urban landscape across the globe. Landscape and Urban Planning, 233(104711), 1–27. https://doi.org/10.1016/j.landurbplan.2023.104711

Debray, H., Lemoine-Rodríguez, R., Wurm, M., Zhu, X., & Taubenböck, H. (2025). Distributions of types of urban fabrics around the world: Beyond narratives of global regularities and local specificities. Applied Geography, 184(103770), 1–23. https://doi.org/10.1016/j.apgeog.2025.103770

Decker, E. H., Kerkhoff, A. J., & Moses, M. E. (2007). Global Patterns of City Size Distributions and Their Fundamental Drivers. PLOS ONE, 2(9)(e934), 1–6. https://doi.org/10.1371/journal.pone.0000934

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society. Series B (Methodological), 39(1), 1–38. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Demuzere, M., Bechtel, B., Middel, A., & Mills, G. (2020). European LCZ map (Version 1) [Dataset]. https://doi.org/10.6084/m9.figshare.13322450.v1

Demuzere, M., Bechtel, B., & Mills, G. (2019). Global transferability of local climate zone models. Urban Climate, 27, 46–63. https://doi.org/10.1016/j.uclim.2018.11.001

Demuzere, M., Kittner, J., Martilli, A., Mills, G., Moede, C., Stewart, I. D., van Vliet, J., & Bechtel, B. (2022). A global map of local climate zones to support earth system modelling and urban-scale environmental science. Earth System Science Data, 14(8), 3835–3873. https://doi.org/10.5194/essd-14-3835-2022

Den Hartog, D. N., & De Hoogh, A. H. B. (1982). Empowering behaviour and leader fairness and integrity: Studying perceptions of ethical leader behaviour from a levels-of-analysis perspective. European Journal of Work and Organizational Psychology, 18(2), 199–230. https://doi.org/10.1080/13594320802362688

Deng, C., Li, C., Zhu, Z., Lin, W., & Xi, L. (2017). Subpixel urban impervious surface mapping: The impact of input Landsat images. ISPRS Journal of Photogrammetry and Remote Sensing, 133, 89–103. https://doi.org/10.1016/j.isprsjprs.2017.09.015

Deng, C., & Wu, C. (2012). BCI: A biophysical composition index for remote sensing of urban environments. Remote Sensing of Environment, 127, 247–259. https://doi.org/10.1016/j.rse.2012.09.009

Deng, C., & Wu, C. (2013a). A spatially adaptive spectral mixture analysis for mapping subpixel urban impervious surface distribution. Remote Sensing of Environment, 133, 62–70. https://doi.org/10.1016/j.rse.2013.02.005

Deng, C., & Wu, C. (2013b). Examining the impacts of urban biophysical compositions on surface urban heat island: A spectral unmixing and thermal mixing approach. Remote Sensing of Environment, 131, 262–274. https://doi.org/10.1016/j.rse.2012.12.020

Deng, C., & Zhu, Z. (2018). Continuous subpixel monitoring of urban impervious surface using Landsat time series. Time Series Analysis with High Spatial Resolution Imagery, 238(110929), 1–21. https://doi.org/10.1016/j.rse.2018.10.011

Denis, G., Claverie, A., Pasco, X., Darnis, J.-P., de Maupeou, B., Lafaye, M., & Morel, E. (2017). Towards disruptions in Earth observation? New Earth Observation systems and markets evolution: Possible scenarios and impacts. Acta Astronautica, 137, 415–433. https://doi.org/10.1016/j.actaastro.2017.04.034

Departamento del Quindío. (2018). Desarrollo de prototipo sistema inteligente de alerta temprana para la prevención de desastres por remoción en masa, avenidas torrenciales e inundación en los municipios de Calarcá, Pijao y Salento, en el departamento del Quindío.

DesInventar. (2022). DesInventar Sendai [Dataset]. https://db.desinventar.org/DesInventar/download.jsp

DeStatis. (2020). ZENSUS 2011—Bevölkerungs- und Wohnungszählung 2011—Ergebnisse des Zensus 2011. https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html?nn=3065474

Deuskar, C. (2015). What does “urban” mean? World Bank Blogs. https://blogs.worldbank.org/en/sustainablecities/what-does-urban-mean

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., Blondel, V. D., & Tatem, A. J. (2014). Dynamic population mapping using mobile phone data. Proceedings of the National Academy of Sciences, 111(45), 15888–15893. https://doi.org/10.1073/pnas.1408439111

Devolder, K., & Harris, J. (2007). The ambiguity of the embryo: Ethical inconsistency in the human embryonic stem cell debate. Metaphilosophy, 38(2–3), 153–169. https://doi.org/10.1111/j.1467-9973.2007.00480.x

Dewan, A. M., & Yamaguchi, Y. (2009). Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. Applied Geography, 29(3), 390–401. https://doi.org/10.1016/j.apgeog.2008.12.005

Dialesandro, J., Brazil, N., Wheeler, S., & Abunnasr, Y. (2021). Dimensions of Thermal Inequity: Neighborhood Social Demographics and Urban Heat in the Southwestern U.S. International Journal of Environmental Research and Public Health, 18(3), 1–15. https://doi.org/10.3390/ijerph18030941

Dibble, J., Prelorendjos, A., Romice, O., Zanella, M., Strano, E., Pagel, M., & Porta, S. (2019). On the origin of spaces: Morphometric foundations of urban form evolution. Environment and Planning B: Urban Analytics and City Science, 46(4), 707–730. https://doi.org/10.1177/2399808317725075

Dickens, B. M. (2005). The Challenges and Opportunities of Ethics. American Journal of Public Health, 95(7), 1.

Dieleman, F., & Wegener, M. (2004). Compact City and Urban Sprawl. Built Environment, 30(4), 308–323. https://doi.org/10.2148/benv.30.4.308.57151

Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., & Schiavina, M. (2021). Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. Delineation of Urban Areas, 125(103312), 1–27. https://doi.org/10.1016/j.jue.2020.103312

Dijkstra, L., & Poelman, H. (2012). Cities in Europe—The new OECD-EC definition (No. 01/2012; Regional Focus). European Commission, Regional and Urban Policy. https://ec.europa.eu/regional_policy/sources/focus/2012_01_city.pdf

Dijkstra, L., & Poelman, H. (2014). Local Climate Zones for Urban Temperature Studies (No. WP 01/2014; Regional and Urban Policy). European Commission, Regional and Urban Policy. https://ec.europa.eu/regional_policy/sources/work/2014_01_new_urban.pdf

Dijkstra, L., Poelman, H., & Veneri, P. (2019). The EU-OECD definition of a functional urban area (No. 2019/11; OECD Regional Development Working Papers). OECD Publishing. https://doi.org/10.1787/d58cb34d-en

Dijstelbloem, H. (2021). Borders as Infrastructure: The Technopolitics of Border Control. The MIT Press. https://doi.org/10.7551/mitpress/11926.001.0001

Dill, D. D. (2009). The Structure of the Academic Profession: Toward a Definition of Ethical Issues. The Journal of Higher Education, 53(3), 255–267. https://doi.org/10.2307/1981746

Ding, J., A, R. van der, Mijling, B., Laat, J. de, Eskes, H., & Boersma, K. F. (2022). NOx emissions in India derived from OMI satellite observations. Atmospheric Environment: X, 14(100174), 1–8. https://doi.org/10.1016/j.aeaoa.2022.100174

Doll, C. N. H., & Pachauri, S. (2010). Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery. The Socio-Economic Transition towards a Hydrogen Economy - Findings from European Research, with Regular Papers, 38(10), 5661–5670. https://doi.org/10.1016/j.enpol.2010.05.014

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.-H., Féménias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J., & Sciarra, R. (2012). The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. The Sentinel Missions - New Opportunities for Science, 120, 37–57. https://doi.org/10.1016/j.rse.2011.07.024

Donnay, J.-P., Barnsley, M. J., & Longley, P. A. (2014). Remote Sensing and Urban Analysis (1st ed.). CRC Press. https://doi.org/10.1201/9781482268119

Dorélien, A., Balk, D., & Todd, M. (2013). What Is Urban? Comparing a Satellite View with the Demographic and Health Surveys. Population and Development Review, 39(3), 413–439. https://doi.org/10.1111/j.1728-4457.2013.00610.x

Dousset, B., Gourmelon, F., Laaidi, K., Zeghnoun, A., Giraudet, E., Bretin, P., Mauri, E., & Vandentorren, S. (2011). Satellite monitoring of summer heat waves in the Paris metropolitan area. International Journal of Climatology, 31(2), 313–323. https://doi.org/10.1002/joc.2222

Dovey, K., & Kamalipour, H. (2017). Informal/Formal Morphologies. In K. Dovey, E. Pafka, & M. Ristic (Eds.), Mapping Urbanities: Morphologies, Flows, Possibilities (1st ed., pp. 223–248). Routledge.

Dovey, K., Oostrum, M. van, Chatterjee, I., & Shafique, T. (2020). Towards a morphogenesis of informal settlements. Habitat International, 104(102240), 1–14. https://doi.org/10.1016/j.habitatint.2020.102240

Dovey Kim, Shafique Tanzil, van Oostrum Matthijs, & Chatterjee Ishita. (2021). Informal settlement is not a euphemism for 'slum': What's at stake beyond the language? International Development Planning Review, 43(2), 139–150. https://doi.org/10.3828/idpr.2020.14

Doxsey-Whitfield, E., MacManus, K., Adamo, S. B., Pistolesi, L., Squires, J., Borkovska, O., & Baptista, S. R. (2015). Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. Papers in Applied Geography, 1(3), 226–234. https://doi.org/10.1080/23754931.2015.1014272

Driscoll, K., & Walker, S. (2014). Big Data, Big Questions| Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data. International Journal of Communication, 8, 1745–1764.

Droin, A., Wurm, M., & Sulzer, W. (2020). Semantic labelling of building types. A comparison of two approaches using Random Forest and Deep Learning. 40. Wissenschaftlich-Technische Jahrestagung Der DGPF in Stuttgart – Publikationen Der DGPF, 29, 527–538. https://dgpf.de/src/tagung/jt2020/proceedings/proceedings/papers/76_KKNP_DGPF2020_Droin_et_al.pdf

Droin, A., Wurm, M., & Taubenböck, H. (2023). The Individual Walkable Neighborhood—Evaluating people-centered spatial units focusing on urban density. Computers, Environment and Urban Systems, 99(101893), 1–10. https://doi.org/10.1016/j.compenvurbsys.2022.101893

Drouyer, S., & de Franchis, C. (2020). Parking occupancy estimation on Sentinel-1 images. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2–2020, 821–828. https://doi.org/10.5194/isprs-annals-V-2-2020-821-2020

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., & Bargellini, P. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. The Sentinel Missions - New Opportunities for Science, 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

du Toit, M. J., & Cilliers, S. S. (2011). Aspects influencing the selection of representative urbanization measures to quantify urban–rural gradients. Landscape Ecology, 26(2), 169–181. https://doi.org/10.1007/s10980-010-9560-4

Duda, K. A., & Jones, B. K. (2011). USGS Remote Sensing Coordination for the 2010 Haiti Earthquake. Photogrammetric Engineering & Remote Sensing, 77(9), 899–907. https://doi.org/10.14358/PERS.77.9.899

Dueker, K. J., & Horton, F. E. (1972). Urban-change detection systems: Remote-sensing inputs. Photogrammetria, 28(3), 89–106. https://doi.org/10.1016/0031-8663(72)90013-0

Duncan, B. N., Yoshida, Y., Foy, B. de, Lamsal, L. N., Streets, D. G., Lu, Z., Pickering, K. E., & Krotkov, N. A. (2013). The observed response of Ozone Monitoring Instrument (OMI) NO2 columns to NOx emission controls on power plants in the United States: 2005–2011. Atmospheric Environment, 81, 102–111. https://doi.org/10.1016/j.atmosenv.2013.08.068

Dunkel. (2021). Mehrheit der Bayern ist für dritten Nationalpark. https://www.sueddeutsche.de/bayern/umfrage-mehrheit-der-bayern-ist-fuer-dritten-nationalpark-1.3902252

Dupras, J., Marull, J., Parcerisas, L., Coll, F., Gonzalez, A., Girard, M., & Tello, E. (2016). The impacts of urban sprawl on ecological connectivity in the Montreal Metropolitan Region. Environmental Science & Policy, 58, 61–73. https://doi.org/10.1016/j.envsci.2016.01.005

E

Ebrahimian, A., Gulliver, J. S., & Wilson, B. N. (2016). Effective impervious area for runoff in urban watersheds. Hydrological Processes, 30(20), 3717–3729. https://doi.org/10.1002/hyp.10839

Eckert, S., & Kohler, S. (2014). Urbanization and Health in Developing Countries: A Systematic Review. World Health & Population, 15(1), 7–20. https://doi.org/10.12927/whp.2014.23722

Eeckhout, J. (2004). Gibrat's Law for (All) Cities. American Economic Review, 94(5), 1429–1451. https://doi.org/10.1257/0002828043052303

Egidi, G., Salvati, L., & Vinci, S. (2020). The long way to tipperary: City size and worldwide urban population trends, 1950–2030. Sustainable Cities and Society, 60(102148), 1–6. https://doi.org/10.1016/j.scs.2020.102148

Ehlers, E. (1993a). Die Stadt des Islaminischen Orients. Modell und Wirklichkeit [The town of the Islamic Orient. Model and reality]. Geographische Rundschau, 45(1), 32–39.

Ehlers, E. (1993b). Die Stadt des islamischen Orients. Modell und Wirklichkeit. Geographische Rundschau, 45(1), 32–39.

Einsiedel, E. F., & Adamson, H. (2012). Stem cell tourism and future stem sell tourists: Policy and ethical implications. Developing World Bioethics, 12(1), 35–44. https://doi.org/10.1111/j.1471-8847.2012.00319.x

Eiras, C. G. S., Souza, J. R. G. de, Freitas, R. D. A. de, Barella, C. F., & Pereira, T. M. (2021). Discriminant analysis as an efficient method for landslide susceptibility assessment in cities with the scarcity of predisposition data. Natural Hazards, 107(2), 1427–1442. https://doi.org/10.1007/s11069-021-04638-4

Eisl, M., Mansberger, G., Matzanetz, P., & Schreilechner, P. (2016). CITIES: Brennpunkte der Menschheit. eoVision.

Ekim, B., Dong, Z., Rashkovetsky, D., & Schmitt, M. (2021). The naturalness index for the identification of natural areas on regional scale. International Journal of Applied Earth Observation and Geoinformation, 105, 1–8. https://doi.org/10.1016/j.jag.2021.102622

Elghazel, H., Aussem, A., & Perraud, F. (2011). Trading-Off Diversity and Accuracy for Optimal Ensemble Tree Selection in Random Forests. In O. Okun, G. Valentini, & M. Re (Eds.), Ensembles in Machine Learning Applications (pp. 169–179). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-22910-7_10

Elhacham, E., Ben-Uri, L., Grozovski, J., Bar-On, Y. M., & Milo, R. (2020). Global human-made mass exceeds all living biomass. Nature, 588(7838), 442–444. https://doi.org/10.1038/s41586-020-3010-5

Eliasson, I. (1992). Infrared thermography and urban temperature patterns. International Journal of Remote Sensing, 13(5), 869–879. https://doi.org/10.1080/01431169208904160

Ellaway, A., Macintyre, S., & Bonnefoy, X. (2005). Graffiti, greenery, and obesity in adults: Secondary analysis of European cross sectional survey. BMJ, 331(7517), 611–612. https://doi.org/10.1136/bmj.38575.664549.F7

Elvidge, C. D. (n.d.). Version 4 DMSP-OLS Nighttime Lights Time Series [Dataset]. Retrieved 9 August 2022, from https://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., & Davis, E. R. (1997). Mapping city lights with nighttime data from the DMSP Operational Linescan System. Photogrammetric Engineering and Remote Sensing, 63(6), 727–734. Scopus.

Elvidge, C. D., Baugh, K. E., Zhizhin, M. N., & Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights. 35. Proceedings of the Asia-Pacific Advanced Network, 62–69. https://doi.org/10.7125/APAN.35.7

Elvidge, C. D., Tuttle, B. T., Sutton, P. C., Baugh, K. E., Howard, A. T., Milesi, C., Bhaduri, B., & Nemani, R. (2007). Global Distribution and Density of Constructed Impervious Surfaces. Sensors, 7(9), 1962–1979. https://doi.org/10.3390/s7091962

Em, P. P. (2018). A Big City as an Independent Central Place System, a Case Study of Moscow. Regional Research of Russia, 8(2), 151–157. https://doi.org/10.1134/S2079970518020041

Engel-Cox, J. A., Holloman, C. H., Coutant, B. W., & Hoff, R. M. (2004). Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. Atmospheric Environment, 38(16), 2495–2509. https://doi.org/10.1016/j.atmosenv.2004.01.039

Engelmeier, T. F., & Gaihre, N. R. (2019). Going the Distance: Off-Grid Lighting Market Dynamics in Papua New Guinea. International Finance Corporation (IFC). https:// www.ifc.org/wps/wcm/connect/region__ext_content/ifc_ external_corporate_site/east+asia+and+the+pacific/resources/ lighting+png+market+report

Erbertseder, T., Bergemann, C., & Baier, F. (2015). Der Wochenrhythmus der Städte – Erfassung anthropogener Aktivitätsmuster aus dem All. In H. Taubenböck, M. Wurm, T. Esch, & S. Dech (Eds.), Globale Urbanisierung: Perspektive aus dem All (pp. 151–158). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-44841-0_16

Erbertseder, T., Taubenböck, H., Esch, T., Gilardi, L., Paeth, H., & Dech, S. (2024). NO2 Air Pollution Trends and Settlement Growth in Megacities. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 12058–12076. https://doi.org/10.1109/JSTARS.2024.3419573

Erbertseder, T., Taubenböck, H., Esch, T., Gilardi, L., Paeth, H., Marconcini, M., & Dech, S. (2023). Earth Observation-based analysis of NO2 pollution and settlement growth in megacities. 2023 Joint Urban Remote Sensing Event (JURSE), 1–4. https://doi.org/10.1109/JURSE57346.2023.10144190

Erbertseder, T., Taubenböck, H., & Meyer-Arnek, J. (2015). Dicke Luft – Stadtregionen als globale Zentren der Luftverschmutzung. In H. Taubenböck, M. Wurm, T. Esch, & S. Dech (Eds.), Globale Urbanisierung: Perspektive aus dem All (pp. 191–203). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-44841-0_20

Erener, A., Mutlu, A., & Sebnem Düzgün, H. (2016). A comparative study for landslide susceptibility mapping using GIS-based multi-criteria decision analysis (MCDA), logistic regression (LR) and association rule mining (ARM). Special Issue on Probabilistic and Soft Computing Methods for Engineering Geology, 203, 45–55. https://doi.org/10.1016/j.enggeo.2015.09.007

ESA - European Space Agency. (2012). Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Service. http://esamultimedia.esa.int/multimedia/publications/SP- 1322_2/offline/download.pdf

Esch, T., Bachofer, F., Heldens, W., Hirner, A., Marconcini, M., Palacios-Lopez, D., Roth, A., Üreyen, S., Zeidler, J., Dech, S., & Gorelick, N. (2018). Where We Live—A Summary of the Achievements and Planned Evolution of the Global Urban Footprint. Remote Sensing, 10(6), 1–18. https://doi.org/10.3390/rs10060895

Esch, T., Brzoska, E., Dech, S., Leutner, B., Palacios-Lopez, D., Metz-Marconcini, A., Marconcini, M., Roth, A., & Zeidler, J. (2022). World Settlement Footprint 3D - A first three-dimensional survey of the global building stock. Remote Sensing of Environment, 270(112877), 1–15. https://doi.org/10.1016/j.rse.2021.112877

Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., & Strano, E. (2017). Breaking new ground in mapping human settlements from space – The Global Urban Footprint. ISPRS Journal of Photogrammetry and Remote Sensing, 134, 30–42. https://doi.org/10.1016/j.isprsjprs.2017.10.012

Esch, T., Himmler, V., Schorcht, G., Thiel, M., Wehrmann, T., Bachofer, F., Conrad, C., Schmidt, M., & Dech, S. (2009). Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. Remote Sensing of Environment, 113(8), 1678–1690. https://doi.org/10.1016/j.rse.2009.03.012

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenbock, H., Muller, A., & Dech, S. (2013). Urban footprint processor-Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. In IEEE Geoscience and Remote Sensing Letters (Vol. 10, Issue 6, pp. 1617–1621). https://doi.org/10.1109/LGRS.2013.2272953

Esch, T., Marconcini, M., Marmanis, D., Zeidler, J., Elsayed, S., Metz, A., Müller, A., & Dech, S. (2014). Dimensioning urbanization – An advanced procedure for characterizing human settlement properties and patterns using spatial network analysis. Applied Geography, 55, 212–228. https://doi.org/10.1016/j.apgeog.2014.09.009

Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felbier, A., Schmidt, M., Mueller, A. A., Thiel, M., & Dech, S. W. (2012). TanDEM-X mission-new perspectives for the inventory and monitoring of global settlement patterns. Journal of Applied Remote Sensing, 6/1(061702), 1–21. https://doi.org/10.1117/1.JRS.6.061702

Esch, T., Üreyen, S., Zeidler, J., Metz–Marconcini, A., Hirner, A., Asamer, H., Tum, M., Böttcher, M., Kuchar, S., Svaton, V., & Marconcini, M. (2018). Exploiting big earth data from space – first experiences with the timescan processing chain. Big Earth Data, 2(1), 36–55. https://doi.org/10.1080/20964471.2018.1433790

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, 226–231.

European Commission. (2008). Why a Covenant of Mayors? https://eu-mayors.ec.europa.eu/en/about

European Commission. (2018). Ethics guidelines for trustworthy AI: High-level expert group on artificial intelligence. https://ec.europa.eu/ futurium/en/ai-alliance-consultation.1.html

European Commission. (2019a). Directive (EU) 2019/790 of the European Parliament and the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC. https://eur-lex.europa.eu/legal-content/ EN/TXT/PDF/?uri=CELEX:32019L0790

European Commission. (2019b). The European Green Deal. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

European Commission. (2020). Developments and Forecasts on Continuing Urbanisation. https://knowledge4policy.ec.europa.eu/foresight/topic/continuing-urbanisation/developments-and-forecasts-on-continuing-urbanisation_en

European Commission: Directorate-General for Environment, Alterra, Eurosite, & PAN Parks Foundation. (2013). Guidelines on wilderness in Natura 2000 – Management of terrestrial wilderness and wild areas within the Natura 2000 network. Publications Office. https://doi.org/doi/10.2779/33572

European Commission, & High-Level Expert Group on AI. (2018). Ethics guidelines for trustworthy AI (Independent High-Level Expert Group on Artificial Intelligence). https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Environment Agency. (2002). Europe's biodiversity—Biogeographical regions and seas (No. EEA Report 1/2002). https://www.eea.europa.eu/publications/report_2002_0524_154909

European Environment Agency. (2022). Air Quality e-reporting (AQ e-reporting) [Dataset]. https://www.eea.europa.eu/en/datahub/ datahubitem-view/3b390c9c-f321-490a-b25a-ae93b2ed80c1

European Environment Agency, & Swiss Federal Office for the Environment (FOEN). (2016). Urban sprawl in Europe – Joint EEA-FOEN report. Publications Office. https://data.europa.eu/doi/10.2800/143470

European Environmental Agency (EEA). (2024). Urban Atlas. https://www.eea.europa. eu/data-and-maps/figures/urban-atlas

European Parliament: Directorate-General for Internal Policies of the Union, & Rosati, E. (2018). The exception for text and data mining (TDM) in the proposed Directive on Copyright in the Digital Single Market – Technical aspects. European Parliament. https://doi.org/10.2861/480649

European Space & Agency. (2019). Blockchain and Earth observation [White Paper]. https://eo4society.esa.int/wp-content/uploads/2019/04/Blockchain-and-Earth-Observation_White-Paper-April-2019.pdf

European Space Agency. (2023). Sentinel-5P TROPOMI user guide. https://sentiwiki.copernicus.eu/web/sentinel-5p

European Space Agency (ESA). (n.d.). GlobCover 2009 (Global Land Cover Map) [Dataset]. Retrieved 9 August 2022, from http://due.esrin.esa.int/page_ globcover.php

Eurostat. (2015). LUCAS 2015 (Land Use/Cover Area Frame Survey) – Technical reference document C1: Instructions for Surveyors (Technical Report. Eurostat, Regional Statistics and Geographic Information). https://ec.europa.eu/eurostat/documents/205002/6786255/LUCAS 2015-C1-Instructions-20150227.pdf

Eurostat. (2018). Urban Audit 2018 [Dataset]. https://ec.europa.eu/eurostat/web/ gisco/geodata/statistical-units/urban-audit

Eurostat. (2021). Applying the degree of urbanisation – A methodological manual to define cities, towns and rural areas for international comparisons – 2021 edition. Publications Office of the European Union. https://doi.org/10.2785/706535

Ewing, R., Meakins, G., Hamidi, S., & Nelson, A. C. (2014). Relationship between urban sprawl and physical activity, obesity, and morbidity – Update and refinement. Health & Place, 26, 118–126. https://doi.org/10.1016/j.healthplace.2013.12.008

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

F

Fan, C., Tian, L., Zhou, L., Hou, D., Song, Y., Qiao, X., & Li, J. (2018). Examining the impacts of urban form on air pollutant emissions: Evidence from China. Journal of Environmental Management, 212, 405–414. https://doi.org/10.1016/j.jenvman.2018.02.001

Fan, Y., Yu, G., & He, Z. (2017). Origin, spatial pattern, and evolution of urban system: Testing a hypothesis of "urban tree". Habitat International, 59, 60–70. https://doi.org/10.1016/j.habitatint.2016.11.012

Fang, C., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. Landscape and Urban Planning, 162, 126–136. https://doi.org/10.1016/j.landurbplan.2017.02.014

Faryadi, S., & Taheri, S. (2009). Interconnections of urban green spaces and environmental quality of Tehran. International Journal of Environmental Research, 3(2), 199–208.

Fauvel, M., Benediktsson, J. A., Chanussot, J., & Sveinsson, J. R. (2008). Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. In IEEE Transactions on Geoscience and Remote Sensing (Vol. 46, Issue 11, pp. 3804–3814). https://doi.org/10.1109/TGRS.2008.922034

Federal Ministry of the Interior and Community (BMI). (2018). Recommendations of the Data Ethics Commission for the Federal Government's Strategy on Artificial Intelligence [Technical Report]. https://www.bmi.bund.de/SharedDocs/downloads/EN/themen/it-digital-policy/recommendations-data-ethics-commission.pdf?__blob=publicationFile&v=3

Feicht, T., & Schmitt, A. (2022). Anamorphic Visualization of the Space-Time-Paradigm by Geodetic Network Adjustment. KN - Journal of Cartography and Geographic Information, 73(1), 39–50. https://doi.org/10.1007/s42489-022-00124-4

Fekete, A., Tzavella, K., Armas, I., Binner, J., Garschagen, M., Giupponi, C., Mojtahed, V., Pettita, M., Schneiderbauer, S., & Serre, D. (2015). Critical Data Source; Tool or Even Infrastructure? Challenges of Geographic Information Systems and Remote Sensing for Disaster Risk Governance. ISPRS International Journal of Geo-Information, 4(4), 1848–1869. https://doi.org/10.3390/ijgi4041848

Felbermayr, G., Gröschl, J., & Steinwachs, T. (2017). The Trade Effects of Border Controls: Evidence from the European Schengen Agreement. JCMS: Journal of Common Market Studies, 56(2), 335–351. https://doi.org/10.1111/jcms.12603

Felson, A. J., Bradford, M. A., & Terway, T. M. (2013). Promoting Earth Stewardship through urban design experiments. Frontiers in Ecology and the Environment, 11(7), 362–367. https://doi.org/10.1890/130061

Feng, J., Glass, T. A., Curriero, F. C., Stewart, W. F., & Schwartz, B. S. (2010). The built environment and obesity: A systematic review of the epidemiologic evidence. Health & Place, 16(2), 175–190. https://doi.org/10.1016/j.healthplace.2009.09.008

Ferrell, O. C., & Fraedrich, J. (2015). Business Ethics: Ethical Decision Making & Cases. Nelson Education.

Field, C. B., Barros, V., Stocker, T. F., & Dahe, Q. (Eds.). (2012). Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/CBO9781139177245

Finke, L. (1994). Landschaftsökologie / Das Geographische Seminar (3rd ed.). Westermann Schulbuchverlag.

Fiorentini, N., & Losa, M. (2025). Road Detection, Monitoring, and Maintenance Using Remotely Sensed Data. Remote Sensing, 17(5)(917), 1–7. https://doi.org/10.3390/rs17050917

Fisher, R. A. (1925). Statistical Methods for Research Workers. Oliver and Boyd.

Fleischmann, M. (2019). momepy: Urban Morphology Measuring Toolkit. Journal of Open Source Software, 4(43)(1807), 1–4. https://doi.org/10.21105/joss.01807

Fleischmann, M., Feliciotti, A., & Kerr, W. (2021). Evolution of Urban Patterns: Urban Morphology as an Open Reproducible Data Science. Geographical Analysis, 54(3), 536–558. https://doi.org/10.1111/gean.12302

Fleischmann, M., Feliciotti, A., Romice, O., & Porta, S. (2020). Morphological tessellation as a way of partitioning space: Improving consistency in urban morphology at the plot scale. Computers, Environment and Urban Systems, 80(101441), 1–13. https://doi.org/10.1016/j.compenvurbsys.2019.101441

Fleischmann, M., Romice, O., & Porta, S. (2021). Measuring urban form: Overcoming terminological inconsistencies for a quantitative and comprehensive morphologic analysis of cities. Environment and Planning B: Urban Analytics and City Science, 48(8), 2133–2150. https://doi.org/10.1177/2399808320910444

Fletcher, T. D., Shuster, W., Hunt, W. F., Ashley, R., Butler, D., Arthur, S., Trowsdale, S., Barraud, S., Semadeni-Davies, A., Bertrand-Krajewski, J.-L., Mikkelsen, P. S., Rivard, G., Uhl, M., Dagenais, D., & Viklander, M. (2014). SUDS, LID, BMPs, WSUD and more – The evolution and application of terminology surrounding urban drainage. Urban Water Journal, 12(7), 525–542. https://doi.org/10.1080/1573062X.2014.916314

Florida, R. (2002). The rise of the creative class: And how it's transforming work, leisure, community, and everyday life. Basic Books.

Florida, R., Gulden, T., & Mellander, C. (2008). The rise of the mega-region. Cambridge Journal of Regions, Economy and Society, 1(3), 459–476. https://doi.org/10.1093/cjres/rsn018

Florida, R., Mellander, C., & Gulden, T. (2012). Global Metropolis: Assessing Economic Activity in Urban Centers Based on Nighttime Satellite Images. The Professional Geographer, 64(2), 178–187. https://doi.org/10.1080/00330124.2011.583590

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Forman, R. T. T. (2014). Urban Ecology: Science of Cities. Cambridge University Press.

Forman, R. T. T., & Alexander, L. E. (1998). Roads and their major ecological effects. In Annual Review of Ecology, Evolution, and Systematics (Vol. 29, Issue Volume 29, 1998, pp. 207–231). Annual Reviews. https://doi.org/10.1146/annurev.ecolsys.29.1.207

Forman, R. T. T., & Godron, M. (1986). Landscape Ecology. John Wiley & Sons.

Forster, B. (1983). Some urban measurements from Landsat data (Sydney, Australia). Photogrammetric Engineering & Remote Sensing, 49(12), 1693–1707.

Forzieri, G., Cescatti, A., e Silva, F. B., & Feyen, L. (2017). Increasing risk over time of weather-related hazards to the European population: A data-driven prognostic study. The Lancet Planetary Health, 1(5), e200–e208. https://doi.org/10.1016/S2542-5196(17)30082-7

Fotheringham, A. S., & Rogerson, P. A. (1993). GIS and spatial analytical problems. International Journal of Geographical Information Systems, 7(1), 3–19. https://doi.org/10.1080/02693799308901936

Fragkias, M., & Seto, K. C. (2009). Evolving rank-size distributions of intra-metropolitan urban clusters in South China. Computers, Environment and Urban Systems, 33(3), 189–199. https://doi.org/10.1016/j.compenvurbsys.2008.08.005

Frank, L. D., & Engelke, P. O. (2001). The Built Environment and Human Activity Patterns: Exploring the Impacts of Urban Form on Public Health. Journal of Planning Literature, 16(2), 202–218. https://doi.org/10.1177/08854120122093339

Frantz, D., Schug, F., Okujeni, A., Navacchi, C., Wagner, W., van der Linden, S., & Hostert, P. (2020). National-scale mapping of building height using Sentinel-1 and Sentinel-2 time series. Remote Sensing of Environment, 252(112128), 1–17. https://doi.org/10.1016/j.rse.2020.112128

Freire, S., Macmanus, K., Pesaresi, M., Doxsey-Whitfield, E., & Mills, J. (2016). Development of new open and free multi-temporal global population grids at 250 m resolution. Geospatial Data in a Changing World. Association of Geographic Information Laboratories in Europe (AGILE) (Organiser). AGILE; 2016. https://publications.jrc.ec.europa.eu/repository/handle/JRC100523

French, S. K., Giacinti, J. A., Robinson, S. J., Pearl, D. L., & Jardine, C. M. (2022). The urban myth: A lack of agreement between definitions of urban environments used in wildlife health research may contribute to inconsistent epidemiological findings. Urban Ecosystems, 25(3), 999–1005. https://doi.org/10.1007/s11252-022-01213-y

Frias-Martinez, V., Soto, V., Hohwald, H., & Frias-Martinez, E. (2012). Characterizing Urban Landscapes Using Geolocated Tweets. 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing, 239–248. https://doi.org/10.1109/SocialCom-PASSAT.2012.19

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: algorithms and early results. The Moderate Resolution Imaging Spectroradiometer (MODIS): A New Generation of Land Surface Monitoring, 83(1), 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Bellouin, N., … Zeng, J. (2021). Global Carbon Budget 2021. Earth System Science Data, 14(4), 1917–2005. https://doi.org/10.5194/essd-14-1917-2022

Friedman, T. L. (2005). The World Is Flat: A Brief History of the Twenty-first Century. Farrar, Straus and Giroux.

Frigui, H., & Krishnapuram, R. (1997). Clustering by competitive agglomeration. Pattern Recognition, 30(7), 1109–1119. https://doi.org/10.1016/S0031-3203(96)00140-9

Fripp, E. (2014). Payments for Ecosystem Services (PES): A practical guide to assessing the feasibility of PES projects. Center for International Forestry Research (CIFOR). https://www.cifor.org/publications/pdf_files/Books/BFripp1401.pdf

Frolking, S., Milliman, T., Seto, K. C., & Friedl, M. A. (2013). A global fingerprint of macro-scale changes in urban structure from 1999 to 2009. Environmental Research Letters, 8(024004), 1–10. https://doi.org/10.1088/1748-9326/8/2/024004

Fu, P., & Weng, Q. (2016). A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. Remote Sensing of Environment, 175, 205–214. https://doi.org/10.1016/j.rse.2015.12.040

Fu, Y., Li, J., Weng, Q., Zheng, Q., Li, L., Dai, S., & Guo, B. (2019). Characterizing the spatial pattern of annual urban growth by using time series Landsat imagery. Science of The Total Environment, 666, 274–284. https://doi.org/10.1016/j.scitotenv.2019.02.178

Fujita, M., Krugman, P., & Mori, T. (1999). On the evolution of hierarchical urban systems. European Economic Review, 43(2), 209–251. https://doi.org/10.1016/S0014-2921(98)00066-X

Fujita, M., & Ogawa, H. (1982). Multiple equilibria and structural transition of non-monocentric urban configurations. Regional Science and Urban Economics, 12(2), 161–196. https://doi.org/10.1016/0166-0462(82)90031-X

Fuller, R. A., & Gaston, K. J. (2009). The scaling of green space coverage in European cities. Biology Letters, 5(3), 352–355. https://doi.org/10.1098/rsbl.2009.0010

G

GADM. (2018). Database of global administrative areas. GADM, Montreal, QC, Canada [Dataset].

Gaitani, N., Burud, I., Thiis, T., & Santamouris, M. (2017). High-resolution spectral mapping of urban thermal properties with Unmanned Aerial Vehicles. Building and Environment, 121, 215–224. https://doi.org/10.1016/j.buildenv.2017.05.027

Gál, T., Bechtel, B., & Unger, J. (2015). Comparison of two different Local Climate Zone mapping methods. ICUC9 - 9th International Conference on Urban Climate Jointly with 12th Symposium on the Urban Environment, 1–6.

Gallo, K. P., & Owen, T. W. (1998). Assessment of urban heat Islands: A multi-sensor perspective for the Dallas-Ft. Worth, USA region. Geocarto International, 13(4), 35–41. https://doi.org/10.1080/10106049809354662

Gallo, K. P., Tarpley, J. D., McNab, A. L., & Karl, T. R. (1995). Assessment of urban heat islands: A satellite perspective. Minimax Workshop, 37(1), 37–43. https://doi.org/10.1016/0169-8095(94)00066-M

Gamperl, M., Singer, J., & Thuro, K. (2021). Internet of Things Geosensor Network for Cost-Effective Landslide Early Warning Systems. Sensors, 21(2609), 1–23. https://doi.org/10.3390/s21082609

Ganguly, Š. (1999). India's Pathway to Pokhran II: The Prospects and Sources of New Delhi's Nuclear Weapons Program. International Security, 23(4), 148–177. https://doi.org/10.1162/isec.23.4.148

Gao, F., Brown de Colstoun, E., Ma, R., Weng, Q., Masek, J. G., Chen, J., Pan, Y., & Song, C. (2012). Mapping impervious surface expansion using medium-resolution satellite image time series: A case study in the Yangtze River Delta, China. International Journal of Remote Sensing, 33(24), 7609–7628. https://doi.org/10.1080/01431161.2012.700424

Gao, F., J. Masek, M. Schwaller, & F. Hall. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. IEEE Transactions on Geoscience and Remote Sensing, 44(8), 2207–2218. https://doi.org/10.1109/TGRS.2006.872081

Garcia, R. A. C., Oliveira, S. C., & Zêzere, J. L. (2016). Assessing population exposure for landslide risk analysis using dasymetric cartography. Natural Hazards and Earth System Sciences, 16(12), 2769–2782. https://doi.org/10.5194/nhess-16-2769-2016

Garcia-López, M.-À. (2010). Population suburbanization in Barcelona, 1991–2005: Is its spatial structure changing? Journal of Housing Economics, 19(2), 119–132. https://doi.org/10.1016/j.jhe.2010.04.002

Garcia-López, M.-À., & Muñiz, I. (2010). Employment Decentralisation: Polycentricity or Scatteration? The Case of Barcelona. Urban Studies, 47(14), 3035–3056. https://doi.org/10.1177/0042098009360229

Garreau, J. (1991). Edge City: Life on the New Frontier. Doubleday.

Gascon, M., Triguero-Mas, M., Martínez, D., Dadvand, P., Rojas-Rueda, D., Plasència, A., & Nieuwenhuijsen, M. J. (2016). Residential green spaces and mortality: A systematic review. Environment International, 86, 60–67. https://doi.org/10.1016/j.envint.2015.10.013

Gatti, A., & Bertolini, A. (2013). Sentinel-2 Products Specification Document. https://earth.esa.int/documents/247904/685211/Sentinel-2+Products+Specification+Document

Gaubatz, P. (1998). Understanding Chinese urban form: Contexts for interpreting continuity and change. Built Environment, 24(4), 251–270.

Gaurav, K., Sinha, R., & Panda, P. K. (2011). The Indus flood of 2010 in Pakistan: A perspective analysis using remote sensing data. Natural Hazards, 59(3), 1815–1826. https://doi.org/10.1007/s11069-011-9869-6

Geddes, J. A., Martin, R. V., Boys, B. L., & Donkelaar, A. van. (2016). Long-Term Trends Worldwide in Ambient NO2 Concentrations Inferred from Satellite Observations. Environmental Health Perspectives, 124(3), 281–289. https://doi.org/10.1289/ehp.1409567

Gehl, J. (2009). Cities for People. Island Press.

Geiß, C., Leichtle, T., Wurm, M., Pelizari, P. A., Standfuß, I., Zhu, X. X., So, E., Siedentop, S., Esch, T., & Taubenböck, H. (2019). Large-Area Characterization of Urban Morphology—Mapping of Built-Up Height and Density Using TanDEM-X and Sentinel-2 Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(8), 2912–2927. https://doi.org/10.1109/JSTARS.2019.2917755

Geiß, C., Priesmeier, P., Aravena Pelizari, P., Soto Calderon, A. R., Schoepfer, E., Riedlinger, T., Villar Vega, M., Santa María, H., Gómez Zapata, J. C., Pittore, M., So, E., Fekete, A., & Taubenböck, H. (2022). Benefits of global earth observation missions for disaggregation of exposure data and earthquake loss modeling: Evidence from Santiago de Chile. Natural Hazards, 119(2), 779–804. https://doi.org/10.1007/s11069-022-05672-6

Geiß, C., Schauß, A., Riedlinger, T., Dech, S., Zelaya, C., Guzmán, N., Hube, M. A., Arsanjani, J. J., & Taubenböck, H. (2017). Joint use of remote sensing data and volunteered geographic information for exposure estimation: Evidence from Valparaíso, Chile. Natural Hazards, 86(1), 81–105. https://doi.org/10.1007/s11069-016-2663-8

Geiß, C., Wurm, M., & Taubenböck, H. (2017). Towards large-area morphologic characterization of urban environments using the TanDEM-X mission and Sentinel-2. 2017 Joint Urban Remote Sensing Event (JURSE), 1–4. https://doi.org/10.1109/JURSE.2017.7924543

Geletič, J., & Lehnert, M. (2016). GIS-based delineation of local climate zones: The case of medium-sized Central European cities. Moravian Geographical Reports, 24(3), 2–12. https://doi.org/10.1515/mgr-2016-0012

Gens, R. (2010). Remote sensing of coastlines: Detection, extraction and monitoring. International Journal of Remote Sensing, 31(7), 1819–1836. https://doi.org/10.1080/01431160902926673

Georg, I., Blaschke, T., & Taubenböck, H. (2016a). A Global Inventory of Urban Corridors Based on Perceptions and Night-Time Light Imagery. ISPRS International Journal of Geo-Information, 5(233), 1–19. https://doi.org/10.3390/ijgi5120233

Georg, I., Blaschke, T., & Taubenböck, H. (2016b). New spatial dimensions of global cityscapes: From reviewing existing concepts to a conceptual spatial approach. Journal of Geographical Sciences, 26(3), 355–380. https://doi.org/10.1007/s11442-016-1273-4

Georg, I., Blaschke, T., & Taubenböck, H. (2018). Are We in Boswash Yet? A Multi-Source Geodata Approach to Spatially Delimit Urban Corridors. ISPRS International Journal of Geo-Information, 7(15), 1–19. https://doi.org/10.3390/ijgi7010015

Georg, I., Blaschke, T., & Taubenböck, H. (2023). Spatial delineation of urban corridors in North America: An approach incorporating fuzziness based on multi-source geospatial data. Cities, 133, 1–16. https://doi.org/10.1016/j.cities.2022.104129

Georgoulias, A. K., van der A, R. J., Stammes, P., Boersma, K. F., & Eskes, H. J. (2019). Trends and trend reversal detection in 2 decades of tropospheric \chemNO_2 satellite observations. Atmospheric Chemistry and Physics, 19(9), 6269–6294. https://doi.org/10.5194/acp-19-6269-2019

Gerundo, R., & Grimaldi, M. (2011). The Measure of Land Consumption Caused by Urban Planning. 2011 International Conference on Green Buildings and Sustainable Cities, 21, 1152–1160. https://doi.org/10.1016/j.proeng.2011.11.2124

Ghamisi, P., Yu, W., Marinoni, A., Gevaert, C. M., Persello, C., Selvakumaran, S., Girotto, M., Horton, B. P., Rufin, P., Hostert, P., Pacifici, F., & Atkinson, P. M. (2025). Responsible Artificial Intelligence for Earth Observation: Achievable and realistic paths to serve the collective good. IEEE Geoscience and Remote Sensing Magazine, 2–26. https://doi.org/10.1109/MGRS.2025.3529726

Ghosh, T., Powell, R. L., Elvidge, C. D., Baugh, K. E., Sutton, P. C., & Anderson, S. (2010). Shedding Light on the Global Distribution of Economic Activity. The Open Geography Journal, 3, 147–160. https://doi.org/10.2174/1874923201003010147

Gibson, J. J. (1979). The Ecological Approach to Visual Perception. Houghton-Mifflin.

Giessner, S., & van Quaquebeke, N. (2010). Using a Relational Models Perspective to Understand Normatively Appropriate Conduct in Ethical Leadership. Journal of Business Ethics, 95(1), 43–55. https://doi.org/10.1007/s10551-011-0790-4

Giffinger, R., Haindlmaier, G., & Kramar, H. (2010). The role of rankings in growing city competition. Urban Research & Practice, 3(3), 299–312. https://doi.org/10.1080/17535069.2010.524420

Gilardi, L., Marconcini, M., Metz-Marconcini, A., Esch, T., & Erbertseder, T. (2023). Long-term exposure and health risk assessment from air pollution: Impact of regional scale mobility. International Journal of Health Geographics, 22(11), 1–14. https://doi.org/10.1186/s12942-023-00333-8

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., Lange, A., & Donovan, R. J. (2005). Increasing walking: How important is distance to, attractiveness, and size of public open space? Active Living Research, 28(2, Supplement 2), 169–176. https://doi.org/10.1016/j.amepre.2004.10.018

Giles-Corti, B., Vernez-Moudon, A., Reis, R., Turrell, G., Dannenberg, A. L., Badland, H., Foster, S., Lowe, M., Sallis, J. F., Stevenson, M., & Owen, N. (2016). City planning and population health: A global challenge. The Lancet, 388(10062), 2912–2924. https://doi.org/10.1016/S0140-6736(16)30066-6

Gill, S. E., Handley, J. F., Ennos, A. R., & Pauleit, S. (2007). Adapting Cities for Climate Change: The Role of the Green Infrastructure. Built Environment, 33(1), 115–133. https://doi.org/doi:10.2148/benv.33.1.115

Giuliani, G., Chatenoux, B., Benvenuti, A., Lacroix, P., Marconcini, M., & Mazzetti, P. (2020). Monitoring land degradation at national level using satellite Earth Observation time-series data to support SDG15 – exploring the potential of data cube. Big Earth Data, 4(1), 3–22. https://doi.org/10.1080/20964471.2020.1711633

Glaeser, E. (2010). Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier. Penguin Press.

Glaeser, E. (2011). Triumph of the City: How Urban Spaces Make Us Human. Pan Macmillan.

Glaeser, E. (2013). Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier (an excerpt). Journal of Economic Sociology, 14(4), 75–94.

Glaeser, E., & Joshi-Ghani, A. (2015). The urban imperative—Towards competitive cities. Oxford University Press.

Glaeser, E. L., & Kahn, M. E. (2004). Sprawl and Urban Growth. In J. V. Henderson & J.-F. Thisse (Eds.), Handbook of Regional and Urban Economics (Vol. 4, pp. 2481–2527). Elsevier. https://doi.org/10.1016/S1574-0080(04)80013-0

Glaeser, E. L., & Kahn, M. E. (2010). The greenness of cities: Carbon dioxide emissions and urban development. Journal of Urban Economics, 67(3), 404–418. https://doi.org/10.1016/j.jue.2009.11.006

Glantz, S. A. (2012). Primer of Biostastistics (7th ed.). McGraw Hill.

Glocker, D. (2018). The Rise of Megaregions: Delineating a new scale of economic geography. OECD Regional Development Working Papers, No. 2018/04, 1–53. https://doi.org/10.1787/f4734bdd-en

Gluch, R., Quattrochi, D. A., & Luvall, J. C. (2006). A multi-scale approach to urban thermal analysis. Remote Sensing of Environment, 104(2), 123–132. https://doi.org/10.1016/j.rse.2006.01.025

Gober, P., Brazel, A., Quay, R., Myint, S., Grossman-Clarke, S., Miller, A., & Rossi, S. (2009). Using Watered Landscapes to Manipulate Urban Heat Island Effects: How Much Water Will It Take to Cool Phoenix? Journal of the American Planning Association, 76(1), 109–121. https://doi.org/10.1080/01944360903433113

Goddard, M. A., Dougill, A. J., & Benton, T. G. (2010). Scaling up from gardens: Biodiversity conservation in urban environments. Trends in Ecology & Evolution, 25(2), 90–98. https://doi.org/10.1016/j.tree.2009.07.016

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., & Schupp, J. (2018). The German Socio-Economic Panel (SOEP). Jahrbücher Für Nationalökonomie Und Statistik, 239(2), 345–360. https://doi.org/doi:10.1515/jbnst-2018-0022

Goebel, J., & Pauer, B. (2014). Datenschutzkonzept zur Nutzung von SOEPgeo im Forschungsdatenzentrum SOEP am DIW Berlin. Zeitschrift Für Amtliche Statistik Berlin-Brandenburg, 3, 42–47.

Goetz, S., Steinberg, D., Dubayah, R., & Blair, B. (2007). Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. Remote Sensing of Environment, 108(3), 254–263. https://doi.org/10.1016/j.rse.2006.11.016

Goldberg, D. L., Anenberg, S. C., Lu, Z., Streets, D. G., Lamsal, L. N., E McDuffie, E., & Smith, S. J. (2021). Urban NOx emissions around the world declined faster than anticipated between 2005 and 2019. Environmental Research Letters, 16(115004), 1–11. https://doi.org/10.1088/1748-9326/ac2c34

Gómez-Baggethun, E., Gren, Å., Barton, D. N., Langemeyer, J., McPhearson, T., O'Farrell, P., Andersson, E., Hamstead, Z., & Kremer, P. (2013). Urban Ecosystem Services. In T. Elmqvist, M. Fragkias, J. Goodness, B. Güneralp, P. J. Marcotullio, R. I. McDonald, S. Parnell, M. Schewenius, M. Sendstad, K. C. Seto, & C. Wilkinson (Eds.), Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities: A Global Assessment (pp. 175–251). Springer Netherlands. https://doi.org/10.1007/978-94-007-7088-1_11

Gong, P., & Howard, P. J. (1992). Frequency-based contextual classification and gray-level vector reduction for land-use identification. Photogrammetric Engineering & Remote Sensing, 58(4), 423–437.

Gong, P., & Howarth, P. J. (1989). Performance analyses of probabilistic relaxation methods for land-cover classification. Remote Sensing of Environment, 30(1), 33–42. https://doi.org/10.1016/0034-4257(89)90045-X

Gong, P., & Howarth, P. J. (1990). The use of structural information for improving land-cover classification accuracies at the rural-urban fringe. Photogrammetric Engineering & Remote Sensing, 56(1), 67–73.

Gong, P., & Howarth, P. J. (1992). Land-use classification of SPOT HRV data using a cover-frequency method. International Journal of Remote Sensing, 13(8), 1459–1471. https://doi.org/10.1080/01431169208904202

Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., & Zhou, Y. (2020). Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. Remote Sensing of Environment, 236(111510), 1–12. https://doi.org/10.1016/j.rse.2019.111510

Gong, P., Li, Z., Huang, H., Sun, G., & Wang, L. (2011). ICESat GLAS data for urban environment monitoring. IEEE Transactions on Geoscience and Remote Sensing, 49(3), 1158–1172. Scopus. https://doi.org/10.1109/TGRS.2010.2070514

Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Ping Suen, H., Guo, J., Xu, N., Li, W., Zhao, Y., … Song, L. (2019). Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. Science Bulletin, 64(6), 370–373. https://doi.org/10.1016/j.scib.2019.03.002

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. GeoJournal, 69(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

Goodchild, M. F. (2013). The quality of big (geo)data. Dialogues in Human Geography, 3(3), 280–284. https://doi.org/10.1177/2043820613513392

Goodfellow, T. (2022). Politics and the Urban Frontier: Transformation and Divergence in Late Urbanizing East Africa. Oxford University Press. https://doi.org/10.1093/oso/9780198853107.001.0001

Goodman, H., Egizi, A., Fonseca, D. M., Leisnham, P. T., & LaDeau, S. L. (2018). Primary blood-hosts of mosquitoes are influenced by social and ecological conditions in a complex urban landscape. Parasites & Vectors, 11(218), 1–10. https://doi.org/10.1186/s13071-018-2779-7

Goodrich, C. G., Gurung, K., & Hamal, M. (2021). Gender Integration in Earth Observation and Geo-information Technology Applications: Correlation and Connections. In B. Bajracharya, R. B. Thapa, & M. A. Matin (Eds.), Earth Observation Science and Applications for Risk Reduction and Enhanced Resilience in Hindu Kush Himalaya Region: A Decade of Experience from SERVIR (pp. 291–306). Springer International Publishing. https://doi.org/10.1007/978-3-030-73569-2_15

Gordon, P., & Cox, W. (2012). Cities in Western Europe and the United States: Do policy differences matter? The Annals of Regional Science, 48(2), 565–594. https://doi.org/10.1007/s00168-011-0495-8

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Big Remotely Sensed Data: Tools, Applications and Experiences, 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. Big Data & Society, 7(1), 1–15. https://doi.org/10.1177/2053951719897945

Gottmann, J. (1957). Megalopolis or the Urbanization of the Northeastern Seaboard. Economic Geography, 33(3), 189–200. https://doi.org/10.2307/142307

Grace, R., Halse, S., Aurite, W., Montarnal, A., & Tapia, A. (2019). Expanding Awareness: Comparing Location, Keyword, and Network Filtering Methods to Collect Hyperlocal Social Media Data. Proceedings of the 52nd Hawaii International Conference on System Sciences, 2699–2708. http://hdl.handle.net/10125/59707

Grafius, D. R., Corstanje, R., & Harris, J. A. (2018). Linking ecosystem services, urban form and green space configuration using multivariate landscape metric analysis. Landscape Ecology, 33(4), 557–573. https://doi.org/10.1007/s10980-018-0618-z

Graham, M., & De Sabbata, S. (2020). The Geographic Contours of Openness. In N. Haddadian, M. L. Smith, & R. K. Seward (Eds.), Making Open Development Inclusive: Lessons from IDRC Research (pp. 1–26). The MIT Press. https://doi.org/10.7551/mitpress/11635.003.0010

Grasso, V. F. (2014). The State of Early Warning Systems. In A. Singh & Z. Zommers (Eds.), Reducing Disaster: Early Warning Systems For Climate Change (pp. 109–125). Springer Netherlands. https://doi.org/10.1007/978-94-017-8598-3_6

Greenhouse Gas Protocol. (2023). Global Protocol for Community-Scale Greenhouse Gas Inventories (p. 190). https://ghgprotocol.org/ghg-protocol-cities

Grekousis, G., & Mountrakis, G. (2015). Sustainable Development under Population Pressure: Lessons from Developed Land Consumption in the Conterminous U.S. PLOS ONE, 10(3)(e0119675), 1–19. https://doi.org/10.1371/journal.pone.0119675

Griffin, E., & Ford, L. (1980). A Model of Latin American City Structure. Geographical Review, 70(4), 397–422. https://doi.org/10.2307/214076

Grimm, N. B., Cook, E. M., Hale, R. L., & Iwaniec, D. M. (2016). A broader framing of ecosystem services in cities. In K. Seto, W. Solecki, & C. Griffith (Eds.), The Routledge Handbook of Urbanization and Global Environmental Change (1st ed., pp. 203–212). Routledge.

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global Change and the Ecology of Cities. Science, 319(5864), 756–760. https://doi.org/10.1126/science.1150195

Grimm, N. B., Pickett, S. T. A., Hale, R. L., & Cadenasso, M. L. (2017). Does the ecological concept of disturbance have utility in urban social–ecological–technological systems? Ecosystem Health and Sustainability, 3(1)(e01255), 1–18. https://doi.org/10.1002/ehs2.1255

Gros, A., Bozonnet, E., & Inard, C. (2014). Cool materials impact at district scale—Coupling building energy and microclimate models. Sustainable Cities and Society, 13, 254–266. https://doi.org/10.1016/j.scs.2014.02.002

Großmann, K., Arndt, T., Haase, A., Rink, D., & Steinführer, A. (2015). The influence of housing oversupply on residential segregation: Exploring the post-socialist city of Leipzig†. Urban Geography, 36(4), 550–577. https://doi.org/10.1080/02723638.2015.1014672

Growe, A., Harrison, J., & Hoyler, M. (2015). Megaregions.Globalization's New Urban Form? Raumforschung Und Raumordnung | Spatial Research and Planning, 73, 305–307. https://doi.org/10.1007/s13147-015-0348-5

Grunewald, K., Richter, B., Meinel, G., Herold, H., & Syrbe, R.-U. (2016). Vorschlag bundesweiter Indikatoren zur Erreichbarkeit öffentlicher Grünflächen. Bewertung der Ökosystemleistung „Erholung in der Stadt" [Proposol for nationwide indicators for "accessibility of urban green spaces"—Evaluation of the ecosystem service "recreation in the city"]. Naturschutz Und Landschaftsplanung, 48(7), 218–226.

Grunewald, K., Xie, G., & Wüstemann, H. (2018). The Multiple Benefits of Urban Green—Ecosystem Services Assessment. In K. Grunewald, J. Li, G. Xie, & L. Kümper-Schlake (Eds.), Towards Green Cities: Urban Biodiversity and Ecosystem Services in China and Germany (pp. 43–104). Springer International Publishing. https://doi.org/10.1007/978-3-319-58223-8_3

Gu, K. (2019). Urban Morphological Regions: Development of an Idea. In V. Oliveira (Ed.), J.W.R. Whitehand and the Historico-geographical Approach to Urban Morphology (pp. 33–46). Springer International Publishing. https://doi.org/10.1007/978-3-030-00620-4_2

Guan, X., Wei, H., Lu, S., Dai, Q., & Su, H. (2018). Assessment on the urbanization strategy in China: Achievements, challenges and reflections. Habitat International, 71, 97–109. https://doi.org/10.1016/j.habitatint.2017.11.009

Guillain, R., Gallo, J. L., & Boiteux-Orain, C. (2006). Changes in Spatial and Sectoral Patterns of Employment in Ile-de-France, 1978-97. Urban Studies, 43(11), 2075–2098. https://doi.org/10.1080/00420980600945203

Gülzau, F., & Mau, S. (2021). Walls, Barriers, Checkpoints, Landmarks, and 'No-Man's-Land:' A Quantitative Typology of Border Control Infrastructure. Historical Social Research, 46(3), 23–48. https://doi.org/10.12759/hsr.46.2021.3.23-48

Gumiran, B., Moncada, F., Gasmen, H., Boyles-Panting, N., & Solidum, R. (2019). Participatory capacities and vulnerabilities assessment: Towards the realisation of community-based early warning system for deep-seated landslides. Jàmbá: Journal of Disaster Risk Studies, 11(1), 1–11. https://doi.org/10.4102/jamba.v11i1.555

Gumuchian, M.-L. (2012). Pro-government forces seize old Gaddafi bastion Bani Walid. Reuters. https://www.reuters.com/article/world/china/pro-government-forces-seize-old-gaddafi-bastion-bani-walid-idUSBRE89N0WQ/

Güneralp, B., Güneralp, İ., & Liu, Y. (2015). Changing global patterns of urban exposure to flood and drought hazards. Global Environmental Change, 31, 217–225. https://doi.org/10.1016/j.gloenvcha.2015.01.002

Güneralp, B., Lwasa, S., Masundire, H., Parnell, S., & Seto, K. C. (2017). Urbanization in Africa: Challenges and opportunities for conservation. Environmental Research Letters, 13(015002), 1–8. https://doi.org/10.1088/1748-9326/aa94fe

Güneralp, B., Reba, M., Hales, B. U., Wentz, E. A., & Seto, K. C. (2020). Trends in urban land expansion, density, and land transitions from 1970 to 2010: A global synthesis. Environmental Research Letters, 15(044015), 1–13. https://doi.org/10.1088/1748-9326/ab6669

Güneralp, B., Zhou, Y., Ürge-Vorsatz, D., Gupta, M., Yu, S., Patel, P. L., Fragkias, M., Li, X., & Seto, K. C. (2017). Global scenarios of urban density and its impacts on building energy use through 2050. Proceedings of the National Academy of Sciences, 114(34), 8945–8950. https://doi.org/10.1073/pnas.1606035114

Günther, A., Van Den Eeckhaut, M., Malet, J.-P., Reichenbach, P., & Hervás, J. (2014). Climate-physiographically differentiated Pan-European landslide susceptibility assessment using spatial multi-criteria evaluation and transnational landslide information. Geomorphology, 224, 69–85. https://doi.org/10.1016/j.geomorph.2014.07.011

Guo, H.-D., Zhang, L., & Zhu, L.-W. (2015). Earth observation big data for climate change research. Special Issue on Advances in Future Earth Research, 6(2), 108–117. https://doi.org/10.1016/j.accre.2015.09.007

Guo, Y., Gasparrini, A., Armstrong, B. G., Tawatsupa, B., Tobias, A., Lavigne, E., Coelho, M. de S. Z. S., Pan, X., Kim, H., Hashizume, M., Honda, Y., Guo, Y.-L. L., Wu, C.-F., Zanobetti, A., Schwartz, J. D., Bell, M. L., Scortichini, M., Michelozzi, P., Punnasiri, K., … Tong, S. (2017). Heat Wave and Mortality: A Multicountry, Multicommunity Study. Environmental Health Perspectives, 125(8), 1–11. https://doi.org/10.1289/EHP1026

Gupta, K., Kumar, P., Pathan, S. K., & Sharma, K. P. (2012). Urban Neighborhood Green Index – A measure of green spaces in urban areas. Landscape and Urban Planning, 105(3), 325–335. https://doi.org/10.1016/j.landurbplan.2012.01.003

Gutiérrez, J. M., Jones, R. G., Narisma, G. T., Alves, L. M., Amjad, M., Gorodetskaya, I. V., Grose, M., Klutse, N. A. B., Krakovska, S., Li, J., Martínez-Castro, D., Mearns, L. O., Mernild, S. H., Ngo-Duc, T., van den Hurk, B., & Yoon, J.-H. (2021). Atlas. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 1927–2058). Cambridge University Press. doi:10.1017/9781009157896.021

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., & Melillo, M. (2020). Geographical landslide early warning systems. Earth-Science Reviews, 200(102973), 1–29. https://doi.org/10.1016/j.earscirev.2019.102973

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K.-T. (2012). Landslide inventory maps: New tools for an old problem. Earth-Science Reviews, 112(1), 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

H

Haack, B. N. (1983). An analysis of thematic mapper simulator data for urban environments. Remote Sensing of Environment, 13(3), 265–275. https://doi.org/10.1016/0034-4257(83)90044-5

Haase, A., Bernt, M., Großmann, K., Mykhnenko, V., & Rink, D. (2016). Varieties of shrinkage in European cities. European Urban and Regional Studies, 23(1), 86–102. https://doi.org/10.1177/0969776413481985

Haase, A., Wolff, M., Špačková, P., & Radzimski, A. (2018). Reurbanisation in Postsocialist Europe – A Comparative View of Eastern Germany, Poland, and the Czech Republic. Omparative Population Studies, 43, 353–390. https://doi.org/10.12765/CPoS-2018-02

Haase, D. (2008). Urban Ecology of Shrinking Cities: An Unrecognized Opportunity? Nature and Culture, 3(1), 1–8. https://doi.org/10.3167/nc.2008.030101

Haase, D., Güneralp, B., Dahiya, B., Bai, X., & Elmqvist, T. (2018). Global Urbanization: Perspectives and Trends. In T. Elmqvist, X. Bai, N. Frantzeskaki, C. Griffith, D. Maddox, T. McPhearson, S. Parnell, P. Romero-Lankao, D. Simon, & M. Watkins (Eds.), Urban Planet: Knowledge towards Sustainable Cities (pp. 19–44). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781316647554.003

Haberl, H., Wiedenhofer, D., Schug, F., Frantz, D., Virág, D., Plutzar, C., Gruhler, K., Lederer, J., Schiller, G., Fishman, T., Lanau, M., Gattringer, A., Kemper, T., Liu, G., Tanikawa, H., van der Linden, S., & Hostert, P. (2021). High-Resolution Maps of Material Stocks in Buildings and Infrastructures in Austria and Germany. Environmental Science & Technology, 55(5), 3368–3379. https://doi.org/10.1021/acs.est.0c05642

Häberle, M., Hoffmann, E. J., & Zhu, X. X. (2022). Can linguistic features extracted from geo-referenced tweets help building function classification in remote sensing? ISPRS Journal of Photogrammetry and Remote Sensing, 188, 255–268. https://doi.org/10.1016/j.isprsjprs.2022.04.006

Häberle, M., Werner, M., & Zhu, X. X. (2019). Building Type Classification from Social Media Texts via Geo-Spatial Textmining. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 10047–10050. https://doi.org/10.1109/IGARSS.2019.8898836

Habraken, N. J. (1987). The Control of Complexity. Places, 4(2), 3–15.

Habraken, N. J. (1998). The structure of the ordinary: Form and control in the built environment (J. Teicher, Ed.). MIT Press.

Habraken, N. J. (2016). Cultivating Complexity: The Need for a Shift in Cognition. In J. Portugali & E. Stolk (Eds.), Complexity, Cognition, Urban Planning and Design (pp. 55–74). Springer International Publishing.

Hadjimitsis, D. G., Retalis, A., & Clayton, C. R. I. (2002). The Assessment of Atmospheric Pollution using Satellite Remote Sensing Technology in Large Cities in the Vicinity of Airports. Water, Air and Soil Pollution: Focus, 2(5), 631–640. https://doi.org/10.1023/A:1021305417004

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. Minds and Machines, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hahn, B. (2014). Die US-amerikanische Stadt im Wandel (2nd ed.). Springer Spektrum Berlin. https://doi.org/10.1007/978-3-662-69931-7

Hajrasoulih, A., del Rio, V., Francis, J., & Edmondson, J. (2018). Urban form and mental wellbeing: Scoping a theoretical framework for action. Journal of Urban Design and Mental Health, 5(10). https://www.urbandesignmentalhealth.com/journal-5---urban-form-and-mental-wellbeing.html

Haken, H., & Portugali, J. (2003). The face of the city is its information. Journal of Environmental Psychology, 23(4), 385–408. https://doi.org/10.1016/S0272-4944(03)00003-3

Hakim, B. S. (2014). Mediterranean Urbanism: Historic Urban / Building Rules and Processes. Springer Dordrecht. https://doi.org/10.1007/978-94-017-9140-3

Hall, J. (2007). The ethical opportunity. Journal of Brand Management, 14(5), 365–367. https://doi.org/10.1057/palgrave.bm.2550089

Hall, P. (1995). Towards a General Urban Theory. In J. F. Brotchie, E. Blakely, P. Hall, & P. Newton (Eds.), Cities in Competition: Productive and Sustainable Cities for the 21st Century (pp. 3–31). Longman Australia.

Hall, P. (2009). Looking Backward, Looking Forward: The City Region of the Mid-21st Century. Regional Studies, 43(6), 803–817. https://doi.org/10.1080/00343400903039673

Hall, P., & Pain, K. (2006). The Polycentric Metropolis: Learning from Mega-City Regions in Europe (1st ed.). Routledge. https://doi.org/10.4324/9781849773911

Hallegatte, S. (2012). A Cost Effective Solution to Reduce Disaster Losses in Developing Countries: Hydro-Meteorological Services, Early Warning, and Evacuation (No. 6058; Policy Research Working Paper). © World Bank. http://hdl.handle.net/10986/9359

Hamilton, F. E. I. (2005). The external forces: Towards globalization and European integration. In F. E. I. Hamilton, K. Dimitrovska Andrews, & N. Pichler-Milanovic (Eds.), Transformation of Cities in Central and Eastern Europe: Towards Globalization (pp. 79–115). United Nations University Press.

Hamilton, F. E. I., Andrews, K. D., Pichler-Milanovic, N., & University, U. N. (2005). Transformation of cities in Central and Eastern Europe: Towards globalization /: edited by F.E. Ian Hamilton, Kaliopa Dimitrovska Andrews, and Natasa Pichler-Milanovic. UN University Press,. https://digitallibrary.un.org/record/559142

Hamin, E. M., & Gurran, N. (2009). Urban form and climate change: Balancing adaptation and mitigation in the U.S. and Australia. Habitat International, 33(3), 238–245. https://doi.org/10.1016/j.habitatint.2008.10.005

Han, R., Cao, H., & Liu, Z. (2018). Studying the urban hierarchical pattern and spatial structure of China using a synthesized gravity model. Science China Earth Sciences, 61(12), 1818–1831. https://doi.org/10.1007/s11430-016-9191-5

Handelman, D. (1987). Myths of Murugan: Asymmetry and Hierarchy in a South Indian Puranic Cosmology. History of Religions, 27(2), 133–170. https://doi.org/10.1086/463110

Hankey, S., & Marshall, J. D. (2017). Urban Form, Air Pollution, and Health. Current Environmental Health Reports, 4(4), 491–503. https://doi.org/10.1007/s40572-017-0167-7

Hänsch, R., Ley, A., & Hellwich, O. (2017). Correct and still wrong: The relationship between sampling strategies and the estimation of the generalization error. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 3672–3675. https://doi.org/10.1109/IGARSS.2017.8127795

Hansen, M. C., Sohlberg, R., Defries, R. S., & Townshend, J. R. G. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. International Journal of Remote Sensing, 21(6–7), 1331–1364. Scopus. https://doi.org/10.1080/014311600210209

Hansen, R., & Pauleit, S. (2014). From Multifunctionality to Multiple Ecosystem Services? A Conceptual Framework for Multifunctionality in Green Infrastructure Planning for Urban Areas. AMBIO, 43(4), 516–529. https://doi.org/10.1007/s13280-014-0510-2

Hansson, S. O. (2021). Risks ethics. Encyclopedia. https://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and-maps/risk-ethics

Hardin, G. (1968). The Tragedy of the Commons. Science, 162(3859), 1243–1248. https://doi.org/10.1126/science.162.3859.1243

Hargittai, E. (2020). Potential Biases in Big Data: Omitted Voices on Social Media. Social Science Computer Review, 38(1), 10–24. https://doi.org/10.1177/0894439318788322

Harlan, S. L., Brazel, A. J., Prashad, L., Stefanov, W. L., & Larsen, L. (2006). Neighborhood microclimates and vulnerability to heat stress. Social Science & Medicine, 63(11), 2847–2863. https://doi.org/10.1016/j.socscimed.2006.07.030

Harris, C. D., & Ullman, E. L. (1945). The Nature of Cities. The ANNALS of the American Academy of Political and Social Science, 242(1), 7–17. https://doi.org/10.1177/000271624524200103

Harrison, J., & Growe, A. (2014). From places to flows? Planning for the new 'regional world' in Germany. European Urban and Regional Studies, 21(1), 21–41. https://doi.org/10.1177/0969776412441191

Hartfield, K. A., Landau, K. I., & Leeuwen, W. J. D. van. (2011). Fusion of High Resolution Aerial Multispectral and LiDAR Data: Land Cover in the Context of Urban Mosquito Habitat. Remote Sensing, 3(11), 2364–2383. https://doi.org/10.3390/rs3112364

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1), 100–108. https://doi.org/10.2307/2346830

Haselsberger, B. (2014). Decoding borders. Appreciating border impacts on space and people. Planning Theory & Practice, 15(4), 505–526. https://doi.org/10.1080/14649357.2014.963652

Hassner, R. E., & Wittenberg, J. (2015). Barriers to Entry: Who Builds Fortified Boundaries and Why? International Security, 40(1), 157–190. https://doi.org/10.1162/ISEC_a_00206

Häussermann, H. (1996). From the Socialist to the Capitalist City: Experiences from Germany. In G. Andrusz, M. Harloe, & I. Szelenyi (Eds.), Cities After Socialism (pp. 214–231). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470712733.ch7

Häußermann, H. (1997). Amerikanisierung der deutschen Städte—Divergenz und Konvergenz. In P. Neitzke, C. Steckeweh, & R. Wustlich (Eds.), Centrum: Jahrbuch Architektur und Stadt 1997 – 1998 (pp. 92–96). Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-83184-2_18

He, B.-J., Zhao, Z.-Q., Shen, L.-D., Wang, H.-B., & Li, L.-G. (2019). An approach to examining performances of cool/hot sources in mitigating/enhancing land surface temperature under different temperature backgrounds based on landsat 8 image. Sustainable Cities and Society, 44, 416–427. https://doi.org/10.1016/j.scs.2018.10.049

He, C., Kumar, R., Tang, W., Pfister, G., Xu, Y., Qian, Y., & Brasseur, G. (2024). Air Pollution Interactions with Weather and Climate Extremes: Current Knowledge, Gaps, and Future Directions. Current Pollution Reports, 10(3), 430–442. https://doi.org/10.1007/s40726-024-00296-9

He, C., Liu, Z., Gou, S., Zhang, Q., Zhang, J., & Xu, L. (2019). Detecting global urban expansion over the last three decades using a fully convolutional network. Environmental Research Letters, 14(034008), 1–16. https://doi.org/10.1088/1748-9326/aaf936

He, C., Shi, Peijun, Xie, D., & Zhao, Y. (2010). Improving the normalized difference built-up index to map urban built-up areas using a semiautomatic segmentation approach. Remote Sensing Letters, 1(4), 213–221. https://doi.org/10.1080/01431161.2010.481681

He, C., Zhou, Y., & Huang, Z. (2016). Fiscal decentralization, political centralization, and land urbanization in China. Urban Geography, 37(3), 436–457. https://doi.org/10.1080/02723638.2015.1063242

Heider, B., Mast, J., Roth, D., Standfuß, I., Siedentop, S., & Taubenböck, H. (2022). Dynamics of intra-urban employment geographies: A comparative study of U.S. and German metropolitan areas. Journal of Urban Affairs, 46(7), 1287–1307. https://doi.org/10.1080/07352166.2022.2122833

Heider, B., & Siedentop, S. (2020). Employment suburbanization in the 21st century: A comparison of German and US city regions. Cities, 104(102802), 1–14. https://doi.org/10.1016/j.cities.2020.102802

Held, G. (2005). Territorium und Großstadt: Die räumliche Differenzierung der Moderne. VS Verlag für Sozialwissenschaften.

Henderson, F. M., Wharton, S. W., & Toll, D. L. (1980). Preliminary restuls of mapping urban land cover with Seasat SAR imagery. Izvestiya VUZ: Radioelektronika, 310–317.

Henderson, J. V., Squires, T., Storeygard, A., & Weil, D. (2017). On the spatial distribution of development. The roles of nature and history. In S. Michalopoulos & E. Papaioannou (Eds.), The Long Economic and Political Shadow of History Volume I. A Global View (pp. 7–18). CEPR Press.

Henderson, M., Yeh, E. T., Gong, P., Elvidge, C., & Baugh, K. (2003). Validation of urban boundaries derived from global night-time satellite imagery. International Journal of Remote Sensing, 24(3), 595–609. Scopus. https://doi.org/10.1080/01431160304982

Henderson, V., Storeygard, A., & Weil, D. N. (2011). A Bright Idea for Measuring Economic Growth. American Economic Review, 101(3), 194–199. https://doi.org/10.1257/aer.101.3.194

Hentschel, T., Hruschka, F., & Priester, M. (2002). Global Report on Artisanal & Small-scale mining (p. 67). International Institute for Environment and Development (IIED) and World Business Council for Sustainable Developme (WBCSD). https://www.iied.org/sites/default/files/pdfs/migrate/G00723.pdf

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2021). The evolution of humanitarian mapping within the OpenStreetMap community. Scientific Reports, 11(3037), 1–15. https://doi.org/10.1038/s41598-021-82404-z

Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J., & Zipf, A. (2023). A spatio-temporal analysis investigating completeness and inequalities of global urban building data in OpenStreetMap. Nature Communications, 14(3985), 1–14. https://doi.org/10.1038/s41467-023-39698-6

Hertzog, A. (2020). The Lagos Abidjan corridor: Migration driven urbanisation in west Africa [Doctoral Thesis, ETH Zürich]. https://doi.org/10.3929/ethz-b-000478236

Heynen, N., Kaika, M., & Swyngedouw, E. (2006). Urban political ecology: Politicizing the production of urban natures. In N. Heynen, M. Kaika, & E. Swyngedouw (Eds.), In the Nature of Cities—Urban Political Ecology and the Politics of Urban Metabolism (pp. 1–20). Routledge. https://doi.org/10.4324/9780203027523

Hidalgo, J., Dumas, G., Masson, V., Petit, G., Bechtel, B., Bocher, E., Foley, M., Schoetter, R., & Mills, G. (2019). Comparison between local climate zones maps derived from administrative datasets and satellite observations. Urban Climate, 27, 64–89. https://doi.org/10.1016/j.uclim.2018.10.004

Higueras, E., Román, E., & Fariña, J. (2021). Guidelines for Healthier Public Spaces for the Elderly Population: Recommendations in the Spanish Context. In J. Martinez, C. A. Mikkelsen, & R. Phillips (Eds.), Handbook of Quality of Life and Sustainability (pp. 35–51). Springer International Publishing. https://doi.org/10.1007/978-3-030-50540-0_3

Hilboll, A., Richter, A., & Burrows, J. P. (2013). Long-term changes of tropospheric NO_2 over megacities derived from multiple satellite instruments. Atmospheric Chemistry and Physics, 13(8), 4145–4169. https://doi.org/10.5194/acp-13-4145-2013

Hirt, S. (2007). Suburbanizing Sofia: Characteristics of Post-Socialist Peri-Urban Change. Urban Geography, 28(8), 755–780. https://doi.org/10.2747/0272-3638.28.8.755

Ho, H. C., Knudby, A., Sirovyak, P., Xu, Y., Hodul, M., & Henderson, S. B. (2014). Mapping maximum urban air temperature on hot summer days. Remote Sensing of Environment, 154, 38–45. https://doi.org/10.1016/j.rse.2014.08.012

Hobson, P. (2019). More than 40 million people work in artisanal mining. Reuters. https://www.reuters.com/article/us-mining-asm-idUSKCN1S025C/

Hodler, R., & Raschky, P. A. (2014). Regional favoritism. The Quarterly Journal of Economics, 129(2), 995–1033. https://doi.org/10.1093/qje/qju004

Hoehner, C. M., Brennan Ramirez, L. K., Elliott, M. B., Handy, S. L., & Brownson, R. C. (2005). Perceived and objective environmental measures and physical activity among urban adults. Active Living Research, 28(2, Supplement 2), 105–116. https://doi.org/10.1016/j.amepre.2004.10.023

Hoffman, J. S., Shandas, V., & Pendleton, N. (2020). The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas. Climate, 8(1), 1–15. https://doi.org/10.3390/cli8010012

Hoffmann, E. J., Wang, Y., Werner, M., Kang, J., & Zhu, X. X. (2019). Model Fusion for Building Type Classification from Aerial and Street View Images. Remote Sensing, 11(1259), 1–20. https://doi.org/10.3390/rs11111259

Hoffmeyer-Zlotnik, J. H. P. (2000). Regionalisierung Sozialwissenschaftlicher Umfragedaten. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-90525-3_1

Hofmeister, B. (1991). Die Stadtstruktur: Ihre Ausprägung in den verschiedenen Kulturräumen der Erde (2nd ed.). Darmstadt: Wiss. Buchges.

Hollenstein, L., & Purves, R. (2010). Exploring place through user-generated content: Using Flickr tags to describe city cores. Journal of Spatial Information Science, 1, 21–48. https://doi.org/10.5311/JOSIS.2010.1.3

Hollis, L. (2013). Cities Are Good for You: The Genius of the Metropolis. Bloomsbury Publishing.

Holmes, G. (2001). The Oxford Illustrated History of Medieval Europe. Oxford University Press.

Hong, S., Hui, E. C., & Lin, Y. (2022). Relationship between urban spatial structure and carbon emissions: A literature review. Ecological Indicators, 144(109456), 1–8. https://doi.org/10.1016/j.ecolind.2022.109456

Hoornweg, D. A., & Bhada-Tata, P. (2012). What a waste: A global review of solid waste management (Vol. 15). Washington, DC: World Bank.

Hornborg, A. (2001). The power of the machine: Global inequalities of economy, technology, and environment. AltaMira Press.

Hospers, G.-J. (2002). Beyond the Blue Banana? Structural change in Europe's geo-economy. Economic Geography, 38(2), 76–85. https://doi.org/10.1007/BF03031774

Howarth, P. J., & Boasson, E. (1983). Landsat digital enhancements for change detection in urban environments. Remote Sensing of Environment, 13(2), 149–160. https://doi.org/10.1016/0034-4257(83)90019-6

Hoyt, H. (1939). The structure and growth of residential neighborhoods in American cities. Washington, DC: Federal Housing Administration.

Hsu, W. (2012). Central Place Theory and City Size Distribution. The Economic Journal, 122(563), 903–932. https://doi.org/10.1111/j.1468-0297.2012.02518.x

Hu, J., Ghamisi, P., & Zhu, X. X. (2018). Feature Extraction and Selection of Sentinel-1 Dual-Pol Data for Global-Scale Local Climate Zone Classification. ISPRS International Journal of Geo-Information, 7(379), 1–20. https://doi.org/10.3390/ijgi7090379

Hu, J., Hong, D., Wang, Y., & Zhu, X. X. (2019). A Comparative Review of Manifold Learning Techniques for Hyperspectral and Polarimetric SAR Image Fusion. Remote Sensing, 11(681), 1–28. https://doi.org/10.3390/rs11060681

Hu, J., Hong, D., & Zhu, X. X. (2019). MIMA: MAPPER-Induced Manifold Alignment for Semi-Supervised Fusion of Optical Image and Polarimetric SAR Data. IEEE Transactions on Geoscience and Remote Sensing, 57(11), 9025–9040. https://doi.org/10.1109/TGRS.2019.2924113

Hu, J., Wang, Y., Taubenböck, H., & Zhu, X. X. (2021). Land consumption in cities: A comparative study across the globe. Cities, 113(103163), 1–17. https://doi.org/10.1016/j.cities.2021.103163

Hu, X., Zhou, Z., Li, H., Hu, Y., Gu, F., Kersten, J., Fan, H., & Klan, F. (2024). Location Reference Recognition from Texts: A Survey and Comparison. ACM Comput. Surv., 56(5), 112:1-112:37. https://doi.org/10.1145/3625819

Hu, Y., & Wang, R.-Q. (2020). Understanding the removal of precise geotagging in tweets. Nature Human Behaviour, 4(12), 1219–1221. https://doi.org/10.1038/s41562-020-00949-x

Huang, B., & Carley, K. M. (2019). A large-scale empirical study of geotagging behavior on Twitter. Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 365–373. https://doi.org/10.1145/3341161.3342870

Huang, B., Zhao, B., & Song, Y. (2018). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sensing of Environment, 214, 73–86. https://doi.org/10.1016/j.rse.2018.04.050

Huang, C., & Xu, N. (2022). Quantifying urban expansion from 1985 to 2018 in large cities worldwide. Geocarto International, 37(27), 18356–18371. https://doi.org/10.1080/10106049.2022.2142957

Huang, D., Xu, S., Sun, J., Liang, S., Song, W., & Wang, Z. (2017). Accuracy assessment model for classification result of remote sensing image based on spatial sampling. Journal of Applied Remote Sensing, 11(4), 1–13. https://doi.org/10.1117/1.JRS.11.046023

Huang, J., Lu, X. X., & Sellers, J. M. (2007). A global comparative analysis of urban form: Applying spatial metrics and remote sensing. Landscape and Urban Planning, 82(4), 184–197. https://doi.org/10.1016/j.landurbplan.2007.02.010

Huang, R., Taubenböck, H., Mou, L., & Zhu, X. X. (2018). Classification of Settlement Types from Tweets Using LDA and LSTM. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 6408–6411. https://doi.org/10.1109/IGARSS.2018.8519240

Huang, R., Taubenböck, H., & Zhu, X. (2018). Towards Urban Settlement Type Mapping With Geotagged Tweets. EARSeL. 5th Joint Workshop "Urban Remote Sensing – Challenges & Solutions", Bochum.

Huang, X., Li, J., Yang, J., Zhang, Z., Li, D., & Liu, X. (2021). 30 m global impervious surface area dynamics and urban expansion pattern observed by Landsat satellites: From 1972 to 2019. Science China Earth Sciences, 64(11), 1922–1933. https://doi.org/10.1007/s11430-020-9797-9

Huang, X., Schneider, A., & Friedl, M. A. (2016). Mapping sub-pixel urban expansion in China using MODIS and DMSP/OLS nighttime lights. Remote Sensing of Environment, 175, 92–108. https://doi.org/10.1016/j.rse.2015.12.042

Huang, X., Wang, C., & Li, Z. (2019). High-resolution population grid in the CONUS using microsoft building footprints: A feasibility study. Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Geospatial Humanities, 19. https://doi.org/10.1145/3356991.3365469

Huggel, C., Khabarov, N., Obersteiner, M., & Ramírez, J. M. (2010). Implementation and integrated numerical modeling of a landslide early warning system: A pilot study in Colombia. Natural Hazards, 52(2), 501–518. https://doi.org/10.1007/s11069-009-9393-0

Hungr, O., Corominas, J., & Eberhardt, E. (2005). Estimating landslide motion mechanism, travel distance and velocity. In O. Hungr, R. Fell, R. Couture, & E. Eberhardt (Eds.), Landslide Risk Management (1st ed., pp. 109–138). CRC Press. https://doi.org/10.1201/9781439833711-7

Hunter, M. R., Gillespie, B. W., & Chen, S. Y.-P. (2019). Urban Nature Experiences Reduce Stress in the Context of Daily Life Based on Salivary Biomarkers. Frontiers in Psychology, 10(722), 1–16. https://doi.org/10.3389/fpsyg.2019.00722

Huntington, S. P. (1997). The clash of civilizations and the remaking of world order. Penguin Books.

Huszar, P., Karlický, J., Marková, J., Nováková, T., Liaskoni, M., & Bartík, L. (2021). The regional impact of urban emissions on air quality in Europe: The role of the urban canopy effects. Atmospheric Chemistry and Physics, 21(18), 14309–14332. https://doi.org/10.5194/acp-21-14309-2021

I

Ianoş, I., Sîrodoev, I., Pascariu, G., & Henebry, G. (2016). Divergent patterns of built-up urban space growth following post-socialist changes. Urban Studies, 53(15), 3172–3188. https://doi.org/10.1177/0042098015608568

Idé, T., Katsuki, T., Morimura, T., & Morris, R. (2017). City-Wide Traffic Flow Estimation From a Limited Number of Low-Quality Cameras. IEEE Transactions on Intelligent Transportation Systems, 18(4), 950–959. https://doi.org/10.1109/TITS.2016.2597160

Ikeda, S. (2017). A City Cannot be a Work of Art. COSMOS + TAXIS, 4(2+3), 79–86.

Imhoff, M. L., Bounoua, L., DeFries, R., Lawrence, W. T., Stutzer, D., Tucker, C. J., & Ricketts, T. (2004). The consequences of urban land transformation on net primary productivity in the United States. Remote Sensing of Environment, 89(4), 434–443. https://doi.org/10.1016/j.rse.2003.10.015

Imhoff, M. L., Lawrence, W. T., Elvidge, C. D., Paul, T., Levine, E., Privalsky, M. V., & Brown, V. (1997). Using nighttime DMSP/OLS images of city lights to estimate the impact of urban land use on soil resources in the United States. Remote Sensing of Environment, 59(1), 105–117. https://doi.org/10.1016/S0034-4257(96)00110-1

Imhoff, M. L., Lawrence, W. T., Stutzer, D. C., & Elvidge, C. D. (1997). A technique for using composite DMSP/OLS 'city lights' satellite data to map urban area. Remote Sensing of Environment, 61(3), 361–370. Scopus. https://doi.org/10.1016/S0034-4257(97)00046-1

Imhoff, M. L., Zhang, P., Wolfe, R. E., & Bounoua, L. (2010). Remote sensing of the urban heat island effect across biomes in the continental USA. Remote Sensing of Environment, 114(3), 504–513. https://doi.org/10.1016/j.rse.2009.10.008

Innes, J. E., Booher, D. E., & Di Vittorio, S. (2010). Strategies for Megaregion Governance. Journal of the American Planning Association, 77(1), 55–67. https://doi.org/10.1080/01944363.2011.533640

Inostroza, L. (2014). Measuring urban ecosystem functions through 'Technomass'—A novel indicator to assess urban metabolism. Contemporary Concepts and Novel Methods Fostering Indicator-Based Approach to Urban Complexities, 42, 10–19. https://doi.org/10.1016/j.ecolind.2014.02.035

Inostroza, L. (2017). Informal urban development in Latin American urban peripheries. Spatial assessment in Bogotá, Lima and Santiago de Chile. Landscape and Urban Planning, 165, 267–279. https://doi.org/10.1016/j.landurbplan.2016.03.021

Inostroza, L. (2018). The circularity of the urban ecosystem material productivity: The transformation of biomass into technomass in Southern Patagonia. Sustainable Cities and Society, 39, 335–343. https://doi.org/10.1016/j.scs.2018.03.001

Inostroza, L., Hamstead, Z., Spyra, M., & Qureshi, S. (2019). Beyond urban–rural dichotomies: Measuring urbanisation degrees in central European landscapes using the technomass as an explicit indicator. Ecological Indicators, 96, 466–476. https://doi.org/10.1016/j.ecolind.2018.09.028

Inostroza, L., & Taubenböck, H. (2024). Searching for the DNA of urbanisation. A material perspective. Cities, 151(105079), 1–10. https://doi.org/10.1016/j.cities.2024.105079

Inostroza, L., & Zepp, H. (2021). The metabolic urban network: Urbanisation as hierarchically ordered space of flows. Cities, 109(103029), 1–8. https://doi.org/10.1016/j.cities.2020.103029

INSPIRE. (2014). INSPIRE - Infrastructure for spatial information in Europe. https://knowledge-base.inspire.ec.europa.eu/index_en

Institut National de L'information Géographique et Forestière. (2022). BD TOPO®. Version 3.0. Descriptif du contenu. https://geoservices.ign.fr/sites/default/files/2022-04/DC_BDTOPO_3-0.pdf

Ioannides, Y., & Skouras, S. (2013). US city size distribution: Robustly Pareto, but only in the tail. Journal of Urban Economics, 73(1), 18–29. https://doi.org/10.1016/j.jue.2012.06.005

IPCC. (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. (Core Writing Team, R. K. Pachauri, & L. A. Meyer, Eds.). https://www.ipcc.ch/report/ar5/syr/

IPCC. (2022). Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (H.-O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama, Eds.). Cambridge University Press. https://doi.org/10.1017/9781009325844

Iturbide, M., Gutiérrez, J. M., Alves, L. M., Bedia, J., Cerezo-Mota, R., Cimadevilla, E., Cofiño, A. S., Di Luca, A., Faria, S. H., Gorodetskaya, I. V., Hauser, M., Herrera, S., Hennessy, K., Hewitt, H. T., Jones, R. G., Krakovska, S., Manzanas, R., Martínez-Castro, D., Narisma, G. T., … Vera, C. S. (2020). An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets. Earth System Science Data, 12(4), 2959–2970. https://doi.org/10.5194/essd-12-2959-2020

Iungman, T., Cirach, M., Marando, F., Pereira Barboza, E., Khomenko, S., Masselot, P., Quijal-Zamorano, M., Mueller, N., Gasparrini, A., Urquiza, J., Heris, M., Thondoo, M., & Nieuwenhuijsen, M. (2023). Cooling cities through urban green infrastructure: A health impact assessment of European cities. The Lancet, 401(10376), 577–589. https://doi.org/10.1016/S0140-6736(22)02585-5

Iungman, T., Khomenko, S., Barboza, E. P., Cirach, M., Gonçalves, K., Petrone, P., Erbertseder, T., Taubenböck, H., Chakraborty, T., & Nieuwenhuijsen, M. (2024). The impact of urban configuration types on urban heat islands, air pollution, CO2 emissions, and mortality in Europe: A data science approach. The Lancet Planetary Health, 8(7), e489–e505. https://doi.org/10.1016/S2542-5196(24)00120-7

J

Jackson, K. T. (1985). Crabgrass Frontier: The Suburbanization of the United States. Oxford University Press.

Jacobs, C., Klok, L., Bruse, M., Cortesão, J., Lenzholzer, S., & Kluck, J. (2020). Are urban water bodies really cooling? Urban Climate, 32(100607), 1–14. https://doi.org/10.1016/j.uclim.2020.100607

Jacobs, J. (1961). The Death and Life of Great American Cities (Vintage Book, Ed.; 1st ed.). New York: Random House.

Jacobs, J. (1970). The economy of cities. Vintage.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. In ACM Computing Surveys (Vol. 31, Issue 3, pp. 264–323). https://doi.org/10.1145/331499.331504

Jain, M., Korzhenevych, A., & Basu, A. M. (2021). Integrating spatial development with infrastructure provision along an envisioned transport corridor: A conceptual framework and its application to India. Land Use Policy, 104(105364), 1–12. https://doi.org/10.1016/j.landusepol.2021.105364

Jain, S., Panda, J., & Kant, S. (2014). Possible Socio-scientific Issues of Land-use and Land-cover Change Impact and Associated Tools of Study with a Special Reference to Delhi-Mumbai Industrial Corridor Region. International Journal of Earth and Atmospheric Science, 1(2), 58–70.

James, P., Berrigan, D., Hart, J. E., Aaron Hipp, J., Hoehner, C. M., Kerr, J., Major, J. M., Oka, M., & Laden, F. (2014). Effects of buffer size and shape on associations between the built environment and energy balance. Health & Place, 27, 162–170. https://doi.org/10.1016/j.healthplace.2014.02.003

Jean, N., Burke, M., Xie, M., Alampay Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. Science, 353(6301), 790–794. https://doi.org/10.1126/science.aaf7894

Jelinski, D. E., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology, 11(3), 129–140. https://doi.org/10.1007/BF02447512

Jellissen, S. M., & Gottheil, F. M. (2013). On the utility of security fences along international borders. Defense & Security Analysis, 29(4), 266–279. https://doi.org/10.1080/14751798.2013.842707

Jenerette, G. D., Harlan, S. L., Buyantuev, A., Stefanov, W. L., Declet-Barreto, J., Ruddell, B. L., Myint, S. W., Kaplan, S., & Li, X. (2016). Micro-scale urban surface temperatures are related to land-cover features and residential heat related health impacts in Phoenix, AZ USA. Landscape Ecology, 31(4), 745–760. https://doi.org/10.1007/s10980-015-0284-3

Jenks, G. F., & Caspall, F. C. (1971). Error on Choroplethic Maps: Definition, Measurement, Reduction. Annals of the Association of American Geographers, 61(2), 217–244. https://doi.org/10.1111/j.1467-8306.1971.tb00779.x

Jerrett, M., Turner, M. C., Beckerman, B. S., C. Arden, P. I., van Donkelaar, A., Randall V, M., Serre, M., Crouse, D., Gapstur, S. M., Krewski, D., Diver, W. R., Coogan, P. F., Thurston, G. D., & Burnett, R. T. (2017). Comparing the Health Effects of Ambient Particulate Matter Estimated Using Ground-Based versus Remote Sensing Exposure Estimates. Environmental Health Perspectives, 125(4), 552–559. https://doi.org/10.1289/EHP575

Jessop, B., Brenner, N., & Jones, M. (2008). Theorizing Sociospatial Relations. Environment and Planning D: Society and Space, 26(3), 389–401. https://doi.org/10.1068/d9107

Jheng, D., Lemoine-Rodríguez, R., Biewer, C., Seto, K. C., & Taubenböck, H. (2025). The Relationship Between Urban Morphology and the Citizens' Social Media Footprints: A case study on Tokyo. 2025 Joint Urban Remote Sensing Event (JURSE), CFP25RSD-ART, 1–4. https://doi.org/10.1109/JURSE60372.2025.11076032

Jiang, B. (2012). Computing the Image of the City. In M. Campagna, A. De Montis, & F. Isola (Eds.), Planning Support Tools: Policy analysis, implementation and evaluation, Proceedings of the 7th Int. Conf. On Informatics and Urban and Regional Planning INPUT 2012 (pp. 111–121). FrancoAngeli.

Jiang, B., Ma, D., Yin, J., & Sandberg, M. (2016). Spatial Distribution of City Tweets and Their Densities. Geographical Analysis, 48(3), 337–351. https://doi.org/10.1111/gean.12096

Jing, W., Yang, Y., Yue, X., & Zhao, X. (2015). Mapping Urban Areas with Integration of DMSP/OLS Nighttime Light and MODIS Data Using Machine Learning Techniques. Remote Sensing, 7(9), 12419–12439. https://doi.org/10.3390/rs70912419

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Johnson, E. A., & Klemens, M. W. (2005). The impacts of sprawl on biodiversity. In E. A. Johnson & M. W. Klemens (Eds.), Nature in fragments (pp. 18–54). Columbia University Press. https://doi.org/10.7312/john12778-004

Johnson, J. M., & Clarke, K. C. (2021). An area preserving method for improved categorical raster resampling. Cartography and Geographic Information Science, 48(4), 292–304. https://doi.org/10.1080/15230406.2021.1892531

Johnson, M. P., Hollander, J., & Hallulli, A. (2014). Maintain, demolish, re-purpose: Policy design for vacant land management using decision models. Vacant Land: The New Urban Green?, 40, 151–162. https://doi.org/10.1016/j.cities.2013.05.005

Joint Research Center (JRC). (2021). GHS-SMOD R2019A [Dataset]. https://ghsl.jrc.ec.europa.eu/ ghs_smod2019.php

Jones, R. (2012). Border Walls: Security and the War on Terror in the United States, India, and Israe. Zed Books.

Joyce, K. E., Belliss, S. E., Samsonov, S. V., McNeill, S. J., & Glassey, P. J. (2009). A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. Progress in Physical Geography: Earth and Environment, 33(2), 183–207. https://doi.org/10.1177/0309133309339563

K

Kabaria, C. W., Molteni, F., Mandike, R., Chacky, F., Noor, A. M., Snow, R. W., & Linard, C. (2016). Mapping intra-urban malaria risk using high resolution satellite imagery: A case study of Dar es Salaam. International Journal of Health Geographics, 15(26), 1–12. https://doi.org/10.1186/s12942-016-0051-y

Kabisch, N., & Haase, D. (2013). Green spaces of European cities revisited for 1990–2006. Landscape and Urban Planning, 110, 113–122. https://doi.org/10.1016/j.landurbplan.2012.10.017

Kabisch, N., Strohbach, M., Haase, D., & Kronenberg, J. (2016). Urban green space availability in European cities. Navigating Urban Complexity: Advancing Understanding of Urban Social – Ecological Systems for Transformation and Resilience, 70, 586–596. https://doi.org/10.1016/j.ecolind.2016.02.029

Kakinuma, K., Puma, M. J., Hirabayashi, Y., Tanoue, M., Baptista, E. A., & Kanae, S. (2020). Flood-induced population displacements in the world. Environmental Research Letters, 15(124029), 1–8. https://doi.org/10.1088/1748-9326/abc586

Kalnay, E., & Cai, M. (2003). Impact of urbanization and land-use change on climate. Nature, 423(6939), 528–531. https://doi.org/10.1038/nature01675

Kaloustian, N., Tamminga, M., & Bechtel, B. (2017). Local climate zones and annual surface thermal response in a Mediterranean city. 2017 Joint Urban Remote Sensing Event (JURSE), 1–4. https://doi.org/10.1109/JURSE.2017.7924597

Kalt, G., Thunshirn, P., Wiedenhofer, D., Krausmann, F., Haas, W., & Haberl, H. (2021). Material stocks in global electricity infrastructures – An empirical analysis of the power sector's stock-flow-service nexus. Resources, Conservation and Recycling, 173(105723), 1–14. https://doi.org/10.1016/j.resconrec.2021.105723

Kamana, A. A., Radoine, H., & Nyasulu, C. (2024). Urban challenges and strategies in African cities – A systematic literature review. City and Environment Interactions, 21(100132), 1–16. https://doi.org/10.1016/j.cacint.2023.100132

Kamwela, V. K., & van Bergeijk, P. (2020). The border walls of (de)globalization. In ISS Working Paper Series / General Series (Vol. 651). http://hdl.handle.net/1765/123704

Kang, C., Ma, X., Tong, D., & Liu, Y. (2012). Intra-urban human mobility patterns: An urban morphology perspective. Physica A: Statistical Mechanics and Its Applications, 391(4), 1702–1717. https://doi.org/10.1016/j.physa.2011.11.005

Kanjir, U. (2019). Detecting migrant vessels in the Mediterranean Sea: Using Sentinel-2 images to aid humanitarian actions. Acta Astronautica, 155, 45–50. https://doi.org/10.1016/j.actaastro.2018.11.012

Kar, A., Varshney, M., & Varshney, M. (2021). The emergence of urban corridors—A brief look in Indian context. International Research Journal of Engineering and Technology (IRJET), 8(11), 1057–1064. https://www.irjet.net/archives/V8/i11/IRJET-V8I11171.pdf

Karami, A., Lundy, M., Webb, F., & Dwivedi, Y. K. (2020). Twitter and Research: A Systematic Literature Review Through Text Mining. IEEE Access, 8, 67698–67717. https://doi.org/10.1109/ACCESS.2020.2983656

Karasiak, N., Dejoux, J.-F., Monteil, C., & Sheeren, D. (2022). Spatial dependence between training and test sets: Another pitfall of classification accuracy assessment in remote sensing. Machine Learning, 111(7), 2715–2740. https://doi.org/10.1007/s10994-021-05972-1

Karimi, N., Ng, K. T. W., & Richter, A. (2022). Development and application of an analytical framework for mapping probable illegal dumping sites using nighttime light imagery and various remote sensing indices. Waste Management, 143, 195–205. https://doi.org/10.1016/j.wasman.2022.02.031

Kasanko, M., Barredo, J. I., Lavalle, C., McCormick, N., Demicheli, L., Sagris, V., & Brezger, A. (2006). Are European cities becoming dispersed? A comparative analysis of 15 European urban areas. Landscape and Urban Planning, 77(1), 111–130. https://doi.org/10.1016/j.landurbplan.2005.02.003

Kato, S., & Yamaguchi, Y. (2007). Estimation of storage heat flux in an urban area using ASTER data. Remote Sensing of Environment, 110(1), 1–17. https://doi.org/10.1016/j.rse.2007.02.011

Kaufmann, R. K., & Seto, K. C. (2001). Change detection, accuracy, and bias in a sequential analysis of Landsat imagery in the Pearl River Delta, China: Econometric techniques. Predicting Land-Use Change, 85(1), 95–105. https://doi.org/10.1016/S0167-8809(01)00190-6

Kaye, J. P., Groffman, P. M., Grimm, N. B., Baker, L. A., & Pouyat, R. V. (2006). A distinct urban biogeochemistry? Trends in Ecology & Evolution, 21(4), 192–199. https://doi.org/10.1016/j.tree.2005.12.006

Kelso, J. A. S., Stolk, E., & Portugali, J. (2016). Self-organization and Design as a Complementary Pair. In J. Portugali & E. Stolk (Eds.), Complexity, Cognition, Urban Planning and Design (pp. 43–53). Springer International Publishing.

Kestur, R., Farooq, S., Abdal, R., Mehraj, E., Omkar Subbaramajois Narasipura, & Meenavathi Mudigere. (2018). UFCN: a fully convolutional neural network for road extraction in RGB imagery acquired by remote sensing from an unmanned aerial vehicle. Journal of Applied Remote Sensing, 12(016020), 1–15. https://doi.org/10.1117/1.JRS.12.016020

Khan, F., Thodi, P., Imtiaz, S., & Abbassi, R. (2016). Real-time monitoring and management of offshore process system integrity. Biotechnology and Bioprocess Engineering / Process Systems Engineering, 14, 61–71. https://doi.org/10.1016/j.coche.2016.08.015

Khanna, P. (2016). A New Map for America. Parak Khanna. https://www.paragkhanna.com/2016-4-16-a-new-map-for-america/

Khomenko, S., Cirach, M., Barrera-Gómez, J., Pereira-Barboza, E., Iungman, T., Mueller, N., Foraster, M., Tonne, C., Thondoo, M., Jephcote, C., Gulliver, J., Woodcock, J., & Nieuwenhuijsen, M. (2022). Impact of road traffic noise on annoyance and preventable mortality in European cities: A health impact assessment. Environment International, 162(107160), 1–13. https://doi.org/10.1016/j.envint.2022.107160

Khomenko, S., Cirach, M., Pereira-Barboza, E., Mueller, N., Barrera-Gómez, J., Rojas-Rueda, D., de Hoogh, K., Hoek, G., & Nieuwenhuijsen, M. (2021). Premature mortality due to air pollution in European cities: A health impact assessment. The Lancet Planetary Health, 5(3), e121–e134. https://doi.org/10.1016/S2542-5196(20)30272-2

Kibert, C. J., Monroe, M. C., Peterson, A. L., Plate, R. R., & Thiele, L. P. (2011). Working Toward Sustainability: Ethical Decision-Making in a Technological World. Wiley.

Kidder, S. Q., & Wu, H.-T. (1987). A multispectral study of the St. Louis area under snow-covered conditions using NOAA-7 AVHRR data. Remote Sensing of Environment, 22(2), 159–172. https://doi.org/10.1016/0034-4257(87)90056-3

Kim, G. H., & Spafford, E. H. (1993). The design and implementation of tripwire: A file system integrity checker. Proceedings of the 2nd ACM Conference on Computer and Communications Security, 18–29. https://doi.org/10.1145/191177.191183

Kim, H. H. (1992). Urban heat island. International Journal of Remote Sensing, 13(12), 2319–2336. https://doi.org/10.1080/01431169208904271

Kim, J. I., Yeo, C. H., & Jin-Hwi, K. (2014). Spatial change in urban employment distribution in Seoul metropolitan city: Clustering, dispersion and general dispersion. International Journal of Urban Sciences, 18(3), 355–372. https://doi.org/10.1080/12265934.2014.916626

Kim, S. (2007). Changes in the nature of urban spatial structure in the United States, 1890-2000. Journal of Regional Science, 47(2), 273–287. https://doi.org/10.1111/j.1467-9787.2007.00509.x

King, L. J. (1985). Central Place Theory (G. I. Thrall, Ed.). Regional Research Institute, West Virginia University.

Kingma, D. P., & Ba, J. (2017). Adam: A Method for Stochastic Optimization. https://arxiv.org/abs/1412.6980

Kingsley, M., & EcoHealth Ontario. (2019). Commentary—Climate change, health and green space co-benefits. Health Promotion and Chronic Disease Prevention in Canada: Research, Policy and Practice, 39(4), 131–135. https://doi.org/10.24095/hpcdp.39.4.04

Kitchin, R., Lauriault, Tracey P., & McArdle, G. (2015). Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. Regional Studies, Regional Science, 2(1), 6–28. https://doi.org/10.1080/21681376.2014.983149

Klein Tank, A. M. G., Zwiers, F. W., & Zhang, X. (2009). Guidelines on Analysis of Extremes in a Changing Climate in Support of Informed Decisions for Adaptation (No. WCDMP-No. 72; Climate Data and Monitoring). World Meteorological Organization. https://www.ecad.eu/documents/WCDMP_72_TD_1500_en_1.pdf

Kloog, I., Koutrakis, P., Coull, B. A., Lee, H. J., & Schwartz, J. (2011). Assessing temporally and spatially resolved PM2.5 exposures for epidemiological studies using satellite aerosol optical depth measurements. Atmospheric Environment, 45(35), 6267–6275. https://doi.org/10.1016/j.atmosenv.2011.08.066

Kloog, I., Stevens, R. G., Haim, A., & Portnov, B. A. (2010). Nighttime light level co-distributes with breast cancer incidence worldwide. Cancer Causes & Control, 21(12), 2059–2068. https://doi.org/10.1007/s10552-010-9624-4

Kloosterman, R. C., & Musterd, S. (2001). The Polycentric Urban Region: Towards a Research Agenda. Urban Studies, 38(4), 623–633. https://doi.org/10.1080/00420980120035259

Klotz, M., Kemper, T., Geiß, C., Esch, T., & Taubenböck, H. (2016). How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. Remote Sensing of Environment, 178, 191–212. https://doi.org/10.1016/j.rse.2016.03.001

Köberl, M., Wurm, M., Droin, A., Garbasevschi, O. M., Dolls, M., & Taubenböck, H. (2024). Liveability in large housing estates in Germany – Identifying differences based on a novel concept for a walkable city. Landscape and Urban Planning, 251(105150), 1–22. https://doi.org/10.1016/j.landurbplan.2024.105150

Kochupillai, M. (2021). Outline of a Novel Approach for Identifying Ethical Issues in Early Stages of AI4EO Research. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 1165–1168. https://doi.org/10.1109/IGARSS47720.2021.9553727

Kochupillai, M., Gallersdörfer, U., Köninger, J., & Beck, R. (2021). Incentivizing research & innovation with agrobiodiversity conserved in situ: Possibilities and limitations of a blockchain-based solution. Journal of Cleaner Production, 309, 1–15. https://doi.org/10.1016/j.jclepro.2021.127155

Kochupillai, M., Kahl, M., Schmitt, M., Taubenböck, H., & Zhu, X. X. (2022). Earth Observation and Artificial Intelligence: Understanding emerging ethical issues and opportunities. IEEE Geoscience and Remote Sensing Magazine, 10(4), 90–124. https://doi.org/10.1109/MGRS.2022.3208357

Kochupillai, M., & Köninger, J. (2022). Cast into the Stones of International Law: A Critique of the UPOV Standards in the Light of Scientific Insights and Policy Shifts toward Agroecology and Natural Farming. In H. Grosse Ruse-Khan & A. Metzger (Eds.), Intellectual Property Ordering beyond Borders (pp. 28–63). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781009071338.004

Kochupillai, M., Lütge, C., & Poszler, F. (2020). Programming Away Human Rights and Responsibilities? "The Moral Machine Experiment" and the Need for a More "Humane" AV Future. NanoEthics, 14(3), 285–299. https://doi.org/10.1007/s11569-020-00374-4

Kochupillai, M., & Taubenböck, H. (2023). Conducting Ethically Mindful Earth Observation Research: The Case of Slum Mapping. IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, 1937–1940. https://doi.org/10.1109/IGARSS52108.2023.10281725

Koehn, D. (2005). Integrity as a Business Asset. Journal of Business Ethics, 58(1), 125–136. https://doi.org/10.1007/s10551-005-1391-x

Kohli, D., Sliuzas, R., Kerle, N., & Stein, A. (2012). An ontology of slums for image-based classification. Computers, Environment and Urban Systems, 36(2), 154–163. https://doi.org/10.1016/j.compenvurbsys.2011.11.001

Kolb, A. (1962). Die Geographie und die Kulturerdteile. In A. Leidlmair (Ed.), Hermann von Wissmann-Festschrift (pp. 42–49).

Kolosov, V., Medvedev, A., & Zotova, M. (2018). Comparing the development of border regions with the use of GIS (the case of Russia). Geographia Polonica, 91(1), 47–61. https://doi.org/10.7163/GPol.0090

Kong, F., Yin, H., James, P., Hutyra, L. R., & He, H. S. (2014). Effects of spatial pattern of greenspace on urban cooling in a large metropolitan area of eastern China. Landscape and Urban Planning, 128, 35–47. https://doi.org/10.1016/j.landurbplan.2014.04.018

Konovalov, I. B., Beekmann, M., Richter, A., Burrows, J. P., & Hilboll, A. (2010). Multi-annual changes of NOx emissions in megacity regions: Nonlinear trend analysis of satellite measurement based estimates. Atmospheric Chemistry and Physics, 10(17), 8481–8498. https://doi.org/10.5194/acp-10-8481-2010

Korte, K. (2021). Filtering or Blocking Mobility? Inequalities, Marginalization, and Power Relations at Fortified Borders. Historical Social Research, 46(3), 49–77. https://doi.org/10.12759/hsr.46.2021.3.49-77

Kostof, S. (1982). Urbanism and Polity: Medieval Siena in Context. In International Laboratory for Architecture and Urban Design Yearbook (pp. 62–73).

Kostof, S. (1991a). Lectures from Architecture 170 class, University of California, Berkeley. https://digicoll.lib.berkeley.edu/record/153507?ln=en

Kostof, S. (1991b). The city shaped: Urban patterns and meanings through history. Thames & Hudson.

Kostof, S. (1992). The city assembled: The elements of urban form through history. Thames & Hudson.

Kraff, N. J., Wurm, M., & Taubenböck, H. (2020a). The dynamics of poor urban areas—Analyzing morphologic transformations across the globe using Earth observation data. Cities, 107(102905), 1–15. https://doi.org/10.1016/j.cities.2020.102905

Kraff, N. J., Wurm, M., & Taubenböck, H. (2020b). Uncertainties of Human Perception in Visual Image Interpretation in Complex Urban Environments. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 4229–4241. https://doi.org/10.1109/JSTARS.2020.3011543

Krapf-Askari, E. (1969). Yoruba Towns and Cities: An Enquiry into the Nature of Urban Social Phenomena (1st ed.). Oxford University Press.

Kraus, S. P., Senger, L. W., & Ryerson, J. M. (1974). Estimating population from photographically determined residential land use types. Remote Sensing of Environment, 3(1), 35–42. https://doi.org/10.1016/0034-4257(74)90036-4

Krausmann, F., Gingrich, S., Eisenmenger, N., Erb, K.-H., Haberl, H., & Fischer-Kowalski, M. (2009). Growth in global materials use, GDP and population during the 20th century. Ecological Economics, 68(10), 2696–2705. https://doi.org/10.1016/j.ecolecon.2009.05.007

Krayenhoff, E. S., Moustaoui, M., Broadbent, A. M., Gupta, V., & Georgescu, M. (2018). Diurnal interaction between urban expansion, climate change and adaptation in US cities. Nature Climate Change, 8(12), 1097–1103. https://doi.org/10.1038/s41558-018-0320-9

Krefis, A. C., Augustin, M., Schlünzen, K. H., Oßenbrügge, J., & Augustin, J. (2018). How Does the Urban Environment Affect Health and Well-Being? A Systematic Review. Urban Science, 2(21), 1–21. https://doi.org/10.3390/urbansci2010021

Krehl, A. (2016). Urban subcentres in German city regions: Identification, understanding, comparison. Papers in Regional Science, 97, 79–105. https://doi.org/10.1111/pirs.12235

Krehl, A., & Siedentop, S. (2019). Towards a typology of urban centers and subcenters – evidence from German city regions. Urban Geography, 40(1), 58–82. https://doi.org/10.1080/02723638.2018.1500245

Krehl, A., Siedentop, S., Taubenböck, H., & Wurm, M. (2016). A Comprehensive View on Urban Spatial Structure: Urban Density Patterns of German City Regions. ISPRS International Journal of Geo-Information, 5(76), 1–21. https://doi.org/10.3390/ijgi5060076

Krey, V., Masera, O., Blanford, G., Bruckner, T., Cooke, R., Fisher-Vanden, K., Haberl, H., Hertwich, E., Kriegler, E., Mueller, D., Paltsev, S., Price, L., Schlömer, S., a Ürge-Vorsatz, D., van Vuuren, D., & Zwickel, T. (2014). Annex II: Metrics & Methodology. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, & J. C. Minx (Eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 1281–1328). Cambridge University Press. https://www.ipcc.ch/report/ar5/wg3/

Krizek, K. J. (2011). Residential Relocation and Changes in Urban Travel: Does Neighborhood-Scale Urban Form Matter? Journal of the American Planning Association, 69(3), 265–281. https://doi.org/10.1080/01944360308978019

Kropf, K. (1993). The definition of built form in urban morphology. University of Birmingham.

Kropf, K. (2009). Aspects of urban form. Urban Morphology, 13(2), 105–120. https://doi.org/10.51347/jum.v13i2.3949

Krueger, R., & Gibbs, D. (2007). The Sustainable Development Paradox: Urban Political Economy in the United States and Europe. Guilford Press.

Kruskal, W. H., & Wallis, W. A. (1952). Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association, 47(260), 583–621. https://doi.org/10.1080/01621459.1952.10483441

Kruspe, A., Häberle, M., Hoffmann, E. J., Rode-Hasinger, S., Abdulahhad, K., & Zhu, X. X. (2021). Changes in Twitter geolocations: Insights and suggestions for future usage. https://arxiv.org/abs/2108.12251

Krylov, A., McCarty, J. L., Potapov, P., Loboda, T., Tyukavina, A., Turubanova, S., & Hansen, M. C. (2014). Remote sensing estimates of stand-replacement fires in Russia, 2002–2011. Environmental Research Letters, 9(105007), 1–8. https://doi.org/10.1088/1748-9326/9/10/105007

Kuang, B., Lu, X., Han, J., Fan, X., & Zuo, J. (2020). How urbanization influence urban land consumption intensity: Evidence from China. Habitat International, 100(102103), 1–8. https://doi.org/10.1016/j.habitatint.2019.102103

Kuang, W., Liu, J., Dong, J., Chi, W., & Zhang, C. (2016). The rapid and massive urban and industrial land expansions in China between 1990 and 2010: A CLUD-based analysis of their trajectories, patterns, and drivers. Landscape and Urban Planning, 145, 21–33. https://doi.org/10.1016/j.landurbplan.2015.10.001

Kuenen, J., Dellaert, S., Visschedijk, A., Jalkanen, J.-P., Super, I., & Denier van der Gon, H. (2022). CAMS-REG-v4: A state-of-the-art high-resolution European emission inventory for air quality modelling. Earth System Science Data, 14(2), 491–515. https://doi.org/10.5194/essd-14-491-2022

Kuffer, M. (2017). Spatial patterns of deprivation in cities of the global south in very high resolution imagery [PhD Thesis - Research UT, graduation UT, University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC)]. https://doi.org/10.3990/1.9789036543699

Kuffer, M., Barros, J., & Sliuzas, R. V. (2014). The development of a morphological unplanned settlement index using very-high-resolution (VHR) imagery. Computers, Environment and Urban Systems, 48, 138–152. https://doi.org/10.1016/j.compenvurbsys.2014.07.012

Kuffer, M., Pfeffer, K., & Sliuzas, R. (2016). Slums from Space—15 Years of Slum Mapping Using Remote Sensing. Remote Sensing, 8(455), 1–29. https://doi.org/10.3390/rs8060455

Kuffer, M., Pfeffer, K., Sliuzas, R., Taubenböck, H., Baud, I., & Maarseveen, M. van. (2018). Capturing the Urban Divide in Nighttime Light Images From the International Space Station. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(8), 2578–2586. https://doi.org/10.1109/JSTARS.2018.2828340

Kuffer, M., Wang, J., Nagenborg, M., Pfeffer, K., Kohli, D., Sliuzas, R., & Persello, C. (2018). The Scope of Earth-Observation to Improve the Consistency of the SDG Slum Indicator. ISPRS International Journal of Geo-Information, 7(428), 1–28. https://doi.org/10.3390/ijgi7110428

Kühnl, M., Sapena, M., & Taubenböck, H. (2021). Categorizing Urban Structural Types using an Object-Based Local Climate Zone Classification Scheme in Medellín, Colombia. In M. Schrenk, V. V. Popovich, P. Zeile, P. Elisei, C. Beyer, J. Ryser, & G. Stöglehner (Eds.), CITIES 20.50 Creating Habitats for the 3rd Millennium: Smart – Sustainable – Climate Neutral. Proceedings of REAL CORP 2021, 26th International Conference on Urban Planning and Regional Development in the Information Society (pp. 173–182). https://archive.corp.at/cdrom2021/papers2021/CORP2021_13.pdf

Kühnl, M., Sapena, M., Wurm, M., Geiß, C., & Taubenböck, H. (2023). Multitemporal landslide exposure and vulnerability assessment in Medellín, Colombia. Natural Hazards, 119(2), 883–906. https://doi.org/10.1007/s11069-022-05679-z

Kumar, N., Chu, A., & Foster, A. (2007). An empirical relationship between PM2.5 and aerosol optical depth in Delhi Metropolitan. Atmospheric Environment, 41(21), 4492–4503. https://doi.org/10.1016/j.atmosenv.2007.01.046

Küpper, P. (2016). Abgrenzung und Typisierung ländlicher Räume. Thünen Working Paper, 68, 1–53. https://doi.org/10.3220/WP1481532921000

Kussul, N., Lavreniuk, M., Kolotii, A., Skakun, S., Rakoid, O., & Shumilo, L. (2019). A workflow for Sustainable Development Goals indicators assessment based on high-resolution satellite data. International Journal of Digital Earth, 13(2), 309–321. https://doi.org/10.1080/17538947.2019.1610807

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software, 82(13), 1–26. https://doi.org/10.18637/jss.v082.i13

L

LaDeau, S. L., Allan, B. F., Leisnham, P. T., & Levy, M. Z. (2015). The ecological foundations of transmission potential and vector-borne disease in urban landscapes. Functional Ecology, 29(7), 889–901. https://doi.org/10.1111/1365-2435.12487

Lambin, E. F., Geist, H., & Rindfuss, R. R. (2006). Introduction: Local Processes with Global Impacts. In E. F. Lambin & H. Geist (Eds.), Land-Use and Land-Cover Change: Local Processes and Global Impacts (pp. 1–8). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-32202-7_1

Lamchin, M., Wang, S. W., Lim, C.-H., Ochir, A., Pavel, U., Gebru, B. M., Choi, Y., Jeon, S. W., & Lee, W.-K. (2020). Understanding global spatio-temporal trends and the relationship between vegetation greenness and climate factors by land cover during 1982–2014. Global Ecology and Conservation, 24, 1–17. https://doi.org/10.1016/j.gecco.2020.e01299

Lane, J. K., Stokes, E. C., Seto, K. C., Thanikachalam, S., Thanikachalam, M., & Bell, M. L. (2017). Associations between Greenness, Impervious Surface Area, and Nighttime Lights on Biomarkers of Vascular Aging in Chennai, India. Environmental Health Perspectives, 125(8), 1–8. https://doi.org/10.1289/EHP541

Lang, R. E., & Dhavale, D. (2005). Beyond Megalopolis: Exploring America's New "Megapolitan" Geography (pp. 1–33). Brookings Mountain West. https://digitalscholarship.unlv.edu/brookings_pubs/38

Lang, R., & Knox, P. K. (2009). The New Metropolis: Rethinking Megalopolis. Regional Studies, 43(6), 789–802. https://doi.org/10.1080/00343400701654251

Lang, W., Lang, H., Hui, E. C. M., Chen, T., Wu, J., & Jahre, M. (2022). Measuring urban vibrancy of neighborhood performance using social media data in Oslo, Norway. Cities, 131(103908), 1–16. https://doi.org/10.1016/j.cities.2022.103908

Lange, K., Richter, A., & Burrows, J. P. (2022). Variability of nitrogen oxide emission fluxes and lifetimes estimated from Sentinel-5P TROPOMI observations. Atmospheric Chemistry and Physics, 22(4), 2745–2767. https://doi.org/10.5194/acp-22-2745-2022

Larkham, P. J. (1992). Organic Thought in Urban Geography: 'The 'Evolution' of towns. Australian Geographical Studies, 30(1), 3–8. https://doi.org/10.1111/j.1467-8470.1992.tb00728.x

Larondelle, N., & Haase, D. (2013). Urban ecosystem services assessment along a rural–urban gradient: A cross-analysis of European cities. Ecological Indicators, 29, 179–190. https://doi.org/10.1016/j.ecolind.2012.12.022

Larondelle, N., Hamstead, Z. A., Kremer, P., Haase, D., & McPhearson, T. (2014). Applying a novel urban structure classification to compare the relationships of urban structure and surface temperature in Berlin and New York City. Applied Geography, 53, 427–437. https://doi.org/10.1016/j.apgeog.2014.07.004

Latham, J., Cumani, R., Rosati, I., & Bloise, M. (2014). Global land cover share (GLC-SHARE) database beta-release version 1.0-2014 (p. 40). FAO. https://www.fao.org/uploads/media/glc-share-doc.pdf

Lawes, R. A., & Wallace, J. F. (2008). Monitoring an invasive perennial at the landscape scale with remote sensing. Ecological Management & Restoration, 9(1), 53–59. https://doi.org/10.1111/j.1442-8903.2008.00387.x

Le Galès, P., & Zagrodzki, M. (2006). Cities are back in town: The US/Europe comparison. Cahier Européen, 5(6).

Le Governement du Grand-Duch′ e de Luxembourg. (2022). Bâtiments 3D 2013 (LOD 1 – Level of Detail 1) [Dataset]. https://data.public.lu/en/datasets/5cecd25a4384b06ad27e5c58/#_

Le, T., Abrahart, R., Aplin, R., & Priestnall, G. (2005). Town centre modelling based on public participation. Proceedings of the CUPUM 5.

Lechner, A. M., Reinke, K. J., Wang, Y., & Bastin, L. (2013). Interactions between landcover pattern and geospatial processing methods: Effects on landscape metrics and classification accuracy. Ecological Complexity, 15, 71–82. https://doi.org/10.1016/j.ecocom.2013.03.003

Lee, C. M., Cable, M. L., Hook, S. J., Green, R. O., Ustin, S. L., Mandl, D. J., & Middleton, E. M. (2015). An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. Remote Sensing of Environment, 167, 6–19. https://doi.org/10.1016/j.rse.2015.06.012

Lee, Y. S. (2018). International isolation and regional inequality: Evidence from sanctions on North Korea. Journal of Urban Economics, 103, 34–51. https://doi.org/10.1016/j.jue.2017.11.002

Lefebvre, A., Sannier, C., & Corpetti, T. (2016). Monitoring Urban Areas with Sentinel-2A Data: Application to the Update of the Copernicus High Resolution Layer Imperviousness Degree. Remote Sensing, 8(606), 1–21. https://doi.org/10.3390/rs8070606

Lefebvre, H. (1974). La production de l'espace. L'Homme et La Société, 31–32, 15–32. https://doi.org/10.3406/homso.1974.1855

Legeby, A., Bergerhauser Pont, M., & Marcus, L. (2015). Streets for co-presence? Mapping potentials. In K. Karimi, L. Vaughan, K. Sailer, G. Palaiologou, & T. Bolton (Eds.), Proceedings of the 10th International Space Syntax Symposium (p. 108:1-108:17).

Lehnert, M., Geletič, J., Husák, J., & Vysoudil, M. (2015). Urban field classification by "local climate zones" in a medium-sized Central European city: The case of Olomouc (Czech Republic). Theoretical and Applied Climatology, 122(3), 531–541. https://doi.org/10.1007/s00704-014-1309-6

Lei, Y., Tian, J., & Lin, L. (2015). A method for automatic delineation of city centers using POI and road networks. Cehui Xuebao/Acta Geodaetica et Cartographica Sinica, 44, 42–48.

Leibrand, A., Sadoff, N., Maslak, T., & Thomas, A. (2019). Using Earth Observations to Help Developing Countries Improve Access to Reliable, Sustainable, and Modern Energy. Frontiers in Environmental Science, Volume 7(123), 1–14. https://doi.org/10.3389/fenvs.2019.00123

Leichtle, T., Kühnl, M., Droin, A., Beck, C., Hiete, M., & Taubenböck, H. (2023). Quantifying urban heat exposure at fine scale—Modeling outdoor and indoor temperatures using citizen science and VHR remote sensing. Urban Climate, 49(101522), 1–25. https://doi.org/10.1016/j.uclim.2023.101522

Lelieveld, J., Pozzer, A., Pöschl, U., Fnais, M., Haines, A., & Münzel, T. (2020). Loss of life expectancy from air pollution compared to other risk factors: A worldwide perspective. Cardiovascular Research, 116(11), 1910–1917. https://doi.org/10.1093/cvr/cvaa025

Lelovics E, Unger J, Gál T, & Gál CV. (2014). Design of an urban monitoring network based on Local Climate Zone mapping and temperature pattern modelling. Climate Research, 60(1), 51–62. https://doi.org/10.3354/cr01220

Lemoine-Rodríguez, R., Biewer, C., Inostroza, L., & Taubenböck, H. (2025). Online Concerns About Urban Heat: Social Media, Temperature, and Urban Form. 2025 Joint Urban Remote Sensing Event (JURSE), CFP25RSD-ART, 1–4. https://doi.org/10.1109/JURSE60372.2025.11075998

Lemoine-Rodríguez, R., Inostroza, L., & Zepp, H. (2020). The global homogenization of urban form. An assessment of 194 cities across time. Landscape and Urban Planning, 204(103949), 1–16. https://doi.org/10.1016/j.landurbplan.2020.103949

Lemoine-Rodríguez, R., Inostroza, L., & Zepp, H. (2022). Does urban climate follow urban form? Analysing intraurban LST trajectories versus urban form trends in 3 cities with different background climates. Science of The Total Environment, 830(154570), 1–13. https://doi.org/10.1016/j.scitotenv.2022.154570

Lemoine-Rodríguez, R., Mast, J., Mühlbauer, M., Mandery, N., Biewer, C., & Taubenböck, H. (2024). The voices of the displaced: Mobility and Twitter conversations of migrants of Ukraine in 2022. Information Processing & Management, 61(103670), 1–22. https://doi.org/10.1016/j.ipm.2024.103670

Levene, H. (1960). Robust tests for equality of variances. In Robust Tests for Equality of Variances (pp. 278–292).

Li, C., Gong, P., Wang, J., Zhu, Z., Biging, G. S., Yuan, C., Hu, T., Zhang, H., Wang, Q., Li, X., Liu, X., Xu, Y., Guo, J., Liu, C., Hackman, K. O., Zhang, M., Cheng, Y., Yu, L., Yang, J., … Clinton, N. (2017). The first all-season sample set for mapping global land cover with Landsat-8 data. Science Bulletin, 62(7), 508–515. https://doi.org/10.1016/j.scib.2017.03.011

Li, D., Bou-Zeid, E., & Oppenheimer, M. (2014). The effectiveness of cool and green roofs as urban heat island mitigation strategies. Environmental Research Letters, 9(055002), 1–16. https://doi.org/10.1088/1748-9326/9/5/055002

Li, F., Yigitcanlar, T., Nepal, M., Nguyen, K., & Dur, F. (2023). Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework. Sustainable Cities and Society, 96(104653), 1–30. https://doi.org/10.1016/j.scs.2023.104653

Li, J., Song, C., Cao, L., Zhu, F., Meng, X., & Wu, J. (2011). Impacts of landscape structure on surface urban heat islands: A case study of Shanghai, China. Remote Sensing of Environment, 115(12), 3249–3263. https://doi.org/10.1016/j.rse.2011.07.008

Li, J., Wang, X., Wang, X., Ma, W., & Zhang, H. (2009). Remote sensing evaluation of urban heat island and its spatial pattern of the Shanghai metropolitan area, China. Ecological Complexity, 6(4), 413–420. https://doi.org/10.1016/j.ecocom.2009.02.002

Li, L., Cheng, W.-Y., Glicksberg, B. S., Gottesman, O., Tamler, R., Chen, R., Bottinger, E. P., & Dudley, J. T. (2015). Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Science Translational Medicine, 7(311), 311ra174-311ra174. https://doi.org/10.1126/scitranslmed.aaa9364

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. Cartography and Geographic Information Science, 40(2), 61–77. https://doi.org/10.1080/15230406.2013.777139

Li, M., Koks, E., Taubenböck, H., & van Vliet, J. (2020). Continental-scale mapping and analysis of 3D building structure. Remote Sensing of Environment, 245(111859), 1–16. https://doi.org/10.1016/j.rse.2020.111859

Li, M., Verburg, P. H., & van Vliet, J. (2022). Global trends and local variations in land take per person. Landscape and Urban Planning, 218(104308), 1–13. https://doi.org/10.1016/j.landurbplan.2021.104308

Li, M., Wang, Y., Rosier, J. F., Verburg, P. H., & van Vliet, J. (2022). Global maps of 3D built-up patterns for urban morphological analysis. International Journal of Applied Earth Observation and Geoinformation, 114(103048), 1–9. https://doi.org/10.1016/j.jag.2022.103048

Li, P., & Cao, X. (2005). Evolution and development of Guangzhou-Hong Kong corridor. Chinese Geographical Science, 15(3), 206–211. https://doi.org/10.1007/s11769-005-0031-9

Li, W., Zhou, W., Bai, Y., Pickett, S. T. A., & Han, L. (2018). The smart growth of Chinese cities: Opportunities offered by vacant land. Land Degradation & Development, 29(10), 3512–3520. https://doi.org/10.1002/ldr.3125

Li, X., & Gong, P. (2016). Urban growth models: Progress and perspective. Science Bulletin, 61(21), 1637–1650. https://doi.org/10.1007/s11434-016-1111-1

Li, X., Gong, P., & Liang, L. (2015). A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data. Remote Sensing of Environment, 166, 78–90. https://doi.org/10.1016/j.rse.2015.06.007

Li, X., Li, X., Li, D., He, X., & Jendryke, M. (2019). A preliminary investigation of Luojia-1 night-time light imagery. Remote Sensing Letters, 10(6), 526–535. https://doi.org/10.1080/2150704X.2019.1577573

Li, X., Wang, C., Zhang, G., Xiao, L., & Dixon, J. (2012). Urbanisation and human health in China: Spatial features and a systemic perspective. Environmental Science and Pollution Research, 19(5), 1375–1384. https://doi.org/10.1007/s11356-011-0718-7

Li, X., Zhou, Y., Asrar, G. R., Imhoff, M., & Li, X. (2017). The surface urban heat island response to urban expansion: A panel analysis for the conterminous United States. Science of The Total Environment, 605–606, 426–435. https://doi.org/10.1016/j.scitotenv.2017.06.229

Li, X., Zhou, Y., Meng, L., Asrar, G., Sapkota, A., & Coates, F. (2019). Characterizing the relationship between satellite phenology and pollen season: A case study of birch. Remote Sensing of Environment, 222, 267–274. https://doi.org/10.1016/j.rse.2018.12.036

Li, Y., Jia, L., Wu, W., Yan, J., & Liu, Y. (2018). Urbanization for rural sustainability – Rethinking China's urbanization strategy. Journal of Cleaner Production, 178, 580–586. https://doi.org/10.1016/j.jclepro.2017.12.273

Li, Y., Martinis, S., & Wieland, M. (2019). Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. ISPRS Journal of Photogrammetry and Remote Sensing, 152, 178–191. https://doi.org/10.1016/j.isprsjprs.2019.04.014

Li, Y., Wang, L., Zhang, L., Liu, M., & Zhao, G. (2019). Monitoring Intra-annual Spatiotemporal Changes in Urban Heat Islands in 1449 Cities in China Based on Remote Sensing. Chinese Geographical Science, 29(6), 905–916. https://doi.org/10.1007/s11769-019-1080-9

Liang, Z., Li, Z., & Ma, Z. (2014). Changing Patterns of the Floating Population in China, 2000–2010. Population and Development Review, 40(4), 695–716. https://doi.org/10.1111/j.1728-4457.2014.00007.x

Liao, W., Pizurica, A., Bellens, R., Gautama, S., & Philips, W. (2015). Generalized graph-based fusion of hyperspectral and LiDAR data using morphological features. In IEEE Geoscience and Remote Sensing Letters (Vol. 12, Issue 3, pp. 552–556). https://doi.org/10.1109/LGRS.2014.2350263

Lichtenberger, E. (1972). Die europäische Stadt: Wesen, Modelle, Probleme. Berichte Zur Raumforschung Und Raumplanung, 16(1), 3–25.

Lin, J., Huang, B., Chen, M., & Huang, Z. (2014). Modeling urban vertical growth using cellular automata—Guangzhou as a case study. Applied Geography, 53, 172–186. https://doi.org/10.1016/j.apgeog.2014.06.007

Lin, J., Wan, H., & Cui, Y. (2020). Analyzing the spatial factors related to the distributions of building heights in urban areas: A comparative case study in Guangzhou and Shenzhen. Sustainable Cities and Society, 52(101854), 1–9. https://doi.org/10.1016/j.scs.2019.101854

Lin, L., Liu, M., Luo, F., Wang, K., Zhang, Q., & Xiang, W.-N. (2012). Comment on "The study of urban metabolism and its applications to urban planning and design" by Kennedy et al. (2011). Environmental Pollution, 167, 184–185. https://doi.org/10.1016/j.envpol.2012.04.011

Linnell, J. D. C., Trouwborst, A., Boitani, L., Kaczensky, P., Huber, D., Reljic, S., Kusak, J., Majic, A., Skrbinsek, T., Potocnik, H., Hayward, M. W., Milner-Gulland, E. J., Buuveibaatar, B., Olson, K. A., Badamjav, L., Bischof, R., Zuther, S., & Breitenmoser, U. (2016). Border Security Fencing and Wildlife: The End of the Transboundary Paradigm in Eurasia? PLOS Biology, 14(6)(e1002483), 1–13. https://doi.org/10.1371/journal.pbio.1002483

Lisini, G., Salentinig, A., Du, P., & Gamba, P. (2018). SAR-Based Urban Extents Extraction: From ENVISAT to Sentinel-1. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(8), 2683–2691. https://doi.org/10.1109/JSTARS.2017.2782180

Liu, C., Huang, X., Zhu, Z., Chen, H., Tang, X., & Gong, J. (2019). Automatic extraction of built-up area from ZY3 multi-view satellite imagery: Analysis of 45 global cities. Remote Sensing of Environment, 226, 51–73. https://doi.org/10.1016/j.rse.2019.03.033

Liu, F., Zhang, Z., Zhao, X., Wang, X., Zuo, L., Wen, Q., Yi, L., Xu, J., Hu, S., & Liu, B. (2019). Chinese cropland losses due to urban expansion in the past four decades. Science of The Total Environment, 650, 847–857. https://doi.org/10.1016/j.scitotenv.2018.09.091

Liu, H., & Weng, Q. (2012). Enhancing temporal resolution of satellite imagery for public health studies: A case study of West Nile Virus outbreak in Los Angeles in 2007. Remote Sensing of Environment, 117, 57–71. https://doi.org/10.1016/j.rse.2011.06.023

Liu, J., & Deng, X. (2010). Progress of the research methodologies on the temporal and spatial process of LUCC. Chinese Science Bulletin, 55(14), 1354–1362. https://doi.org/10.1007/s11434-009-0733-y

Liu, J., Liu, M., Zhuang, D., Zhang, Z., & Deng, X. (2003). Study on spatial pattern of land-use change in China during 1995–2000. Science in China Series D: Earth Sciences, 46(4), 373–384. https://doi.org/10.1360/03yd9033

Liu, J., Xu, Q., Yi, J., & Huang, X. (2022). Analysis of the heterogeneity of urban expansion landscape patterns and driving factors based on a combined Multi-Order Adjacency Index and Geodetector model. Ecological Indicators, 136(108655), 1–15. https://doi.org/10.1016/j.ecolind.2022.108655

Liu, J., Zhan, J., & Deng, X. (2005). Spatio-temporal Patterns and Driving Forces of Urban Land Expansion in China during the Economic Reform Era. AMBIO: A Journal of the Human Environment, 34(6), 450–455. https://doi.org/10.1579/0044-7447-34.6.450

Liu, X., Huang, Y., Xu, X., Li, X., Li, X., Ciais, P., Lin, P., Gong, K., Ziegler, A. D., Chen, A., Gong, P., Chen, J., Hu, G., Chen, Y., Wang, S., Wu, Q., Huang, K., Estes, L., & Zeng, Z. (2020). High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. Nature Sustainability, 3(7), 564–570. https://doi.org/10.1038/s41893-020-0521-x

Liu, X., & Sweeney, J. (2012). Modelling the impact of urban form on household energy demand and related CO2 emissions in the Greater Dublin Region. Energy Policy, 46, 359–369. https://doi.org/10.1016/j.enpol.2012.03.070

Liu, X., Wu, X., Li, X., Xu, X., Liao, W., Jiao, L., Zeng, Z., Chen, G., & Li, X. (2025). Global Mapping of Three-Dimensional Urban Structures Reveals Escalating Utilization in the Vertical Dimension and Pronounced Building Space Inequality. Engineering, 47, 86–99. https://doi.org/10.1016/j.eng.2024.01.025

Liu, Y., Huang, H., Yan, L., Liu, X., & Zhang, Z. (2016). Influence of suspended kelp culture on seabed sediment composition in Heini Bay, China. Estuarine, Coastal and Shelf Science, 181, 39–50. https://doi.org/10.1016/j.ecss.2016.07.017

Liu, Z., He, C., Zhou, Y., & Wu, J. (2014). How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. Landscape Ecology, 29(5), 763–771. https://doi.org/10.1007/s10980-014-0034-y

Lloyd, D. A. (2009). GIS-based Applications and the English Planning System: Modeling Town Centers and Housing Potential. In H. Hagen, S. Guhathakurta, & G. Steinebach (Eds.), Visualizing Sustainable Planning (pp. 51–61). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-88203-9_3

Lo, C. P., & Welch, R. (1977). Chinese urban population estimates. Annals of the Association of American Geographers, 67(2), 246–253. https://doi.org/10.1111/j.1467-8306.1977.tb01137.x

Locke, D. H., Hall, B., Grove, J. M., Pickett, S. T. A., Ogden, L. A., Aoki, C., Boone, C. G., & O'Neil-Dunne, J. P. M. (2021). Residential housing segregation and urban tree canopy in 37 US Cities. Npj Urban Sustainability, 1(15), 1–9. https://doi.org/10.1038/s42949-021-00022-0

Lomborg, S., & Bechmann, A. (2014). Using APIs for Data Collection on Social Media. The Information Society, 30(4), 256–265. https://doi.org/10.1080/01972243.2014.915276

López-Goyburu, P., & García-Montero, L. G. (2018). The urban-rural interface as an area with characteristics of its own in urban planning: A review. Sustainable Cities and Society, 43, 157–165. https://doi.org/10.1016/j.scs.2018.07.010

Loveland, T. R., Reed, B. C., Ohlen, D. O., Brown, J. F., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing, 21(6–7), 1303–1330. Scopus. https://doi.org/10.1080/014311600210191

Löw, M. (2018). Soziologie der Städte. Suhrkamp Verlag.

Lowe, M., Adlakha, D., Sallis, J. F., Salvo, D., Cerin, E., Moudon, A. V., Higgs, C., Hinckson, E., Arundel, J., Boeing, G., Liu, S., Mansour, P., Gebel, K., Puig-Ribera, A., Mishra, P. B., Bozovic, T., Carson, J., Dygrýn, J., Florindo, A. A., … Giles-Corti, B. (2022). City planning policies to support health and sustainability: An international comparison of policy indicators for 25 cities. The Lancet Global Health, 10(6), e882–e894. https://doi.org/10.1016/S2214-109X(22)00069-9

Lu, D. F., & Dang, A. R. (2016). New Villages on the Urban Fringe: Spatial Planning, Lifestyle Changes, and Health Implications. In Urbanization and Public Health in China (pp. 197–212). https://doi.org/10.1142/9781783268559_0010

Lu, H., Caverlee, J., & Niu, W. (2015). BiasWatch: A Lightweight System for Discovering and Tracking Topic-Sensitive Opinion Bias in Social Media. Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, 213–222. https://doi.org/10.1145/2806416.2806573

Lu, X., Ye, X., Zhou, M., Zhao, Y., Weng, H., Kong, H., Li, K., Gao, M., Zheng, B., Lin, J., Zhou, F., Zhang, Q., Wu, D., Zhang, L., & Zhang, Y. (2021). The underappreciated role of agricultural soil nitrogen oxide emissions in ozone pollution regulation in North China. Nature Communications, 12(5021), 1–9. https://doi.org/10.1038/s41467-021-25147-9

Luck, M., & Wu, J. (2002). A gradient analysis of urban landscape pattern: A case study from the Phoenix metropolitan region, Arizona, USA. Landscape Ecology, 17(4), 327–339. https://doi.org/10.1023/A:1020512723753

Luetge, C., & Mukerji, N. (Eds.). (2016). Order Ethics: An Ethical Framework for the Social Market Economy (1st ed.). Springer Cham. https://doi.org/10.1007/978-3-319-33151-5

Lüscher, P., & Weibel, R. (2012). Exploiting empirical knowledge for automatic delineation of city centres from large-scale topographic databases. Computers, Environment and Urban Systems, 37, 18–34. https://doi.org/10.1016/j.compenvurbsys.2012.07.001

Lussier, B., Hartmann, N. N., & Bolander, W. (2019). Curbing the Undesirable Effects of Emotional Exhaustion on Ethical Behaviors and Performance: A Salesperson–Manager Dyadic Approach. Journal of Business Ethics, 169(4), 747–766. https://doi.org/10.1007/s10551-019-04271-z

Lv, Y., Lan, Z., Kan, C., & Zheng, X. (2021). Polycentric Urban Development and its Determinants in China: A Geospatial Big Data Perspective. Geographical Analysis, 53(3), 520–542. https://doi.org/10.1111/gean.12236

Lynch, K. (1960). The image of the city. MIT Press.

Lynch, K. (1981). Good city form. MIT Press.

Lynch, K., & Rodwin, L. (1958). A Theory of Urban Form. Journal of the American Institute of Planners, 24(4), 201–214. https://doi.org/10.1080/01944365808978281

M

MacDonald, C., & Marcoux, A. (n.d.). Ethical Theory: Overview. Concise Encyclopedia of Business Ethics. Retrieved 4 January 2023, from https://conciseencyclopedia.org/ entries/ethical-theory/

Machault, V., Vignolles, C., Pagès, F., Gadiaga, L., Gaye, A., Sokhna, C., Trape, J.-F., Lacaux, J.-P., & Rogier, C. (2010). Spatial heterogeneity and temporal evolution of malaria transmission risk in Dakar, Senegal, according to remotely sensed environmental data. Malaria Journal, 9(252), 1–14. https://doi.org/10.1186/1475-2875-9-252

Madelin, M., Grasland, C., Mathian, E., Sanders, L., & Vincent, J.-M. (2009). Das „MAUP ": Modifiable Areal Unit-Problem oder Fortschritt? (No. 10; Informationen Zur Raumentwicklung, pp. 645–660).

Madlener, R., & Sunak, Y. (2011). Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management? Sustainable Cities and Society, 1(1), 45–53. https://doi.org/10.1016/j.scs.2010.08.006

Mahendra, A., & Seto, K. C. (2019). Upward and outward growth: Managing urban expansion for more equitable cities in the global south. World Resources Institute.

Mahtta, R., Fragkias, M., Güneralp, B., Mahendra, A., Reba, M., Wentz, E. A., & Seto, K. C. (2022). Urban land expansion: The role of population and economic growth for 300+ cities. Npj Urban Sustainability, 2(5), 1–11. https://doi.org/10.1038/s42949-022-00048-y

Mahtta, R., Mahendra, A., & Seto, K. C. (2019). Building up or spreading out? Typologies of urban growth across 478 cities of 1 million+. Environmental Research Letters, 14(124077), 1–14. https://doi.org/10.1088/1748-9326/ab59bf

Malczewski, J. (2004). GIS-based land-use suitability analysis: A critical overview. Progress in Planning, 62(1), 3–65. https://doi.org/10.1016/j.progress.2003.09.002

Malik, M., Lamba, H., Nakos, C., & Pfeffer, J. (2015). Population Bias in Geotagged Tweets. Proceedings of the International AAAI Conference on Web and Social Media, 9(4), 18–27. https://doi.org/10.1609/icwsm.v9i4.14688

Malizia, N. (2013). The Effect of Data Inaccuracy on Tests of Space-Time Interaction. Transactions in GIS, 17(3), 426–451. https://doi.org/10.1111/j.1467-9671.2012.01350.x

Malyuta, D., Brommer, C., Hentzen, D., Stastny, T., Siegwart, R., & Brockers, R. (2020). Long-duration fully autonomous operation of rotorcraft unmanned aerial systems for remote-sensing data acquisition. Journal of Field Robotics, 37(1), 137–157. https://doi.org/10.1002/rob.21898

Mangin, D., & Panerai, P. (1999). Projet urbain. Parenthèses. https://editionsparentheses.com/IMG/pdf/p604_projet_urbain.pdf

Mangin, J. (2014). The New Exclusionary Zoning. Stanford Law and Policy Review, 25(91), 91–120.

Maniadaki, M., Papathanasopoulos, A., Mitrou, L., & Maria, E.-A. (2021). Reconciling Remote Sensing Technologies with Personal Data and Privacy Protection in the European Union: Recent Developments in Greek Legislation and Application Perspectives in Environmental Law. Laws, 10(33), 1–19. https://doi.org/10.3390/laws10020033

Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and Health Impacts of Air Pollution: A Review. Frontiers in Public Health, 8(14), 1–13. https://doi.org/10.3389/fpubh.2020.00014

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics, 18(1), 50–60. https://doi.org/10.1214/aoms/1177730491

Manoli, G., Fatichi, S., Schläpfer, M., Yu, K., Crowther, T. W., Meili, N., Burlando, P., Katul, G., & Zeid, E. B. (2020). Reply to Martilli et al. (2020): Summer average urban-rural surface temperature differences do not indicate the need for urban heat reduction (No. OSF). https://doi.org/10.31219/osf.io/mwpna

Marando, F., Heris, M. P., Zulian, G., Udías, A., Mentaschi, L., Chrysoulakis, N., Parastatidis, D., & Maes, J. (2022). Urban heat island mitigation by green infrastructure in European Functional Urban Areas. Sustainable Cities and Society, 77(103564), 1–15. https://doi.org/10.1016/j.scs.2021.103564

Marceau, D. J. (1999). The Scale Issue in the Social and Natural Sciences. Canadian Journal of Remote Sensing, 25(4), 347–356. https://doi.org/10.1080/07038992.1999.10874734

Marchezini, V., Horita, F. E. A., Matsuo, P. M., Trajber, R., Trejo-Rangel, M. A., & Olivato, D. (2018). A Review of Studies on Participatory Early Warning Systems (P-EWS): Pathways to Support Citizen Science Initiatives. Frontiers in Earth Science, Volume 6(184), 1–18. https://doi.org/10.3389/feart.2018.00184

Marconcini, M., Gorelick, N., Metz-Marconcini, A., & Esch, T. (2018). Mapping the Global Settlement Growth from 1985 to 2015 – the World Settlement Footprint Evolution Dataset. AGU Fall Meeting. https://agu.confex.com/agu/fm18/meetingapp.cgi/Paper/408596

Marconcini, M., Metz-Marconcini, A., Esch, T., & Gorelick, N. (2021). Understanding Current Trends in Global Urbanisation—The World Settlement Footprint Suite. GI_Forum 2021, Volume 9, Issue 1, Volume 9(1), 33–38. https://doi.org/10.1553/giscience2021_01_s33

Marconcini, M., Metz-Marconcini, A., Üreyen, S., Palacios-Lopez, D., Hanke, W., Bachofer, F., Zeidler, J., Esch, T., Gorelick, N., Kakarla, A., Paganini, M., & Strano, E. (2020). Outlining where humans live, the World Settlement Footprint 2015. Scientific Data, 7(242), 1–14. https://doi.org/10.1038/s41597-020-00580-5

Marcotullio, P. J., & Solecki, W. (2013). What Is a City? An Essential Definition for Sustainability. In C. G. Boone & M. Fragkias (Eds.), Urbanization and Sustainability: Linking Urban Ecology, Environmental Justice and Global Environmental Change (pp. 11–25). Springer Netherlands. https://doi.org/10.1007/978-94-007-5666-3_2

Margulies, J. D., Magliocca, Nicholas R., Schmill, Matthew D., & Ellis, E. C. (2016). Ambiguous Geographies: Connecting Case Study Knowledge with Global Change Science. Annals of the American Association of Geographers, 106(3), 572–596. https://doi.org/10.1080/24694452.2016.1142857

Mariathasan, V., Bezuidenhoudt, E., & Olympio, K. R. (2019). Evaluation of Earth Observation Solutions for Namibia's SDG Monitoring System. Remote Sensing, 11(1612), 1–29. https://doi.org/10.3390/rs11131612

Marshall, M. G., & Elzinga-Marshall, G. (2017). Global Report 2017: Conflict, Governance and State Fragility. https://coilink.org/20.500.12592/4rmrkg

Marshall, V. J., Cadenasso, M. L., McGrath, B. P., & Pickett, S. T. A. (2020). Patch Atlas—Integrating Design Practices and Ecological Knowledge for Cities as Complex Systems. Yale University Press.

Marti, R., Li, Z., Catry, T., Roux, E., Mangeas, M., Handschumacher, P., Gaudart, J., Tran, A., Demagistri, L., Faure, J.-F., Carvajal, J. J., Drumond, B., Xu, L., Herbreteau, V., Gurgel, H., Dessay, N., & Gong, P. (2020). A Mapping Review on Urban Landscape Factors of Dengue Retrieved from Earth Observation Data, GIS Techniques, and Survey Questionnaires. Remote Sensing, 12(932), 1–82. https://doi.org/10.3390/rs12060932

Martilli, A., Krayenhoff, E. S., & Nazarian, N. (2020). Is the Urban Heat Island intensity relevant for heat mitigation studies? Urban Climate, 31(100541), 1–4. https://doi.org/10.1016/j.uclim.2019.100541

Martilli, A., Roth, M., Chow, W. T. L., Demuzere, M., Lipson, M., Krayenhoff, E. S., Sailor, D., Nazarian, N., Voogt, J., Wouters, H., Middel, A., Stewart, I. D., Bechtel, B., Christen, A., & Hart, M. A. (2020). Summer average urban-rural surface temperature differences do not indicate the need for urban heat reduction. OSF. https://doi.org/10.31219/osf.io/8gnbf

Martin, M. E., Newman, S. D., Aber, J. D., & Congalton, R. G. (1998). Determining Forest Species Composition Using High Spectral Resolution Remote Sensing Data. Remote Sensing of Environment, 65(3), 249–254. https://doi.org/10.1016/S0034-4257(98)00035-2

Martinuzzi, S., Gould, W. A., & Ramos González, O. M. (2007). Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data. Landscape and Urban Planning, 79(3), 288–297. https://doi.org/10.1016/j.landurbplan.2006.02.014

Mason, D. C., Schumann, G. J.-P., Neal, J. C., Garcia-Pintado, J., & Bates, P. D. (2012). Automatic near real-time selection of flood water levels from high resolution Synthetic Aperture Radar images for assimilation into hydraulic models: A case study. Remote Sensing of Environment, 124, 705–716. https://doi.org/10.1016/j.rse.2012.06.017

Masoudi, M., & Tan, P. Y. (2019). Multi-year comparison of the effects of spatial pattern of urban green spaces on urban land surface temperature. Landscape and Urban Planning, 184, 44–58. https://doi.org/10.1016/j.landurbplan.2018.10.023

Massaro, E., Schifanella, R., Piccardo, M., Caporaso, L., Taubenböck, H., Cescatti, A., & Duveiller, G. (2023). Spatially-optimized urban greening for reduction of population exposure to land surface temperature extremes. Nature Communications, 14(2903), 1–10. https://doi.org/10.1038/s41467-023-38596-1

Mast, J., Lemoine-Rodríguez, R., Rittlinger, V., Mühlbauer, M., Biewer, C., Geiß, C., & Taubenböck, H. (2025a). Geospatiality: The effect of topics on the presence of geolocation in English text data. International Journal of Geographical Information Science, 0(0), 1–32. https://doi.org/10.1080/13658816.2025.2460051

Mast, J., Lemoine-Rodríguez, R., Rittlinger, V., Mühlbauer, M., Biewer, C., Geiß, C., & Taubenböck, H. (2025b). Geospatiality: The effect of topics on the presence of geolocation in English text data. International Journal of Geographical Information Science, 1–32. https://doi.org/10.1080/13658816.2025.2460051

Mast, J., Sapena, M., Debray, H., Inkoom, J. N., Lemoine-Rodríguez, R., Geiß, C., & Taubenböck, H. (2025a). The digital urban frontier: Disparities in social media activity between consolidated and newly urbanized areas in Africa. Applied Geography, 181(103687), 1–15. https://doi.org/10.1016/j.apgeog.2025.103687

Mast, J., Sapena, M., Debray, H., Inkoom, J. N., Lemoine-Rodríguez, R., Geiß, C., & Taubenböck, H. (2025b). The digital urban frontier: Disparities in social media activity between consolidated and newly urbanized areas in Africa. Applied Geography, 181(103687), 1–15. https://doi.org/10.1016/j.apgeog.2025.103687

Mast, J., Sapena, M., Mühlbauer, M., Biewer, C., & Taubenböck, H. (2024). The migrant perspective: Measuring migrants' movements and interests using geolocated tweets. Population, Space and Place, 30(e2732), 1–17. https://doi.org/10.1002/psp.2732

Masucci, A. P., Arcaute, E., Hatna, E., Stanilov, K., & Batty, M. (2015). On the problem of boundaries and scaling for urban street networks. Journal of The Royal Society Interface, 12(111), 1–7. https://doi.org/10.1098/rsif.2015.0763

Mau, S. (2021). Sorting Machines: The Reinvention of the Border in the 21st Century (1st ed., Vol. 4600). Polity Press. https://doi.org/10.17104/9783406775772

Mau, S., Gülzau, Fabian, Laube, Lena, & Zaun, N. (2015). The Global Mobility Divide: How Visa Policies Have Evolved over Time. Journal of Ethnic and Migration Studies, 41(8), 1192–1213. https://doi.org/10.1080/1369183X.2015.1005007

Maxwell, A. E., Warner, T. A., & Guillén, L. A. (2021). Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review. Remote Sensing, 13(13). https://doi.org/10.3390/rs13132450

Maxwell, E. L., & Riordan, C. J. (1980). Urban area change detection procedures with remote sensing data (Contractor Report (CR) No. 19830009666). NASA. https://ntrs.nasa.gov/citations/19830009666

McCoy, M. D. (2017). Geospatial Big Data and archaeology: Prospects and problems too great to ignore. Archaeological GIS Today: Persistent Challenges, Pushing Old Boundaries, and Exploring New Horizons, 84, 74–94. https://doi.org/10.1016/j.jas.2017.06.003

McDonald, J., Wise, M., & Harris, P. J. (2008). The Health Impacts of the Urban Form: A Review of Reviews (Centre for Health Equity Training, Research and Evaluation). University of New South Wales.

McGarigal, K., Tagil, S., & Cushman, S. A. (2009). Surface metrics: An alternative to patch metrics for the quantification of landscape structure. Landscape Ecology, 24(3), 433–450. https://doi.org/10.1007/s10980-009-9327-y

McGee, T. G. (2014). The Emergence of Desakota Regions in Asia: Expanding a Hypothesis. In N. Brenner (Ed.), Towards a Study of Planetary Urbanization (pp. 121–137). JOVIS. https://doi.org/doi:10.1515/9783868598933-010

McGrath, B. (2018). Intersecting disciplinary frameworks: The architecture and ecology of the city. Ecosystem Health and Sustainability, 4(6), 148–159. https://doi.org/10.1080/20964129.2018.1482730

McHale, M. R., Pickett, S. T. A., Barbosa, O., Bunn, D. N., Cadenasso, M. L., Childers, D. L., Gartin, M., Hess, G. R., Iwaniec, D. M., McPhearson, T., Peterson, M. N., Poole, A. K., Rivers, L., Shutters, S. T., & Zhou, W. (2015). The New Global Urban Realm: Complex, Connected, Diffuse, and Diverse Social-Ecological Systems. Sustainability, 7(5), 5211–5240. https://doi.org/10.3390/su7055211

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. https://arxiv.org/abs/1802.03426

McManus, P. (2012). Measuring Urban Sustainability: The potential and pitfalls of city rankings. Australian Geographer, 43(4), 411–424. https://doi.org/10.1080/00049182.2012.731301

McMillen, D. P. (2001). Nonparametric Employment Subcenter Identification. Journal of Urban Economics, 50(3), 448–473. https://doi.org/10.1006/juec.2001.2228

McMillen, D. P. (2004). Employment Densities, Spatial Autocorrelation, and Subcenters in Large Metropolitan Areas. Journal of Regional Science, 44(2), 225–244. https://doi.org/10.1111/j.0022-4146.2004.00335.x

McMillen, D. P. (2006). Testing for Monocentricity. In A Companion to Urban Economics (pp. 128–140). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470996225.ch8

McNeill, D. (1999). Globalization and the European city. Cities, 16(3), 143–148. https://doi.org/10.1016/S0264-2751(99)00012-8

McPhearson, T., Haase, D., Kabisch, N., & Gren, Å. (2016). Advancing understanding of the complex nature of urban systems. Navigating Urban Complexity: Advancing Understanding of Urban Social – Ecological Systems for Transformation and Resilience, 70, 566–573. https://doi.org/10.1016/j.ecolind.2016.03.054

McPhearson, T., Pickett, S. T. A., Grimm, N. B., Niemelä, J., Alberti, M., Elmqvist, T., Weber, C., Haase, D., Breuste, J., & Qureshi, S. (2016). Advancing Urban Ecology toward a Science of Cities. BioScience, 66(3), 198–212. https://doi.org/10.1093/biosci/biw002

Meerow, S., Newell, J. P., & Stults, M. (2016). Defining urban resilience: A review. Landscape and Urban Planning, 147, 38–49. https://doi.org/10.1016/j.landurbplan.2015.11.011

Meijers, E. (2008). Measuring Polycentricity and its Promises. European Planning Studies, 16(9), 1313–1323. https://doi.org/10.1080/09654310802401805

Meijers, E. J. (2013). Metropolitan Labor Productivity and Urban Spatial Structure. In J. Klaesson, B. Johansson, & C. Karlsson (Eds.), Metropolitan Regions: Knowledge Infrastructures of the Global Economy (pp. 141–166). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-32141-2_7

Meili, N., Manoli, G., Burlando, P., Carmeliet, J., Chow, W. T. L., Coutts, A. M., Roth, M., Velasco, E., Vivoni, E. R., & Fatichi, S. (2021). Tree effects on urban microclimate: Diurnal, seasonal, and climatic temperature differences explained by separating radiation, evapotranspiration, and roughness effects. Urban Forestry & Urban Greening, 58(126970), 1–13. https://doi.org/10.1016/j.ufug.2020.126970

Melchiorri, M., Florczyk, A. J., Freire, S., Schiavina, M., Pesaresi, M., & Kemper, T. (2018). Unveiling 25 Years of Planetary Urbanization with Remote Sensing: Perspectives from the Global Human Settlement Layer. Remote Sensing, 10(768), 1–19. https://doi.org/10.3390/rs10050768

Melchiorri, M., Pesaresi, M., Florczyk, A. J., Corbane, C., & Kemper, T. (2019). Principles and Applications of the Global Human Settlement Layer as Baseline for the Land Use Efficiency Indicator—SDG 11.3.1. ISPRS International Journal of Geo-Information, 8(96), 1–19. https://doi.org/10.3390/ijgi8020096

Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-Time Light Data: A Good Proxy Measure for Economic Activity? PLOS ONE, 10(10)(e0139779), 1–18. https://doi.org/10.1371/journal.pone.0139779

Meng, Q., Gao, J., Zhang, L., Hu, X., Qian, J., & Jancsó, T. (2024). Coupled cooling effects between urban parks and surrounding building morphologies based on the microclimate evaluation framework integrating remote sensing data. Sustainable Cities and Society, 102(105235), 1–12. https://doi.org/10.1016/j.scs.2024.105235

Meng, Q., Zhang, L., Sun, Z., Meng, F., Wang, L., & Sun, Y. (2018). Characterizing spatial and temporal trends of surface urban heat island effect in an urban main built-up area: A 12-year case study in Beijing, China. Remote Sensing of Environment, 204, 826–837. https://doi.org/10.1016/j.rse.2017.09.019

Mercer Consulting. (2019). Quality of living worldwide city rankings. Mercer Survey. https://www.mercer.com/insights/total-rewards/talent-mobility-insights/quality-of-living-city-ranking/

Mertes, C. M., Schneider, A., Sulla-Menashe, D., Tatem, A. J., & Tan, B. (2015). Detecting change in urban areas at continental scales with MODIS data. Remote Sensing of Environment, 158, 331–347. https://doi.org/10.1016/j.rse.2014.09.023

Metzger, N., Vargas-Muñoz, J. E., Daudt, R. C., Kellenberger, B., Whelan, T. T.-T., Ofli, F., Imran, M., Schindler, K., & Tuia, D. (2022). Fine-grained population mapping from coarse census counts and open geodata. Scientific Reports, 12(20085), 1–14. https://doi.org/10.1038/s41598-022-24495-w

Miatto, A., Emami, N., Goodwin, K., West, J., Taskhiri, M. S., Wiedmann, T., & Schandl, H. (2024). Australia's circular economy metrics and indicators. Journal of Industrial Ecology, 28(2), 216–231. https://doi.org/10.1111/jiec.13458

Michalopoulos, S., & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa. The Quarterly Journal of Economics, 129(1), 151–213. https://doi.org/10.1093/qje/qjt029

Microsoft. (2020). Building Footprints—Bing Maps [Dataset]. https://www.microsoft.com/en-us/maps/building-footprints

Microsoft. (2022). Building Footprints—Bing Maps [Dataset]. https://www.microsoft.com/en-us/maps/building-footprints

Middel, A., Häb, K., Brazel, A. J., Martin, C. A., & Guhathakurta, S. (2014). Impact of urban form and design on mid-afternoon microclimate in Phoenix Local Climate Zones. Landscape and Urban Planning, 122, 16–28. https://doi.org/10.1016/j.landurbplan.2013.11.004

Mihelic, K. K., Lipicnik, B., & Tekavcic, M. (2010). Ethical Leadership. International Journal of Management & Information Systems (IJMIS), 14(5), 31–41. https://doi.org/10.19030/ijmis.v14i5.11

Milesi, C., Elvidge, C. D., Nemani, R. R., & Running, S. W. (2003). Assessing the impact of urban land development on net primary productivity in the southeastern United States. Urban Remote Sensing, 86(3), 401–410. https://doi.org/10.1016/S0034-4257(03)00081-6

Miller, B. C. (2021). Cluster-based Delineation of Megaregions in the United States: Identifying administrative boundaries that reflect meta-communities to improve the effectiveness and efficiency of government. Georgia Institute of Technology.

Mills, G., Bechtel, B., Ching, J., See, L., Feddema, J., Foley, M., Alexander, P., & O'Connor, M. (2015). An Introduction to the WUDAPT project. ICUC9 - 9th International Conference on Urban Climate Jointly with 12th Symposium on the Urban Environment, 1–6. https://www.wudapt.org/wp-content/uploads/2015/05/Mills_etal_ICUC9.pdf

Min, B., & Gaba, K. M. (2014). Tracking Electrification in Vietnam Using Nighttime Lights. Remote Sensing, 6(10), 9511–9529. https://doi.org/10.3390/rs6109511

Min, B., Gaba, K. M., Ousmane, F. S., & Agalassou, A. (2013). Detection of rural electrification in Africa using DMSP-OLS night lights imagery. International Journal of Remote Sensing, 34(22), 8118–8141. https://doi.org/10.1080/01431161.2013.833358

Mishra, S., Stumpf, R. P., Schaeffer, B. A., Werdell, P. J., Loftin, K. A., & Meredith, A. (2019). Measurement of Cyanobacterial Bloom Magnitude using Satellite Remote Sensing. Scientific Reports, 9(18310), 1–17. https://doi.org/10.1038/s41598-019-54453-y

Mobaraki, A., & Oktay Vehbi, B. (2022). A Conceptual Model for Assessing the Relationship between Urban Morphology and Sustainable Urban Form. Sustainability, 14(5), 1–20. https://doi.org/10.3390/su14052884

Mohammad, M. A., Olesen, K. K. W., Koul, S., Gale, C. P., Rylance, R., Jernberg, T., Baron, T., Spaak, J., James, S., Lindahl, B., Maeng, M., & Erlinge, D. (2022). Development and validation of an artificial neural network algorithm to predict mortality and admission to hospital for heart failure after myocardial infarction: A nationwide population-based study. The Lancet Digital Health, 4(1), e37–e45. https://doi.org/10.1016/S2589-7500(21)00228-4

Molina, M. J., & Molina, L. T. (2004). Megacities and Atmospheric Pollution. Journal of the Air & Waste Management Association, 54(6), 644–680. https://doi.org/10.1080/10473289.2004.10470936

Molinier, M., & Kilpi, J. (2019). Avoiding Overfitting When Applying Spectral-Spatial Deep Learning Methods on Hyperspectral Images with Limited Labels. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 5049–5052. https://doi.org/10.1109/IGARSS.2019.8900328

Mondini, A. C., Guzzetti, F., Reichenbach, P., Rossi, M., Cardinali, M., & Ardizzone, F. (2011). Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. Remote Sensing of Environment, 115(7), 1743–1757. https://doi.org/10.1016/j.rse.2011.03.006

Monkkonen, P. (2019). The Elephant in the Zoning Code: Single Family Zoning in the Housing Supply Discussion. Housing Policy Debate, 29(1), 41–43. https://doi.org/10.1080/10511482.2018.1506392

Montello, D. R., Goodchild, M. F., Gottsegen, J., & Fohl, P. (2003). Where's downtown?: Behavioral methods for determining referents of vague spatial queries. Spatial Cognition and Computation, 3(2–3), 185–204. https://doi.org/10.1080/13875868.2003.9683761

Montero, G., Tannier, C., & Thomas, I. (2021). Delineation of cities based on scaling properties of urban patterns: A comparison of three methods. International Journal of Geographical Information Science, 35(5), 919–947. https://doi.org/10.1080/13658816.2020.1817462

Montgomery, C. (2013). Happy City: Transforming Our Lives Through Urban Design. Penguin Books.

Moore, M. (2003). A World without Walls: Freedom, Development, Free Trade and Global Governance. Cambridge University Press.

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., Counsell, C. W. W., Dietrich, B. S., Johnston, E. T., Louis, L. V., Lucas, M. P., McKenzie, M. M., Shea, A. G., Tseng, H., Giambelluca, T. W., Leon, L. R., Hawkins, E., & Trauernicht, C. (2017). Global risk of deadly heat. Nature Climate Change, 7(7), 501–506. https://doi.org/10.1038/nclimate3322

Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2021). Metropolitan areas in the world. Delineation and population trends. Journal of Urban Economics, 125(103242), 1–22. https://doi.org/10.1016/j.jue.2020.103242

Morgenpost. (2016). Das sind Deutschlands grünste Großstädte. Berliner Morgenpost. https://interaktiv.morgenpost.de/gruenste-staedte-deutschlands/

Moroni, S., & Cozzolino, S. (2019). Action and the city. Emergence, complexity, planning. Cities, 90, 42–51. https://doi.org/10.1016/j.cities.2019.01.039

Moudon, A. V. (1997). Urban Morphology as an emerging interdisciplinary field. Urban Morphology, 1(1), 3–10. https://doi.org/10.51347/jum.v1i1.4047

Mouratidis, K. (2018). Built environment and social well-being: How does urban form affect social life and personal relationships? Cities, 74, 7–20. https://doi.org/10.1016/j.cities.2017.10.020

Mouratidis, K., & Poortinga, W. (2020). Built environment, urban vitality and social cohesion: Do vibrant neighborhoods foster strong communities? Landscape and Urban Planning, 204(103951), 1–9. https://doi.org/10.1016/j.landurbplan.2020.103951

Mück, M., Klotz, M., & Taubenböck, H. (2017). Validation of the DLR Global Urban Footprint in rural areas: A case study for Burkina Faso. 2017 Joint Urban Remote Sensing Event (JURSE), 1–4. https://doi.org/10.1109/JURSE.2017.7924618

Mueller, H., Groeger, A., Hersh, J., Matranga, A., & Serrat, J. (2021). Monitoring war destruction from space using machine learning. Proceedings of the National Academy of Sciences, 118(23), 1–9. https://doi.org/10.1073/pnas.2025400118

Mueller, N., Daher, C., Rojas-Rueda, D., Delgado, L., Vicioso, H., Gascon, M., Marquet, O., Vert, C., Martin, I., & Nieuwenhuijsen, M. (2021). Integrating health indicators into urban and transport planning: A narrative literature review and participatory process. International Journal of Hygiene and Environmental Health, 235(113772), 1–13. https://doi.org/10.1016/j.ijheh.2021.113772

Muhar, A., Raymond, C. M., van den Born, R. J. G., Bauer, N., Böck, K., Braito, M., Buijs, A., Flint, C., de Groot, W. T., Ives, C. D., Mitrofanenko, T., Plieninger, T., Tucker, C., & van Riper, C. J. (2018). A model integrating social-cultural concepts of nature into frameworks of interaction between social and natural systems. Journal of Environmental Planning and Management, 61(5–6), 756–777. https://doi.org/10.1080/09640568.2017.1327424

Mukhopadhyay, C. (2018). Is the Delhi Mumbai Industrial Corridor (DMIC) an Emerging 'Megaregion' in India? Planning Theory & Practice, 19(2), 305–309. https://doi.org/10.1080/14649357.2017.1392136

Müller, I., Erbertseder, T., & Taubenböck, H. (2022). Tropospheric NO2: Explorative analyses of spatial variability and impact factors. Remote Sensing of Environment, 270(112839), 1–13. https://doi.org/10.1016/j.rse.2021.112839

Müller, I., Taubenböck, H., Kuffer, M., & Wurm, M. (2020). Misperceptions of Predominant Slum Locations? Spatial Analysis of Slum Locations in Terms of Topography Based on Earth Observation Data. Remote Sensing, 12(2474), 1–19. https://doi.org/10.3390/rs12152474

Mumford, L. (1934). Technics and civilization. The University of Chicago Press.

Mumford, L. (1961). The city in history: Its origins, its transformations, and its prospects. Secker & Warburg.

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., Huckle, R., Lacan, A., Grzegorski, M., Holdak, A., Kokhanovsky, A., Livschitz, J., & Eisinger, M. (2016). The GOME-2 instrument on the Metop series of satellites: Instrument design, calibration, and level 1 data processing \hack\newline – an overview. Atmospheric Measurement Techniques, 9(3), 1279–1301. https://doi.org/10.5194/amt-9-1279-2016

Murai, S. (1974). Estimation of population density in Tokyo districts from ERTS-1 data. Proc of the Int Symp on Remote Sensing of Environ, 1, 13–22.

Murphy, R. E. (1971). The Central Business District – A study in urban geopgraphy. Transaction Publishers.

Murphy, R., & Vance, J. J. E. (1954). Delimiting the CBD. Economic Geography, 30, 189–222.

Musil, J. (1993). Changing Urban Systems in Post-communist Societies in Central Europe: Analysis and Prediction. Urban Studies, 30(6), 899–905. https://doi.org/10.1080/00420989320080841

Mutiibwa, D., Strachan, S., & Albright, T. (2015). Land Surface Temperature and Surface Air Temperature in Complex Terrain. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(10), 4762–4774. https://doi.org/10.1109/JSTARS.2015.2468594

Mwakesi, I., Wahome, R., & Ichang'i, D. (2020). Mining impact on communities' livelihoods: A case study of Taita Taveta County, Kenya. AIMS Environmental Science, 7(3), 286–301. https://doi.org/10.3934/environsci.2020018

Myers, G. A. (2011). African Cities: Alternative visions of urban theory and practice. Zed Books.

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. Object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment, 115(5), 1145–1161. https://doi.org/10.1016/j.rse.2010.12.017

Mykhnenko, V., & Turok, I. (2008). East European Cities—Patterns of Growth and Decline, 1960–2005. International Planning Studies, 13(4), 311–342. https://doi.org/10.1080/13563470802518958

N

Nagendra, H., Lucas, R., Honrado, J. P., Jongman, R. H. G., Tarantino, C., Adamo, M., & Mairota, P. (2013). Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. Biodiversity Monitoring, 33, 45–59. https://doi.org/10.1016/j.ecolind.2012.09.014

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. J. R. Soc. Interface, 14(20170213), 1–11. https://doi.org/10.1098/rsif.2017.0213

Nassehi, A. (2002). Dichte Räume. Städte als Synchronisations- und Inklusionsmaschinen. In M. Löw (Ed.), Differenzierungen des Städtischen (pp. 211–232). Opladen: Leske + Budrich.

Nassehi, A. (2024). Kritik der großen Geste. Wirtschaftswoche. https://www.wiwo.de/politik/deutschland/essay-kritik-der-grossen-geste/29917012.html

National Academies of Sciences, Engineering, and Medicine. (2016). Rationalizing Rural Area Classifications for the Economic Research Service: A Workshop Summary. National Academies Press.

Nchake, M. A., & Shuaibu, M. (2022). Investment in ICT infrastructure and inclusive growth in Africa. Scientific African, 17(e01293), 1–16. https://doi.org/10.1016/j.sciaf.2022.e01293

Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., & Hossard, L. (2018). Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. Remote Sensing, 10(1217), 1–16. https://doi.org/10.3390/rs10081217

Nechyba, T. J., & Walsh, R. P. (2004). Urban Sprawl. Journal of Economic Perspectives, 18(4), 177–200. https://doi.org/10.1257/0895330042632681

Nelson, G. D., & Rae, A. (2016). An Economic Geography of the United States: From Commutes to Megaregions. PLOS ONE, 11(11)(e0166083), 1–23. https://doi.org/10.1371/journal.pone.0166083

Neuman, M. (2000). Regional design: Recovering a great landscape architecture and urban planning tradition. Landscape and Urban Planning, 47(3), 115–128. https://doi.org/10.1016/S0169-2046(99)00079-1

Newman, D. (2006). The lines that continue to separate us: Borders in our `borderless' world—David Newman, 2006. Progress in Human Geography, 30(2), 143–161. https://doi.org/10.1191/0309132506ph599xx

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. Contemporary Physics, 46(5), 323–351. https://doi.org/10.1080/00107510500052444

Ng, A., Jordan, M., & Weiss, Y. (2001). On Spectral Clustering: Analysis and an algorithm. In T. Dietterich, S. Becker, & Z. Ghahramani (Eds.), Advances in Neural Information Processing Systems (Vol. 14, pp. 849–856). MIT Press. https://proceedings.neurips.cc/paper_files/paper/2001/file/801272ee79cfde7fa5960571fee36b9b-Paper.pdf

Ngom, N. M., Mbaye, M., Baratoux, D., Baratoux, L., Catry, T., Dessay, N., Faye, G., Sow, E. H., & Delaitre, E. (2020). Mapping Artisanal and Small-Scale Gold Mining in Senegal Using Sentinel 2 Data. GeoHealth, 4(12), 1–21. https://doi.org/10.1029/2020GH000310

Nguyen, T. T., Ngo, H. H., Guo, W., & Wang, X. C. (2020). A new model framework for sponge city implementation: Emerging challenges and future developments. Journal of Environmental Management, 253(109689), 1–14. https://doi.org/10.1016/j.jenvman.2019.109689

Nichol, J. E., Fung, W. Y., Lam, K., & Wong, M. S. (2009). Urban heat island diagnosis using ASTER satellite images and 'in situ' air temperature. Atmospheric Research, 94(2), 276–284. https://doi.org/10.1016/j.atmosres.2009.06.011

Nichol, J. E., King, B., & Ding, X. (2013). Sustainable urbanization. International Journal of Remote Sensing, 34(3), 755–758. https://doi.org/10.1080/01431161.2013.721284

Nicolau, R., David, J., Caetano, M., & Pereira, J. M. C. (2019). Ratio of land consumption rate to population growth rate-analysis of different formulations applied to mainland Portugal. In ISPRS International Journal of Geo-Information (Vol. 8, Issue 10, pp. 1–21). https://doi.org/10.3390/ijgi8010010

Nielsen, M. B., & Jensen, M. B. (2015). Land cover in single-family housing areas and how it correlates with urban form. Urban Ecosystems, 18(4), 1103–1123. https://doi.org/10.1007/s11252-015-0471-7

Nieuwenhuijsen, M. J. (2020). Urban and transport planning pathways to carbon neutral, liveable and healthy cities; A review of the current evidence. Environment International, 140(105661), 1–10. https://doi.org/10.1016/j.envint.2020.105661

Nieuwenhuijsen, M. J. (2021). New urban models for more sustainable, liveable and healthier cities post covid19; reducing air pollution, noise and heat island effects and increasing green space and physical activity. Environment International, 157(106850), 1–8. https://doi.org/10.1016/j.envint.2021.106850

Nieuwenhuijsen, M., & Khreis, H. (2019). Urban and Transport Planning, Environment and Health. In M. Nieuwenhuijsen & H. Khreis (Eds.), Integrating Human Health into Urban and Transport Planning: A Framework (pp. 3–16). Springer International Publishing. https://doi.org/10.1007/978-3-319-74983-9_1

Niinemets, Ü., & Peñuelas, J. (2008). Gardening and urban landscaping: Significant players in global change. Trends in Plant Science, 13(2), 60–65. https://doi.org/10.1016/j.tplants.2007.11.009

Nikoobakht, S., Azarafza, M., Akgün, H., & Derakhshani, R. (2022). Landslide Susceptibility Assessment by Using Convolutional Neural Network. Applied Sciences, 12(5992), 1–22. https://doi.org/10.3390/app12125992

Nita, S. (2017). Free movement of people within regional integration processes: A comparative view. In S. Nita, A. Pécoud, P. De Lombaerde, P. de Guchteneire, K. Neyts, & J. Gartland (Eds.), Migration, free movement and regional integration (pp. 3–44). UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000368002

Nitsch, V. (2005). Zipf zipped. Journal of Urban Economics, 57(1), 86–100. https://doi.org/10.1016/j.jue.2004.09.002

NOAA (National Oceanic and Atmospheric Administration). (n.d.). Global Distribution and Density of Constructed Impervious Surfaces [Dataset]. Retrieved 9 August 2022, from https://www.ngdc.noaa.gov/eog/dmsp/download_global_isa.html

Noizet, H., & Clémençon, A.-S. (2020). Faire ville: Entre planifié et impensé, la fabrique ordinaire des formes urbaines. Presses universitaires de Vincennes.

Nolan, L. B. (2015). Slum Definitions in Urban India: Implications for the Measurement of Health Inequalities. Population and Development Review, 41(1), 59–84. https://doi.org/10.1111/j.1728-4457.2015.00026.x

Nordhaus, W., & Chen, X. (2015). A sharper image? Estimates of the precision of nighttime lights as a proxy for economic statistics1. Journal of Economic Geography, 15(1), 217–246. https://doi.org/10.1093/jeg/lbu010

Nyamekye, C., Ghansah, B., Agyapong, E., & Kwofie, S. (2021). Mapping changes in artisanal and small-scale mining (ASM) landscape using machine and deep learning algorithms. - A proxy evaluation of the 2017 ban on ASM in Ghana. Environmental Challenges, 3, 1–15. https://doi.org/10.1016/j.envc.2021.100053

O

Oberhauser, D. (2019). Blockchain for Environmental Governance: Can Smart Contracts Reinforce Payments for Ecosystem Services in Namibia? Frontiers in Blockchain, 2(21), 1–17. https://doi.org/10.3389/fbloc.2019.00021

O'Brien, D. T., Farrell, C., & Welsh, B. C. (2019). Looking Through Broken Windows: The Impact of Neighborhood Disorder on Aggression and Fear of Crime Is an Artifact of Research Design. In Annual Review of Criminology (Vol. 2, Issue Volume 2, 2019, pp. 53–71). Annual Reviews. https://doi.org/10.1146/annurev-criminol-011518-024638

Oda, T., Maksyutov, S., & Andres, R. J. (2018). The Open-source Data Inventory for Anthropogenic CO_2, version 2016 (ODIAC2016): A global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. Earth System Science Data, 10(1), 87–107. https://doi.org/10.5194/essd-10-87-2018

OECD. (2017a). Land-Use Planning Systems in the OECD. Country Fact Sheets (OECD Regional Development Studies). OECD Publishing, Paris. https://doi.org/10.1787/9789264268579-en

OECD. (2017b). The Governance of Land Use in OECD Countries: Policy Analysis and Recommendations (OECD Regional Development Studies). OECD Publishing, Paris. https://doi.org/10.1787/9789264268609-en

OECD. (2018). Rethinking Urban Sprawl: Moving Towards Sustainable Cities. OECD Publishing, Paris. https://doi.org/10.1787/9789264189881-en

OECD. (2025). Africa's Urbanisation Dynamics 2025: Planning for Urban Expansion. West African Studies, OECD Publishing, Paris. https://doi.org/10.1787/2a47845c-en

Office of the Geographer and Global Issues at the U.S. Department of State. (2021). Large Scale International Boundaries (Version 10.0) [Dataset]. https://catalog.data.gov/dataset/large-scale-international-boundaries

Ohmae, K. (1990). The Borderless World: Power and Strategy in the Interlinked Economy: Ohmae, Kenichi: Amazon.de: Books. HarperCollins Publishers Ltd.

Oke, T. R. (1973). City size and the urban heat island. Atmospheric Environment (1967), 7(8), 769–779. https://doi.org/10.1016/0004-6981(73)90140-6

Oke, T. R. (1982). The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society, 108(455), 1–24. https://doi.org/10.1002/qj.49710845502

Oke, T. R. (20006). Initial guidance to obtain representative meteorological observations at urban sites (No. WMO/TD-No. 1250; Instruments and Observing Methods Report No. 81, p. 51). WMO. https://library.wmo.int/idurl/4/35333

Oke, T. R., Mills, G., Christen, A., & Voogt, J. A. (2017). Urban Climates. Cambridge University Press.

Okoh, G., & Hilson, G. (2011). Poverty and livelihood diversification: Exploring the linkages between smallholder farming and artisanal mining in rural Ghana. Journal of International Development, 23(8), 1100–1114. https://doi.org/10.1002/jid.1834

Oliveira, V. (2016). Urban Morphology. An Introduction to the Study of the Physical Form of Cities. Springer Cham. https://doi.org/10.1007/978-3-319-32083-0

Olsen, J. R., Nicholls, N., Moon, G., Pearce, J., Shortt, N., & Mitchell, R. (2019). Which urban land covers/uses are associated with residents' mortality? A cross-sectional, ecological, pan-European study of 233 cities. BMJ Open, 9(11)(e033623), 1–11. https://doi.org/10.1136/bmjopen-2019-033623

Openshaw, S. (1984). The modifiable areal unit problem (No. 38; Concepts and Techniques in Modern Geography). Geo Books.

OpenStreetMap contributors. (2015). Nominatim [Dataset]. https://nominatim. openstreetmap.org/ui/search.html

OpenStreetMap (OSM). (2021). Land polygons [Dataset]. https://osmdata.openstreetmap.de/data/land-polygons.html

Ortman, S. G., Smith, M. E., Lobo, J., & Bettencourt, L. M. A. (n.d.). Why Archaeology Is Necessary for a Theory of Urbanization. Journal of Urban Archaeology, 1, 151–167. https://doi.org/10.1484/J.JUA.5.120914

OSM. (2018). OpenStreetMap [Dataset]. https://www.openstreetmap.org

Ospina-Gutiérrez, J. P., & Aristizábal, E. (2021). Application of Artificial Intelligence and machine learning techniques for landslide susceptibility assessment. Revista Mexicana de Ciencias Geológicas, 38(1), 43–54. https://doi.org/10.22201/cgeo.20072902e.2021.1.1605

Ostrom, E. (1990). Governing the commons: The evolution of institutions for collective action. Cambridge University Press. https://doi.org/10.1017/CBO9780511807763

Ott, T. (2001). From Concentration to De-concentration—Migration Patterns in the Post-socialist City. Cities, 18(6), 403–412. https://doi.org/10.1016/S0264-2751(01)00032-4

Ottensmann, J. R. (2020). Density and Sprawl in the Older and Newer Parts of Large Urban Areas in the U.S., 1970-2020. https://ssrn.com/abstract=4483579 or http://dx.doi.org/10.2139/ssrn.4483579

Owen, K. K., & Wong, D. W. (2013). An approach to differentiate informal settlements using spectral, texture, geomorphology and road accessibility metrics. Applied Geography, 38, 107–118. https://doi.org/10.1016/j.apgeog.2012.11.016

P

Paasi, A. (2019). Borderless worlds and beyond: Challenging the state-centric cartographies. In A. Paasi, E.-K. Prokkola, J. Saarinen, & K. Zimmerbauer (Eds.), Borderless Worlds for Whom? (1st ed., pp. 21–36). Routledge. https://doi.org/10.4324/9780429427817-2

Pacione, M. (2009). Urban geography: A global perspective (3rd ed.). Routledge.

Pahl, R. E. (1966). The rural-urban continuum. Sociologia Ruralis, 6(3), 299–329. https://doi.org/10.1111/j.1467-9523.1966.tb00537.x

Palacio Cordoba, J., Mergili, M., & Aristizábal, E. (2020). Probabilistic landslide susceptibility analysis in tropical mountainous terrain using the physically based r.slope.stability model. Natural Hazards and Earth System Sciences, 20(3), 815–829. https://doi.org/10.5194/nhess-20-815-2020

Palme, M., Inostroza, L., & Salvati, A. (2018). Technomass and cooling demand in South America: A superlinear relationship? Building Research & Information, 46(8), 864–880. https://doi.org/10.1080/09613218.2018.1483868

Panerai, P., Depaule, J.-C., & Demorgon, M. (1999). Analyse urbaine. Parenthèses. https://editionsparentheses.com/IMG/pdf/p603_analyse_urbaine.pdf

Park, S., Hepcan, Ç. C., Hepcan, Ş., & Cook, E. A. (2014). Influence of urban form on landscape pattern and connectivity in metropolitan regions: A comparative case study of Phoenix, AZ, USA, and Izmir, Turkey. Environmental Monitoring and Assessment, 186(10), 6301–6318. https://doi.org/10.1007/s10661-014-3855-x

Parmehr, E. G., Amati, M., Taylor, E. J., & Livesley, S. J. (2016). Estimation of urban tree canopy cover using random point sampling and remote sensing methods. Urban Forestry & Urban Greening, 20, 160–171. https://doi.org/10.1016/j.ufug.2016.08.011

Parr, J. B. (2007). Spatial Definitions of the City: Four Perspectives. Urban Studies, 44(2), 381–392. https://doi.org/10.1080/00420980601075059

Parr, J. B. (2014). The Regional Economy, Spatial Structure and Regional Urban Systems. Regional Studies, 48(12), 1926–1938. https://doi.org/10.1080/00343404.2013.799759

Parry, L., Barlow, J., & Pereira, H. (2014). Wildlife Harvest and Consumption in Amazonia's Urbanized Wilderness. Conservation Letters, 7(6), 565–574. https://doi.org/10.1111/conl.12151

Parselia, E., Kontoes, C., Tsouni, A., Hadjichristodoulou, C., Kioutsioukis, I., Magiorkinis, G., & Stilianakis, N. I. (2019). Satellite Earth Observation Data in Epidemiological Modeling of Malaria, Dengue and West Nile Virus: A Scoping Review. Remote Sensing, 11(1862), 1–40. https://doi.org/10.3390/rs11161862

Pascal, M., Corso, M., Chanel, O., Declercq, C., Badaloni, C., Cesaroni, G., Henschel, S., Meister, K., Haluza, D., Martin-Olmedo, P., & Medina, S. (2013). Assessing the public health impacts of urban air pollution in 25 European cities: Results of the Aphekom project. Science of The Total Environment, 449, 390–400. https://doi.org/10.1016/j.scitotenv.2013.01.077

Pataki, D. E., Alberti, M., Cadenasso, M. L., Felson, A. J., McDonnell, M. J., Pincetl, S., Pouyat, R. V., Setälä, H., & Whitlow, T. H. (2021). The Benefits and Limits of Urban Tree Planting for Environmental and Human Health. Frontiers in Ecology and Evolution, Volume 9(603757), 1–9. https://doi.org/10.3389/fevo.2021.603757

Patil, K. R., & Heinrichs, B. (2022). Verifiability as a Complement to AI Explainability: A Conceptual Proposal. https://philsci-archive.pitt.edu/20297/1/Verification%20as%20a%20Complement%20to%20AI%20Explainability%23final.pdf

Patino, J. E., Duque, J. C., Pardo-Pascual, J. E., & Ruiz, L. A. (2014). Using remote sensing to assess the relationship between crime and the urban layout. Applied Geography, 55, 48–60. https://doi.org/10.1016/j.apgeog.2014.08.016

Patz, J. A. (2002). A human disease indicator for the effects of recent global climate change. Proceedings of the National Academy of Sciences, 99(20), 12506–12508. https://doi.org/10.1073/pnas.212467899

Paulsen, K. (2012). Yet even more evidence on the spatial size of cities: Urban spatial expansion in the US, 1980–2000. Regional Science and Urban Economics, 42(4), 561–568. https://doi.org/10.1016/j.regsciurbeco.2012.02.002

Paulson, M. L., Olivier, K. N., & Holland, S. M. (2012). Pulmonary non-tuberculous mycobacterial infection in congenital contractural arachnodactyly [Case study]. The International Journal of Tuberculosis and Lung Disease, 16(4), 561–563. https://doi.org/doi:10.5588/ijtld.11.0301

PDOK. (2022). 3D Basisvoorziening (OGC API) [Dataset]. https://api.pdok.nl/kadaster/3d-basisvoorziening/ogc/v1

Pécoud, A., & de Guchteneire, P. (2006). International migration, border controls and human rights: Assessing the relevance of a right to mobility. Journal of Borderlands Studies, 21(1), 69–86. https://doi.org/10.1080/08865655.2006.9695652

PeiSer, R. (2001). Decomposing urban sprawl. Town Planning Review, 72(3), 275–298. https://doi.org/10.3828/tpr.2001.72.3.275

Peiser, R. B. (1989). Density and Urban Sprawl. Land Economics, 65(3), 193–204.

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540(7633), 418–422. https://doi.org/10.1038/nature20584

Pellikka, P., & Rees, W. G. (2009). Remote Sensing of Glaciers: Techniques for Topographic, Spatial and Thematic Mapping of Glaciers. Taylor & Francis Ltd.

Peltier, W. R. (2004). Global glacial isostasy and the surface of the ice-age earth: The ICE-5G (VM2) model and grace. In Annual Review of Earth and Planetary Sciences (Vol. 32, Issue Volume 32, 2004, pp. 111–149). Annual Reviews. https://doi.org/10.1146/annurev.earth.32.082503.144359

Pereira, G., Foster, S., Martin, K., Christian, H., Boruff, B. J., Knuiman, M., & Giles-Corti, B. (2012). The association between neighborhood greenness and cardiovascular disease: An observational study. BMC Public Health, 12(466), 1–9. Scopus. https://doi.org/10.1186/1471-2458-12-466

Perera, N. G. R., & Emmanuel, R. (2018). A "Local Climate Zone" based approach to urban planning in Colombo, Sri Lanka. ICUC9: The 9th International Conference on Urban Climate, 23, 188–203. https://doi.org/10.1016/j.uclim.2016.11.006

Pérez-Invernón, F. J., Huntrieser, H., Erbertseder, T., Loyola, D., Valks, P., Liu, S., Allen, D. J., Pickering, K. E., Bucsela, E. J., Jöckel, P., van Geffen, J., Eskes, H., Soler, S., Gordillo-Vázquez, F. J., & Lapierre, J. (2022). Quantification of lightning-produced NO_x over the Pyrenees and the Ebro Valley by using different TROPOMI-NO_2 and cloud research products. Atmospheric Measurement Techniques, 15(11), 3329–3351. https://doi.org/10.5194/amt-15-3329-2022

Perks, R., Kelly, J., Constantian, S., & Pham, P. (2018). Resources and resourcefulness: Gender, human rights and resilience in Artisanal Mining towns of Eastern Congo. In K. Lahiri-Dutt (Ed.), Between the Plough and the Pick: Informal, artisanal and small-scale mining in the contemporary world (pp. 209–231). ANU Press. http://dx.doi.org/10.22459/BPP.03.2018.10

Pesaresi, M., Corbane, C., Julea, A., Florczyk, A. J., Syrris, V., & Soille, P. (2016). Assessment of the Added-Value of Sentinel-2 for Detecting Built-up Areas. Remote Sensing, 8(299), 1–18. https://doi.org/10.3390/rs8040299

Pesaresi, M., Corbane, C., Ren, C., & Edward, N. (2021). Generalized Vertical Components of built-up areas from global Digital Elevation Models by multi-scale linear regression modelling. PLOS ONE, 16(2)(e0244478), 1–33. https://doi.org/10.1371/journal.pone.0244478

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Freire, S., Haag, F., Halkia, M., Julea, A. M., Kemper, T., & Soille, P. (2015). Global human settlement analysis for disaster risk reduction. 40(7W3), 837–843. Scopus. https://doi.org/10.5194/isprsarchives-XL-7-W3-837-2015

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A.J., Freire, S., Halkia, M., Julea, A., Kemper, T., Soille, P., & Syrris, V. (2016). Operating procedure for the production of the global human settlement layer from Landsat data of the epochs 1975, 1990, 2000, and 2014. Publications Office. https://doi.org/10.2788/253582

Pesaresi, M., Ehrlich, D., Florczyk, A., Freire, S., Julea, A., Kemper, T., Soille, P., & Syrris, V. (2015). GHS-BUILT R2015B - GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014)—OBSOLETE RELEASE. [Dataset]. European Commission, Joint Research Centre (JRC). http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmt_globe_r2015b

Pesaresi, M., Gerhardinger, A., & Kayitakire, F. (2008). A robust built-up area presence index by anisotropic rotation-invariant textural measure. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1(3), 180–192. Scopus. https://doi.org/10.1109/JSTARS.2008.2002869

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T., Lu, L., Marin-Herrera, M. A., Ouzounis, G. K., Scavazzon, M., Soille, P., Syrris, V., & Zanchetta, L. (2013). A global human settlement layer from optical HR/VHR RS data: Concept and first results. In IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (Vol. 6, Issue 5, pp. 2102–2131). https://doi.org/10.1109/JSTARS.2013.2271445

Pettit, C., Shi, Y., Han, H., Rittenbruch, M., Foth, M., Lieske, S., van den Nouwelant, R., Mitchell, P., Leao, S., Christensen, B., & Jamal, M. (2020). A new toolkit for land value analysis and scenario planning. Environment and Planning B: Urban Analytics and City Science, 47(8), 1490–1507. https://doi.org/10.1177/2399808320924678

Petutschnig, A., Havas, C. R., Resch, B., Krieger, V., & Ferner, C. (2020). Exploratory Spatiotemporal Language Analysis of Geo-Social Network Data for Identifying Movements of Refugees. GI_Forum 2020, Volume 8, Issue 1, 137–152.

Pflugmacher, D., Rabe, A., Peters, M., & Hostert, P. (2019). Mapping pan-European land cover using Landsat spectral-temporal metrics and the European LUCAS survey. Remote Sensing of Environment, 221, 583–595. https://doi.org/10.1016/j.rse.2018.12.001

Pickett, S. T. A., Cadenasso, M. L., Grove, J. M., Nilon, C. H., Pouyat, R. V., Zipperer, W. C., & Costanza, R. (2001). Urban Ecological Systems: Linking Terrestrial Ecological, Physical, and Socioeconomic Components of Metropolitan Areas1. In Annual Review of Ecology, Evolution, and Systematics (Vol. 32, pp. 127–157). Annual Reviews. https://doi.org/10.1146/annurev.ecolsys.32.081501.114012

Pickett, S. T. A., & Zhou, W. (2015). Global urbanization as a shifting context for applying ecological science toward the sustainable city. Ecosystem Health and Sustainability, 1(1), 1–15. https://doi.org/10.1890/EHS14-0014.1

Pietrzak, M. B. (2014). The Modifiable Areal Unit Problem? Analysis of Correlation and Regression. Equilibrium. Quarterly Journal of Economics and Economic Policy, 9(4), 113–131. https://doi.org/10.12775/EQUIL.2014.028

Pinker, S. (2018). Aufklärung jetzt: Für Vernunft, Wissenschaft, Humanismus und Fortschritt. Eine Verteidigung. Fischer.

Pont, M., & Olsson, J. (2018). Typology based on three density variables central to Spacematrix using cluster analysis. 24th ISUF International Conference. Book of Papers. Editorial Universitat Politècnica de València, 1337–1348. https://doi.org/10.4995/ISUF2017.2017.5319

Popescu, G. (2012). Bordering and Ordering the Twenty-first Century: Understanding Borders. Rowman & Littlefield Publishers.

Portugali, J. (2016). What Makes Cities Complex? In J. Portugali & E. Stolk (Eds.), Complexity, Cognition, Urban Planning and Design (pp. 3–19). Springer International Publishing.

Portugali, J., & Stolk, E. (Eds.). (2016). Complexity, Cognition, Urban Planning and Design. Springer International Publishing. https://doi.org/10.1007/978-3-319-32653-5

Potere, D., Schneider, A., Angel, S., & Civco, D. L. (2009). Mapping urban areas on a global scale: Which of the eight maps now available is more accurate? International Journal of Remote Sensing, 30(24), 6531–6558. https://doi.org/10.1080/01431160903121134

Pourghasemi, H. R., Sadhasivam, N., Amiri, M., Eskandari, S., & Santosh, M. (2021). Landslide susceptibility assessment and mapping using state-of-the art machine learning techniques. Natural Hazards, 108(1), 1291–1316. https://doi.org/10.1007/s11069-021-04732-7

Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. Remote Sensing of Environment, 106(2), 253–267. https://doi.org/10.1016/j.rse.2006.09.005

Prakash, M., Ramage, S., Kavvada, A., & Goodman, S. (2020). Open Earth Observations for Sustainable Urban Development. Remote Sensing, 12(1646), 1–16. https://doi.org/10.3390/rs12101646

Priemus, H., & Zonneveld, W. (2003). What are corridors and what are the issues? Introduction to special issue: The governance of corridors. The Governance of Corridors, 11(3), 167–177. https://doi.org/10.1016/S0966-6923(03)00028-0

Prieur-Richard, A.-H., Walsh, B., Craig, M., Melamed, M. L., Colbert, M., Pathak, M., Connors, S., Bai, X., Barau, A., Bulkeley, H., Cleugh, H., Cohen, M., Colenbrander, S., Dodman, D., Dhakal, S., Dawson, R., Espey, J., Greenwalt, J., Kurian, P., … Ürge-Vorsatz, D. (2018). Extended Version: Global Research and Action Agenda on Cities and Climate Change Science. 242–261. https://doi.org/10.13140/RG.2.2.10315.44323

Protected Planet. (n.d.). Discover the World's Protected Areas. Retrieved 22 September 2022, from https://www.protectedplanet.net

Pu, R., & Landry, S. (2012). A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. Remote Sensing of Environment, 124, 516–533. https://doi.org/10.1016/j.rse.2012.06.011

Pudifoot, B., Cárdenas, M. L., Buytaert, W., Paul, J. D., Narraway, C. L., & Loiselle, S. (2021). When It Rains, It Pours: Integrating Citizen Science Methods to Understand Resilience of Urban Green Spaces. Frontiers in Water, Volume 3(654493), 1–12. https://doi.org/10.3389/frwa.2021.654493

Purwani, S., Nahar, J., & Twining, C. (2017). Analyzing bin-width effect on the computed entropy. AIP Conference Proceedings, 1868(040008), 1–6. https://doi.org/10.1063/1.4995123

Pushkarenko, Y., & Zaslavskyi, V. (2024). Research on the state of areas in Ukraine affected by military actions based on remote sensing data and deep learning architectures. Radioelectronic and Computer Systems, 2024(2), 5–18. https://doi.org/10.32620/reks.2024.2.01

Python Software. (2020). 3.7.17 Documentation [Computer software]. https://docs.python.org/3.7/

Q

Qi, M., Xu, C., Zhang, W., Demuzere, M., Hystad, P., Lu, T., James, P., Bechtel, B., & Hankey, S. (2024). Mapping urban form into local climate zones for the continental US from 1986–2020. Scientific Data, 11(195), 1–14. https://doi.org/10.1038/s41597-024-03042-4

Qian, Y., Zhou, W., Li, W., & Han, L. (2015). Understanding the dynamic of greenspace in the urbanized area of Beijing based on high resolution satellite images. Urban Forestry & Urban Greening, 14(1), 39–47. https://doi.org/10.1016/j.ufug.2014.11.006

Qin, J., Fang, C., Wang, Y., Li, G., & Wang, S. (2015). Evaluation of three-dimensional urban expansion: A case study of Yangzhou City, Jiangsu Province, China. Chinese Geographical Science, 25(2), 224–236. https://doi.org/10.1007/s11769-014-0728-8

Qiu, C., Mou, L., Schmitt, M., & Zhu, X. X. (2019). Local climate zone-based urban land cover classification from multi-seasonal Sentinel-2 images with a recurrent residual network. ISPRS Journal of Photogrammetry and Remote Sensing, 154, 151–162. https://doi.org/10.1016/j.isprsjprs.2019.05.004

Qiu, C., Schmitt, M., & Zhu, X. X. (2019). Fusing Multi-Seasonal Sentinel-2 Images with Residual Convolutional Neural Networks for Local Climate Zone-Derived Urban Land Cover Classification. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 5037–5040. https://doi.org/10.1109/IGARSS.2019.8898223

Qiu, C., Tong, X., Schmitt, M., Bechtel, B., & Zhu, X. X. (2020). Multilevel Feature Fusion-Based CNN for Local Climate Zone Classification From Sentinel-2 Images: Benchmark Results on the So2Sat LCZ42 Dataset. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 2793–2806. https://doi.org/10.1109/JSTARS.2020.2995711

Quan, J., Zhan, W., Ma, T., Du, Y., Guo, Z., & Qin, B. (2018). An integrated model for generating hourly Landsat-like land surface temperatures over heterogeneous landscapes. Remote Sensing of Environment, 206, 403–423. https://doi.org/10.1016/j.rse.2017.12.003

Quattrochi, D. A. (1985). An initial analysis of Landsat-4 Thematic Mapper data for the discrimination of agricultural, forested wetlands, and urban land cover (USA). Landsat-4 Science Characterization Early Results. Vol. 4: Applications. Symposium, Greenbelt, February 1983, 131–152.

Quattrochi, D. A., & Luvall, J. C. (2004). Thermal remote sensing in land surface processing. CRC Press.

R

Rabari, C., & Storper, M. (2015). The digital skin of cities: Urban theory and research in the age of the sensored and metered city, ubiquitous computing and big data. Cambridge Journal of Regions, Economy and Society, 8(1), 27–42. https://doi.org/10.1093/cjres/rsu021

Raciti, S. M., Hutyra, L. R., Rao, P., & Finzi, A. C. (2012). Inconsistent definitions of "urban" result in different conclusions about the size of urban carbon and nitrogen stocks. Ecological Applications, 22(3), 1015–1035. https://doi.org/10.1890/11-1250.1

Rademacher, A., Cadenasso, M. L., & Pickett, S. T. A. (2019). From feedbacks to coproduction: Toward an integrated conceptual framework for urban ecosystems. Urban Ecosystems, 22(1), 65–76. https://doi.org/10.1007/s11252-018-0751-0

Rahman, Md. M., Najaf, P., Fields, M. G., & Thill, J.-C. (2022). Traffic congestion and its urban scale factors: Empirical evidence from American urban areas. International Journal of Sustainable Transportation, 16(5), 406–421. https://doi.org/10.1080/15568318.2021.1885085

Rainforth, T., & Wood, F. (2015). Canonical Correlation Forests. arXiv reprint. https://doi.org/10.48550/arXiv.1507.05444

Rajasekar, U., & Weng, Q. (2009). Urban heat island monitoring and analysis using a non-parametric model: A case study of Indianapolis. ISPRS Journal of Photogrammetry and Remote Sensing, 64(1), 86–96. https://doi.org/10.1016/j.isprsjprs.2008.05.002

Ralphs, M., & Wyatt, P. (2003). GIS in Land and Property Management. Taylor & Francis.

Ramaswami, A., Russell, A. G., Culligan, P. J., Sharma, K. R., & Kumar, E. (2016). Meta-principles for developing smart, sustainable, and healthy cities. Science, 352(6288), 940–943. https://doi.org/10.1126/science.aaf7160

Rao, P. K. (1972). Remote Sensing of Urban Heat Islands from an Environmental Satellite. Bulletin of the American Meteorological Society, 53, 647–648.

Rashid, B., & Rehmani, M. H. (2016). Applications of wireless sensor networks for urban areas: A survey. Journal of Network and Computer Applications, 60, 192–219. https://doi.org/10.1016/j.jnca.2015.09.008

Rashid, M. (2017). The Geometry of Urban Layouts: A Global Comparative Study. Springer International Publishing. https://doi.org/10.1007/978-3-319-30750-3

Ratti, C., Frenchman, D., Pulselli, R. M., & Williams, S. (2006). Mobile Landscapes: Using Location Data from Cell Phones for Urban Analysis. Environment and Planning B: Planning and Design, 33(5), 727–748. https://doi.org/10.1068/b32047

Raucoules, D., Le Cozannet, G., Wöppelmann, G., de Michele, M., Gravelle, M., Daag, A., & Marcos, M. (2013). High nonlinear urban ground motion in Manila (Philippines) from 1993 to 2010 observed by DInSAR: Implications for sea-level measurement. Remote Sensing of Environment, 139, 386–397. https://doi.org/10.1016/j.rse.2013.08.021

Raymond, C., Matthews, T., & Horton, R. M. (2020). The emergence of heat and humidity too severe for human tolerance. Science Advances, 6(19)(eaaw1838), 1–8. https://doi.org/10.1126/sciadv.aaw1838

Reba, M., & Seto, K. C. (2020). A systematic review and assessment of algorithms to detect, characterize, and monitor urban land change. Remote Sensing of Environment, 242, 1–20. https://doi.org/10.1016/j.rse.2020.111739

Reddy, G. P. O., Kumar, N., Sahu, N., & Singh, S. K. (2018). Evaluation of automatic drainage extraction thresholds using ASTER GDEM and Cartosat-1 DEM: A case study from basaltic terrain of Central India. The Egyptian Journal of Remote Sensing and Space Science, 21(1), 95–104. https://doi.org/10.1016/j.ejrs.2017.04.001

Redepenning, Ma., & Hefner, C. (2018). Stadt und Land als Klassiker der Raumwissenschaften. In A. Hlukhovych, B. Bauer, K. Beuter, K. Lindner, & S. Vogt (Eds.), Kultur und kulturelle Bildung: Interdisziplinare ¨ Verortungen – Lehrerinnen- und Lehrerbildung – Perspektiven für die Schule (pp. 77–103). University of Bamberg Press. https://doi.org/10.20378/irb-46501

Regional Plan Association (RPA). (2007). Northeast Megaregion 2050: A Common Future (p. 28). New York Regional Plan Association. http://www.rpa.org/2007/11/northeast-megaregion-2050-a-common-future.html

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Reinoso Jerez, C. A. (2013). Diseño y construcción de un prototipo de sistema de sensores inalámbricos para alerta temprana de deslaves. Pontificia Universidad Católica de Ecuador, Quito, Ecuador.

Reis, C., & Lopes, A. (2019). Evaluating the Cooling Potential of Urban Green Spaces to Tackle Urban Climate Change in Lisbon. Sustainability, 11(2480), 1–17. https://doi.org/10.3390/su11092480

Ren, C., Cai, M., Li, X., Zhang, L., Wang, R., Xu, Y., & Ng, E. (2019). Assessment of Local Climate Zone Classification Maps of Cities in China and Feasible Refinements. Scientific Reports, 9(18848), 1–11. https://doi.org/10.1038/s41598-019-55444-9

Ren, C., Wang, R., Cai, M., Xu, Y., Zheng, Y., & Ng, E. (2016). The Accuracy of LCZ maps Generated by the World Urban Database and Access Portal Tools (WUDAPT) Method: A Case Study of Hong Kong. 1–11.

Rendgen, S. (2020). Information Graphics. Taschen Verlag.

Rendtorff, J. D. (2019). Business Ethics and Social Responsibility in Sustainable Development. Emerald Group Publishing. http://dx.doi.org/10.1108/9781789734539

Rengert, G., & Pelfrey, W. (1997). Cognitive mapping of the city center: Comparative perceptions of dangerous places. In D. Weisburd & T. McEwen (Eds.), Crime Mapping and Crime Prevention (pp. 193–217). Criminal Justice Press.

Resch, B., Summa, A., Sagl, G., Zeile, P., & Exner, J.-P. (2015). Urban Emotions—Geo-Semantic Emotion Extraction from Technical Sensors, Human Sensors and Crowdsourced Data. In G. Gartner & H. Huang (Eds.), Progress in Location-Based Services 2014 (pp. 199–212). Springer International Publishing. https://doi.org/10.1007/978-3-319-11879-6_14

Resonance Consultancy. (2020). Ranking of the best cities in the world. https://www.best cities.org/rankings/worlds-best-cities/

Revi, A., Satterthwaite, D. E., Aragón-Durand, F., Corfee-Morlot, J., Kiunsi, R. B. R., Pelling, M., Roberts, D. C., & Solecki, W. (2014). Urban Areas. In C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea, & L. L. White (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 535–612). Cambridge University Press.

Rezaei, N., & Millard-Ball, A. (2023). Urban form and its impacts on air pollution and access to green space: A global analysis of 462 cities. PLOS ONE, 18(1)(e0278265), 1–26. https://doi.org/10.1371/journal.pone.0278265

Richards, D. R., & Belcher, R. N. (2020). Global Changes in Urban Vegetation Cover. Remote Sensing, 12(23), 1–16. https://doi.org/10.3390/rs12010023

Richter, A., & Burrows, J. P. (2002). Tropospheric NO2 from GOME measurements. Advances in Space Research, 29(11), 1673–1683. https://doi.org/10.1016/S0273-1177(02)00100-X

Richter, A., Burrows, J. P., Nüß, H., Granier, C., & Niemeier, U. (2005). Increase in tropospheric nitrogen dioxide over China observed from space. Nature, 437(7055), 129–132. https://doi.org/10.1038/nature04092

Richter, B., Behnisch, M., & Grunewald, K. (2017). Messansatz zur Grünflächenversorgung von Einwohnern auf Stadt- und Stadtteilebene (No. 73; Flächennutzungsmonitoring IX - Nachhaltigkeit Der Siedlungs- Und Verkehrsentwicklung?, pp. 229–239). IÖR Schriften. https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-213014

Richter, B., Grunewald, K., & Meinel, G. (2016). Urbane Grünflächenausstattung und deren Erreichbarkeit – Indikatordesign und empirischer Städtevergleich (No. 69; Flächennutzungsmonitoring VIII - Flächensparen – Ökosystemleistungen – Handlungsstrategien, pp. 229–239). IÖR Schriften. https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-172615

Ridd, M. K. (1995). Exploring a V-I-S (vegetation-impervious surface-soil) model for urban ecosystem analysis through remote sensing: Comparative anatomy for cities†. International Journal of Remote Sensing, 16(12), 2165–2185. https://doi.org/10.1080/01431169508954549

Riguelle, F., Thomas, I., & Verhetsel, A. (2007). Measuring urban polycentrism: A European case study and its implications. Journal of Economic Geography, 7(2), 193–215. https://doi.org/10.1093/jeg/lbl025

Riitters, K. H., O'Neill, R. V., Hunsaker, C. T., Wickham, J. D., Yankee, D. H., Timmins, S. P., Jones, K. B., & Jackson, B. L. (1995). A factor analysis of landscape pattern and structure metrics. Landscape Ecology, 10(1), 23–39. https://doi.org/10.1007/BF00158551

Riofrio, G. (2003). Urban slums report: The case of Lima, Peru. In United Nations Human Settlements Programme (UN-Habitat) (Ed.), Global Report on Human Settlements, The Challenge of Slums (pp. 195–228).

Rittweger, J., Gilardi, L., Baltruweit, M., Dally, S., Erbertseder, T., Mittag, U., Naeem, M., Schmid, M., Schmitz, M.-T., Wüst, S., Dech, S., Jordan, J., Antoni, T., & Bittner, M. (2022). Temperature and particulate matter as environmental factors associated with seasonality of influenza incidence – an approach using Earth observation-based modeling in a health insurance cohort study from Baden-Württemberg (Germany). Environmental Health, 21(131), 1–13. https://doi.org/10.1186/s12940-022-00927-y

Roberts, D. A., Quattrochi, D. A., Hulley, G. C., Hook, S. J., & Green, R. O. (2012). Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the Hyperspectral Infrared Imager (HyspIRI) Decadal Survey mission. Remote Sensing of Urban Environments, 117, 83–101. https://doi.org/10.1016/j.rse.2011.07.021

Robertson, C., & Feick, R. (2016). Bumps and bruises in the digital skins of cities: Unevenly distributed user-generated content across US urban areas. Cartography and Geographic Information Science, 43(4), 283–300. https://doi.org/10.1080/15230406.2015.1088801

Robin, E. (2021). City Benchmarking, Globalized Urban Scholarship and the View from Above: Reflections on a Few Absences. International Journal of Urban and Regional Research, 45(2), 378–380. Scopus. https://doi.org/10.1111/1468-2427.12976

Roca Cladera, J., Marmolejo Duarte, C. R., & Moix, M. (2009). Urban Structure and Polycentrism: Towards a Redefinition of the Sub-centre Concept. Urban Studies, 46(13), 2841–2868. https://doi.org/10.1177/0042098009346329

Rode, P., Keim, C., Robazza, G., Viejo, P., & Schofield, J. (2014). Cities and Energy: Urban Morphology and Residential Heat-Energy Demand. Environment and Planning B: Planning and Design, 41(1), 138–162. https://doi.org/10.1068/b39065

Röder, A., Pröpper, M., Stellmes, M., Schneibel, A., & Hill, J. (2015). Assessing urban growth and rural land use transformations in a cross-border situation in Northern Namibia and Southern Angola. Land Use Policy, 42, 340–354. https://doi.org/10.1016/j.landusepol.2014.08.008

Rodrigue, J.-P. (2004). Freight, gateways and mega-urban regions: The logistical integration of the Bostwash corridor. Tijdschrift Voor Economische En Sociale Geografie, 95(2), 147–161. https://doi.org/10.1111/j.0040-747X.2004.t01-1-00297.x

Rodríguez-Álvarez, J. (2016). Urban Energy Index for Buildings (UEIB): A new method to evaluate the effect of urban form on buildings' energy demand. Landscape and Urban Planning, 148, 170–187. https://doi.org/10.1016/j.landurbplan.2016.01.001

Rogers, D., & Tsirkunov, V. (2011). Global assessment report on disaster risk reduction: Costs and benefits of early warning systems. Washington, DC: World Bank. http://documents.worldbank.org/curated/en/609951468330279598

Román, M. O., Stokes, E. C., Shrestha, R., Wang, Z., Schultz, L., Carlo, E. A. S., Sun, Q., Bell, J., Molthan, A., Kalb, V., Ji, C., Seto, K. C., McClain, S. N., & Enenkel, M. (2019). Satellite-based assessment of electricity restoration efforts in Puerto Rico after Hurricane Maria. PLOS ONE, 14(6)(e0218883), 1–22. https://doi.org/10.1371/journal.pone.0218883

Román, M. O., Wang, Z., Sun, Q., Kalb, V., Miller, S. D., Molthan, A., Schultz, L., Bell, J., Stokes, E. C., Pandey, B., Seto, K. C., Hall, D., Oda, T., Wolfe, R. E., Lin, G., Golpayegani, N., Devadiga, S., Davidson, C., Sarkar, S., … Masuoka, E. J. (2018). NASA's Black Marble nighttime lights product suite. Remote Sensing of Environment, 210, 113–143. https://doi.org/10.1016/j.rse.2018.03.017

Romero, V., Solís, E., & De Ureña, J. M. (2014). Beyond the metropolis: New employment centers and historic administrative cities in the Madrid global city region. Urban Geography, 35(6), 889–915. https://doi.org/10.1080/02723638.2014.939538

Rosentreter, J., Hagensieker, R., & Waske, B. (2020). Towards large-scale mapping of local climate zones using multitemporal Sentinel 2 data and convolutional neural networks. Remote Sensing of Environment, 237(111472), 1–12. https://doi.org/10.1016/j.rse.2019.111472

Ross, C. L. (2011). Megaregions: Literature Review of Organizational Structures and Finance of Multijurisdictional Initiatives and the Implications for Megaregion Transportation. United States. Federal Highway Administration. Office of Planning, Environment, and Realty. https://rosap.ntl.bts.gov/view/dot/50809

Ross, C. L., Barringer, J., & Jiawen, Y. (2008). Megaregions: Literature review of the implications for US infrastructure investment and transportation planning [Technical Report]. Georgia Institute of Technology. https://rosap.ntl.bts.gov/view/dot/50811/dot_50811_DS1.pdf

Ross, C. L., Barringer, J., & Yang, J. (2008). Megaregions: Literature Review of the Implications for U.S. Infrastructure Investment and Transportation Planning (dot:50811; pp. 1–102). https://rosap.ntl.bts.gov/view/dot/50811

Roth, C., Kang, S. M., Batty, M., & Barthélemy, M. (2011). Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows. PLOS ONE, 6(1)(e15923), 1–8. https://doi.org/10.1371/journal.pone.0015923

Roth, M., Oke, T. R., & Emery, W. J. (1989). Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology. International Journal of Remote Sensing, 10(11), 1699–1720. https://doi.org/10.1080/01431168908904002

Rowe, C., & Koetter, F. (1978). Collage City. MIT Press.

Roy, D. P., Wulder, M. A., Loveland, T. R., C.E., W., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf, C. B., Schott, J. R., Sheng, Y., Vermote, E. F., Belward, A. S., Bindschadler, R., Cohen, W. B., Gao, F., … Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment, 145, 154–172. https://doi.org/10.1016/j.rse.2014.02.001

Royston, J. P. (1982). An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. Journal of the Royal Statistical Society. Series C (Applied Statistics), 31(2), 115–124. JSTOR. https://doi.org/10.2307/2347973

Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2011). The Area and Population of Cities: New Insights from a Different Perspective on Cities. American Economic Review, 101(5), 2205–2225. https://doi.org/10.1257/aer.101.5.2205

Ruiz Peña, G. L., Navarro Alarcón, S. del R., Gamboa Rodríguez, C. A., Ramírez Hernández, K. C., Camargo Holguín, B. L., Trejos González, G. A., & Pérez Cerón, R. (2017). Las amenazas por movimientos en masa de Colombia, una visión a escala 1:100.000. Servicio Geológico Colombiano. https://doi.org/10.32685/9789589952887

Rumbold, J. M. M., & Pierscionek, B. K. (2018). What Are Data? A Categorization of the Data Sensitivity Spectrum. Big Data Centric Computational Intelligence in Bioinformatics and Healthcare, 12, 49–59. https://doi.org/10.1016/j.bdr.2017.11.001

Russell, R. F. (2001). The role of values in servant leadership. Leadership & Organization Development Journal, 22(2), 76–84. https://doi.org/10.1108/01437730110382631

Rustiadi, E., Pravitasari, A. E., Setiawan, Y., Mulya, S. P., Pribadi, D. O., & Tsutsumida, N. (2021). Impact of continuous Jakarta megacity urban expansion on the formation of the Jakarta-Bandung conurbation over the rice farm regions. Cities, 111(103000), 1–19. https://doi.org/10.1016/j.cities.2020.103000

S

Saaty, T. L. (1980). The analytic hierarchy process: Planning, priority setting, resource allocation. McGraw Hill Higher Education.

Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. Annals of Behavioral Medicine, 25(2), 80–91. https://doi.org/10.1207/S15324796ABM2502_03

Sailer-Fliege, U. (1999). Characteristics of post-socialist urban transformation in East Central Europe. GeoJournal, 49(1), 7–16. https://doi.org/10.1023/A:1006905405818

Salat, S. (2011a). Cities and Forms – On sustainable urbanism. Hermann.

Salat, S. (2011b). Les villes et les formes: Sur l'urbanisme durable. Hermann.

Salat, S., & Bourdic, L. (2012). Urban Complexity, Scale Hierarchy, Energy Efficiency And Economic Value Creation. The Sustainable City VII, 1, 97–107. http://dx.doi.org/10.2495/SC120091

Salat, S., Bourdic, L., & Labbe, F. (2014). Breaking Symmetries and Emerging Scaling Urban Structures: A Morphological Tale of 3 Cities: Paris, New York and Barcelona. Archnet-IJAR: International Journal of Architectural Research, 8(2), 77–93. https://doi.org/10.26687/archnet-ijar.v8i2.445

Salgueiro Barrio, R., & O'Shea, C. (2022). Testing planetary urbanisation: Siberia's trans-scalar spatial regime of oil production. City, Territory and Architecture, 9(23), 1–15. https://doi.org/10.1186/s40410-022-00168-z

Sallis, J. F., Cerin, E., Conway, T. L., Adams, M. A., Frank, L. D., Pratt, M., Salvo, D., Schipperijn, J., Smith, G., Cain, K. L., Davey, R., Kerr, J., Lai, P.-C., Mitáš, J., Reis, R., Sarmiento, O. L., Schofield, G., Troelsen, J., Dyck, D. V., … Owen, N. (2016). Physical activity in relation to urban environments in 14 cities worldwide: A cross-sectional study. The Lancet, 387(10034), 2207–2217. https://doi.org/10.1016/S0140-6736(15)01284-2

Salvati, L., & Carlucci, M. (2015). Land-use structure, urban growth, and periurban landscape: A multivariate classification of the European cities. Environment and Planning B, 42(5), 801–829. https://doi.org/10.1068/b120059p

Salvati, L., Zambon, I., Chelli, F. M., & Serra, P. (2018). Do spatial patterns of urbanization and land consumption reflect different socioeconomic contexts in Europe? Science of The Total Environment, 625, 722–730. https://doi.org/10.1016/j.scitotenv.2017.12.341

Salvia, M., Reckien, D., Pietrapertosa, F., Eckersley, P., Spyridaki, N.-A., Krook-Riekkola, A., Olazabal, M., De Gregorio Hurtado, S., Simoes, S. G., Geneletti, D., Viguié, V., Fokaides, P. A., Ioannou, B. I., Flamos, A., Csete, M. S., Buzasi, A., Orru, H., de Boer, C., Foley, A., … Heidrich, O. (2021). Will climate mitigation ambitions lead to carbon neutrality? An analysis of the local-level plans of 327 cities in the EU. Renewable and Sustainable Energy Reviews, 135(110253), 1–14. https://doi.org/10.1016/j.rser.2020.110253

Salvucci, G., & Salvati, L. (2022). Official Statistics, Building Censuses, and OpenStreetMap Completeness in Italy. ISPRS International Journal of Geo-Information, 11(29), 1–12. https://doi.org/10.3390/ijgi11010029

Samad, A., Kiseleva, O., Holst, C. C., Wegener, R., Kossmann, M., Meusel, G., Fiehn, A., Erbertseder, T., Becker, R., Roiger, A., Stanislawsky, P., Klemp, D., Emeis, S., Kalthoff, N., & Vogt, U. (2023). Meteorological and air quality measurements in a city region with complex terrain: Influence of meteorological phenomena on urban climate. Meteorologische Zeitschrift, 32(4), 293–315. https://doi.org/10.1127/metz/2023/1124

Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V., & Woolmer, G. (2002). The Human Footprint and the Last of the Wild: The human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. BioScience, 52(10), 891–904. https://doi.org/10.1641/0006-3568(2002)052[0891:THFATL]2.0.CO;2

Sanderson, R., Franklin, R., MacKinnon, D., & Matthews, J. (2024). Left out and invisible? : Exploring social media representation of 'left behind places'. GeoJournal, 89(37), 1–20. https://doi.org/10.1007/s10708-024-11010-z

Sapena, M., Gamperl, M., Kühnl, M., Garcia-Londoño, C., Singer, J., & Taubenböck, H. (2023). Cost estimation for the monitoring instrumentation of landslide early warning systems. Natural Hazards and Earth System Sciences, 23(12), 3913–3930. https://doi.org/10.5194/nhess-23-3913-2023

Sapena, M., Kühnl, M., Wurm, M., Patino, J. E., Duque, J. C., & Taubenböck, H. (2022). Empiric recommendations for population disaggregation under different data scenarios. PLOS ONE, 17(9)(e0274504), 1–29. https://doi.org/10.1371/journal.pone.0274504

Sapena, M., Mast, J., Schoepfer, E., & Taubenböck, H. (2024). Do-it-Yourself Built-Up Mapping Tool: A Practical Cloud-Based Solution Using Sentinel Imagery for Mapping Urban Expansion in Africa. SSRN Scholarly Paper, Rochester, NY. https://doi.org/ 10.2139/ssrn.4762416

Sapena, M., Ruiz, L. A., & Taubenböck, H. (2020). Analyzing Links between Spatio-Temporal Metrics of Built-Up Areas and Socio-Economic Indicators on a Semi-Global Scale. ISPRS International Journal of Geo-Information, 9(436), 1–22. https://doi.org/10.3390/ijgi9070436

Sapena, M., Wurm, M., Taubenböck, H., Tuia, D., & Ruiz, L. A. (2021). Estimating quality of life dimensions from urban spatial pattern metrics. Computers, Environment and Urban Systems, 85(101549), 1–11. https://doi.org/10.1016/j.compenvurbsys.2020.101549

Sassen, S. (2002). Global Networks, Linked Cities. Routledge.

Sassen, S. (2019). Cities in a World Economy (Fifth Edition). SAGE Publications, Inc. https://doi.org/10.4135/9781071872710

Saunders, D. (2010). Arrival cities. Blessing.

Savtchenko, A., Ouzounov, D., Ahmad, S., Acker, J., Leptoukh, G., Koziana, J., & Nickless, D. (2004). Terra and Aqua MODIS products available from NASA GES DAAC. Advances in Space Research, 34(4), 710–714. https://doi.org/10.1016/j.asr.2004.03.012

Šćepanović, S., Bogucka, E. P., Quercia, D., & Nattero, C. (2023). Responsible AI for Earth Observation: Attitudes Among Experts. IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, 1934–1936. https://doi.org/10.1109/IGARSS52108.2023.10282983

Šćepanović, S., Bogucka, E., & Quercia, D. (2025). AI in the city: Impact assessment of artificial intelligence uses in earth observation. Urban Informatics, 4(10), 1–30. https://doi.org/10.1007/s44212-025-00074-y

Schaap, M., Apituley, A., Timmermans, R. M. A., Koelemeijer, R. B. A., & de Leeuw, G. (2009). Exploring the relation between aerosol optical depth and PM2.5 at Cabauw, the Netherlands. Atmospheric Chemistry and Physics, 9(3), 909–925. https://doi.org/10.5194/acp-9-909-2009

Schabenberger, O., & Gotway, C. A. (2005). Statistical Methods for Spatial Data Analysis. Taylor & Francis.

Schaefer, M., Ebrahimi Salari, H., Köckler, H., & Thinh, N. X. (2021). Assessing local heat stress and air quality with the use of remote sensing and pedestrian perception in urban microclimate simulations. Science of The Total Environment, 794(148709), 1–15. https://doi.org/10.1016/j.scitotenv.2021.148709

Schiavina, M., Freire, S., & MacManus, K. (2021). GHS-POP R2019A - GHS population grid multitemporal (1975-1990-2000-2015)—OBSOLETE RELEASE. European Commission, Joint Research Centre (JRC). [Dataset]. https://doi.org/10.2905/0C6B9751-A71F-4062-830B-43C9F432370F. PID: http://data.europa.eu/89h/0c6b9751-a71f-4062-830b-43c9f432370f

Schiavina, M., Freire, S., Carioli, A., & Kytt, M. (2023). GHS-POP R2023A - GHS population grid multitemporal (1975-2030) [Dataset]. European Commission, Joint Research Centre (JRC). https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE. PID: http://data.europa.eu/89h/2ff68a52-5b5b-4a22-8f40-c41da8332cfe

Schiavina, M., Melchiorri, M., Freire, S., Florio, P., Ehrlich, D., Tommasi, P., Pesaresi, M., & Kemper, T. (2022). Land use efficiency of functional urban areas: Global pattern and evolution of development trajectories. Habitat International, 123(102543), 1–11. https://doi.org/10.1016/j.habitatint.2022.102543

Schiavina, M., Moreno-Monroy, A., Maffenini, L., & Veneri, P. (2019). GHS-FUA R2019A - GHS functional urban areas, derived from GHS-UCDB R2019A (2015) [Dataset]. European Commission, Joint Research Centre (JRC). https://doi.org/10.2905/347F0337-F2DA-4592-87B3-E25975EC2C95

Schiller, R. (2001). The Dynamics of Property Location. Spon Press.

Schimel, D. S. (1995). Terrestrial biogeochemical cycles: Global estimates with remote sensing. Remote Sensing of Land Surface for Studies of Global Chage, 51(1), 49–56. https://doi.org/10.1016/0034-4257(94)00064-T

Schlüter, O. (1899). Über den Grundriß der Städte. Zeitschrift Der Gesellschaft Für Erdkunde Zu Berlin, 34(6), 446–461.

Schmidt, S. (2011). Sprawl Without Growth in Eastern Germany. Urban Geography, 32(1), 105–128. https://doi.org/10.2747/0272-3638.32.1.105

Schmidt, S., & Buehler, R. (2007). The Planning Process in the US and Germany: A Comparative Analysis. International Planning Studies, 12(1), 55–75. https://doi.org/10.1080/13563470701346592

Schmidt, S., Fina, S., & Siedentop, S. (2015). Post-socialist Sprawl: A Cross-Country Comparison. European Planning Studies, 23(7), 1357–1380. https://doi.org/10.1080/09654313.2014.933178

Schmitt, A. (2016). Multiscale and Multidirectional Multilooking for SAR Image Enhancement. IEEE Transactions on Geoscience and Remote Sensing, 54(9), 5117–5134. https://doi.org/10.1109/TGRS.2016.2555624

Schmitt, A., Sieg, T., Wurm, M., & Taubenböck, H. (2018). Investigation on the separability of slums by multi-aspect TerraSAR-X dual-co-polarized high resolution spotlight images based on the multi-scale evaluation of local distributions. International Journal of Applied Earth Observation and Geoinformation, 64, 181–198. https://doi.org/10.1016/j.jag.2017.09.006

Schmitt, A., Uth, P., Standfuß, I., Heider, B., Siedentop, S., & Taubenböck, H. (2023). Quantitative assessment and comparison of urban patterns in Germany and the United States. Computers, Environment and Urban Systems, 100(101920), 1–18. https://doi.org/10.1016/j.compenvurbsys.2022.101920

Schmitt, A., Wendleder, A., & Hinz, S. (2014). The Kennaugh element framework for multi-scale, multi-polarized, multi-temporal and multi-frequency SAR image preparation. ISPRS Journal of Photogrammetry and Remote Sensing, 102, 122–139. https://doi.org/10.1016/j.isprsjprs.2015.01.007

Schmitt, A., Wendleder, A., Kleynmans, R., Hell, M., Roth, A., & Hinz, S. (2020). Multi-Source and Multi-Temporal Image Fusion on Hypercomplex Bases. Remote Sensing, 12(943), 1–37. https://doi.org/10.3390/rs12060943

Schmitt, M., Ahmadi, S. A., & Hänsch, R. (2021). There is No Data Like More Data—Current Status of Machine Learning Datasets in Remote Sensing. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 1206–1209. https://doi.org/10.1109/IGARSS47720.2021.9555129

Schmitt, M., Hughes, L. H., Qiu, C., & Zhu, X. X. (2019). Aggregating cloud-free Sentinel-2 images with Google Earth Engine. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W7, 145–152. https://doi.org/10.5194/isprs-annals-IV-2-W7-145-2019

Schneider, A., Friedl, M. A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. Environmental Research Letters, 4(044003), 1–11. https://doi.org/10.1088/1748-9326/4/4/044003

Schneider, A., & Woodcock, C. E. (2008). Compact, Dispersed, Fragmented, Extensive? A Comparison of Urban Growth in Twenty-five Global Cities using Remotely Sensed Data, Pattern Metrics and Census Information. Urban Studies, 45(3), 659–692. https://doi.org/10.1177/0042098007087340

Schneider, P., Lahoz, W. A., & van der A, R. (2015). Recent satellite-based trends of tropospheric nitrogen dioxide over large urban agglomerations worldwide. Atmospheric Chemistry and Physics, 15(3), 1205–1220. https://doi.org/10.5194/acp-15-1205-2015

Schneider, P., & van der A, R. J. (2012). A global single-sensor analysis of 2002–2011 tropospheric nitrogen dioxide trends observed from space. Journal of Geophysical Research: Atmospheres, 117(D16309), 1–17. https://doi.org/10.1029/2012JD017571

Schollerer, L., Schmitt, A., Wendleder, A., & Rogginger, S. (2002). Schritthaltende Baufallerkundung aus dem All mit frei verfügbaren Satellitenaufnahmen. Zfv, 3, 168–180. https://doi.org/10.12902/zfv-0391-2022

Schon, J., & Leblang, D. (2021). Why Physical Barriers Backfire: How Immigration Enforcement Deters Return and Increases Asylum Applications. Comparative Political Studies, 54(14), 2611–2652. https://doi.org/10.1177/00104140211024282

Schott, J. R., Gerace, A., Woodcock, C. E., Wang, S., Zhu, Z., Wynne, R. H., & Blinn, C. E. (2016). The impact of improved signal-to-noise ratios on algorithm performance: Case studies for Landsat class instruments. Remote Sensing of Environment, 185, 37–45. https://doi.org/10.1016/j.rse.2016.04.015

Schubert, J. E., Sanders, B. F., Smith, M. J., & Wright, N. G. (2008). Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. Advances in Water Resources, 31(12), 1603–1621. https://doi.org/10.1016/j.advwatres.2008.07.012

Schumann, G. J.-P., Neal, J. C., Mason, D. C., & Bates, P. D. (2011). The accuracy of sequential aerial photography and SAR data for observing urban flood dynamics, a case study of the UK summer 2007 floods. Remote Sensing of Environment, 115(10), 2536–2546. https://doi.org/10.1016/j.rse.2011.04.039

Schwaab, J., Meier, R., Mussetti, G., Seneviratne, S., Bürgi, C., & Davin, E. L. (2021). The role of urban trees in reducing land surface temperatures in European cities. Nature Communications, 12(6763), 1–11. https://doi.org/10.1038/s41467-021-26768-w

Schwartz, M. S. (2013). Developing and sustaining an ethical corporate culture: The core elements. Business Horizons, 56(1), 39–50. https://doi.org/10.1016/j.bushor.2012.09.002

Schwarz, N. (2010). Urban form revisited—Selecting indicators for characterising European cities. Landscape and Urban Planning, 96(1), 29–47. https://doi.org/10.1016/j.landurbplan.2010.01.007

Schwarz, N., Lautenbach, S., & Seppelt, R. (2011). Exploring indicators for quantifying surface urban heat islands of European cities with MODIS land surface temperatures. Remote Sensing of Environment, 115(12), 3175–3186. https://doi.org/10.1016/j.rse.2011.07.003

Scott, A. J., & Storper, M. (2015). The Nature of Cities: The Scope and Limits of Urban Theory. International Journal of Urban and Regional Research, 39(1), 1–15. https://doi.org/10.1111/1468-2427.12134

Scullion, J., Thomas, C. W., Vogt, K. A., Pérez-Maqueo, O., & Logsdon, M. G. (2011). Evaluating the environmental impact of payments for ecosystem services in Coatepec (Mexico) using remote sensing and on-site interviews. Environmental Conservation, 38(4), 426–434. Cambridge Core. https://doi.org/10.1017/S037689291100052X

SEDAC. (2018). Gridded population of the world (GPW) v4 [Dataset]. https://www.earthdata.nasa.gov/data/projects/gpw

Seele, P., & Gatti, L. (2017). Greenwashing Revisited: In Search of a Typology and Accusation-Based Definition Incorporating Legitimacy Strategies. Business Strategy and the Environment, 26(2), 239–252. https://doi.org/10.1002/bse.1912

Segoni, S., Serengil, Y., & Aydin, F. (2023). A prototype landslide early warning system in Rize (Turkey): Analyzing recent impacts to design a safer future. Landslides, 20(3), 683–694. https://doi.org/10.1007/s10346-022-01988-3

Şen, İ. (2024). A decolonial analysis of urban and regional planning education in the global south. http://hdl.handle.net/11527/26452

Senaratne, H., Mühlbauer, M., Kiefl, R., Cárdenas, A., Prathapan, L., Riedlinger, T., Biewer, C., & Taubenböck, H. (2023). The Unseen—An Investigative Analysis of Thematic and Spatial Coverage of News on the Ongoing Refugee Crisis in West Africa. ISPRS International Journal of Geo-Information, 12(4), 1–26. https://doi.org/10.3390/ijgi12040175

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S. M., Wehner, M., & Zhou, B. (2023). Weather and Climate Extreme Events in a Changing Climate. In V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou (Eds.), Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 1513–1766). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781009157896.013

Sennett, R. (2018a). Building and Dwelling: Ethics for the City (1st edition). Allen Lane.

Sennett, R. (2018b). Die offene Stadt. Eine Ethik des Bauens und Bewohnens. Hanser Berlin.

Sert, J. L. (1952). Centres of Community Life. In J. Tyrwhitt, J. Lluis Sert, & E. N. Rogers (Eds.), The Heart of the City: Towards the humanisation of urban life (pp. 3–16). Pellegrini and Cudahy.

Seto, K. C., Dhakal, S., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMaho, J., Müller, D. B., Murakami, J., Nagendra, H., & Ramaswami, A. (2014). Human Settlements, Infrastructure, and Spatial Planning. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, & J. C. Minx (Eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (pp. 923–1000). Cambridge University Press.

Seto, K. C., & Fragkias, M. (2005). Quantifying Spatiotemporal Patterns of Urban Land-use Change in Four Cities of China with Time Series Landscape Metrics. Landscape Ecology, 20(7), 871–888. https://doi.org/10.1007/s10980-005-5238-8

Seto, K. C., Fragkias, M., Güneralp, B., & Reilly, M. K. (2011). A Meta-Analysis of Global Urban Land Expansion. PLOS ONE, 6(8)(e23777), 1–9. https://doi.org/10.1371/journal.pone.0023777

Seto, K. C., Golden, J. S., Alberti, M., & Turner, B. L. (2017). Sustainability in an urbanizing planet. Proceedings of the National Academy of Sciences, 114(34), 8935–8938. https://doi.org/10.1073/pnas.1606037114

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. Proceedings of the National Academy of Sciences, 109(40), 16083–16088. https://doi.org/10.1073/pnas.1211658109

Seto, K. C., & Ramankutty, N. (2016). Hidden linkages between urbanization and food systems. Science, 352(6288), 943–945. https://doi.org/10.1126/science.aaf7439

Seto, K. C., & Reba, M. (2018). Cities Unseen – New Visions of an Urban Planet. Yale University Press.

Seto, K. C., Sánchez-Rodríguez, R., & Fragkias, M. (2010). The New Geography of Contemporary Urbanization and the Environment. Annual Review of Environment and Resources, 35(Volume 35), 167–194. https://doi.org/10.1146/annurev-environ-100809-125336

Seto, K. C., Woodcock, C. E., Song, C., Huang, X., Lu, J., & Kaufmann, R. K. (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. International Journal of Remote Sensing, 23(10), 1985–2004. https://doi.org/10.1080/01431160110075532

Seto, K., Solecki, W., & Griffith, C. (2015). The Routledge Handbook of Urbanization and Global Environmental Change (1st ed.). Routledge. https://doi.org/10.4324/9781315849256

Seto, T., Kanasugi, H., & Nishimura, Y. (2020). Quality Verification of Volunteered Geographic Information Using OSM Notes Data in a Global Context. ISPRS International Journal of Geo-Information, 9(372), 1–19. https://doi.org/10.3390/ijgi9060372

Sexton, J. O., Song, X.-P., Huang, C., Channan, S., Baker, M. E., & Townshend, J. R. (2013). Urban growth of the Washington, D.C.–Baltimore, MD metropolitan region from 1984 to 2010 by annual, Landsat-based estimates of impervious cover. Remote Sensing of Environment, 129, 42–53. https://doi.org/10.1016/j.rse.2012.10.025

Shachar, A., Fine, S., Huber, J., Keitner, C. I., Lori, N., Mau, S., & Volpp, L. (2020). The shifting border: Legal cartographies of migration and mobility: Ayelet Shachar in dialogue (1st ed.). Manchester University Press.

Shahtahmassebi, A. R., Li, C., Fan, Y., Wu, Y., lin, Y., Gan, M., Wang, K., Malik, A., & Blackburn, G. A. (2021). Remote sensing of urban green spaces: A review. Urban Forestry & Urban Greening, 57(126946), 1–15. https://doi.org/10.1016/j.ufug.2020.126946

Shankar, S. S. R. (n.d.). The Three Shaktis (Powers). Wisdom by Gurudev. Retrieved 4 January 2023, from https://wisdom.srisriravishankar.org/dynamics-of-three-shaktis/

Shankar, S. S. R. (2007). Universal declaration for human values (p. 2). International Association for Human Values. https://www.iahv.org/us-en/wp-content/themes/IAHV/PDF/Universal-Declaration-of-Human-Values.pdf

Shannon, C. E. (1948). A mathematical theory of communication. The Bell System Technical Journal, 27(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shao, Z., Sumari, N. S., Portnov, A., Ujoh, F., Musakwa, W., & Mandela, P. J. (2020). Urban sprawl and its impact on sustainable urban development: A combination of remote sensing and social media data. Geo-Spatial Information Science, 24(2), 241–255. https://doi.org/10.1080/10095020.2020.1787800

Shapiro, S. (1998). Places and Spaces: The Historical Interaction of Technology, Home, and Privacy. The Information Society, 14(4), 275–284. https://doi.org/10.1080/019722498128728

Shashoua, M., Basu, S. R., & UN.ESCAP. (2015). Polarizing world: GDP, development and beyond. United Nations. https://repository.unescap.org/handle/20.500.12870/1276

She, Q., Peng, X., Xu, Q., Long, L., Wei, N., Liu, M., Jia, W., Zhou, T., Han, J., & Xiang, W. (2017). Air quality and its response to satellite-derived urban form in the Yangtze River Delta, China. Ecological Indicators, 75, 297–306. https://doi.org/10.1016/j.ecolind.2016.12.045

Shea, K. M., Truckner, R. T., Weber, R. W., & Peden, D. B. (2008). Climate change and allergic disease. Journal of Allergy and Clinical Immunology, 122(3), 443–453. https://doi.org/10.1016/j.jaci.2008.06.032

Shearmur, R. (2017). Urban bias in innovation studies. In H. Bathelt, P. Cohendet, S. Henn, & L. Simon (Eds.), The Elgar Companion to Innovation and Knowledge Creation (pp. 440–456). https://doi.org/10.4337/9781782548522.00037

Shearmur, R., & Alvergne, C. (2003). Regional planning policy and the location of employment in the Ile-de-France: Does policy matter? In Urban Affairs Review (Vol. 39, Issue 1, pp. 3–31). https://doi.org/10.1177/1078087403253557

Shearmur, R., & Coffey, W. J. (2002). A Tale of Four Cities: Intrametropolitan Employment Distribution in Toronto, Montreal, Vancouver, and Ottawa – Hull, 1981 – 1996. Environment and Planning A: Economy and Space, 34(4), 575–598. https://doi.org/10.1068/a33214

Shepherd, J. M. (2005). A Review of Current Investigations of Urban-Induced Rainfall and Recommendations for the Future. Earth Interactions, 9(12), 1–27. https://doi.org/10.1175/EI156.1

Sheppard, E. (2000). Socialist cities? Urban Geography, 21(8), 758–763. https://doi.org/10.2747/0272-3638.21.8.758

Shi, K., Huang, C., Yu, B., Yin, B., Huang, Y., & Wu, J. (2014). Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. Remote Sensing Letters, 5(4), 358–366. https://doi.org/10.1080/2150704X.2014.905728

Shi, L., Liu, F., Zhang, Z., Zhao, X., Liu, B., Xu, J., Wen, Q., Yi, L., & Hu, S. (2015). Spatial differences of coastal urban expansion in China from 1970s to 2013. Chinese Geographical Science, 25(4), 389–403. https://doi.org/10.1007/s11769-015-0765-y

Shi, L., Taubenböck, H., Zhang, Z., Liu, F., & Wurm, M. (2019). Urbanization in China from the end of 1980s until 2010 – spatial dynamics and patterns of growth using EO-data. International Journal of Digital Earth, 12(1), 78–94. https://doi.org/10.1080/17538947.2017.1400599

Shi, L., Wurm, M., Huang, X., Zhong, T., Leichtle, T., & Taubenböck, H. (2020). Urbanization that hides in the dark – Spotting China's "ghost neighborhoods" from space. Landscape and Urban Planning, 200(103822), 1–14. https://doi.org/10.1016/j.landurbplan.2020.103822

Shi, L., Wurm, M., Huang, X., Zhong, T., & Taubenböck, H. (2020). Measuring the spatial hierarchical urban system in China in reference to the Central Place Theory. Habitat International, 105(102264), 1–13. https://doi.org/10.1016/j.habitatint.2020.102264

Shi, L., Zhang, Z., Liu, F., Zhao, X., Wang, X., Liu, B., Hu, S., Wen, Q., Zuo, L., Yi, L., & Xu, J. (2016). City size distribution and its spatiotemporal evolution in China. Chinese Geographical Science, 26(6), 703–714. https://doi.org/10.1007/s11769-016-0832-z

Shipley, R., & Utz, S. (2012). Making it Count: A Review of the Value and Techniques for Public Consultation. Journal of Planning Literature, 27(1), 22–42. https://doi.org/10.1177/0885412211413133

SIATA. (2023). Sistema de Alerta Temprana de Medellín y el Valle de Aburrá [Dataset]. https://siata.gov.co/siata_nuevo/

Sibley, L. M., & Weiner, J. P. (2011). An evaluation of access to health care services along the rural-urban continuum in Canada. BMC Health Services Research, 11(20), 1–11. https://doi.org/10.1186/1472-6963-11-20

Siedentop, S. (2005). Urban Sprawl—verstehen, messen, steuern: Ansatzpunkte für ein empirisches Mess- und Evaluationskonzept der urbanen Siedlungsentwicklung. DISP, 160, 23–35.

Siedentop, S., & Fina, S. (2010). Monitoring urban sprawl in Germany: Towards a GIS-based measurement and assessment approach. Journal of Land Use Science, 5(2), 73–104. https://doi.org/10.1080/1747423X.2010.481075

Siedentop, S., & Fina, S. (2010). Urban Sprawl beyond Growth: The Effect of Demographic Change on Infrastructure Costs. Flux, 79–80(1), 90–100. Cairn.info. https://doi.org/10.3917/flux.079.0090

Siedentop, S., & Fina, S. (2012). Who Sprawls Most? Exploring the Patterns of Urban Growth across 26 European Countries. Environment and Planning A, 44(11), 2765–2784. https://doi.org/10.1068/a4580

Siedentop, S., Kausch, S., Einig, K., & Gössel, J. (2003). Siedlungsstrukturelle Veränderungen im Umland von Agglomerationsräumen (Forschungen Heft 114). Bundesamt für Bauwesen und Raumordnung.

Sieverts, T. (1997). Zwischenstadt: Zwischen Ort und Welt Raum und Zeit Stadt und Land. Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-663-11933-3

Silva, C. P., García, C. E., Estay, S. A., & Barbosa, O. (2015). Bird Richness and Abundance in Response to Urban Form in a Latin American City: Valdivia, Chile as a Case Study. PLOS ONE, 10(9)(e0138120), 1–16. https://doi.org/10.1371/journal.pone.0138120

Silvern, R. F., Jacob, D. J., Mickley, L. J., Sulprizio, M. P., Travis, K. R., Marais, E. A., Cohen, R. C., Laughner, J. L., Choi, S., Joiner, J., & Lamsal, L. N. (2019). Using satellite observations of tropospheric \chemNO_2 columns to infer long-term trends in US \chemNO_\mathitx emissions: The importance of accounting for the free tropospheric \chemNO_2 background. Atmospheric Chemistry and Physics, 19(13), 8863–8878. https://doi.org/10.5194/acp-19-8863-2019

SIMMA. (2022). Sistema de Información de Movimientos en Masa [Dataset]. https://simma.sgc.gov.co/

Simmons, B. A. (2019). Border Rules. International Studies Review, 21(2), 256–283. https://doi.org/10.1093/isr/viz013

Simmons, B. A., & Kenwick, M. R. (2021). Border Orientation in a Globalizing World. American Journal of Political Science, 66(4), 853–870. https://doi.org/10.1111/ajps.12687

Simon, D. (2008). Urban environments: Issues on the peri-urban fringe. In Annual Review of Environment and Resources (Vol. 33, pp. 167–185). https://doi.org/10.1146/annurev.environ.33.021407.093240

Simpson, E. H. (1949). Measurement of Diversity. Nature, 163(4148), 688–688. https://doi.org/10.1038/163688a0

Singer, J., & Thuro, K. (2006). Development of a continuous 3Dmonitoring system for unstable slopes using Time Domain Reflectometry. IAEG2006, Paper number 495, 1–7. https://media.geolsoc.org.uk/iaeg2006/PAPERS/IAEG_495.PDF

Singer, J., Thuro, K., Gamperl, M., Breuninger, T., & Menschik, B. (2021). Technical Concepts for an Early Warning System for Rainfall Induced Landslides in Informal Settlements. In N. Casagli, V. Tofani, K. Sassa, P. T. Bobrowsky, & K. Takara (Eds.), Understanding and Reducing Landslide Disaster Risk: Volume 3 Monitoring and Early Warning (pp. 209–215). Springer International Publishing. https://doi.org/10.1007/978-3-030-60311-3_24

Singer, P. (2022). ethics. Encyclopedia Britannica. https://www.britannica.com/topic/ ethics-philosophy

Singh, G., Memoli, F., & Carlsson, G. (2007). Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. In M. Botsch, R. Pajarola, B. Chen, & M. Zwicker (Eds.), Eurographics Symposium on Point-Based Graphics (pp. 91–100). The Eurographics Association. https://doi.org//10.2312/SPBG/SPBG07/091-100

Singh, H., Singh, D., Singh, S. K., & Shukla, D. N. (2017). Assessment of river water quality and ecological diversity through multivariate statistical techniques, and earth observation dataset of rivers Ghaghara and Gandak, India. International Journal of River Basin Management, 15(3), 347–360. https://doi.org/10.1080/15715124.2017.1300159

Sirko, W., Kashubin, S., Ritter, M., Annkah, A., Bouchareb, Y. S. E., Dauphin, Y., Keysers, D., Neumann, M., Cisse, M., & Quinn, J. (2021). Continental-Scale Building Detection from High Resolution Satellite Imagery. https://arxiv.org/abs/2107.12283

Skilodimou, H. D., Bathrellos, G. D., Chousianitis, K., Youssef, A. M., & Pradhan, B. (2019). Multi-hazard assessment modeling via multi-criteria analysis and GIS: a case study. Environmental Earth Sciences, 78(47), 1–21. https://doi.org/10.1007/s12665-018-8003-4

Skole, D., & Tucker, C. (1993). Tropical Deforestation and Habitat Fragmentation in the Amazon: Satellite Data from 1978 to 1988. Science, 260(5116), 1905–1910. https://doi.org/10.1126/science.260.5116.1905

Sloterdijk, P. (2013). Ausgewählte Übertreibungen. Gespräche und Interviews 1993-2012. Suhrkamp.

Sloterdijk, P. (2018). Neue Zeilen und Tage – Notizen 2011-2013. Suhrkamp.

Small, C. (2001). Estimation of urban vegetation abundance by spectral mixture analysis. International Journal of Remote Sensing, 22(7), 1305–1334. https://doi.org/10.1080/01431160151144369

Small, C., Elvidge, C. D., Balk, D., & Montgomery, M. (2011). Spatial scaling of stable night lights. Remote Sensing of Environment, 115(2), 269–280. https://doi.org/10.1016/j.rse.2010.08.021

Small, C., Pozzi, F., & Elvidge, C. D. (2005). Spatial analysis of global urban extent from DMSP-OLS night lights. Remote Sensing of Environment, 96(3), 277–291. https://doi.org/10.1016/j.rse.2005.02.002

Small, C., & Sousa, D. (2016). Humans on Earth: Global extents of anthropogenic land cover from remote sensing. Anthropocene, 14, 1–33. https://doi.org/10.1016/j.ancene.2016.04.003

Small, C., Sousa, D., Yetman, G., Elvidge, C., & MacManus, K. (2018). Decades of urban growth and development on the Asian megadeltas. Global and Planetary Change, 165, 62–89. https://doi.org/10.1016/j.gloplacha.2018.03.005

Smigiel, C., & Brade, I. (2011). Suburbanisierung im östlichen Europa im Zeitalter neoliberaler Stadtentwicklung. disP - The Planning Review, 47(185), 27–43. https://doi.org/10.1080/02513625.2011.10557131

Smith, M. E. (2015). From Prehistoric Villages to Cities: Settlement Aggregation and Community Transformation by Jennifer Birch, ed. American Anthropologist, 117(1), 178–179. https://doi.org/10.1111/aman.12196

Smith, P., Gould, M., & Manley, E. (2025). The Rise and Fall of Twitter Data in Urban Analytics. Findings, 1–10. https://doi.org/10.32866/001c.133245

Socio-economic Data and Applications Center (SEDAC). (2021). Global Rural-Urban Mapping Project (GRUMP). V1 [Dataset]. https://sedac.ciesin.columbia.edu/data/collection/ grump-v1/methods

Sohn, C., & Licheron, J. (2018). The multiple effects of borders on metropolitan functions in Europe. Regional Studies, 52(11), 1512–1524. https://doi.org/10.1080/00343404.2017.1410537

Soja, E. W., & Kanai, J. M. (2008). The Urbanization of the World. In R. Burdett & D. Sudjic (Eds.), Towards a Study of Planetary Urbanization (pp. 54–69). Phaidon Press.

Soliman, A., Soltani, K., Yin, J., Padmanabhan, A., & Wang, S. (2017). Social sensing of urban land use based on analysis of Twitter users' mobility patterns. PLOS ONE, 12(7)(e0181657), 1–16. https://doi.org/10.1371/journal.pone.0181657

Solomon, A. (1998). Draft declaration of human values. Humanist Canada, 31(3), 1–11.

Solow, R. M. (1973). On Equilibrium Models of Urban Location. In M. Parkin (Ed.), Essays in modern economics (pp. 2–16). Longman Group Ltd.

Somers, M. J. (2001). Ethical Codes of Conduct and Organizational Context: A Study of the Relationship Between Codes of Conduct, Employee Behavior and Organizational Values. Journal of Business Ethics, 30(2), 185–195. https://doi.org/10.1023/A:1006457810654

Soo, K. T. (2005). Zipf's Law for cities: A cross-country investigation. Regional Science and Urban Economics, 35(3), 239–263. https://doi.org/10.1016/j.regsciurbeco.2004.04.004

Sorichetta, A., Hornby, G. M., Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. In Scientific Data (Vol. 2, Issue 150045, pp. 1–12). https://doi.org/10.1038/sdata.2015.45

Spivey, C. (2008). The Mills—Muth Model of Urban Spatial Structure: Surviving the Test of Time? Urban Studies, 45(2), 295–312. https://doi.org/10.1177/0042098007085964

Staab, J., Schady, A., Weigand, M., Lakes, T., & Taubenböck, H. (2021). Predicting traffic noise using land-use regression—A scalable approach. Journal of Exposure Science & Environmental Epidemiology, 32(2), 232–243. https://doi.org/10.1038/s41370-021-00355-z

Standord Dictionary. (2015). Relativism. https://plato.stanford.edu/entries/relativism/

Stanilov, K., & Sýkora, L. (2014). Confronting Suburbanization. Urban Decentralization in Postsocialist Central and Eastern Europe. Wiley-Blackwell.

Stavrakou, T., Müller, J.-F., Bauwens, M., Boersma, K. F., & van Geffen, J. (2020). Satellite evidence for changes in the NO2 weekly cycle over large cities. Scientific Reports, 10(10066), 1–9. https://doi.org/10.1038/s41598-020-66891-0

Steiger, E., Westerholt, R., Resch, B., & Zipf, A. (2015). Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. Computers, Environment and Urban Systems, 54, 255–265. https://doi.org/10.1016/j.compenvurbsys.2015.09.007

Steinberger, J. K., Krausmann, F., & Eisenmenger, N. (2010). Global patterns of materials use: A socioeconomic and geophysical analysis. Ecological Economics, 69(5), 1148–1158. https://doi.org/10.1016/j.ecolecon.2009.12.009

Steiner, F. (2014). Frontiers in urban ecological design and planning research. Landscape and Urban Planning, 125, 304–311. https://doi.org/10.1016/j.landurbplan.2014.01.023

Stephenne, N., Magoni, R., & Laneve, G. (2009). From Real Time Border Monitoring to a Permeability Model. In B. Jasani, M. Pesaresi, S. Schneiderbauer, & G. Zeug (Eds.), Remote Sensing from Space: Supporting International Peace and Security (pp. 239–259). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8484-3_15

Stevens, F. R., Gaughan, A. E., Linard, C., & Tatem, A. J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. PLOS ONE, 10(2)(e0107042), 1–22. https://doi.org/10.1371/journal.pone.0107042

Stevenson, M., Thompson, J., de Sá, T. H., Ewing, R., Mohan, D., McClure, R., Roberts, I., Tiwari, G., Giles-Corti, B., Sun, X., Wallace, M., & Woodcock, J. (2016). Land use, transport, and population health: Estimating the health benefits of compact cities. The Lancet, 388(10062), 2925–2935. https://doi.org/10.1016/S0140-6736(16)30067-8

Stewart, I. D. (2011). A systematic review and scientific critique of methodology in modern urban heat island literature. International Journal of Climatology, 31(2), 200–217. https://doi.org/10.1002/joc.2141

Stewart, I. D., Krayenhoff, E. S., Voogt, J. A., Lachapelle, J. A., Allen, M. A., & Broadbent, A. M. (2021). Time Evolution of the Surface Urban Heat Island. Earth's Future, 9(e2021EF002178), 1–33. https://doi.org/10.1029/2021EF002178

Stewart, I. D., & Oke, T. R. (2012). Local climate zones for urban temperature studies. Bulletin of the American Meteorological Society, 93(12), 1879–1900. Scopus. https://doi.org/10.1175/BAMS-D-11-00019.1

Stewart, I. D., Oke, T. R., & Krayenhoff, E. S. (2014). Evaluation of the 'local climate zone' scheme using temperature observations and model simulations. International Journal of Climatology, 34(4), 1062–1080. https://doi.org/10.1002/joc.3746

Stewart, I., & Oke, T. (2009). Classifying urban climate field sites by "local climate zones": The case of Nagano, Japan. 1–5. http://www.ide.titech.ac.jp/~icuc7/extended_abstracts/pdf/385055-1-090515165722-002.pdf

Stewart, I., & Oke, T. (2011). Local Climate Zones: Origins, development, and application to urban heat island studies. Annual Meeting of the American Association of Geographers.

Stiglitz, J. E. (2017). The overselling of globalization. Business Economics, 52(3), 129–137. https://doi.org/10.1057/s11369-017-0047-z

Stokes, E. C., & Seto, K. C. (2016). Climate change and urban land systems: Bridging the gaps between urbanism and land science. Journal of Land Use Science, 11(6), 698–708. https://doi.org/10.1080/1747423X.2016.1241316

Stokes, E. C., & Seto, K. C. (2019a). Characterizing and measuring urban landscapes for sustainability. Environmental Research Letters, 14(045002), 1–14. https://doi.org/10.1088/1748-9326/aafab8

Stokes, E. C., & Seto, K. C. (2019b). Characterizing urban infrastructural transitions for the Sustainable Development Goals using multi-temporal land, population, and nighttime light data. Remote Sensing of Environment, 234(111430), 1–11. https://doi.org/10.1016/j.rse.2019.111430

Stone, B., Hess, J. J., & Frumkin, H. (2010). Urban form and extreme heat events: Are sprawling cities more vulnerable to climate change than compact cities? Environmental Health Perspectives, 118(10), 1425–1428. https://doi.org/10.1289/ehp.0901879

Storper, M., & Scott, A. J. (2015). Current debates in urban theory: A critical assessment. Urban Studies, 53(6), 1114–1136. https://doi.org/10.1177/0042098016634002

Streiffer, R. (2007). At the Edge of Humanity: Human Stem Cells, Chimeras, and Moral Status. Journal of Philosophical Research, 32(9999), 63–83. https://doi.org/10.5840/jpr_2007_7

Streutker, D. R. (2002). A remote sensing study of the urban heat island of Houston, Texas. International Journal of Remote Sensing, 23(13), 2595–2608. https://doi.org/10.1080/01431160110115023

Streutker, D. R. (2003). Satellite-measured growth of the urban heat island of Houston, Texas. Remote Sensing of Environment, 85(3), 282–289. https://doi.org/10.1016/S0034-4257(03)00007-5

Strubelt, W., Dosch, F., & Meinel, G. (2021). Die Gestalt des Raumes Landschaften Deutschlands als Abbilder der Gesellschaft. Wasmuth & Zohlen Verlag.

Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., Baraldi, A., & Blaschke, T. (2020). Big Earth data: Disruptive changes in Earth observation data management and analysis? International Journal of Digital Earth, 13(7), 832–850. https://doi.org/10.1080/17538947.2019.1585976

Sukhanov, S., Tankoyeu, I., Louradour, J., Heremans, R., Trofimova, D., & Debes, C. (2017). Multilevel ensembling for local climate zones classification. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1201–1204. https://doi.org/10.1109/IGARSS.2017.8127173

Sun, L., Chen, J., Li, Q., & Huang, D. (2020). Dramatic uneven urbanization of large cities throughout the world in recent decades. Nature Communications, 11(5366), 1–9. https://doi.org/10.1038/s41467-020-19158-1

Sun, Y., Fan, H., Li, M., & Zipf, A. (2015). Identifying the city center using human travel flows generated from location-based social networking data. Environment and Planning B: Planning and Design, 43(3), 480–498. https://doi.org/10.1177/0265813515617642

Sun, Z., Du, W., Jiang, H., Weng, Q., Guo, H., Han, Y., Xing, Q., & Ma, Y. (2022). Global 10-m impervious surface area mapping: A big earth data based extraction and updating approach. International Journal of Applied Earth Observation and Geoinformation, 109(102800), 1–16. https://doi.org/10.1016/j.jag.2022.102800

Sutton, P. C., & Costanza, R. (2002). Global estimates of market and non-market values derived from nighttime satellite imagery, land cover, and ecosystem service valuation. Ecological Economics, 41(3), 509–527. https://doi.org/10.1016/S0921-8009(02)00097-6

Sutton, P., Roberts, D., Elvidge, C., & Baugh, K. (2001). Census from Heaven: An estimate of the global human population using night-time satellite imagery. International Journal of Remote Sensing, 22(16), 3061–3076. Scopus. https://doi.org/10.1080/01431160010007015

Sýkora, L., & Bouzarovski, S. (2012). Multiple Transformations: Conceptualising the Post-communist Urban Transition. Urban Studies, 49(1), 43–60. https://doi.org/10.1177/0042098010397402

Sýkora, L., & Stanilov, K. (2014). The Challenge of Postsocialist Suburbanization. In K. Stanilov & L. Sýkora (Eds.), Confronting Suburbanization (pp. 1–32). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118295861.ch1

T

Taalab, K., Cheng, T., & Zhang, Y. (2018). Mapping landslide susceptibility and types using Random Forest. Big Earth Data, 2(2), 159–178. https://doi.org/10.1080/20964471.2018.1472392

Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). Environmental DNA: For Biodiversity Research and Monitoring. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001

Tai, M. C.-T. (2013). Western or Eastern principles in globalized bioethics? An Asian perspective view. Tzu Chi Medical Journal, 25(1), 64–67. https://doi.org/10.1016/j.tcmj.2012.05.004

Talukder, S., Capon, A., Nath, D., Kolb, A., Jahan, S., & Boufford, J. (2015). Urban health in the post-2015 agenda. The Lancet, 385(9970), 769. https://doi.org/10.1016/S0140-6736(15)60428-7

Tanikawa, H., & Hashimoto, S. (2009). Urban stock over time: Spatial material stock analysis using 4d-GIS. Building Research & Information, 37(5–6), 483–502. https://doi.org/10.1080/09613210903169394

Tanikawa, H., Fishman, T., Okuoka, K., & Sugimoto, K. (2015). The Weight of Society Over Time and Space: A Comprehensive Account of the Construction Material Stock of Japan, 1945–2010. Journal of Industrial Ecology, 19(5), 778–791. https://doi.org/10.1111/jiec.12284

Tarboton, D. G., Bras, R. L., & Rodriguez-Iturbe, I. (1991). On the extraction of channel networks from digital elevation data. Hydrological Processes, 5(1), 81–100. https://doi.org/10.1002/hyp.3360050107

Tatem, A. J. (2017). WorldPop, open data for spatial demography. Scientific Data, 4(170004), 1–4. https://doi.org/10.1038/sdata.2017.4

Taubenböck, H. (2019). Remote Sensing for the Analysis of Global Urbanization. DLR Research Report 2019-10. Habilitation Thesis, University Würzburg.

Taubenböck, H. (2021). How We Live and What That Means—A Character Study with Data from Space. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 1190–1193. https://doi.org/10.1109/IGARSS47720.2021.9553084

Taubenböck, H., Bauer, P., Geiß, C., & Wurm, M. (2017). Mega-regions in China. 2017 Joint Urban Remote Sensing Event (JURSE), 1–4. https://doi.org/10.1109/JURSE.2017.7924603

Taubenböck, H., Debray, H., Qiu, C., Schmitt, M., Wang, Y., & Zhu, X. X. (2020). Seven city types representing morphologic configurations of cities across the globe. Cities, 105(102814), 1–15. https://doi.org/10.1016/j.cities.2020.102814

Taubenböck, H., Droin, A., Standfuß, I., Dosch, F., Sander, N., Milbert, A., Eichfuss, S., & Wurm, M. (2022). To be, or not to be 'urban'? A multi-modal method for the differentiated measurement of the degree of urbanization. Computers, Environment and Urban Systems, 95(101830), 1–12. https://doi.org/10.1016/j.compenvurbsys.2022.101830

Taubenböck, H., Esch, T., Felbier, A., Roth, A., & Dech, S. (2011). Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data. In IEEE Geoscience and Remote Sensing Letters (Vol. 8, Issue 2, pp. 278–282). https://doi.org/10.1109/LGRS.2010.2069083

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. Remote Sensing of Environment, 117, 162–176. https://doi.org/10.1016/j.rse.2011.09.015

Taubenböck, H., Ferstl, J., & Dech, S. (2017). Regions Set in Stone—Delimiting and Categorizing Regions in Europe by Settlement Patterns Derived from EO-Data. ISPRS International Journal of Geo-Information, 6(55), 1–27. https://doi.org/10.3390/ijgi6020055

Taubenböck, H., Gerten, C., Rusche, K., Siedentop, S., & Wurm, M. (2019). Patterns of Eastern European urbanisation in the mirror of Western trends – Convergent, unique or hybrid? Environment and Planning B: Urban Analytics and City Science, 46(7), 1206–1225. https://doi.org/10.1177/2399808319846902

Taubenböck, H., Kehrer, J., & Wurm, M. (2015). Zu Stein gewordene Philosophien – die Morphologie geplanter Wohnviertel. In H. Taubenböck, M. Wurm, T. Esch, & S. Dech (Eds.), Globale Urbanisierung: Perspektive aus dem All (pp. 135–147). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-44841-0_15

Taubenböck, H., Klotz, M., Wurm, M., Schmieder, J., Wagner, B., Wooster, M., Esch, T., & Dech, S. (2013). Delineation of Central Business Districts in mega city regions using remotely sensed data. Remote Sensing of Environment, 136, 386–401. https://doi.org/10.1016/j.rse.2013.05.019

Taubenböck, H., & Kraff, N. J. (2014). The physical face of slums: A structural comparison of slums in Mumbai, India, based on remotely sensed data. Journal of Housing and the Built Environment, 29(1), 15–38. https://doi.org/10.1007/s10901-013-9333-x

Taubenböck, H., Kraff, N. J., & Wurm, M. (2018). The morphology of the Arrival City—A global categorization based on literature surveys and remotely sensed data. Applied Geography, 92, 150–167. https://doi.org/10.1016/j.apgeog.2018.02.002

Taubenböck, H., Mast, J., Geiß, C., Wurm, M., Esch, T., & Seto, K. C. (2024). Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios. Cities, 151(105117), 1–31. https://doi.org/10.1016/j.cities.2024.105117

Taubenböck, H., Mast, J., Lemoine Rodríguez, R., Debray, H., Wurm, M., & Geiß, C. (2025). Was global urbanization from 1985 to 2015 efficient in terms of land consumption? Habitat International, 160(103397), 1–15. https://doi.org/10.1016/j.habitatint.2025.103397

Taubenböck, H., Murawski, M., & Wurm, M. (2018). Large Housing Estates – Analysing the Morphologic Similarities and Differences of a Specific Town Planning Concept. Proceedings of 23rd International Conference on Urban Planning, Regional Development and Information, 275–284. http://repository.corp.at/id/eprint/433

Taubenböck, H., Otto, C., Gülzau, F., & Mau, S. (2023). Border regions across the globe: Analyzing border typologies, economic and political disparities, and development dynamics. Applied Geography, 151(102866), 1–15. https://doi.org/10.1016/j.apgeog.2022.102866

Taubenböck, H., Post, J., Roth, A., Zosseder, K., Strunz, G., & Dech, S. (2008). A conceptual vulnerability and risk framework as outline to identify capabilities of remote sensing. Natural Hazards and Earth System Sciences, 8(3), 409–420. https://doi.org/10.5194/nhess-8-409-2008

Taubenböck, H., Reiter, M., Dosch, F., Leichtle, T., Weigand, M., & Wurm, M. (2021). Which city is the greenest? A multi-dimensional deconstruction of city rankings. Computers, Environment and Urban Systems, 89, 101687. https://doi.org/10.1016/j.compenvurbsys.2021.101687

Taubenböck, H., Staab, J., Zhu, X. X., Geiß, C., Dech, S., & Wurm, M. (2018). Are the Poor Digitally Left Behind? Indications of Urban Divides Based on Remote Sensing and Twitter Data. ISPRS International Journal of Geo-Information, 7(304), 1–18. https://doi.org/10.3390/ijgi7080304

Taubenböck, H., Standfuß, I., Klotz, M., & Wurm, M. (2016). The Physical Density of the City—Deconstruction of the Delusive Density Measure with Evidence from Two European Megacities. ISPRS International Journal of Geo-Information, 5(206), 1–24. https://doi.org/10.3390/ijgi5110206

Taubenböck, H., Standfuß, I., Wurm, M., Krehl, A., & Siedentop, S. (2017). Measuring morphological polycentricity—A comparative analysis of urban mass concentrations using remote sensing data. Computers, Environment and Urban Systems, 64, 42–56. https://doi.org/10.1016/j.compenvurbsys.2017.01.005

Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J., & Dech, S. (2019). A new ranking of the world's largest cities—Do administrative units obscure morphological realities? Remote Sensing of Environment, 232(111353), 1–14. https://doi.org/10.1016/j.rse.2019.111353

Taubenböck, H., & Wiesner, M. (2015). The spatial network of megaregions—Types of connectivity between cities based on settlement patterns derived from EO-data. Computers, Environment and Urban Systems, 54, 165–180. https://doi.org/10.1016/j.compenvurbsys.2015.07.001

Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., & Dech, S. (2014). New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. Applied Geography, 47, 137–153. https://doi.org/10.1016/j.apgeog.2013.12.002

Taubenböck, H., & Wurm, M. (2015). Ich weiß, dass ich nichts weiß – Bevölkerungsschätzung in der Megacity Mumbai. In H. Taubenböck, M. Wurm, T. Esch, & S. Dech (Eds.), Globale Urbanisierung: Perspektive aus dem All (pp. 171–178). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-44841-0_18

Taubenböck, H., Wurm, M., Geiß, C., Dech, S., & Siedentop, S. (2019). Urbanization between compactness and dispersion: Designing a spatial model for measuring 2D binary settlement landscape configurations. International Journal of Digital Earth, 12(6), 679–698. https://doi.org/10.1080/17538947.2018.1474957

Taylor, P. J., Catalano, G., & Walker, D. R. F. (2002). Measurement of the World City Network. Urban Studies, 39(13), 2367–2376. https://doi.org/10.1080/00420980220080011

Taylor, P. J., Hoyler, M., & Verbruggen, R. (2010). External Urban Relational Process: Introducing Central Flow Theory to Complement Central Place Theory. Urban Studies, 47(13), 2803–2818. https://doi.org/10.1177/0042098010377367

Taylor, P., & Pickles, J. (1995). Geographic information systems and geography. In Ground Truth: The Social Implications of Geographic Information Systems (pp. 51–56). Guilford Press.

Taylor, Z. (2013). Rethinking planning culture: A new institutionalist approach. Town Planning Review, 84(6), 683–702. https://doi.org/10.3828/tpr.2013.36

Thai, H. M. H., Stevens, Q., & Rogers, J. (2019). The influence of organic urban morphologies on opportunities for home-based businesses within inner-city districts in Hanoi, Vietnam. Journal of Urban Design, 24(6), 926–946. https://doi.org/10.1080/13574809.2018.1554995

Thai, H., Stevens, Q., & Rogers, J. (2020). The influences of formally planned urban morphology on home-based economic opportunities and the economically-driven self-organisation of urban form. Proceedings of the 12th Space Syntax Symposium.

The Economics of Ecosystems and Biodiversity (TEEB). (2011). TEEB manual for cities: Ecosystem services in urban management (TEEB Manual for CITIES: Ecosystem Services in Urban Management). www.teebweb.org

The Economist Intelligence Unit. (2019). Global liveability ranking 2019. www.eiu.com

The World Bank. (2021). GDP per capita, PPP (constant 2017 international $) [Dataset]. https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD?view=chart

Therien, J.-P. (1999). Beyond the North-South divide: The two tales of world poverty. Third World Quarterly, 20(4), 723–742. https://doi.org/10.1080/01436599913523

Thiebes, B., Bell, R., Glade, T., Jäger, S., Mayer, J., Anderson, M., & Holcombe, L. (2014). Integration of a limit-equilibrium model into a landslide early warning system. Landslides, 11(5), 859–875. https://doi.org/10.1007/s10346-013-0416-2

Thomas, C. J., & Bromley, R. D. F. (2003). Retail revitalization and small town centres: The contribution of shopping linkages. Applied Geography, 23(1), 47–71. https://doi.org/10.1016/S0143-6228(02)00068-1

Thomas, I., Frankhauser, P., & Badariotti, D. (2012). Comparing the fractality of European urban neighbourhoods: Do national contexts matter? Journal of Geographical Systems, 14(2), 189–208. https://doi.org/10.1007/s10109-010-0142-4

Thomas, R., Schmidt, S., & Siedentop, S. (2022). Toward comparative polycentricity scores: Assessing variations in regional delineation and subcenter identification. Environment and Planning B: Urban Analytics and City Science, 49(6), 1597–1611. https://doi.org/10.1177/23998083211065765

Thompson, J., Stevenson, M., Wijnands, J. S., Nice, K. A., Aschwanden, G. D., Silver, J., Nieuwenhuijsen, M., Rayner, P., Schofield, R., Hariharan, R., & Morrison, C. N. (2020). A global analysis of urban design types and road transport injury: An image processing study. The Lancet Planetary Health, 4(1), e32–e42. https://doi.org/10.1016/S2542-5196(19)30263-3

Thrower, NormanJ. W. (1970). Annals map supplement number twelve: Land use in the southwestern United States—From Gemini and Apollo-imagery. Annals of the Association of American Geographers, 60(1), 208–209. https://doi.org/10.1111/j.1467-8306.1970.tb00714.x

Thuro, K., Singer, J., Festl, J., Wunderlich, T., Wasmeier, P., Reith, C., Heunecke, O., Glabsch, J., & Schuhbäck, S. (2010). New landslide monitoring techniques – developments and experiences of the alpEWAS project. 4(2), 69–90. https://doi.org/10.1515/jag.2010.008

Thuro, K., Singer, J., Menschik, B., Breuninger, T., & Gamperl, M. (2020). Development of an early warning system for landslides in the tropical Andes (Medellín; Colombia). Geomechanics and Tunnelling, 13(1), 103–115. https://doi.org/10.1002/geot.201900071

Thuro, K., Wunderlich, Th., Heunecke, O., Singer, J., Wasmeier, P., Schuhbäck, St., Festl, J., Reith, Ch., & Glabsch, J. (2014). Low Cost 3D Early Warning System for Alpine Instable Slopes: The Aggenalm Landslide Monitoring System. In F. Wenzel & J. Zschau (Eds.), Early Warning for Geological Disasters: Scientific Methods and Current Practice (pp. 289–306). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12233-0_15

Thurstain-Goodwin, M., & Unwin, D. (2000). Defining and Delineating the Central Areas of Towns for Statistical Monitoring Using Continuous Surface Representations. Transactions in GIS, 4(4), 305–317. https://doi.org/10.1111/1467-9671.00058

Tian, Y., Tsendbazar, N.-E., van Leeuwen, E., Fensholt, R., & Herold, M. (2022). A global analysis of multifaceted urbanization patterns using Earth Observation data from 1975 to 2015. Landscape and Urban Planning, 219(104316), 1–13. https://doi.org/10.1016/j.landurbplan.2021.104316

Tiangco, M., Lagmay, A. M. F., & Argete, J. (2008). ASTER-based study of the night-time urban heat island effect in Metro Manila. International Journal of Remote Sensing, 29(10), 2799–2818. https://doi.org/10.1080/01431160701408360

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2), 411–423. https://doi.org/10.1111/1467-9868.00293

Tombolini, I., Munafò, M., & Salvati, L. (2016). Soil sealing footprint as an indicator of dispersed urban growth: A multivariate statistics approach. Urban Research & Practice, 9(1), 1–15. https://doi.org/10.1080/17535069.2015.1037340

Tong, S., Prior, J., McGregor, G., Shi, X., & Kinney, P. (2021). Urban heat: An increasing threat to global health. BMJ, 375(2467), 1–5. https://doi.org/10.1136/bmj.n2467

Tong, X., Hong, Z., Liu, S., Zhang, X., Xie, H., Li, Z., Yang, S., Wang, W., & Bao, F. (2012). Building-damage detection using pre- and post-seismic high-resolution satellite stereo imagery: A case study of the May 2008 Wenchuan earthquake. ISPRS Journal of Photogrammetry and Remote Sensing, 68, 13–27. https://doi.org/10.1016/j.isprsjprs.2011.12.004

Tong, X.-Y., Xia, G.-S., Lu, Q., Shen, H., Li, S., You, S., & Zhang, L. (2020). Land-cover classification with high-resolution remote sensing images using transferable deep models. Remote Sensing of Environment, 237(111322), 1–20. https://doi.org/10.1016/j.rse.2019.111322

Tonkiss, F. (2013). Cities by design: The social life of urban form. Polity Press.

Tooke, T. R., Coops, N. C., Goodwin, N. R., & Voogt, J. A. (2009). Extracting urban vegetation characteristics using spectral mixture analysis and decision tree classifications. Remote Sensing of Environment, 113(2), 398–407. https://doi.org/10.1016/j.rse.2008.10.005

Törnberg, P. (n.d.). The urban attention economy: How our cities are rebuilt for social media. Inside Media: A Journal of Cities and Culture, 8(1). https://www.media polisjournal.com/2023/03/urban-attention-economy/

Törnberg, P., & Uitermark, J. (2022). Urban Mediatization and Planetary Gentrification: The Rise and Fall of a Favela across Media Platforms. City & Community, 21(4), 340–361. https://doi.org/10.1177/15356841211068521

Tosics, I. (2005). City development in Central and Eastern Europe since 1990: The impacts of internal forces. In F. E. I. Hamilton, K. Dimitrovska Andrews, & N. Pichler-Milanovic (Eds.), Transformation of Cities in Central and Eastern Europe: Towards Globalization (pp. 44–78). United Nations University Press.

Tratalos, J., Fuller, R. A., Warren, P. H., Davies, R. G., & Gaston, K. J. (2007). Urban form, biodiversity potential and ecosystem services. Landscape and Urban Planning, 83(4), 308–317. https://doi.org/10.1016/j.landurbplan.2007.05.003

Trip, J. J. (2003). The Corridor as a transport network. In W. Zonneveld & J. J. Trip (Eds.), Megacorridors in North West Europe. Investigating a new transnational planning concept (pp. 17–23). Delft University Press.

Troped, P. J., Wilson, J. S., Matthews, C. E., Cromley, E. K., & Melly, S. J. (2010). The Built Environment and Location-Based Physical Activity. American Journal of Preventive Medicine, 38(4), 429–438. https://doi.org/10.1016/j.amepre.2009.12.032

Trouwborst, A., Fleurke, F., & Dubrulle, J. (2016). Border Fences and their Impacts on Large Carnivores, Large Herbivores and Biodiversity: An International Wildlife Law Perspective. Review of European, Comparative & International Environmental Law, 25(3), 291–306. https://doi.org/10.1111/reel.12169

Tsai, Y.-H. (2005). Quantifying Urban Form: Compactness versus 'Sprawl'. Urban Studies, 42(1), 141–161. https://doi.org/10.1080/0042098042000309748

Tuholske, C., Caylor, K., Funk, C., Verdin, A., Sweeney, S., Grace, K., Peterson, P., & Evans, T. (2021). Global urban population exposure to extreme heat. Proceedings of the National Academy of Sciences, 118(41), 1–9. https://doi.org/10.1073/pnas.2024792118

Tuia, D., Roscher, R., Wegner, J. D., Jacobs, N., Zhu, X., & Camps-Valls, G. (2021). Toward a Collective Agenda on AI for Earth Science Data Analysis. IEEE Geoscience and Remote Sensing Magazine, 9(2), 88–104. https://doi.org/10.1109/MGRS.2020.3043504

Tuia, D., Schindler, K., Demir, B., Zhu, X. X., Kochupillai, M., Džeroski, D., van Rijn, J. N., Hoos, H. H., Del Frate, F., Datcu, M., Markl, V., Le Saux, B., Schneider, R., & Camps-Valls, G. (2024). Artificial Intelligence to Advance Earth Observation: A review of models, recent trends, and pathways forward. IEEE Geoscience and Remote Sensing Magazine, 2–24. https://doi.org/10.1109/MGRS.2024.3425961

Tuia, D., Volpi, M., Trolliet, M., & Camps-Valls, G. (2014). Semisupervised Manifold Alignment of Multimodal Remote Sensing Images. IEEE Transactions on Geoscience and Remote Sensing, 52(12), 7708–7720. https://doi.org/10.1109/tgrs.2014.2317499

Turner, J. C. (1967). Barriers and Channels for Housing Development in Modernizing Countries. Journal of the American Institute of Planners, 33(3), 167–181. https://doi.org/10.1080/01944366708977912

Turner, V. K., Rogers, M. L., Zhang, Y., Middel, A., Schneider, F. A., Ocón, J. P., Seeley, M., &, Dialesandro, J. (2022). More than surface temperature: Mitigating thermal exposure in hyper-local land system. Journal of Land Use Science, 17(1), 79–99. https://doi.org/10.1080/1747423X.2021.2015003

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A. K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L. P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., & Woodcock, C. (2015). Free and open-access satellite data are key to biodiversity conservation. Biological Conservation, 182, 173–176. https://doi.org/10.1016/j.biocon.2014.11.048

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., & Steininger, M. (2003). Remote sensing for biodiversity science and conservation. Trends in Ecology & Evolution, 18(6), 306–314. https://doi.org/10.1016/S0169-5347(03)00070-3

Twitter. (2022). Twitter API for academic research | products [Computer software]. https://developer. twitter.com/en/products/twitter-api/academic-research

U

Uchimura, T., Towhata, I., Wang, L., Nishie, S., Yamaguchi, H., Seko, I., & Qiao, J. (2015). Precaution and early warning of surface failure of slopes using tilt sensors. Special Issue on the Six International Symposium on Deformation Characteristics of Geomaterials IS-Buenos Aires2015, 55(5), 1086–1099. https://doi.org/10.1016/j.sandf.2015.09.010

Unger, J., Lelovics, E., & Gál, T. (2014). Local Climate Zone mapping using GIS methods in Szeged. Hungarian Geographical Bulletin, 63(1), 29–41. https://doi.org/10.15201/hungeobull.63.1.3

UN-Habitat. (2013). Prosperity of Cities: State of the World's Cities 2012/2013. Routledge for and on behalf of the United Nations Human Settlements Programme. https://unhabitat.org/prosperity-of-cities-state-of-the-worlds-cities-20122013

UN-Habitat, & ECA. (2008). The state of African cities: A framework for addressing urban challenges in Africa (p. 206). Nairobi: United Nations Human Settlement Programme. https://hdl.handle.net/10855/15223

United Nations. (1966). International Covenant on Civil and Political Rights (No. General Assembly resolution 2200A (XXI)). https:// www.ohchr.org/en/professionalinterest/pages/ccpr.aspx

United Nations. (1973). Urban Land Policies and Land-use Control Measures: Volume VI. Northern America (p. 174). Department of Economic and Social Affairs.

United Nations. (1999). Declaration on the right and responsibility of individuals, groups and organs of society to promote and protect universally recognized human rights and fundamental freedoms (No. A/RES/53/144). https://digitallibrary.un.org/record/265855?v=pdf#files

United Nations. (2015a). Resolution adopted by the General Assembly on 25 September 2015. United Nations. https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf

United Nations. (2015b). The Paris Agreement. https://unfccc.int/sites/default/files/english_paris_agreement.pdf

United Nations. (2016a). The Millennium Development Goals Report 2015. United Nations. https://www.un-ilibrary.org/content/books/9789210574662

United Nations. (2016b). The World's Cities in 2016. United Nations. https://doi.org/10.18356/8519891f-en

United Nations. (2018). The World's Cities in 2018. United Nations. https://doi.org/10.18356/c93f4dc6-en

United Nations. (2025). Big Data for Sustainable Development. https://www.un.org/en/global-issues/big-data-for-sustainable-development

United Nations Department of Economic and Social Affairs. (2016). United Nations Demographic Yearbook 2015. United Nations. https://doi.org/10.18356/6a66c8df-en-fr

United Nations, Department of Economic and Social Affairs, Population Division. (2014). World Urbanization Prospects: The 2014 Revision (No. ST/ESA/SER.A/366). https://digitallibrary.un.org/record/826634/files/WUP2014-Report.pdf

United Nations, Department of Economic and Social Affairs, Population Division. (2017). World Urbanization Prospects: The 2017 Revision (No. ST/ESA/SER.A/399; Volume I: Comprehensive Tables). Department of Economic and Social Affairs, Population Division. https://population.un.org/wpp/assets/Files/WPP2017_Volume-I_Comprehensive-Tables.pdf

United Nations, Department of Economic and Social Affairs, Population Division. (2018). World Urbanization Prospects: The 2018 Revision (No. ST/ESA/SER.A/420). United Nations. https://population.un.org/wup/assets/WUP2018-Report.pdf

United Nations, Department of Economic and Social Affairs, Population Division. (2019). World Urbanization Prospects 2019: Highlights (No. ST/ESA/SER.A/423). United Nations. https://population.un.org/wpp/assets/Files/WPP2019_Highlights.pdf

United Nations, Department of Economic and Social Affairs, Population Division. (2020). Demographic Yearbook, 70th issue. United Nations. https://unstats.un.org/unsd/demographic-social/products/dyb/dybsets/2019.pdf

United Nations Habitat. (2016). Goal 11. Make cities and human settlements inclusive, safe, resilient and sustainable. https://unstats.un.org/sdgs/metadata/?Text=&Goal=&Target=11.3

United Nations Human Settlements Programme (UN-HABITAT). (2008). State of the world's cities 2010/2011: Bridging The Urban Divide. London & Sterling, VA: Earthscan. https://sustainabledevelopment.un.org/ content/documents/745habitat.pdf

United Nations Human Settlements Programme (UN-HABITAT). (2013). State of the world's cities 2012/2013: Prosperity of cities. https://sustainabledevelopment.un.org/ content/documents/745habitat.pdf

United Nations Statistical Office. (1982). Standard country or area codes for statistical use (M49) [Dataset]. https://unstats.un.org/unsd/methodology/m49/

Universidad Nacional de Colombia. (2009). Amenaza, Vulnerabilidad y riesgo por movimientos en masa, avenidas torrenciales e inundaciones en el valle de aburrá, Formulación de propuestas de gestión, informe final, convenio 4800002397 de 2007, File Rep.

Urry, J. (1999). Sociology Beyond Societies: Mobilities for the Twenty-First Century. Routledge.

USGS. (2023). Landsat Data Access. https://www.usgs. gov/landsat-missions/landsat-data-access

USGS. (2025). National Land Cover Database (NLCD) [Dataset]. https://www.usgs.gov/node/279743

Uth, P. (2019). Development of a globally suitable estimation approach for the fully automated determination of city centres and borders applied to Germany and the USA. In German: Entwicklung eines global anwendbaren Ausgleichungsverfahrens zur vollautomatischen Ermittlung von Stadtzentren und Stadtgrenzen am Beispiel von Deutschland und den USA [Master's thesis, Hochschule München].

V

Valks, P., Pinardi, G., Richter, A., Lambert, J.-C., Hao, N., Loyola, D., Van Roozendael, M., & Emmadi, S. (2011). Operational total and tropospheric NO_2 column retrieval for GOME-2. Atmospheric Measurement Techniques, 4(7), 1491–1514. https://doi.org/10.5194/amt-4-1491-2011

Vallet, E. (2016). Borders, Fences and Walls: State of Insecurity? Routledge.

van der A, R. J., Eskes, H. J., Boersma, K. F., van Noije, T. P. C., Van Roozendael, M., De Smedt, I., Peters, D. H. M. U., & Meijer, E. W. (2008). Trends, seasonal variability and dominant NOx source derived from a ten year record of NO2 measured from space. Journal of Geophysical Research: Atmospheres, 113(D04302), 1–12. https://doi.org/10.1029/2007JD009021

van Donkelaar, A., Randall, M. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., & Villeneuve, P. J. (2010). Global Estimates of Ambient Fine Particulate Matter Concentrations from Satellite-Based Aerosol Optical Depth: Development and Application. Environmental Health Perspectives, 118(6), 847–855. https://doi.org/10.1289/ehp.0901623

Van Houtum, H., & Legendijk, A. (2001). Contextualising regional identity and imagination on the construction of polycentric urban regions: The cases of the Ruhr area and the Basque country. Urban Studies, 38(4), 747–767. Scopus. https://doi.org/10.1080/00420980120035321

van Meeteren, M., & Poorthuis, A. (2018). Christaller and "big data": Recalibrating central place theory via the geoweb. Urban Geography, 39(1), 122–148. https://doi.org/10.1080/02723638.2017.1298017

van Vliet, J. (2019). Direct and indirect loss of natural area from urban expansion. Nature Sustainability, 2(8), 755–763. https://doi.org/10.1038/s41893-019-0340-0

Vance, J. E. (1964). Geography and urban evolution in the San Francisco Bay Area.

Vardoulakis, S., & Kinney, P. (2019). Grand Challenges in Sustainable Cities and Health. Frontiers in Sustainable Cities, 1(7), 1–5. https://doi.org/10.3389/frsc.2019.00007

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A., & Watson, J. E. M. (2016). Global terrestrial Human Footprint maps for 1993 and 2009. Scientific Data, 3(1), 1–20. https://doi.org/10.1038/sdata.2016.67

Verbesselt, J., Zeileis, A., & Herold, M. (2012). Near real-time disturbance detection using satellite image time series. Remote Sensing of Environment, 123, 98–108. https://doi.org/10.1016/j.rse.2012.02.022

Verhezen, P. (2008). The (Ir)relevance of Integrity in Organizations. Public Integrity, 10(2), 133–149. https://doi.org/10.2753/PIN1099-9922100203

Vernon, V., & Zimmermann, K. F. (2019). Walls and Fences: A Journey Through History and Economics (Working Paper No. 330). GLO Discussion Paper. https://www.econstor.eu/handle/10419/193640

Vincent, J. (2022). What a machine learning tool that turns Obama white can (and can't) tell us about AI bias. The Verge. https://www.theverge.com/21298762/face -depi xelizer-ai-machine-learning-tool-pulse-stylegan -obama-bias

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human Domination of Earth's Ecosystems. Science, 277(5325), 494–499. https://doi.org/10.1126/science.277.5325.494

Vogt, P., Riitters, K., Rambaud, P., d'Annunzio, R., Lindquist, E., & Pekkarinen, A. (2022). GuidosToolbox Workbench: Spatial analysis of raster maps for ecological applications. Ecography, 2022(e05864), 1–7. https://doi.org/10.1111/ecog.05864

Voigt, C., Lelieveld, J., Schlager, H., Schneider, J., Curtius, J., Meerkötter, R., Sauer, D., Bugliaro, L., Bohn, B., Crowley, J. N., Erbertseder, T., Groß, S., Hahn, V., Li, Q., Mertens, M., Pöhlker, M. L., Pozzer, A., Schumann, U., Tomsche, L., … Rapp, M. (2022). Cleaner Skies during the COVID-19 Lockdown. Bulletin of the American Meteorological Society, 103(8), E1796–E1827. https://doi.org/10.1175/BAMS-D-21-0012.1

Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., & H. Mehl. (2007). Satellite Image Analysis for Disaster and Crisis-Management Support. IEEE Transactions on Geoscience and Remote Sensing, 45(6), 1520–1528. https://doi.org/10.1109/TGRS.2007.895830

Vojinovic, Z., Seyoum, S. D., Mwalwaka, J. M., & Price, R. K. (2011). Effects of model schematisation, geometry and parameter values on urban flood modelling. Water Science and Technology, 63(3), 462–467. https://doi.org/10.2166/wst.2011.244

Vojtek, M., Vojteková, J., Costache, R., Pham, Q. B., Lee, S., Arshad, Arfan, Sahoo, Satiprasad, Linh, N. T. T., & Anh, D. T. (2021). Comparison of multi-criteria-analytical hierarchy process and machine learning-boosted tree models for regional flood susceptibility mapping: A case study from Slovakia. Geomatics, Natural Hazards and Risk, 12(1), 1153–1180. https://doi.org/10.1080/19475705.2021.1912835

von Borries, F. (2024). Architektur im Anthropozän. Suhrkamp.

von Luxburg, U. (2007). A Tutorial on Spectral Clustering. Statistics and Computing, 17(4), 395–416.

Von Szombathely, M., Albrecht, M., Antanaskovic, D., Augustin, J., Augustin, M., Bechtel, B., Bürk, T., Fischereit, J., Grawe, D., Hoffmann, P., Kaveckis, G., Krefis, A. C., Oßenbrügge, J., Scheffran, J., & Schlünzen, K. H. (2017). A Conceptual Modeling Approach to Health-Related Urban Well-Being. Urban Science, 1(17), 1–18. https://doi.org/10.3390/urbansci1020017

Voogt, J. A., & Oke, T. R. (2003). Thermal remote sensing of urban climates. Remote Sensing of Environment, 86(3), 370–384. https://doi.org/10.1016/S0034-4257(03)00079-8

W

Wakode, H. B., Baier, K., Jha, R., & Azzam, R. (2018). Impact of urbanization on groundwater recharge and urban water balance for the city of Hyderabad, India. International Soil and Water Conservation Research, 6(1), 51–62. https://doi.org/10.1016/j.iswcr.2017.10.003

Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Transactions on Geoscience and Remote Sensing, 34(4), 892–905. https://doi.org/10.1109/36.508406

Wang, C., & Mahadevan, S. (2009). A General Framework for Manifold Alignment. AAAI Fall Symposium: Manifold Learning and Its Applications, 53–58. https://www.researchgate.net/publication/228753529_A_General_Framework_for_Manifold_Alignment

Wang, F., Wang, Z., Yang, H., Zhao, Y., Li, Z., & Wu, J. (2018). Capability of Remotely Sensed Drought Indices for Representing the Spatio–Temporal Variations of the Meteorological Droughts in the Yellow River Basin. Remote Sensing, 10(1834), 1–18. https://doi.org/10.3390/rs10111834

Wang, J., & Christopher, S. A. (2003). Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies. Geophysical Research Letters, 30(21), 1–4. https://doi.org/10.1029/2003GL018174

Wang, J., Lin, Y., Glendinning, A., & Xu, Y. (2018). Land-use changes and land policies evolution in China's urbanization processes. Land Use Policy, 75, 375–387. https://doi.org/10.1016/j.landusepol.2018.04.011

Wang, L., Li, C., Ying, Q., Cheng, X., Wang, X., Li, X., Hu, L., Liang, L., Yu, L., Huang, H., & Gong, P. (2012). China's urban expansion from 1990 to 2010 determined with satellite remote sensing. Chinese Science Bulletin, 57(22), 2802–2812. https://doi.org/10.1007/s11434-012-5235-7

Wang, R., Cai, M., Ren, C., Bechtel, B., Xu, Y., & Ng, E. (2019). Detecting multi-temporal land cover change and land surface temperature in Pearl River Delta by adopting local climate zone. Urban Climate, 28(100455), 1–18. https://doi.org/10.1016/j.uclim.2019.100455

Wang, S., Cebula, R., Liu, X., & Foley, M. (2020). Housing prices and urban land use efficiency. Applied Economics Letters, 28, 1–4. https://doi.org/10.1080/13504851.2020.1803476

Wang, S., Liu, X., Zhou, C., Hu, J., & Ou, J. (2017). Examining the impacts of socioeconomic factors, urban form, and transportation networks on CO2 emissions in China's megacities. Applied Energy, 185, 189–200. https://doi.org/10.1016/j.apenergy.2016.10.052

Wang, S., So, E., & Smith, P. (2015). Detecting tents to estimate the displaced populations for post-disaster relief using high resolution satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 36, 87–93. https://doi.org/10.1016/j.jag.2014.11.013

Wang, W., & Stewart, K. (2015). Spatiotemporal and semantic information extraction from Web news reports about natural hazards. Computers, Environment and Urban Systems, 50, 30–40. https://doi.org/10.1016/j.compenvurbsys.2014.11.001

Wang, Y., Huang, C., Feng, Y., Zhao, M., & Gu, J. (2020). Using Earth Observation for Monitoring SDG 11.3.1-Ratio of Land Consumption Rate to Population Growth Rate in Mainland China. Remote Sensing, 12(357), 1–21. https://doi.org/10.3390/rs12030357

Watkins, K. (2021). Houston's 'Father of Environmental Justice' On Where He Sees The Movement Headed. Houston Public Media, Houston, TX, USA. https:// www.houstonpublicmedia.org/articles/news/energy-environment/2021/04/26/396685/what-is-environmental-justice -what-is-the-future/

Wattles, J. (1996). The Golden Rule. Oxford University Press Inc.

Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Belesova, K., Berry, H., Bouley, T., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., Chambers, J., Daly, M., Dasandi, N., Davies, M., Depoux, A., Dominguez-Salas, P., Drummond, P., Ebi, K. L., … Costello, A. (2018). The 2018 report of the Lancet Countdown on health and climate change: Shaping the health of nations for centuries to come. The Lancet, 392(10163), 2479–2514. https://doi.org/10.1016/S0140-6736(18)32594-7

Weatherhead, E. C., Reinsel, G. C., Tiao, G. C., Meng, X.-L., Choi, D., Cheang, W.-K., Keller, T., DeLuisi, J., Wuebbles, D. J., Kerr, J. B., Miller, A. J., Oltmans, S. J., & Frederick, J. E. (1998). Factors affecting the detection of trends: Statistical considerations and applications to environmental data. Journal of Geophysical Research: Atmospheres, 103(D14), 17149–17161. https://doi.org/10.1029/98JD00995

Weber, C. (2001). Urban agglomeration delimitation using remote sensing data. In J.-P. Donnay, M. J. Barnsley, & P. A. Longley (Eds.), Remote Sensing and Urban Analysis (pp. 145–159). Taylor & Francis.

Wei, C., Blaschke, T., Kazakopoulos, P., Taubenböck, H., & Tiede, D. (2017). Is Spatial Resolution Critical in Urbanization Velocity Analysis? Investigations in the Pearl River Delta. Remote Sensing, 9(80), 1–25. https://doi.org/10.3390/rs9010080

Wei, C., Taubenböck, H., & Blaschke, T. (2017). Measuring urban agglomeration using a city-scale dasymetric population map: A study in the Pearl River Delta, China. Habitat International, 59, 32–43. https://doi.org/10.1016/j.habitatint.2016.11.007

Weigand, M., Staab, J., Wurm, M., & Taubenböck, H. (2020). Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data. International Journal of Applied Earth Observation and Geoinformation, 88(102065), 1–9. https://doi.org/10.1016/j.jag.2020.102065

Weigand, M., Wurm, M., Dech, S., & Taubenböck, H. (2019). Remote Sensing in Environmental Justice Research—A Review. ISPRS International Journal of Geo-Information, 8(20), 1–28. https://doi.org/10.3390/ijgi8010020

Weismann, D., Möckel, M., Paeth, H., & Slagman, A. (2023). Modelling variations of emergency attendances using data on community mobility, climate and air pollution. Scientific Reports, 13(20595), 1–13. https://doi.org/10.1038/s41598-023-47857-4

Wellman, B. (2001). Physical Place and Cyberplace: The Rise of Personalized Networking. International Journal of Urban and Regional Research, 25(2), 227–252. https://doi.org/10.1111/1468-2427.00309

Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. ISPRS Journal of Photogrammetry and Remote Sensing, 64(4), 335–344. https://doi.org/10.1016/j.isprsjprs.2009.03.007

Weng, Q. (2012). Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends. Remote Sensing of Urban Environments, 117, 34–49. https://doi.org/10.1016/j.rse.2011.02.030

Weng, Q. (2019). Techniques and Methods in Urban Remote Sensing. John Wiley & Sons.

Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sensing of Environment, 89(4), 467–483. https://doi.org/10.1016/j.rse.2003.11.005

Wentrup, R., Ström, P., & Nakamura, H. R. (2016). Digital oases and digital deserts in Sub-Saharan Africa. Journal of Science and Technology Policy Management, 7(1), 77–100. https://doi.org/10.1108/JSTPM-03-2015-0013

Wentz, E. A., Anderson, S., Fragkias, M., Netzband, M., Mesev, V., Myint, S. W., Quattrochi, D., Rahman, A., & Seto, K. C. (2014). Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing. Remote Sensing, 6(5), 3879–3905. https://doi.org/10.3390/rs6053879

Wentz, E. A., York, A. M., Alberti, M., Conrow, L., Fischer, H., Inostroza, L., Jantz, C., Pickett, S. T. A., Seto, K. C., & Taubenböck, H. (2018). Six fundamental aspects for conceptualizing multidimensional urban form: A spatial mapping perspective. Landscape and Urban Planning, 179, 55–62. https://doi.org/10.1016/j.landurbplan.2018.07.007

Werthmann, C. (2023). Reporting from the front. Natural Hazards, 119(2), 1145–1148. https://doi.org/10.1007/s11069-023-05815-3

Werthmann, C., & Echeverri, A. (2013). Rehabitar La Montaña: Estrategias y procesos para un hábitat sostenible en las laderas de Medellín. Universidad EAFIT, File Rep.

Werthmann, C., Sapena, M., Kühnl, M., Singer, J., Garcia, C., Breuninger, T., Gamperl, M., Menschik, B., Schäfer, H., Schröck, S., Seiler, L., Thuro, K., & Taubenböck, H. (2024). Insights into the development of a landslide early warning system prototype in an informal settlement: The case of Bello Oriente in Medellín, Colombia. Natural Hazards and Earth System Sciences, 24(5), 1843–1870. https://doi.org/10.5194/nhess-24-1843-2024

West, G. (2017). Scale: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies. Penguin Press.

Wetherley, E. B., Roberts, D. A., & McFadden, J. P. (2017). Mapping spectrally similar urban materials at sub-pixel scales. Remote Sensing of Environment, 195, 170–183. https://doi.org/10.1016/j.rse.2017.04.013

Whebell, C. F. J. (1969). Corridors: A theory of urban systems. Annals of the Association of American Geographers, 59(1), 1–26. https://doi.org/10.1111/j.1467-8306.1969.tb00655.x

Wheeler, S. (2009). Regions, Megaregions, and Sustainability. Regional Studies, 43(6), 863–876. https://doi.org/10.1080/00343400701861344

White, J. M., & Kitchin, R. (2021). For or Against 'The Business of Benchmarking'? International Journal of Urban and Regional Research, 45(2), 385–388. Scopus. https://doi.org/10.1111/1468-2427.12978

White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., Bone, A., Depledge, M. H., & Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. Scientific Reports, 9(7730), 1–11. https://doi.org/10.1038/s41598-019-44097-3

White, R. W. (1977). Dynamic Central Place Theory: Results of a Simulation Approach. Geographical Analysis, 9(3), 226–243. https://doi.org/10.1111/j.1538-4632.1977.tb00576.x

Whitehand, J. W. R. (2001). British urban morphology: The Conzenian tradition. Urban Morphology, 5(2), 103–109. https://doi.org/10.51347/jum.v5i2.3896

Whyte, W. H. (1988). City: Rediscovering the Center. Doubleday.

Wiedenhofer, D., Steinberger, J. K., Eisenmenger, N., & Haas, W. (2015). Maintenance and Expansion: Modeling Material Stocks and Flows for Residential Buildings and Transportation Networks in the EU25. Journal of Industrial Ecology, 19(4), 538–551. https://doi.org/10.1111/jiec.12216

Wikipedia. (n.d.-a). List of countries and dependencies by population. Wikipedia - The Free Encyclopedia. Retrieved 16 December 2021, from https:// en.wikipedia.org/wiki/List_of_countries_and_dependencies _by_population

Wikipedia. (n.d.-b). List of countries by Human Development Index. Wikipedia - The Free Encyclopedia. Retrieved 16 December 2021, from https:// en.wikipedia.org/wiki/List_of_countries_by_Human_ Development_Index

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. Biometrics Bulletin, 1(6), 80–83. https://doi.org/10.2307/3001968

Wilson, C. E., Hunt, W. F., Winston, R. J., & Smith, P. K. (2013). A Comparison of Runoff Quality and Quantity from a Urban Commercial Infill Low Impact Development and a Conventional Development. In World Environmental and Water Resources Congress 2013 (pp. 2910–2923). https://doi.org/10.1061/9780784412947.288

Wilson, M. L., Renne, E., Roncoli, C., Agyei-Baffour, P., & Tenkorang, E. Y. (2015). Integrated Assessment of Artisanal and Small-Scale Gold Mining in Ghana—Part 3: Social Sciences and Economics. International Journal of Environmental Research and Public Health, 12(7), 8133–8156. https://doi.org/10.3390/ijerph120708133

Wilson, R. (n.d.). A categorised list of links to over 500 sites providing freely available geographic datasets. Free GIS Data. Retrieved 16 December 2021, from https://freegisdata.rtwilson. com/

Wirth, L. (1969). Urbanism as a way of life. In R. Sennett (Ed.), Classic essays on the culture of cities (pp. 143–164). Appleton-Century-Crofts.

WMO. (2018). Multi-hazard Early Warning Systems: A Checklist (p. 20). https://library.wmo.int/idurl/4/55893

Wójcik, T., & Towpasz, K. (2019). Occurrence of Gentiana cruciata in dry grassland (Festuco-Brometea) in Kołaczyce (Strzyżowskie Foothills). Ecological Questions, 30(1), 9–19. https://doi.org/10.12775/EQ.2019.001

Wolch, J. R., Byrne, J., & Newell, J. P. (2014). Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'. Landscape and Urban Planning, 125, 234–244. https://doi.org/10.1016/j.landurbplan.2014.01.017

Wolff, M., Haase, D., & Haase, A. (2018). Compact or spread? A quantitative spatial model of urban areas in Europe since 1990. PLOS ONE, 13(2)(e0192326), 1–22. https://doi.org/10.1371/journal.pone.0192326

Wong, P.-H. (2020). Democratizing Algorithmic Fairness. Philosophy & Technology, 33(2), 225–244. https://doi.org/10.1007/s13347-019-00355-w

Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., Vogelmann, J., Wulder, M. A., & Wynne, R. (2008). Free Access to Landsat Imagery. Science, 320(5879), 1011–1011. https://doi.org/10.1126/science.320.5879.1011a

Woodcock, C. E., & Strahler, A. H. (1987). The factor of scale in remote sensing. Remote Sensing of Environment, 21(3), 311–332. https://doi.org/10.1016/0034-4257(87)90015-0

World Bank. (2009). World Development Report 2009: Reshaping Economic Geography. © World Bank. http://hdl.handle.net/10986/5991. License: CC BY 3.0 IGO

World Bank. (2012). Disaster Risk Management in Latin America and the Caribbean Region: GFDRR Country Notes. © World Bank. https://doi.org/10.1596/27336

World Bank. (2022a). New World Bank country classifications by income level: 2022-2023. World Bank Blogs. https://blogs.worldbank.org/en/opendata/new-world-bank-country-classifications-income-level-2022-2023

World Bank. (2022b). Overview. World Bank. https://www.worldbank.org/en/topic/urbandevelopment/overview

World Bank. (2022c). Urban development overview. World Bank. https://www.worldbank.org/en/topic/urbandevelopment/overview

World Bank. (2023). World Bank Group country classifications by income level for FY24. https://blogs.worldbank.org/en/opendata/new-world-bank-group-country-classifications-income-level-fy24#_ftn1

World Bank. (2025). Urban population (% of total population) [Dataset]. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS

World development indicators. (2005). Washington, DC: World Bank. http://documents.worldbank.org/curated/en/947951468140975423

World Health Organization. (2010). Urban planning, environment and health: From evidence to policy action. In Urban planning, environment and health: From evidence to policy action. World Health Organization; CABI Databases. http://www.euro.who.int/__data/assets/pdf_file/0004/114448/E93987.pdf

World Health Organization. (2021). WHO global air quality guidelines. Particulate matter (PM 2.5 and PM 10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. https://www.who.int/publications/i/item/9789240034228

World Health Organization (WHO). (2016). Urban green spaces and health. WHO Regional Office for Europe. https://iris.who.int/bitstream/handle/10665/345751/WHO-EURO-2016-3352-43111-60341-eng.pdf

World Meterological Organization (WMO). (n.d.). List of all Satellites. Observing Systems Capability Analysis and Review Tool (OSCAR). Retrieved 16 December 2021, from https://space.oscar.wmo.int/satellites

WorldPop. (2025). Open spatial demographic data and research [Dataset]. https://www.worldpop.org/

Wrbka, T., Erb, K.-H., Schulz, N. B., Peterseil, J., Hahn, C., & Haberl, H. (2004). Linking pattern and process in cultural landscapes. An empirical study based on spatially explicit indicators. Land Use and Sustainability Indicators, 21(3), 289–306. https://doi.org/10.1016/j.landusepol.2003.10.012

Wu, B., Zhang, M., Zeng, H., Tian, F., Potgieter, A. B., Qin, X., Yan, N., Chang, S., Zhao, Y., Dong, Q., Boken, V., Plotnikov, D., Guo, H., Wu, F., Zhao, H., Deronde, B., Tits, L., & Loupian, E. (2023). Challenges and opportunities in remote sensing-based crop monitoring: A review. National Science Review, 10(4)(nwac290), 1–17. https://doi.org/10.1093/nsr/nwac290

Wu, C., Wang, J., Wang, M., Biljecki, F., & Kraak, M.-J. (2025). Formalising the urban pattern language: A morphological paradigm towards understanding the multi-scalar spatial structure of cities. Cities, 161(105854), 1–15. https://doi.org/10.1016/j.cities.2025.105854

Wulder, M. A., & Coops, N. C. (2014). Satellites: Make Earth observations open access. Nature, 513(7516), 30–31. https://doi.org/10.1038/513030a

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J. D., Hostert, P., Hughes, M. J., … Zhu, Z. (2019). Current status of Landsat program, science, and applications. Remote Sensing of Environment, 225, 127–147. https://doi.org/10.1016/j.rse.2019.02.015

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Landsat Legacy Special Issue, 122, 2–10. https://doi.org/10.1016/j.rse.2012.01.010

Wulder, M. A., White, J. C., Loveland, T. R., Woodcock, C. E., Belward, A. S., Cohen, W. B., Fosnight, E. A., Shaw, J., Masek, J. G., & Roy, D. P. (2016). The global Landsat archive: Status, consolidation, and direction. Landsat 8 Science Results, 185, 271–283. https://doi.org/10.1016/j.rse.2015.11.032

Wurm, M., d'Angelo, P., Reinartz, P., & Taubenböck, H. (2014). Investigating the Applicability of Cartosat-1 DEMs and Topographic Maps to Localize Large-Area Urban Mass Concentrations. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(10), 4138–4152. https://doi.org/10.1109/JSTARS.2014.2346655

Wurm, M., Goebel, J., Wagner, G. G., Weigand, M., Dech, S., & Taubenböck, H. (2021). Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception. Environment and Planning B: Urban Analytics and City Science, 48(2), 265–279. https://doi.org/10.1177/2399808319869341

Wurm, M., Stark, T., Zhu, X. X., Weigand, M., & Taubenböck, H. (2019). Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing, 150, 59–69. https://doi.org/10.1016/j.isprsjprs.2019.02.006

Wurm, M., & Taubenböck, H. (2018). Detecting social groups from space – Assessment of remote sensing-based mapped morphological slums using income data. Remote Sensing Letters, 9(1), 41–50. https://doi.org/10.1080/2150704X.2017.1384586

Wurm, M., Taubenböck, H., Schardt, M., Esch, T., & Dech, S. (2011). Object-based image information fusion using multisensor earth observation data over urban areas. International Journal of Image and Data Fusion, 2(2), 121–147. https://doi.org/10.1080/19479832.2010.543934

Wurm, M., Taubenböck, H., Weigand, M., & Schmitt, A. (2017). Slum mapping in polarimetric SAR data using spatial features. Remote Sensing of Environment, 194, 190–204. https://doi.org/10.1016/j.rse.2017.03.030

Wüstemann, H., Kalisch, D., & Kolbe, J. (2017). Access to urban green space and environmental inequalities in Germany. Landscape and Urban Planning, 164, 124–131. https://doi.org/10.1016/j.landurbplan.2017.04.002

X

Xian, G., & Homer, C. (2010). Updating the 2001 National Land Cover Database Impervious Surface Products to 2006 using Landsat Imagery Change Detection Methods. Remote Sensing of Environment, 114(8), 1676–1686. https://doi.org/10.1016/j.rse.2010.02.018

Xiao, Q., Ustin, S. L., & McPherson, E. G. (2004). Using AVIRIS data and multiple-masking techniques to map urban forest tree species. International Journal of Remote Sensing, 25(24), 5637–5654. https://doi.org/10.1080/01431160412331291224

Xiao, R., Ouyang, Z., Zheng, H., Li, W., Schienke, E. W., & Wang, X. (2007). Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. Journal of Environmental Sciences (China), 19(2), 250–256. https://doi.org/10.1016/s1001-0742(07)60041-2

Xie, Y., & Weng, Q. (2016). Updating urban extents with nighttime light imagery by using an object-based thresholding method. Remote Sensing of Environment, 187, 1–13. https://doi.org/10.1016/j.rse.2016.10.002

Xin, X., Liu, B., Di, K., Zhu, Z., Zhao, Z., Liu, J., Yue, Z., & Zhang, G. (2017). Monitoring urban expansion using time series of night-time light data: A case study in Wuhan, China. International Journal of Remote Sensing, 38(21), 6110–6128. https://doi.org/10.1080/01431161.2017.1312623

Xu, G., Dong, T., Cobbinah, P. B., Jiao, L., Sumari, N. S., Chai, B., & Liu, Y. (2019). Urban expansion and form changes across African cities with a global outlook: Spatiotemporal analysis of urban land densities. Journal of Cleaner Production, 224, 802–810. https://doi.org/10.1016/j.jclepro.2019.03.276

Xu, H. (2008). A new index for delineating built-up land features in satellite imagery. International Journal of Remote Sensing, 29(14), 4269–4276. https://doi.org/10.1080/01431160802039957

Xu, L., Wang, X., Liu, J., He, Y., Tang, J., Nguyen, M., & Cui, S. (2019). Identifying the trade-offs between climate change mitigation and adaptation in urban land use planning: An empirical study in a coastal city. Environment International, 133(105162), 1–11. https://doi.org/10.1016/j.envint.2019.105162

Xu, M., He, C., Liu, Z., & Dou, Y. (2016). How Did Urban Land Expand in China between 1992 and 2015? A Multi-Scale Landscape Analysis. PLOS ONE, 11(5)(e0154839), 1–19. https://doi.org/10.1371/journal.pone.0154839

Xu, Y., Ren, C., Cai, M., Edward, N. Y. Y., & Wu, T. (2017). Classification of Local Climate Zones Using ASTER and Landsat Data for High-Density Cities. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(7), 3397–3405. https://doi.org/10.1109/JSTARS.2017.2683484

Xu, Z., Chen, J., Xia, J., Du, P., Zheng, H., & Gan, L. (2018). Multisource Earth Observation Data for Land-Cover Classification Using Random Forest. IEEE Geoscience and Remote Sensing Letters, 15(5), 789–793. https://doi.org/10.1109/LGRS.2018.2806223

Y

Yaaba Baah-Ennumh, T., & Forson, J. A. (2017). The impact of artisanal small-scale mining on sustainable livelihoods. World Journal of Entrepreneurship, Management and Sustainable Development, 13(3), 204–222. https://doi.org/10.1108/WJEMSD-09-2016-0042

Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., & Dickinson, R. (2013). The role of satellite remote sensing in climate change studies. Nature Climate Change, 3(10), 875–883. https://doi.org/10.1038/nclimate1908

Yang, J., Jin, S., Xiao, X., Jin, C., Xia, J. (Cecilia), Li, X., & Wang, S. (2019). Local climate zone ventilation and urban land surface temperatures: Towards a performance-based and wind-sensitive planning proposal in megacities. Sustainable Cities and Society, 47(101487), 1–11. https://doi.org/10.1016/j.scs.2019.101487

Yang, J., Sun, J., Ge, Q., & Li, X. (2017). Assessing the impacts of urbanization-associated green space on urban land surface temperature: A case study of Dalian, China. Urban Forestry & Urban Greening, 22, 1–10. https://doi.org/10.1016/j.ufug.2017.01.002

Yang, J., Wang, Y., Xiu, C., Xiao, X., Xia, J., & Jin, C. (2020). Optimizing local climate zones to mitigate urban heat island effect in human settlements. Journal of Cleaner Production, 275(123767), 1–12. https://doi.org/10.1016/j.jclepro.2020.123767

Yang, J., Zhan, Y., Xiao, X., Xia, J. C., Sun, W., & Li, X. (2020). Investigating the diversity of land surface temperature characteristics in different scale cities based on local climate zones. Urban Climate, 34(100700), 1–12. https://doi.org/10.1016/j.uclim.2020.100700

Yang, L., Huang, C., Homer, C. G., Wylie, B. K., & Coan, M. J. (2003). An approach for mapping large-area impervious surfaces: Synergistic use of Landsat-7 ETM+ and high spatial resolution imagery. Canadian Journal of Remote Sensing, 29(2), 230–240. https://doi.org/10.5589/m02-098

Yang, T., & Hillier, B. (2007). The fuzzy boundary: The spatial definition of urban areas. Proceedings of 6th International Space Syntax Symposium, 091/01-091/16. http://spacesyntaxistanbul.itu.edu.tr/papers/longpapers/091%20-%20Yang%20Hillier.pdf

Yang, Y., Xiao, P., Feng, X., & Li, H. (2017). Accuracy assessment of seven global land cover datasets over China. ISPRS Journal of Photogrammetry and Remote Sensing, 125, 156–173. https://doi.org/10.1016/j.isprsjprs.2017.01.016

Yang, Y.-Y., & Lusk, M. G. (2018). Nutrients in Urban Stormwater Runoff: Current State of the Science and Potential Mitigation Options. Current Pollution Reports, 4(2), 112–127. https://doi.org/10.1007/s40726-018-0087-7

Ye, T., Zhao, N., Yang, X., Ouyang, Z., Liu, X., Chen, Q., Hu, K., Yue, W., Qi, J., Li, Z., & Jia, P. (2019). Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model. Science of The Total Environment, 658, 936–946. https://doi.org/10.1016/j.scitotenv.2018.12.276

Yin, J., & Chi, G. (2021). Characterizing People's Daily Activity Patterns in the Urban Environment: A Mobility Network Approach with Geographic Context-Aware Twitter Data. Annals of the American Association of Geographers, 111(7), 1967–1987. https://doi.org/10.1080/24694452.2020.1867498

Yokoya, N., Ghamisi, P., & Xia, J. (2017). Multimodal, multitemporal, and multisource global data fusion for local climate zones classification based on ensemble learning. 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1197–1200. https://doi.org/10.1109/IGARSS.2017.8127172

Yoo, C., Han, D., Im, J., & Bechtel, B. (2019). Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using Landsat images. ISPRS Journal of Photogrammetry and Remote Sensing, 157, 155–170. https://doi.org/10.1016/j.isprsjprs.2019.09.009

You, H. (2016). Quantifying megacity growth in response to economic transition: A case of Shanghai, China. Habitat International, 53, 115–122. https://doi.org/10.1016/j.habitatint.2015.11.001

Young, R. F. (2009). Interdisciplinary foundations of urban ecology. Urban Ecosystems, 12(3), 311–331. https://doi.org/10.1007/s11252-009-0095-x

Youngmin, K. (n.d.). Wang Yangming (1472—1529). Internet Encyclopedia of Philosophy (IEP). Retrieved 27 September 2022, from https://iep.utm.edu/wangyang/

Yu, B., Liu, H., Wu, J., Hu, Y., & Zhang, L. (2010). Automated derivation of urban building density information using airborne LiDAR data and object-based method. Climate Change and Spatial Planning, 98(3), 210–219. https://doi.org/10.1016/j.landurbplan.2010.08.004

Yu, D., & Coulthard, T. J. (2015). Evaluating the importance of catchment hydrological parameters for urban surface water flood modelling using a simple hydro-inundation model. Journal of Hydrology, 524, 385–400. https://doi.org/10.1016/j.jhydrol.2015.02.040

Yu, D., & Fang, C. (2023). Urban Remote Sensing with Spatial Big Data: A Review and Renewed Perspective of Urban Studies in Recent Decades. Remote Sensing, 15(5), 1–34. https://doi.org/10.3390/rs15051307

Yu, W., Ai, T., & Shao, S. (2015). The analysis and delimitation of Central Business District using network kernel density estimation. Journal of Transport Geography, 45, 32–47. https://doi.org/10.1016/j.jtrangeo.2015.04.008

Yuan, F., & Bauer, M. E. (2007). Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. Remote Sensing of Environment, 106(3), 375–386. https://doi.org/10.1016/j.rse.2006.09.003

Yuan, F., Sawaya, K. E., Loeffelholz, B. C., & Bauer, M. E. (2005). Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. Remote Sensing of Environment, 98(2), 317–328. https://doi.org/10.1016/j.rse.2005.08.006

Z

Zadeh, L. A. (1965). Fuzzy sets. Information and Control, 8(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Zambon, I., Serra, P., & Salvati, L. (2019). The (Evolving) urban footprint under sequential building cycles and changing socio-demographic contexts. Environmental Impact Assessment Review, 75, 27–36. https://doi.org/10.1016/j.eiar.2018.11.003

Zelnik-Manor, L., & Perona, P. (2005). Self-Tuning Spectral Clustering. In L. Saul, Y. Weiss, & L. Bottou (Eds.), Advances in Neural Information Processing Systems (Vol. 17, pp. 1–8). MIT Press. https://proceedings.neurips.cc/paper_files/paper/2004/file/40173ea48d9567f1f393b20c855bb40b-Paper.pdf

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. International Journal of Remote Sensing, 24(3), 583–594. https://doi.org/10.1080/01431160304987

Zhang, G., Ghamisi, P., & Zhu, X. X. (2019). Fusion of Heterogeneous Earth Observation Data for the Classification of Local Climate Zones. IEEE Transactions on Geoscience and Remote Sensing, 57(10), 7623–7642. https://doi.org/10.1109/TGRS.2019.2914967

Zhang, G. J., Cai, M., & Hu, A. (2013). Energy consumption and the unexplained winter warming over northern Asia and North America. Nature Climate Change, 3(5), 466–470. https://doi.org/10.1038/nclimate1803

Zhang, L., & Weng, Q. (2016). Annual dynamics of impervious surface in the Pearl River Delta, China, from 1988 to 2013, using time series Landsat imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 113, 86–96. https://doi.org/10.1016/j.isprsjprs.2016.01.003

Zhang, L., Xia, G.-S., Wu, T., Lin, L., & Tai, X. C. (2016). Deep Learning for Remote Sensing Image Understanding. Journal of Sensors, 2016(1), 1–2. https://doi.org/10.1155/2016/7954154

Zhang, M., & Zhao, P. (2018). The determinants of informal housing price in Beijing: Village power, informal institutions, and property security. Cities, 77, 117–129. https://doi.org/10.1016/j.cities.2018.01.018

Zhang, Q., Schaaf, C., & Seto, K. C. (2013). The Vegetation Adjusted NTL Urban Index: A new approach to reduce saturation and increase variation in nighttime luminosity. Remote Sensing of Environment, 129, 32–41. https://doi.org/10.1016/j.rse.2012.10.022

Zhang, Q., & Seto, K. C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. Remote Sensing of Environment, 115(9), 2320–2329. https://doi.org/10.1016/j.rse.2011.04.032

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., & Schneider, A. (2004). The footprint of urban climates on vegetation phenology. Geophysical Research Letters, 31(12), 1–4. https://doi.org/10.1029/2004GL020137

Zhang, Y., Zhang, H., & Lin, H. (2014). Improving the impervious surface estimation with combined use of optical and SAR remote sensing images. Remote Sensing of Environment, 141, 155–167. https://doi.org/10.1016/j.rse.2013.10.028

Zhang, Z., Liu, F., Zhao, X., Wang, X., Shi, L., Xu, J., Yu, S., Wen, Q., Zuo, L., Yi, L., Hu, S., & Liu, B. (2018). Urban Expansion in China Based on Remote Sensing Technology: A Review. Chinese Geographical Science, 28(5), 727–743. https://doi.org/10.1007/s11769-018-0988-9

Zhang, Z., Wang, X., Zhao, X., Liu, B., Yi, L., Zuo, L., Wen, Q., Liu, F., Xu, J., & Hu, S. (2014). A 2010 update of National Land Use/Cover Database of China at 1:100000 scale using medium spatial resolution satellite images. Remote Sensing of Environment, 149, 142–154. https://doi.org/10.1016/j.rse.2014.04.004

Zhao, L., Lee, X., Smith, R. B., & Oleson, K. (2014). Strong contributions of local background climate to urban heat islands. Nature, 511(7508), 216–219. https://doi.org/10.1038/nature13462

Zhao, T., Brown, D. G., & Bergen, K. M. (2007). Increasing Gross Primary Production (GPP) in the Urbanizing Landscapes of Southeastern Michigan. Photogrammetric Engineering & Remote Sensing, 73(10), 1159–1167. https://doi.org/doi:10.14358/PERS.73.10.1159

Zhao, X., Yu, B., Liu, Y., Yao, S., Lian, T., Chen, L., Yang, C., Chen, Z., & Wu, J. (2018). NPP-VIIRS DNB Daily Data in Natural Disaster Assessment: Evidence from Selected Case Studies. Remote Sensing, 10(1526), 1–25. https://doi.org/10.3390/rs10101526

Zheng, Y., Ren, C., Xu, Y., Wang, R., Ho, J., Lau, K., & Ng, E. (2018). GIS-based mapping of Local Climate Zone in the high-density city of Hong Kong. Urban Climate, 24, 419–448. https://doi.org/10.1016/j.uclim.2017.05.008

Zheng, Z., Wang, J., Shan, B., He, Y., Liao, C., Gao, Y., & Yang, S. (2020). A New Model for Transfer Learning-Based Mapping of Burn Severity. Remote Sensing, 12(708), 1–18. https://doi.org/10.3390/rs12040708

Zhong, C., Müller Arisona, S., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. International Journal of Geographical Information Science, 28(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521

Zhou, B., Rybski, D., & Kropp, J. P. (2017). The role of city size and urban form in the surface urban heat island. Scientific Reports, 7(4791), 1–9. https://doi.org/10.1038/s41598-017-04242-2

Zhou, D., Zhao, S., Zhang, L., & Liu, S. (2016). Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities. Remote Sensing of Environment, 176, 272–281. https://doi.org/10.1016/j.rse.2016.02.010

Zhou, W., Huang, G., & Cadenasso, M. L. (2011). Does spatial configuration matter? Understanding the effects of land cover pattern on land surface temperature in urban landscapes. Landscape and Urban Planning, 102(1), 54–63. https://doi.org/10.1016/j.landurbplan.2011.03.009

Zhou, W., Pickett, S. T. A., & Cadenasso, M. L. (2017). Shifting concepts of urban spatial heterogeneity and their implications for sustainability. Landscape Ecology, 32(1), 15–30. https://doi.org/10.1007/s10980-016-0432-4

Zhou, X., & Zhang, L. (2016). Crowdsourcing functions of the living city from Twitter and Foursquare data. Cartography and Geographic Information Science, 43(5), 393–404. https://doi.org/10.1080/15230406.2015.1128852

Zhou, Y., Li, X., Asrar, G. R., Smith, S. J., & Imhoff, M. (2018). A global record of annual urban dynamics (1992–2013) from nighttime lights. Remote Sensing of Environment, 219, 206–220. https://doi.org/10.1016/j.rse.2018.10.015

Zhou, Y., Smith, S. J., Elvidge, C. D., Zhao, K., Thomson, A., & Imhoff, M. (2014). A cluster-based method to map urban area from DMSP/OLS nightlights. Remote Sensing of Environment, 147, 173–185. https://doi.org/10.1016/j.rse.2014.03.004

Zhou, Y., Smith, S. J., Zhao, K., Imhoff, M., Thomson, A., Bond-Lamberty, B., Asrar, G. R., Zhang, X., He, C., & Elvidge, C. D. (2015). A global map of urban extent from nightlights. Environmental Research Letters, 10(054011), 1–11. https://doi.org/10.1088/1748-9326/10/5/054011

Zhou, Y., & Weng, Q. (2024). Building up a data engine for global urban mapping. Remote Sensing of Environment, 311(114242), 1–23. https://doi.org/10.1016/j.rse.2024.114242

Zhou, Y., Weng, Q., Gurney, K. R., Shuai, Y., & Hu, X. (2012). Estimation of the relationship between remotely sensed anthropogenic heat discharge and building energy use. ISPRS Journal of Photogrammetry and Remote Sensing, 67, 65–72. https://doi.org/10.1016/j.isprsjprs.2011.10.007

Zhu, J., & Sun, Y. (2017). Building an Urban Spatial Structure from Urban Land Use Data: An Example Using Automated Recognition of the City Centre. ISPRS International Journal of Geo-Information, 6(122), 1–24. https://doi.org/10.3390/ijgi6040122

Zhu, L., & Lerman, K. (2016). Attention Inequality in Social Media. https://arxiv.org/abs/1601.07200

Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010). An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sensing of Environment, 114(11), 2610–2623. https://doi.org/10.1016/j.rse.2010.05.032

Zhu, X. X., Hu, J., Qiu, C., Shi, Y., Kang, J., Mou, L., Bagheri, H., Haberle, M., Hua, Y., Huang, R., Hughes, L., Li, H., Sun, Y., Zhang, G., Han, S., Schmitt, M., & Wang, Y. (2019). So2Sat LCZ42: A Benchmark Data Set for the Classification of Global Local Climate Zones [Software and Data Sets]. IEEE Geoscience and Remote Sensing Magazine, 8(3), 76–89. https://doi.org/10.1109/MGRS.2020.2964708

Zhu, X. X., Qiu, C., Hu, J., Shi, Y., Wang, Y., Schmitt, M., & Taubenböck, H. (2022). The urban morphology on our planet – Global perspectives from space. Remote Sensing of Environment, 269(112794), 1–11. https://doi.org/10.1016/j.rse.2021.112794

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Zhu, X. X., Wang, Y., Kochupillai, M., Werner, M., Häberle, M., Hoffmann, E. J., Taubenböck, H., Tuia, D., Levering, A., Jacobs, N., Kruspe, A., & Abdulahhad, K. (2022). Geoinformation Harvesting From Social Media Data: A community remote sensing approach. IEEE Geoscience and Remote Sensing Magazine, 10(4), 150–180. https://doi.org/10.1109/MGRS.2022.3219584

Zhu, Z. (2017). Change detection using landsat time series: A review of frequencies, preprocessing, algorithms, and applications. ISPRS Journal of Photogrammetry and Remote Sensing, 130, 370–384. https://doi.org/10.1016/j.isprsjprs.2017.06.013

Zhu, Z., Fu, Y., Woodcock, C. E., Olofsson, P., Vogelmann, J. E., Holden, C., Wang, M., Dai, S., & Yu, Y. (2016). Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). Landsat 8 Science Results, 185, 243–257. https://doi.org/10.1016/j.rse.2016.03.036

Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. Remote Sensing of Environment, 144, 152–171. https://doi.org/10.1016/j.rse.2014.01.011

Zhu, Z., Woodcock, C. E., Rogan, J., & Kellndorfer, J. (2012). Assessment of spectral, polarimetric, temporal, and spatial dimensions for urban and peri-urban land cover classification using Landsat and SAR data. Remote Sensing of Environment, 117, 72–82. https://doi.org/10.1016/j.rse.2011.07.020

Zhu, Z., Wulder, M. A., Roy, D. P., Woodcock, C. E., Hansen, M. C., Radeloff, V. C., Healey, S. P., Schaaf, C., Hostert, P., Strobl, P., Pekel, J.-F., Lymburner, L., Pahlevan, N., & Scambos, T. A. (2019). Benefits of the free and open Landsat data policy. Remote Sensing of Environment, 224, 382–385. https://doi.org/10.1016/j.rse.2019.02.016

Zhu, Z., Zhang, J., Yang, Z., Aljaddani, A. H., Cohen, W. B., Qiu, S., & Zhou, C. (2019). Continuous monitoring of land disturbance based on Landsat time series. Time Series Analysis with High Spatial Resolution Imagery, 238(111116), 1–18. https://doi.org/10.1016/j.rse.2019.03.009

Zhu, Z., Zhou, Y., Seto, K. C., Stokes, E. C., Deng, C., Pickett, S. T. A., & Taubenböck, H. (2019). Understanding an urbanizing planet: Strategic directions for remote sensing. Remote Sensing of Environment, 228, 164–182. https://doi.org/10.1016/j.rse.2019.04.020

Zimmer, M. (2018). Addressing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity. Social Media + Society, 4(2), 1–11. https://doi.org/10.1177/2056305118768300

Zipf, G. K. (1941). National Unity and Disunity—The Nation as a Bio-Social Organism. The Principia Press, Inc. https://babel.hathitrust.org/cgi/pt?id=mdp.39015057175484&view=1up&seq=5

Ziter, C. D., Pedersen, E. J., Kucharik, C. J., & Turner, M. G. (2019). Reply to Drescher: Interdisciplinary collaboration is essential to understand and implement climate-resilient strategies in cities. Proceedings of the National Academy of Sciences, 116(52), 26155–26156. https://doi.org/10.1073/pnas.1918746116

Zook, M. A. (2001). Old Hierarchies or New Networks of Centrality?: The Global Geography of the Internet Content Market. American Behavioral Scientist, 44(10), 1679–1696. https://doi.org/10.1177/00027640121958113

Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B. A., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2017). Ten simple rules for responsible big data research. PLOS Computational Biology, 13(3)(e1005399), 1–10. https://doi.org/10.1371/journal.pcbi.1005399

Zukunftsinstitut. (2021). Urbanisierung: Die Stadt von morgen. https://www.zukunftsinstitut.de/zukunftsthemen/urbanisierung-die-stadt-von-morgen