

Heidelberg University
Faculty of Chemistry and Earth Sciences
Institute of Geography

Master Thesis

Exploring city patterns globally
the intra-urban morphology through
the scope of unsupervised learning

First Supervisor: Apl. Prof. Dr. Sven Lautenbach, Heidelberg University
Second Supervisor: Dr. Hannes Taubenböck, German Aerospace Center

*A thesis in partial fulfillment of the requirements
for the degree of Master of Science*

submitted March 2022
by
Matthias Gassilloud

Matriculation-No.: 3543420
e-mail: matthias.gass@posteo.de

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

This paper was not previously presented to another examination board and has not been published.

Heidelberg, 24.03.2022

M. Gaschler

Acknowledgements

I would like to thank my examiners for the supervision of this master thesis, Apl. Prof. Dr. Sven Lautenbach from the Heidelberg University and Dr. Hannes Taubenböck from the German Aerospace Center.

I especially would like to thank Henri Debray from the German Aerospace Center for his profound introduction to urban morphology, his guidance in this work, his everlasting support and belief in me.

I would like to acknowledge the German Aerospace Center for supporting this master thesis and providing the data.

I would like to express my gratitude towards my family for their unconditional support and encouragement throughout my studies and the process of this master thesis.

Further, I thank my proofreaders Julio Garbers, Katharina Heß, Mathieu Pinger, Fabian Thomas and Julian Zimmermann.

And lastly thanks to Mathieu, Jules, Katharina, Franzi, Julio, Fabian, David, Louisa, Stefanie and all my other friends and fellow students for their support throughout my studies.

Abstract

Cities are complex systems with a unique composition of diverse elements and their relationships. Throughout history, humans form and shape cities by a range of functional, social, economical and political interactions. This diversity is reflected in the formation of individual built and non-built environments. However, similar elements and features form patterns that can be observed among multiple cities. City models try to understand the underlying processes that manifest into spatial patterns of urban form, but are often limited by a regional context and lack of comparable data. This master thesis aims to explore the urban morphology in a comparable framework on a global scale with the use of new consistent datasets such as the Local Climate Zones (LCZs) to describe the urban morphology of cities and the Morphological Urban Areas (MUAs) to delineate urban agglomerations. A search of urban morphological patterns is conducted without prior knowledge on subsets of 1523 cities. With state of the art methods of unsupervised learning 138 clusters of urban morphological patterns are found. The patterns show urban morphological configurations with similar statistical and spatial characteristics. A similarity metric is developed to compare cities based on the found patterns. Grouping similar cities leads to the formation of clusters which are partially congruent with geographic regions. The results of this work show that the formation of patterns with similar urban morphological configurations is linked to the geographic location. This master thesis is a first step towards a comprehensive knowledge on the formation of urban morphological configurations and contributes to a better understanding of cities.

Zusammenfassung

Städte sind einzigartige, komplexe Systeme, die sich aus unterschiedlichsten Elementen und ihren Beziehungen zusammensetzen. Seit jeher formen und gestalten Menschen ihre Städte durch eine Vielzahl verschiedener funktionaler, sozialer, wirtschaftlicher und politischer Interaktionen. Diese Vielfalt spiegelt sich in den individuellen Ausprägungen von bebauten und unbebauten Gebieten wieder. Jedoch lassen sich ähnliche Elemente und Muster in mehreren Städten beobachten, deren zugrundeliegende Prozesse mit Hilfe von Stadtmodellen untersucht werden. Die Aussagekraft von Modellen ist jedoch häufig auf einen regionalen Kontext beschränkt oder durch den Mangel an vergleichbaren Datensätzen limitiert. Mit dieser Masterarbeit soll die urbane Morphologie mit Hilfe neuer konsistenter Datensätze auf einer globalen Ebene untersucht werden. Verwendet werden hierfür die Local Climate Zones (LCZs) zum Beschreiben der urbanen Morphologie sowie die Morphological Urban Areas (MUAs) zur Definition von urbanem Raum. Stadtmorphologische Muster werden mit neuen Methoden zum unüberwachten Lernen in Teilausschnitten von 1523 Städten gesucht. Insgesamt werden 138 verschiedene stadtmorphologische Muster gefunden, die sowohl ähnliche statistische Charakteristiken als auch ähnliche räumliche Anordnungen aufweisen. Zudem wird in dieser Arbeit ein Ähnlichkeitsmaß von Städten auf Grundlage der gefundenen stadtmorphologischen Muster entwickelt. Beim Gruppieren ähnlicher Städte entstehen Gebiete, die teilweise eine hohe Übereinstimmung mit geographischen Regionen aufweisen. Die Ergebnisse dieser Arbeit zeigen, dass die Bildung von stadtmorphologischen Mustern mit der geographischen Lage zusammenhängt. Damit leistet diese Arbeit einen Beitrag zur Untersuchung von urbanen morphologischen Konfigurationen und trägt zum besseren Verständnis von Städten bei.

Contents

1	Introduction	1
2	State of the Art	6
2.1	The Grasp of a City	6
2.2	Approaches in Urban Geography	7
2.3	Structural City Models	9
2.4	Exploring Urban Morphology	14
2.5	Urban Morphology through the Scope of Earth Observation	15
3	Framework and Research Questions	20
4	Methodology	21
4.1	Theoretical Background	21
4.1.1	Conceptualization	21
4.1.2	Introduction to Deep Learning in Computer Vision	24
4.1.3	Methods for Unsupervised Image Clustering	29
4.2	Data Basis	33
4.2.1	Local Climate Zones	33
4.2.2	Morphological Urban Areas	33
4.3	Methods	34
4.3.1	Data Preparation	34
4.3.2	Semantic Clustering by Adopting Nearest Neighbors	38
4.3.3	Robust learning for Unsupervised Clustering	45
4.3.4	Distance Calculation with Hungarian Matching	50
4.3.5	Clustering Cities based on Urban Morphological Configurations	52
4.3.6	Clustering Cities with a 18-dimensional Feature Space	57
5	Results	59
5.1	Patterns of Intra-urban Morphological Configurations	59
5.2	Clustering Cities based on Urban Morphological Patterns	63
5.3	Ablation Studies on Clustering Cities.	69
6	Discussion	75
6.1	Understanding Patterns of Urban Morphological Configurations	75
6.2	Formation of Geographic Clusters with similar Cities	80
7	Conclusion and Outlook	87
	References	89
	Appendix	100

List of Figures

1	Development of Urban and Total Population from 1950 to 2050	2
2	Number of active Earth Observation (EO) Satellites from 1959 to 2019	5
3	Palmanova - 'ideal city' of the Renaissance	9
4	The city as a Tree and Semi-lattice	10
5	Concentric Zones Model	11
6	Sectoral City Model	11
7	Multiple Nuclei Model	12
8	The 17 Local Climate Zones (LCZs)	17
9	Typical Workflow in Pattern Recognition	22
10	Threshold Logic Unit (TLU) of a Perceptron	25
11	Biological Neurons in the Cerebral Cortex	25
12	Multilayer Perceptron (MLP)	26
13	Convolutional Layer	28
14	Max-Pooling Layer	28
15	Example of a Convolutional Neural Network (CNN) Architecture	28
16	Concept of Residual Blocks	29
17	SCAN example Clusters of ImageNet-1000	31
18	Mannheim with Administrative Area, Morphological Urban Area (MUA) and Local Climate Zones (LCZs)	34
19	General Overview of the Methodological Components	35
20	Local Climate Zones (LCZs) - Edge Removal	35
21	Guangzhou LCZ Raster Tiles and buffered MUA	36
22	Guangzhou consistent LCZs Data and buffered MUA	37
23	Example of extracted Patches	38
24	Adapted ResNet-18 with a Contrastive Head	39
25	SimCLR Workflow	40
26	Adapted ResNet-18 with a Cluster Head	41
27	Mining Nearest Neighbors	42
28	SCAN Workflow	42
29	Sample Confidence - SCAN intermediate Results	44
30	Self-label Workflow	44
31	Sample Confidence - Self-label intermediate Results	45
32	RUC Workflow	47
33	Sample Confidence - RUC final Results	49
34	Distance Calculation via Hungarian Matching	51
35	Multidimensional Scaling for 1523 Cities	53
36	Multidimensional Scaling for 110 Cities	54
37	Gap Statistic Result for HM_{1523}	56
38	Gap Statistic Result for HM_{110}	56
39	Gap Statistic Result for F_{1523}	57
40	Urban Morphological Configurations Result - Example 1	61

41	Urban Morphological Configurations Result - Example 2	61
42	Urban Morphological Configurations Result - Example 3	62
43	Urban Morphological Configurations Result - Example 4	62
44	Overview of the Seven Clusters Result $C_{HM_{1523}}$	64
45	Cluster $C_{HM_{1523}}(1)$	65
46	Cluster $C_{HM_{1523}}(2)$	65
47	Cluster $C_{HM_{1523}}(3)$	66
48	Cluster $C_{HM_{1523}}(4)$	67
49	Cluster $C_{HM_{1523}}(5)$	67
50	Cluster $C_{HM_{1523}}(6)$	68
51	Cluster $C_{HM_{1523}}(7)$	68
52	Overview of the Seven Clusters Result $C_{HM_{110}}$	70
53	Overview of the 29 Clusters Result $C_{F_{1523}}$	73
54	Overview of the Seven Clusters Result $C_{F_{1523}}^{*7}$	74
55	Exemplary Transformations with T_{con}	100
56	Exemplary Transformations with T_{scan}	101
57	Exemplary Transformations with T_{self}	101
58	Urban Morphological Configurations Result - Example 5	103
59	Urban Morphological Configurations Result - Example 6	103
60	Urban Morphological Configurations Result - Example 7	104
61	Urban Morphological Configurations Result - Example 8	104
62	Urban Morphological Configurations Result - Example 9	105
63	Urban Morphological Configurations Result - Example 10	105
64	Urban Morphological Configurations Result - Example 11	106
65	Urban Morphological Configurations Result - Example 12	106
66	Cities assigned to Cluster $C_{HM_{1523}}(1)$	107
67	Cities assigned to Cluster $C_{HM_{1523}}(2)$	107
68	Cities assigned to Cluster $C_{HM_{1523}}(3)$	108
69	Cities assigned to Cluster $C_{HM_{1523}}(4)$	108
70	Cities assigned to Cluster $C_{HM_{1523}}(5)$	109
71	Cities assigned to Cluster $C_{HM_{1523}}(6)$	109
72	Cities assigned to Cluster $C_{HM_{1523}}(7)$	110

List of Tables

1	Transformations T_{con} for the Pretext Task	100
2	Transformations T_{scan} for Clustering	100
3	Transformations T_{self} for Self-labeling	101
4	Transformations T_{weak} used in RUC	101
5	Transformations T_{strong} used in RUC	102
6	Randaugment Parameters	102
7	Overview of Data Base and Results	111

List of Acronyms

ANN artificial neural network

BCE before the common era

BN batch normalization

CBD Central Business District

CNN convolutional neural network

EO Earth Observation

GPU graphical processing unit

GUF Global Urban Footprint

LCZ Local Climate Zone

MDS multi-dimensional scaling

MLP multilayer perceptron

MoCo Momentum Contrast for Unsupervised Visual Representation Learning

MSI morphological settlement index

MUA Morphological Urban Area

NFL no free lunch

OSM Open Street Map

PCA principal component analysis

ReLU Rectified Linear Units

ResNet residual neural network

RUC Robust learning for Unsupervised Clustering

S2 Sentinel-2

SAR spaceborne synthetic aperture radar

SCAN Semantic Clustering by Adopting Nearest neighbors

SimCLR Simple Framework for Contrastive Learning of Visual Representations

TLU threshold logic unit

UHI Urban Heat Island

US United States

WUDAPT World Urban Database and Portal Tool

1 Introduction

When walking through a city, what are the perceptions one experiences? In a central business district, it might be the noise of traffic, horns honking and the hammering of a nearby construction site. Perhaps one might smell the fumes of the bypassing cars and see the tall buildings arising next to the street far up, leaving only a small trace of sky visible. Changing the location to a crowded shopping mall, it is the indistinctive chattering of the bypassing people, entering and leaving the different shops. Continuing to a calm residential area, one might be surrounded by single family houses, each with a garden. In these different scenarios, everyone will have a very subjective impression of the surroundings. One might be overwhelmed by the experience of all the individual elements, but everyone will subconsciously structure, order and evaluate the world around them.

From past experiences, a feeling for the physical environment emerges. When visiting a different country, or maybe another continent, one might encounter fundamentally different ways of living, but notice also some similarities and familiar elements. Spatial patterns emerge and form neighborhoods, commercial centers and industrial areas, similar across multiple cities. Observing solely the single elements, each building, each place and each city is unique. But when comparing them on a larger scale, they seem to follow a general order, somehow shaped and regularized by an invisible force. To uncover these principles of order has been endeavored by urban geography for a long time. Many models have been proposed to explain spatial patterns, yet they are often qualitative and restricted to a regional or cultural context. The recent abundant amounts of available geodata support a rise of quantitative approaches and the generation of new theories. This master thesis aims to explore the intra-urban configurations on a global scale and support an empirical approach to unveil the unwritten rules that seem to shape cities throughout the world.

But where to begin? For millennia humans develop and transform land and build artificial structures (Goh et al. 2016). It is assumed that the neolithic revolution starting 10.000 before the common era (BCE) initiated the transition from hunter-gatherers to agriculture and livestock breeding (Weisdorf 2005). Along this, the first settlements arose which were the foundations for a development to smaller villages and cities with an increasing population. The first traces of the oldest cities like Ur and Uruk located in the valley between Tigris and the Euphrates date back 5000 BCE (Nissen 1988). Further ancient cities like Memphis and Theben in Egypt have more than 5000 years of history, but also along the Indus river several ancient cities, which are around 4500 years old, can be found (Wright 2010).

The extensive history of cities shows how this form of social spatial organization became an essential part for human society and its development. With the continuous increase of the world population, the number and size of agglomerations of human settlements grew throughout the world. This trend of a growing world population continues until the current day (see figure 1). A city itself grows with an increasing number of citizens but also with an expansion of urban areas, both in vertical and horizontal dimensions through new and higher buildings (Gaebe 2004). These processes are included in the concept of 'urbanization'. Furthermore, the term urbanization

describes in its wider sense not only the growth of cities but also the diffusion of urban functions into formerly rural areas. Even though cities are inherent to our societies and a majority of the world population now lives in urban areas, the capability of understanding and managing them in a scientific way remains limited due to their complex and interdependent facets (Bettencourt 2013).

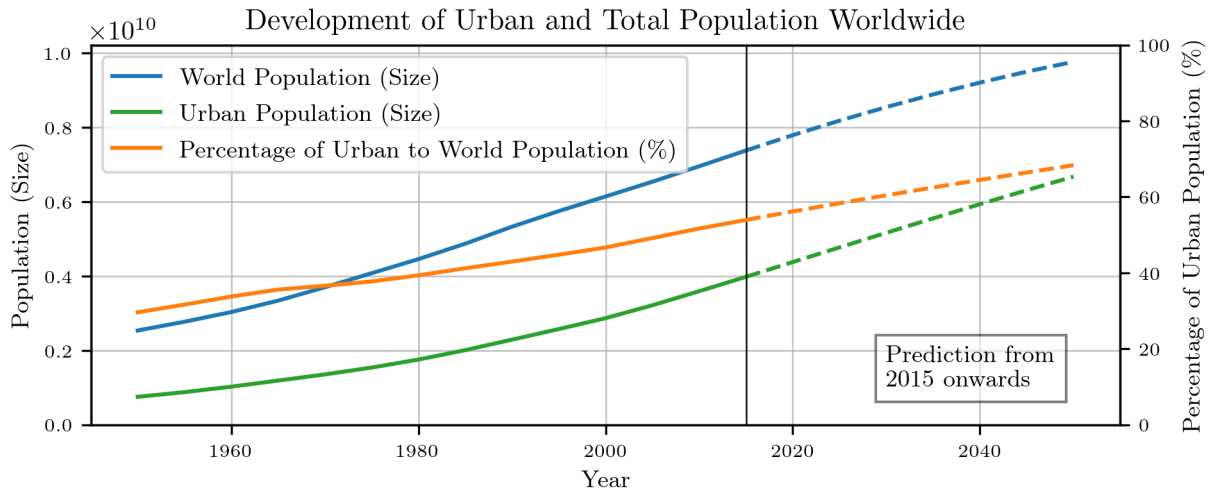


Figure 1: Development of urban and total population worldwide from 1950 to 2050 (United Nations 2018) (own illustration). The data from 2015 to 2050 are predicted. The world population is continuously growing and is predicted to be near 10 billions in 2050 (blue line). Simultaneously, the population living in urban areas is constantly increasing (green line). The share of the people living in urban areas (orange line) rose from around 30% in 1950 to 54% in 2015 and is expected to reach around 68% in 2050.

Urban areas are indeed one of the most important and complex ecosystems on our planet: They interact with their surroundings in many ways and have a strong influence on environmental variables (Gamba 2013). The ongoing urbanization increases the dependency on energy, forms socio-spatial inequalities and is a concern regarding sustainability and environment issues (Barthelemy 2016). Many effects like the change of previous surface covers to impermeable materials, a concentration of activities that create pollution of the surrounding soil, air and water come typically along with the growth of cities. This results in soil degradation and has an influence on the local and global climate (Bechtel, Alexander et al. 2015; Gamba 2013). So on one side, cities have a high impact on the bio-physical environment and are (among other causes) a major driver “of global environmental change” (Bechtel, Alexander et al. 2015, p. 200). But on the other side, they are particularly vulnerable to those changes and for example affected by increasing air temperatures, rising sea levels and heavy rain events (Bechtel, Alexander et al. 2015; Hunt and Watkiss 2011). Therefore, it is of high relevance to gain consistent knowledge about urban areas and their resulting structures, since policy makers look for new paradigms in urban planning (Barthelemy 2016). Hence, the development of urban agglomerations and the driving forces that shape them are an object of research in multiple scientific disciplines.

Human settlements do not emerge randomly, they are initially founded for spatial, political, social, economical or cultural reasons and functions (Gaebe 2004). Yet, their formation and development can be ambiguous and range from the un-regularized interaction of their individual

inhabitants up to a completely planned and supervised process led by superior authorities. This results in a broad variety of shapes and forms of built structures a city is composed of. It is this physical appearance that determines how cities are mainly perceived (Taubenböck, Debray et al. 2020). The urban environment is experienced through the composition and interaction of many individual elements. Streets of different shapes and sizes divide street blocks (Oliveira 2016) and define accessibility within a city. Residential buildings, working spaces, commercial buildings and factories form the urban structure and are a visible and material expression of economic and social structures and different forms of living. The morphological appearance reflects only a small part of the complex system of a city but still determines many routines and functions of it (Taubenböck, Debray et al. 2020). So is it the form that shapes our modes of living, or vice versa, is it “form (ever) follows function” (Sullivan 1896, p. 5), the famous mantra of modern design theory? Either way is true, but the (built) environment is not understood as given, it is something that is created by us humans (Marcus et al. 2017). Further, nothing is ever constant in a city. It is continuously changed, reshaped and adapted by the demands and needs of their inhabitants. This process is not randomly driven but rather reflects the political, social, economical and cultural key events that occur throughout human history (Fleischmann, Feliciotti and Kerr 2021). Changes can be registered in the constantly evolving urban form of buildings, plots and streets but also in the functional, social and cultural relationships. Over time, the functions a city fulfills can shift and lead to a change of status, importance and attraction. Yet their physical form is usually more enduring than their primordial urban functions (Gaebe 2004). But why do spatial patterns emerge exactly on this physical registration plate of urban processes?

The urban morphology can be seen as the product of “a collective behavior of constituents coupled to each other” (Goh et al. 2016, p. 1). Those specific configurations are the unique result of a long term shaping process and reflect in a condensed way all those processes that are part of the city. Despite the great variety and differences of urban forms, statistical regularities can be observed (Barthelemy 2016) and similar patterns can be found throughout different cities (Faßmann 2009). The reasons for this universality are shared fundamental processes that result in a spatial organization of physical form and functions (Barthelemy 2016) and a collective behavior with statistical regularities produced by systems with many components (D’Acci 2019). Hence, exploring and analyzing these patterns helps to find and understand the profound underlying processes that shape a city. The exploration of urban structures which are a result of fundamental social processes allows to build a foundation for further analyses in many different scientific fields, ranging from the understanding of city development to finding the driving social and economic forces of a city. Moreover, “a better understanding of urban forms helps to make effective provisions [...] on urban form issues, which is among the most urgent needs [...]” (D’Acci 2019, p. 3). Approaching these intra-urban configurations with a scientific point of view leads to the field of geography in which this master thesis is written.

Geography aims to analyze spatial structures and constructions to find regularities and arranged patterns and to gain knowledge about ‘space’ for a profound understanding of today’s world (Gebhardt, Glaser et al. 2020). As a subfield of geography, urban geography analyzes the

environment of cities on different scales, ranging from a single object to districts, cities and city systems. The scientific field moves between an idiographic urbanism which explores and describes specific, unique cases and a nomothetic urbanism whose goal is to find and generate universal statements about the structure and development of urban settlements (Faßmann 2009). Uncovering the general laws that underlie the shaping process of cities and lead to the formation of re-occurring patterns induces a certain degree of generalization.

A possible way to describe spatial patterns is to construct regions of homogeneous characteristics and separate them through boundaries. The earth shows only few clear natural borders that are experienced by humans in their daily lives (such as coastlines, rivers and mountain chains). This means that all efforts to construct regions are intellectual abstractions of the reality, constructions through the scope of theoretical perspectives and methodological techniques. Geographic regions are not exact, neutral or objective. They are contextual, can change throughout the history and underlie constant negotiations. Yet, this form of structuring and classifying space has the purpose to give orientation (Gebhardt, Glaser et al. 2020) and reduce the complexity of the world surrounding us (Fleischmann, Feliciotti, Romice et al. 2021). Through its simplification, it allows us to understand better the basic elements and investigate their spatial arrangement on a more abstract level. The creation of spatial order allows the development of models and formation of paradigms. Those paradigms define not only the research questions and theory but also the methods used to answer them (Reuber and Gebhardt 2020). Urban geography started to investigate patterns of urban form with comparative studies, yet they largely relied on qualitative approaches and set the focus on functional aspects of the city. But this left out the physical environment, the “setting of all lived urban experience” (Fleischmann, Feliciotti and Kerr 2021, p. 2). Further, it hinders the development of replicable results and scalable data-driven approaches. These shortcomings are slowly overcome with a range of new methods and statistical-mathematical and analytical approaches developed in the previous years (Gebhardt, Glaser et al. 2020).

Mathematical methods have been proven to be useful in the investigation of cities to generate scientific theories of urban forms in a formal, deductive and on human behavior based manner (D’Acci 2019).

“Cities are ultimately the result of human behavior, whose understanding might be reductive if solely framed within a strictly ‘mathematical’ confine.” (D’Acci 2019, p. 1)

Yet, D’Acci sees mathematics as the universal language between humans and nature to understand the complex phenomena of this world. A recent game changer are the huge amount of data that come with the continuing technological developments (Barthelemy 2016). Improvements of sensors and capturing systems, the increase of storage capacity and computational processing power allow more sophisticated data driven analyses. Especially in the field of earth observation, huge progress was made in capturing images of the earth surface on a global scale and in increasing resolution.

This global 'objective' view by the lenses of spectral satellite measurement systems has been proven to be very successful in mapping the surface of the world. In the past decades, the number of earth observation satellites and the amount of satellite imagery constantly increased (see figure 2).

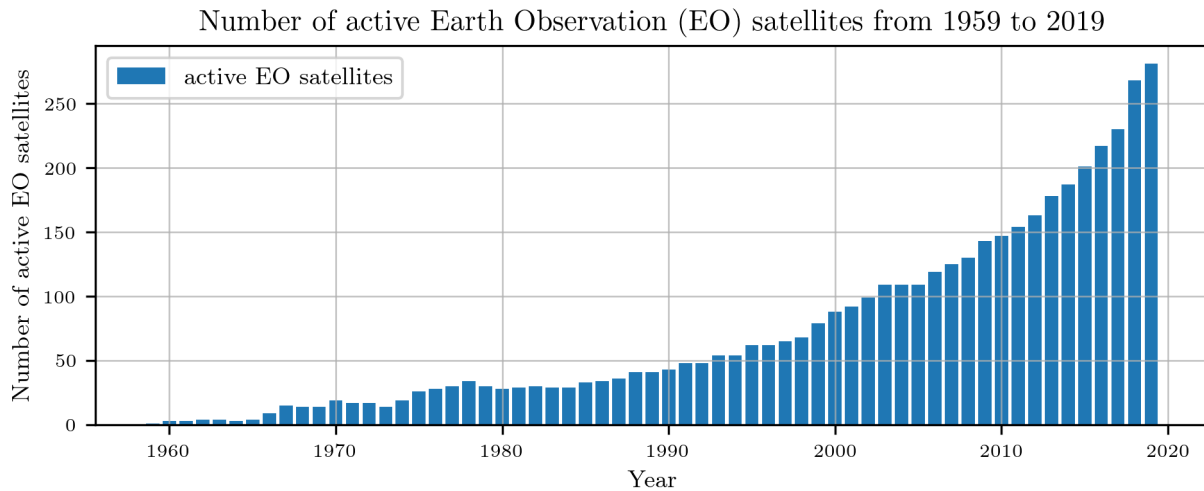


Figure 2: Number of active Earth Observation (EO) satellites from 1959 to 2019 (World Meteorological Organization 2022) (own illustration). There is a continuous increase of active EO satellites which capture data of electromagnetic reflections of the earth. Especially in the past 20 years the amount of new satellites increased exponentially.

Further, the improved resolution made it possible to capture smaller scale structures which are especially interesting for the exploration of cities. Government-funded and free satellite imagery like Landsat and Sentinel-2 increased the availability and use in scientific studies. The progress on the field of computer vision and machine learning opened up new possibilities of applications and analyses in remote sensing. This provides the opportunity for unprecedented analytical views on cities on a global scale. Everyone has very 'subjective' impressions of a city. With a systematic data-driven methodology it is now possible to conduct comparable studies on a global scale. Yet, the exploration of urban configurations with Earth Observation (EO) data just started in the recent years (Barthelemy 2016).

This master thesis proceeds as follows: Chapter 2 provides a theoretical background and outlines the scientific gap in research. In chapter 3 the framework and research questions to explore city patterns are formulated. Chapter 4 presents the methodology which includes explorative approaches of unsupervised learning. The main results are reported in chapter 5 and discussed in chapter 6. A conclusion is given in chapter 7.

2 State of the Art

This chapter begins presenting the challenge of finding a universal definition of a city. It continues with different approaches in urban geography, each examining urban areas with their respective point of view. With this background, a brief look at structural city models throughout the history is taken. Diving into the more specific field of urban morphology, in which this master thesis is located, the current approaches for the investigation of urban structures are introduced and the scientific gaps are presented. In conclusion, a scalable, comparable and data-driven research framework is constructed to investigate the intra-urban configurations of cities without prior knowledge on urban form.

2.1 The Grasp of a City

What is a city? The simplest questions are often the most difficult to answer. In today's language, fundamentally different terms and definitions of a city can be found. In historical terms, a city received the privileges of borough rights by the current ruler. In Germany, those rights are now given by the federal state authority based on the number of inhabitants, type of settlements and economic power. This leads to other possible definitions of cities regarding statistical numbers such as defining and exceeding certain demographic thresholds like the number of inhabitants or population density (Faßmann 2009; Heineberg 2017). These thresholds cannot be scientifically justified and vary in space and time, yet they allow to measure a certain perception of what is regarded as a city. Other approaches define and classify cities regarding their functions, sociological structure, economic power, centrality and infrastructure development (Faßmann 2009; Heineberg 2017; Gaebe 2004). These manifold domains are all part of the phenomenon 'city' and make it difficult to find a common understanding.

Yet, cities and urban settlements differentiate from rural areas through a row of the above mentioned characteristics. But where does a city begin and where does the rural space end? There are transitions between urban and rural settlements which make a distinct differentiation impossible (Gaebe 2004). Also Gebhardt and Reuber (2020) question the dichotomous thinking of 'urban vs. rural' categories that influenced geographical debates in the previous decades. They state that the distinctions are vanishing nowadays which makes it even harder to find clear boundaries that enclose a city.

Therefore a 'city' cannot be defined exactly, neither in the field of urban geography nor in other interdisciplinary fields, and especially not on an international or global scale (Heineberg 2017). Attempts of finding a clear definition of a 'city' will only produce unclear results (Faßmann 2009). The criteria of defining a 'city' vary in different cultural areas across the earth, dependent on their specific status of development (Heineberg 2017). Finding a definition of the term 'city' that is valid for all times, cultures and regions remains fiction and therefore can only be very superficial (Faßmann 2009). Subsequently, there are also no universal definitions of the terms 'urban' or 'urban population' in the scientific discourse (Taubenböck, Weigand et al. 2019).

There might not be a universal valid term for defining and explaining such a complex system like a city. Nevertheless, everyone experiences a city through the interaction of all its agents forming it. A city is not homogeneous, it can be differentiated by its physical, social and functional structure (Faßmann 2009). On smaller scales, those characteristics can be aggregated to form residential, economical and social intra-urban sub-units (Gaebe 2004). Even though the same problem of defining boundaries and thresholds to form these units arise, they form specific patterns that shape and differentiate the city. These spatial patterns are the result of the organization of a city and reflect their “own topographical, historical, economical, and political environments” (Goh et al. 2016, p. 1).

2.2 Approaches in Urban Geography

As complex and ambiguous cities are, as many approaches and paradigms exist to investigate them within the field of urban geography. Throughout the literature, the typification and classification of cities and their elements can be found as an attempt to simplify the vast variety of urban settlements. Each approach investigates cities with different objectives, methods and on different scales. Even though there is no clear separation between the approaches and there is a variety of intersections regarding the methods and considered variables used in scientific research, it is possible to roughly differentiate the following paradigms:

The *cultural-historical* and *morphogenetic approach* analyzes and investigates the plots and elevations of cities. The focus is set on architectural and physical structures, but also on the historical development and historico-genetic deduction of a city (Faßmann 2009). The physical form is seen as a registration plate of the social history, as being shaped by functions and sociological phenomena. Hence, cities or their elements can be classified by their time of foundation and their different historico-cultural impact on the physical structure (cf. Schlüter 1899; cf. Hassinger 1916; cf. M.R.G. Conzen 1981).

The *comparative* and *cultural-genetic approach* investigates urban structures in different cultural areas. It is often combined with the cultural-historical and morphogenetic approach, exploring the genesis of urban form in different cultural areas. The cultural-genetic approach assumes influential factors that force a homogenization of cities and allow a designation of *cultural areas* (cf. Passarge 1930; cf. Kolb 1962; cf. Hofmeister 1996). It is assumed that for whole regions and continents a dominant *culture* exists which has an influence on the society and urban areas. Cities are supposed to have more similarities than differences inside their respective cultural areas due to external and political influences (Gaebe 2004). But it is not the cultural area itself that shapes a city, it is the result of economic, social and political development processes. Yet, these processes are influenced by social constructs and underlie cultural traditions. In reality, these cultural areas can neither be defined exactly, nor are they homogeneous. Therefore this approach is rarely used due to the strong essentialistic interpretation of the term culture (Faßmann 2009).

Functional approaches deal with the non-visible functions of a city which include for example commerce, service facilities and cultural linkages (Gebhardt, Glaser et al. 2020). These functions can often be measured via statistical metrics. A typification of cities can therefore be performed through the scope of political, cultural and economic or infrastructural functions.

Other paradigms are the *social-geographic urbanism* which looks at the socio-spatial structure of cities (cf. Park, Burgess and McKenzie [1925] 1967; cf. Hoyt 1939; cf. Harris and Ullman 1945) or the *behavioral urbanism* to study interactions between the physical environment and humans. Approaches that can be more intuitively understood are classifications and typifications of cities based on statistical values, for example the city population or size (cf. Taubenböck, Weigand et al. 2019).

The term of '*urban morphology*' is as ambiguous as the grasp of a city. Again, it is difficult to find common definitions and approaches. The word '*urban*' originates from the Latin word '*urbs*', the '*city*', and '*urbanus*', which means '*belonging to the city*'. Johann Wolfgang von Goethe created the word '*morphology*' with the ancient greek suffix '*logía*' as '*the study of a subject*' and '*morphé*' as '*form*' (cf. Goethe [1817] 1954; cf. Goethe [1817] 1995). So '*urban morphology*' as the '*study of urban forms*' focuses on the physical elements that shape and structure a city (D'Acci 2019). The goal "[...] is to identify the repeating patterns in the structure, formation and transformation of the built environment to help comprehend how the elements work together [...]" (Kropf 2014, p. 41). The fundamental elements in urban morphology include a range from streets and plots to buildings and their aggregation to urban tissues (Oliveira 2016; Fleischmann, Feliciotti and Kerr 2021).

In this field of research, multiple influential schools of thought investigate the urban morphology, each of them with different approaches and objectives. The *Italian School*, founded by Muratori in the 1940s and later continued by Caniggia and Maffei (1979), stands for a process-typological approach with a historical and architectural point of view. It is an observation on the level of individual buildings with a strong focus on Italian cities such as Venice. A building can only be identified in the context of its neighborhood (Oliveira 2019). With this approach Muratori (1959) created the concept of urban tissue (Sadeghi and Li 2019). The *British School*, centered around M.R.G. Conzen (1960), follows a historico-geographical and morphogenetic approach to understand the process that creates patterns of building form and landuse. It is an analysis combining town plans, building forms and functions through different periods of time. The work was carried on by Whitehand with the *Birmingham School*. The Versailles National School of Landscape Architecture, considered as the *French School*, was founded by the architects Philippe Panerai, Jean Castex and the sociologist Jean-Charles Depaule. They provide a historical description of French cities on a larger scale (compared to the British and Italian School) and analyze urbanization processes and urban design in the context of social practices (cf. Castex, Depaule and Panerai 1977).

In conclusion, many different scientific approaches and paradigms exist in urban geography. Paradigms are neither correct nor false, they denote different research questions, methodological approaches and interpretations (Faßmann 2009). The schools of thought in urban morphology are largely driven to understand building types on a fine-grained scale within an architectural background, limited to a few cities in their local context. They describe and explain urban form through their objective experience and thus with a cultural bias. To overcome these shortcomings with a global empirical approach, this master thesis aims to examine urban form on a coarser scale and thus not on the level of individual buildings. With this background, this master thesis takes a brief look at different structural city models throughout the history in the following sub-chapter.

2.3 Structural City Models

The previously mentioned paradigms and theoretical possibilities of city typification provide a framework for the analysis of intra-urban patterns. City models try to describe a city and to understand the underlying processes that form and shape it. They can also be used in planning and development to pursue and implement certain visions of a city. City models are developed on a high level of abstraction to generalize and reduce the ambiguity of these complex systems. Homogeneous regions sharing a set of similar characteristics can be defined and create spatial-arranged patterns. The implicit idea is that intra-urban patterns are the result of theoretical principles of order and not produced randomly (Faßmann 2009). The concept of these models is to formalize these spatial patterns and understand the rules that lead to their formation. Considering only a few variables will never be enough to capture the complex reality of a city, yet it allows the quantitative analysis and large-scale characterization of urban systems (Barthelemy 2016). Dependent on the research question, the respective framework is adapted to a specific scale and hence cannot provide a “grand unified theory of cities” (Barthelemy 2016, p. 27).

Early historical city models are known since the reconstruction of the ancient Greek city of Milet based on the *hippodamic scheme* after its destruction in 497 BCE. The model named after Hippodamus was not only used in antique Greek cities, but also throughout the Mediterranean area with the planning of Roman cities and up to the foundation of colonial cities in Latin America in the 16th century (Heineberg 2017). Another well known example of a historical city model is the *ideal city* developed by Leon Battista Alberti in the 15th century during the renaissance. Its key element regarding the symmetrical shape of physical structures can still be found throughout European cities like Zamość and

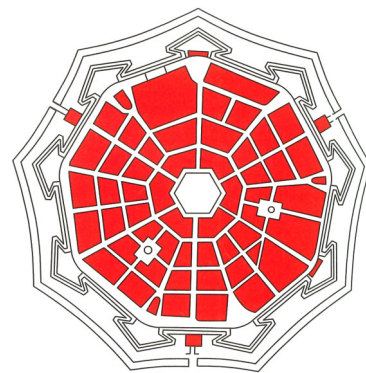


Figure 3: Palmanova, considered as an 'ideal city' in the Italian renaissance. The plot of this planned city shows a radial street network, a circular fortified wall and symmetric built elements. Reprinted from Heineberg (2020, p. 850), adapted after Grassnick and Hofrichter (1982).

Palmanova (see figure 3). Concepts and ideals for cities continued for example with the *Garden Cities* to create green belts around cities by Howard ([1902] 1944) and the utopian planning of the *Ville Contemporaine* by Le Corbusier (1924). Up to the current day, models play an important role in urban design and planning.

Other city models deal less with the planning of a city but more with the modelling, and therefore understanding, of its status quo. First studies with a mathematical approach investigated the relationship of space and economic factors in cities (Barthelemy 2016). Mentionable is Thünen (2014) with his concentric ring model of agricultural land use published 1826 in *Der isolierte Staat*.

The *central place theory* developed by Christaller (1933) deals with hierarchical city systems and the theoretical spatial and statistical distribution of different city sizes and their functions. Christaller's theory models the importance of a city and the influence on its surroundings and is still used in regional planning policies for a comprehensive supply with goods and functions. Based on this work, Lösch (1940) tries to explain with the *theory of market places* the spatial distribution of production sites and their product specialization. The concept of central places finds also application in the analysis of intra-urban patterns. For example Alexander (1988) finds *artificial cities* to be build as a hierarchical tree diagram (see figure 4, Diagram C and D), while *natural grown cities* are structured as a semi-lattice (see figure 4, Diagram A and B). But the description of Alexander goes beyond that of central places, it is a general idea on how elements of a city interact with each other. A natural city consists of a set of units which are in complex relationships to each other. Artificial cities fail to be conceptualized this way.

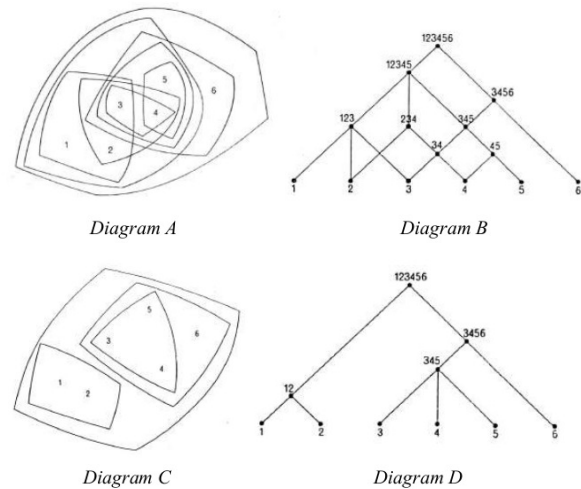


Figure 4: Diagramm A: Set of units with boundaries (Semi-Lattice); Diagramm B: Units in an arranged order (Semi-Lattice); Diagramm C: Set of units with boundaries (Tree); Diagramm D: Units in an arranged order (Tree). *Natural grown cities* form a semi-lattice while *artificial cities* form a tree. Reprinted from Alexander (2015, pp. 6, 8).

An early work to mention in the field of urban morphology is '*Deutsche Stadtanlagen*' published by Fritz (1894), a comparative analysis and classification of 200 German cities based on their respective city plans. Schlüter (1899) further developed the work of Fritz in his article '*Über den Grundriss der Städte*' including the identification of different units that form the city center. The historico-geographical approach was later consolidated by Conzen (1960) who provides with his work a framework for the investigation of urban forms (Oliveira 2016). For Conzen the climax of the investigation on urban form was the distinction of areas with homogeneous morphological characteristics (Whitehand 2001).

A model regarding transportation is for instance the *star theory* by Hurd (1903) which describes the star-shaped growth of a city along its main infrastructure routes in the age of trams and before the widely use of cars. It is seen by Hofmeister (1996) as the predecessor of the models developed by the Chicago School around Park, Burgess and McKenzie.

The *Chicago School of Social Ecology* developed three classic city models in the beginning of the 20th century. The fundament for these models was the city of Chicago, which had to deal with an increasing number of social problems during this time. The School of Social Ecology focuses on the society and its living environment. The three classic models explain the development and arrangement of functional and socio-spatial intra-urban units. The principles are based on Charles Darwin's inspired idea of competition between species. However, instead of 'surviving' it is about the competition on available space on the free land market. This competition enforces the division of labour and, if it endures long enough, forms homogeneous units inside a city (Faßmann 2009). The city is composed by those homogeneous units which are the object of interest in Social Ecology.

The first model was the *concentric zones model* developed by Park, Burgess and McKenzie ([1925] 1967). It represents the ideal state of American cities between the two world wars. According to the model, the city develops over time different homogeneous zones surrounding the city center. In the middle of the city, the Central Business District (CBD) is surrounded by a mixed commercial/residential zone, a working class residential zone, a mid-class residential zone and a commuter zone (see figure 5). Quinn (1940) criticizes the model of Park, Burgess and McKenzie for the predicted symmetry and homogeneity of the assumed zones that do not exist in reality.

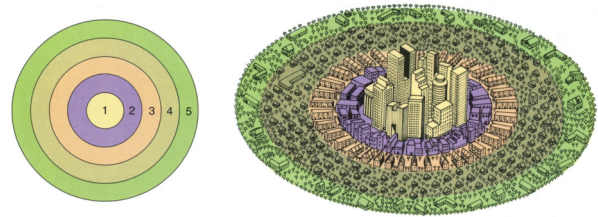


Figure 5: Concentric zones model developed by Park, Burgess and McKenzie ([1925] 1967) and reprinted from (Rubenstein 1999, p. 453). The CBD (1) is surrounded by a zone of transition (2), a working class residential zone (3), a mid-class residential zone (4) and a commuter zone (5).

The second model is the *sectoral model* developed by Hoyt (1939). It is based on the empirical research on rental-fees of 30 United States (US) -American cities (Heineberg 2017). The model divides the city into a pattern of relatively homogeneous sectors emerging from the city center (Heineberg 2017) (see figure 6). Hoyt confirms the assumption of Park, Burgess and McKenzie that the social gradient

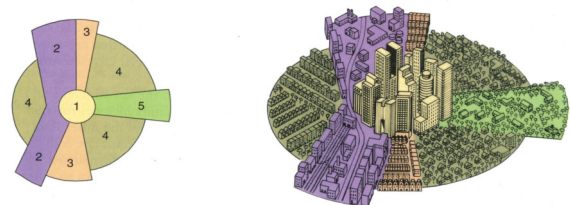


Figure 6: Sectoral model developed by Hoyt (1939) and reprinted from Rubenstein (1999, p. 453). From the CBD (1) emerge the industry (2), low-class residential (3), middle-class residential (4) and high-class residential (5) in a sectoral shape.

is increasing from the city center to the periphery. According to Heineberg (2017), the sectoral model works well for working class settlements and industrial areas. Even though the model is better adapted to the US-American society, it is still far away from reality (Hofmeister 1996).

The third important model of the Chicago School of Social Ecology is the *multiple nuclei model* by Harris and Ullman (1945). It assumes multiple homogeneous settlement nuclei inside a city (see figure 7). It hypothesizes that the growth of a city leads to an increasing specialization and number of nuclei. This model fits better to the reality of city structures and moves away from the strict geometry of the concentric zones or sectoral model (Faßmann 2009). Furthermore, the model paves the way for the polycentric city, from which it is only a small step to multi-centered city regions and the *central place theory* (Zonneveld 2005).

The three classical models describe the situation of US-American cities in the period between the world wars, before the process of sub-urbanization started. A combination of two approaches, the functional and the socio-spatial, are used for the typification of intra-urban structures (Heineberg 2017). The models are criticized for deficits in theory and a lack of empirical verification of their claimed universality. They are not sufficient to explain the increasing diffusion of city boundaries and intra-urban fragmentation of specialized functions (Heineberg 2020). Thus they can only make an explanatory contribution in capitalistic states with a free market economy and little influence by city planning (Heineberg 2017). Despite the critics, they provide a starting point for the development of newer, more complex models in other cultural regions. Especially the combination of all three models has been shown to be relevant. Furthermore, they fulfill an important didactic function for understanding and visualizing urban structures (Heineberg 2020).

The models of *social area analysis* which follow up the concept of *natural areas* are seen in contrast to the above mentioned models of Social Ecology (Hofmeister 1996). These natural areas (cf. Hatt 1946) are physical and morphological boundaries or barriers that separate areas from each other. This causes a development of similar behaviors and a residential population with cultural, social and economic homogeneous characteristics. The social area concept was further developed by Shevky and Williams (1949) and Shevky and Bell (1955) by considering statistical features of social, economic and ethical status for the formation of units. Of course, the selection and categorization of those variables to map the reality offers opportunity for criticism. The antinomic models of social area analysis and of the Chicago School of Social Ecology reflect two

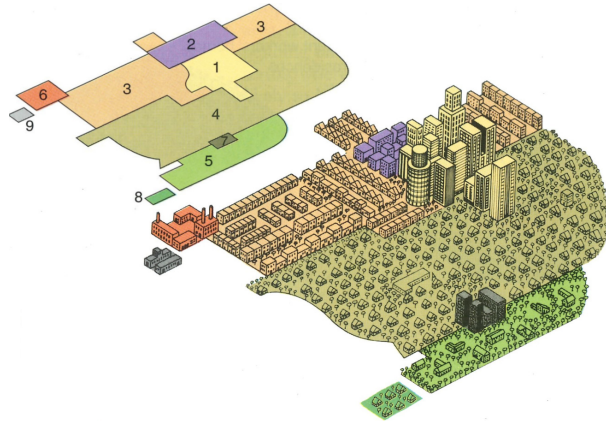


Figure 7: Multiple nuclei model developed by Harris and Ullman (1945) and reprinted from Rubenstein (1999, p. 454). Over time, homogeneous nuclei develop within a city. A distinction is made between the CBD (1), light manufacturing (2), low-class residential (3), middle-class residential (4), high-class residential (5), heavy manufacturing (6), outlying CBD (7), residential suburbs (8) and industrial suburbs (9).

main ideas in urban geography, whether the 'physical form defines the functions of a city' or 'the functions of a city define the physical form'.

Other models like the *land use theory* and *bid rent models*, as a part of urban economics, consider prices for land, rental fees and transportation costs and are influenced by the models of Thünen (2014) and Christaller (1933). The bid rent models explain different land use in countries with a free market, driven by the overbidding of competing commercial, residential and industrial agents. The restriction of available land, which is always the case in central areas, combined with an increasing demand will rise the ground rents and exclude less bidding agents (Heineberg 2017). An example is the monocentric city model developed by Alonso (1964) with a CBD in its center and functions modelling the price and supply of real estate. Further models by Mills (1967) and Muth (1969) were aggregated and formalized by Wheaton (1974) in the so-called *Alonso-Mills-Muth model* and include transportation costs and the value of alternative land uses (Kulish, Richards and Gillitzer 2012). Prices and functions determine the urban form. This results in the construction of tall buildings with small units close to the city center and more spacious buildings with smaller height further away from the city center. Criticized is the use of a monocentric model which often does not represent today's city structures (cf. Adolphson 2009).

The cultural-genetic approach dates back to Passarge (1930) who differentiated between eight specific cultural city types in his book '*Stadtlandschaften der Erde*'. It is based on large *cultural areas* for which a respective specific cultural-genetic type exists (Heineberg 2017). The concept of cultural areas was further developed by Kolb (1962) who created ten different cultural areas and Huntington (1993) with his classification of *civilizations*. Hofmeister (1996) developed 12 cultural city types in his book *Die Stadtstruktur*, first published in 1980. He modified the intercultural comparison of city structures also with the use of the historico-genetic approach. Models of the pre-industrial city (Sjoberg 1960) to the industrial (Wirth 1938) and the post-industrial or after Vance (1971) post-capitalistic city are used to compare and determine specific phases of city development on a global scale (Hofmeister 1996). It is based under the conception that, even though processes of urbanization and internal differentiation of cities are similar, each culture provides a different starting point and premises. Furthermore, the reaction of the population on similar processes result in different outcomes (Hofmeister 1996). Hofmeister concludes in defining the European, Russian, Chinese, Middle Eastern, Indian, Southeast Asian, Tropical African Latin American, Anglo American, South African, Australian-New Zealand and Japanese city. But again, like many approaches within the cultural-genetic paradigm, those models are subject to controversial discussions (see cultural-genetic approach in chapter 2.2).

Structural city models develop social and economic functions, regionalize by aggregating similar elements to homogeneous areas, build hierarchies and typify, all with the goal to either explore and explain certain phenomena or provide planning schemes of a city. They largely rely on subjective ideas or qualitative data that are hardly scalable for a global approach. The models are based only little on consistent empirical data and are hardly comparable (Taubenböck, Debray et al. 2020). There is a lack of consistent parameters and standardized methods to characterize

cities on a global scale (Bechtel, Alexander et al. 2015). The scarcity and inconsistency of available data results in an absence of globally comparative studies (Taubenböck, Debray et al. 2020). After this broad overview on structural city models, the master thesis continues exploring the field of urban morphology for data-driven approaches that go beyond the scope of the ideas presented so far.

2.4 Exploring Urban Morphology

Urban morphological approaches compare and examine patterns of spatial form. It is a study of urban tissues which are areas with a homogeneous morphology and functions (see chapter 2.2). Cities can be seen as composed by a set of different physical elements. It is these fundamental elements that are similar in each city, yet their diverse composition creates different patterns and forms distinct urban regions (Oliveira 2016). The complexity of cities is hence reflected by the unique configuration of their urban patterns (Taubenböck, Debray et al. 2020). To deal with the complexity of cities, the basic elements can be aggregated and a hierarchical view of cities can be constructed (Oliveira 2016). Different attempts to classify the basic elements of a city (such as a CBD or residential areas for example) and to arrange them to structural models or hierarchical views have been conducted in the past (see chapter 2.3). The urban morphology can be explored on the scale of individual building types up to the aggregation of larger areas and cities. The rise of geographic data science in combination with the increasing availability of geodata allowed data-driven morphometric studies (Fleischmann, Feliciotti, Romice et al. 2021).

The study of urban morphology through the measurement of metric parameters focused for a long time on connectivity assessments such as the *space syntax* developed by Hillier (1996). Representations of space are retrieved from street networks and their spatial relations are investigated. *Place syntax* added accessibility and socio-economic attraction measurements to the parameters (Stähle, Marcus and Karlström 2005). A study of Marcus et al. (2017) performs a step towards an architectural model of pure spatial form by combining the components of space and place syntax. Yet, this study and the approaches using space and place syntax focus mainly on functional variables and street networks while leaving out most of the physical characteristics of buildings. Berghauser-Pont and Haupt (2010) set with the creation of *space matrix* the focus on the density of urban blocks. Based on those variables Berghauser-Pont and Olsson (2017) created a building typology and clustered neighborhoods with multiple buildings, still leaving aside the concepts of geometry and urban fabric. Moving towards building characteristics, Dibble et al. (2019) derive morphometric parameters of urban areas enclosed by main streets and measure their similarity with a hierarchical clustering. On the level of individual buildings Fleischmann, Feliciotti and Kerr (2021) analyze in their study different historical periods of city development on the base of morphometric characters. In a next step they develop the morphometric characterization of buildings to a numerical taxonomy of urban form (Fleischmann, Feliciotti, Romice et al. 2021). With a hierarchical clustering of individual buildings they are capable to derive urban types. It is the development towards a reproducible and unsupervised numerical classification of urban patterns. Still, the use of fine-grained data on the level of

individual buildings impedes these morphometric studies for a global comparative approach.

More general classification schemes focus on a medium scale and can be generated based on EO data such as Urban Structural Types (Lehner and Blaschke 2019) or the Local Climate Zones (LCZs) developed by Stewart and Oke (2012). The exploration of urban morphology based on global available EO data will be the subject of the following subchapter.

2.5 Urban Morphology through the Scope of Earth Observation

The rise of free EO data promoted their increased usage in scientific studies. The global coverage allows a consistent generation of knowledge, decoupled from manual data acquisition as used previously in urban morphological studies. It allows to continue the mathematical approaches in urban morphology. Yet, the use of EO data for a worldwide investigation of urban areas started in the recent years.

Early studies attempted to generate knowledge on human settlements based on EO data with land cover masks mapping urban areas (cf. Small and Sousa 2016). Global products were developed from spectral EO imagery such as the 'MODIS 500m Map of Global extent' (Schneider, Friedl and Potere 2010) or from spaceborne synthetic aperture radar (SAR) data like the Global Urban Footprint (GUF) with a resolution of 0.4 arcsec (~ 12 meter) (Esch et al. 2013). Yet, they provide only a binary mask without any differentiation, internal structure or texture of urban areas (Bechtel, Alexander et al. 2015). Urban footprints are used in studies on urban growth (Taubenböck, Esch et al. 2012) or landscape configurations (Taubenböck, Wurm et al. 2019) but are of course thematically constrained (Taubenböck, Debray et al. 2020). Also comparative analyses of urban morphology based on EO data are often limited only to a few cities (cf. Seto and Fragkias 2005; cf. Cai, Huang and Song 2017).

The challenges for an empirical investigation of cities and their intra-urban structures on a global scale are manifold. A common definition of a city is a difficult task (see chapter 2.1) and determining their spatial boundaries poses another challenge. Finding boundaries of urban areas is important more than ever since these administrative and political units of cities are rather arbitrary and do not represent their spatial extend adequately (Taubenböck, Weigand et al. 2019). Furthermore, the subjective cognitive perception of urban morphology makes it hard to find a global applicable classification system. Hence, there is clearly a need for a comparable framework to assess intra-urban structures and for a common definition of urban areas. Several studies overcame these challenges with the help of global EO data and development of classification frameworks on urban form.

In 2012, Stewart and Oke developed a land cover classification scheme, called the Local Climate Zones (LCZs), with the purpose of standardizing temperature studies in urban areas. To investigate urban meteorological effects such as the Urban Heat Island (UHI) effect, a consistent knowledge of site properties is necessary to standardize measurements and allow intersite

comparison, since the main causes of UHI are structural and land cover differences (Stewart and Oke 2012). The LCZs are defined “as regions of uniform surface cover, structure, material and human activity that span hundreds of meters to several kilometers” (Stewart and Oke 2012, p. 1884). The developed classification scheme takes several parameters into account such as the surface structure, surface cover, building materials and a row of further qualitative and quantitative attributes (Stewart and Oke 2012). The result is a distinction between 17 LCZ classes, from which ten classes are built types and seven are non-built types (see figure 8).

Even though the classification scheme was developed for UHI studies, it also provides a solid framework for the description of urban configurations. The LCZs provide comprehensive information on their physical characteristics such as urban structures and the urban fabric (Bechtel, Foley et al. 2015). Therefore, it is no wonder the LCZs soon found their usage in the field of urban morphology, as similar attempts of categorization have been made in the past (see for example the structural city models of the Chicago School of Social Ecology 2.3). Of course, like all classifications, the LCZs are restricted in their function (Stewart and Oke 2012) and imply homogeneous structures (Bechtel, Alexander et al. 2015). Yet, they rely on “universal, standardized and measurable parameters of urban form” (Taubenböck, Debray et al. 2020, p. 2) and hence allow a global comparative approach of urban morphology.

First attempts for mapping LCZs with the use of EO data were conducted by Bechtel, Alexander et al. (2015) which showed to be a challenging task due to the different spectral properties in the distinct regions of the world. They provide a conceptual classification protocol and publish results of tested cities in the World Urban Database and Portal Tool (WUDAPT) (Ching et al. 2018), but the community-driven mapping approach is not suited for large scale mapping.

After a series of further mapping attempts (cf. Yokoya et al. 2018; cf. Hu, Ghamisi and Zhu 2018) the work of Qiu, Mou et al. (2019) and Qiu, Schmitt and Zhu (2019) achieved a reliable mapping of LCZs on a global scale with a resolution of $100m * 100m$. By combining a recurrent network and a residual neural network (ResNet) the LCZ classification of Sentinel-2 (S2) multi-seasonal imagery achieved an accuracy of 86.7% (Qiu, Schmitt and Zhu 2019). Even though there is room left for improvement regarding the achieved accuracy and the limited separability of specific classes, it is the first global classification product with structural information on urban morphology that goes beyond a binary mask. With this mapping a comparison of urban morphology on a global scale is possible, yet an applicable delineation of urban areas is missing. The built LCZ classes 1-10 can somehow be seen as urban areas, however they do not define the border between a city and its surroundings.

The question where settlements change from urban to rural was tackled by Taubenböck, Weigand et al. (2019) with the development of Morphological Urban Areas (MUAs) based on EO data. They understand urban areas as “places dominated by the built environment” (Taubenböck, Weigand et al. 2019, p. 2) which include roads, buildings and other human-built elements. By combining GUF data, information on the building density and Open Street Map (OSM) data,

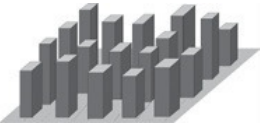
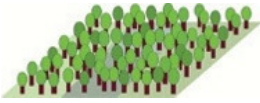

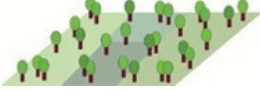

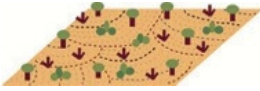


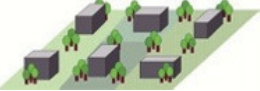
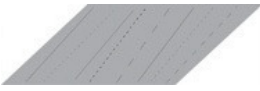
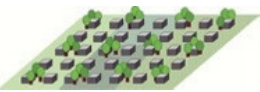

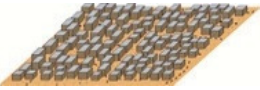




Built types	Non-built types
 LCZ 1 Compact high-rise	 LCZ 11 Dense trees
 LCZ 2 Compact midrise	 LCZ 12 Scattered trees
 LCZ 3 Compact low-rise	 LCZ 13 Bush, scrub
 LCZ 4 Open high-rise	 LCZ 14 Low plants
 LCZ 5 Open midrise	 LCZ 15 Bare rock or paved
 LCZ 6 Open low-rise	 LCZ 16 Bare soil or sand
 LCZ 7 Lightweight low-rise	 LCZ 17 Water
 LCZ 8 Large low-rise	
 LCZ 9 Sparsely built	Local Climate Zones Built types represented as LCZ 1-10.
 LCZ 10 Heavy industry	Local Climate Zones Land cover types represented as LCZ 11-17.

Figure 8: Local Climate Zones (LCZs) with 17 classes defined by Stewart and Oke (2012). The LCZs characterize specific configurations of the built-environment which is represented by LCZ 1-10 and the non-built environment represented by LCZ 11-17. Visualization of the LCZs reprinted from Stewart and Oke (2012, p. 1885).

the morphological settlement index (MSI) was developed to define a cut-off within a sectoral monocentric city model, determining the MUAs boundary (Taubenböck, Weigand et al. 2019). The method was applied to all cities with more than 300.000 inhabitants (representing the status in 2015, cf. United Nations 2015), resulting in 1692 MUAs worldwide (Taubenböck, Weigand

et al. 2019). This consistent method provides a comparable database on a global scale respecting the reality of urban extents compared to the arbitrary administrative units. This includes the merge of multiple cities to urban agglomerations, if their theoretical boundaries (based on the developed model) are overlapping. The usage of a monocentric city model limits the adequate delineation of polycentric cities, yet it is suitable enough for a simplified urban spatial model.

The LCZs describing urban morphological characteristics and the MUAs delineating city boundaries, both in a comparable way, provide a baseline for the investigation of urban morphological features across the world. Taubenböck, Debray et al. (2020) aim to examine the intra-urban morphological attributes and group cities with similar spatial configurations. They developed a framework for an empirical study based on the LCZs classified by Qiu, Schmitt and Zhu (2019) and the MUAs of 110 cities, a representative selection across the globe (Taubenböck, Debray et al. 2020). The authors understand the LCZs as the basic elements that form the urban morphological configuration of cities through their unique composition. Hence, looking for similarities and differences between cities can be conducted on the spatial and statistical distribution of the LCZs. Therefore, they cluster cities on the basis of the statistical proportions and of the spatial location of the individual LCZ classes inside the MUAs. In their study they create a feature space based on the LCZs inside the MUAs for each city and cluster the cities based on this feature space in a second step. Two methods for creating a feature space are applied:

- 1) A 17-dimensional feature space is created with the relative coverage of the 17 LCZs.
- 2) A 1700-dimensional feature space is created with the relative coverage of the 17 LCZs in 100 relative concentric zones around the city center.

Further, they took the size of the MUAs as an additional feature into account, resulting in four different feature spaces. With an unsupervised approach, a clustering with the k-means and expectation-maximization methods were performed and the optimal number of clusters determined via the gap statistic algorithm. As a result, an average number of seven clusters was found for all possible variations on the choice of feature space and cluster algorithm, even though the spatial compositions varied slightly (Taubenböck, Debray et al. 2020). Moreover, the seven city clusters found showed similarities with the theorized cultural areas from Huntington (1993). The authors conclude that the seven city clusters represent different intra-urban morphological configurations that can be found around the world. The city clusters form spatial arrangements that show partially similarities with geographic-cultural areas as proposed in the literature (see cultural-genetic approach in chapter 2.3). On the other hand, some clusters that span multiple continents suggest similar urban configurations that are not spatially restricted to the cultural areas (Taubenböck, Debray et al. 2020).

This study comes with uncertainties regarding the database and the selection of cities used for the analysis. The distribution of city sizes is not equal among the continents, since only a representative choice on the largest cities per continent is used in the study. Since the urban morphological configurations might be influenced by the city size, this could lead to a

bias regarding the spatial arrangement of clusters similar to the geographic-cultural areas. To overcome this, a following study with a larger database is of interest. The selection of the feature space is a reasonable, yet limited approach for the investigation of urban morphology. The statistical distribution of the LCZ classes describe the city as one object, but they do not reflect the spatial arrangement. Adding the geographic element of monocentric rings to the feature space brings in a spatial component, still the statistical distribution of the LCZs is the key feature inside those rings. Considering this, two cities might have a similar statistical distribution and thus a similar feature space, but the spatial arrangement of the LCZ classes can be fundamentally different. Further, the monocentric rings do not always represent the true spatial structure of cities. Even though the MUAs are developed on a monocentric city model, they result sometimes in the formation of urban agglomerations from merged MUAs (cf. Taubenböck, Weigand et al. 2019). In consequence, those merged MUAs are composed by multiple monocentric models and hence limit a comparison of cities based on a monocentric ring feature space (see for example the Pearl River Delta in Taubenböck, Weigand et al. (2019)).

Nevertheless, the study of Taubenböck, Debray et al. (2020) is a first step towards a consistent, scalable and unsupervised comparative approach of urban morphological configurations around the globe. They found spatial clusters of cities with similar urban configurations using a quantitative data-driven approach. An investigation on why the spatial patterns appear is promising to give interesting insights and might even lead to the development of new theories and paradigms in the field of urban geography. Due to the controversy about cultural areas, a comparison with geographical areas might be more expedient. So far, patterns of intra-urban configurations (represented by the LCZs) are investigated mostly in their statistical, less in their spatial composition. This master thesis aims to overcome some of these weaknesses with a new method and an extensive database.

3 Framework and Research Questions

In the previous chapter, a brief overview on the current state of the art was presented. New solid datasets like the LCZs and MUAs were created and provide the basis for a comparable research framework in the field of urban morphology. First approaches in a global comparison of morphological configurations were conducted and give hints on the formation of geographical clusters of cities with similar characteristics (cf. Taubenböck, Debray et al. 2020). Yet, the limitations of the previous studies call for a more sophisticated methodology and a larger database.

This master thesis aims to go one step further in the investigation of urban morphological configurations. The LCZs classification developed by Stewart and Oke (2012) is a good approach in the description of urban structures and can be seen as the basic units composing a city. This master thesis focuses not only on the statistical distribution of classes, but also on their spatial arrangement inside a city and the formation of morphological patterns. A morphological pattern can be seen as a re-occurring spatial arrangement of similar urban configurations, represented by the LCZs, that can be found in multiple cities. This results in a more significant meaning of morphological configuration, since it brings in the key component of any geographic analysis: *space*. Once patterns of morphological configurations are found, a comparative analysis of cities based on these results is more meaningful. With the LCZs mapping by Qiu, Schmitt and Zhu (2019) and the MUAs developed by Taubenböck, Weigand et al. (2019), this master thesis intends to apply new experimental approaches of unsupervised learning to investigate urban morphological configurations. The objective is a comparative analysis on a larger database of 1569 urban agglomerations which include all the MUAs provided by Taubenböck, Weigand et al. (2019).

Within this framework following research questions are formulated for this master thesis:

- 1) Is it possible to find patterns of urban morphological configurations based on the LCZs across the world?
- 2) And if yes, do cities form geographical clusters when comparing them based on the found patterns?

The framework of this master thesis is further expanded by two ablation studies. The formation of geographical city clusters is examined additionally by (i) clustering only the 110 MUAs used in the study of Taubenböck, Debray et al. (2020) and (ii) clustering all MUAs with the methodology developed by Taubenböck, Debray et al. (2020).

4 Methodology

This chapter begins with an introduction to the theoretical background on the methods used in this master thesis. In the second part, the data used for this work are presented. In the third part, the methods and analyses of this master thesis are explained in detail.

4.1 Theoretical Background

To begin the theoretical background, some general thoughts on the framework implementation with the use of machine learning methods are made. Following, the fundamental concepts of neuronal networks used in computer vision are explained. The theoretical background of the methodology is concluded with a short review of the state of the art methods on unsupervised image clustering.

4.1.1 Conceptualization

In the research framework, the aim for finding *urban morphological patterns* as re-occurring spatial arrangements is formulated. But from which perspective can this issue be tackled and how exactly is it possible to find those *patterns*? To answer these questions, some theoretical considerations have to be made.

A *pattern* found in the objects of investigation induces a certain degree of *similarity* among the elements belonging to it. Further, elements that are part of a found pattern are more *similar* to each other than they are to elements belonging to a different pattern. This leads to the necessity of making comparisons and comes to a result of *similarity* or *dissimilarity*. To do so, in a first step the characteristics of an object need to be described through elementary properties (Sonka, Hlaváč and Boyle 2008). Second, based on these properties a comparison and the search of patterns can be conducted.

It is clear, that a manual comparison for such a large database is not feasible. To process such large amounts of data and to eliminate the 'subjective' view and errors produced by humans, this task is best to be performed by a more 'objective' computer algorithm. The fast evolving field of machine learning analyzes data through mathematical algorithms to detect regularities or abnormalities, patterns and perform classification, clustering and regression tasks. Automatic pattern detection can be conducted in two ways:

(i) A syntactical approach describes a pattern through the inter-relationship of sub-patterns. With it comes the need for a formal grammar, a set of rules, which define the membership of a pattern. Within the framework of this master thesis, this could be achieved by creating a syntax including the type of LCZs as the sub-elements and the spatial organization as their relationship. But the problem of finding and defining patterns is just postponed to the element of sub-patterns. A city can be divided into subsets which can be divided further up to the atomic

elements of the single LCZ classes. Building manually a typology from those basic elements up to the aggregation level of cities is not practicable. A formal qualitative description of a pattern requires a profound empirical knowledge on urban morphological patterns which does not exist to this date. In conclusion, this approach is not suited without a prior knowledge on urban patterns.

(ii) A statistical approach calculates the likelihood for an object to belong to one or to another pattern. It has no pre-built rules like the syntactical approach, instead it is a mathematical function that assigns the object to a category, based on its properties. Considering the aim of this master thesis, the properties of cities (or subsets of cities) can be measured and patterns may be found through a mathematical model. This approach is better suited, since it does not require a prior knowledge on the searched patterns. Yet, it requires an adequate data description and careful model selection to achieve the desirable results. Therefore, a closer look at the possibilities of statistical pattern recognition is taken.

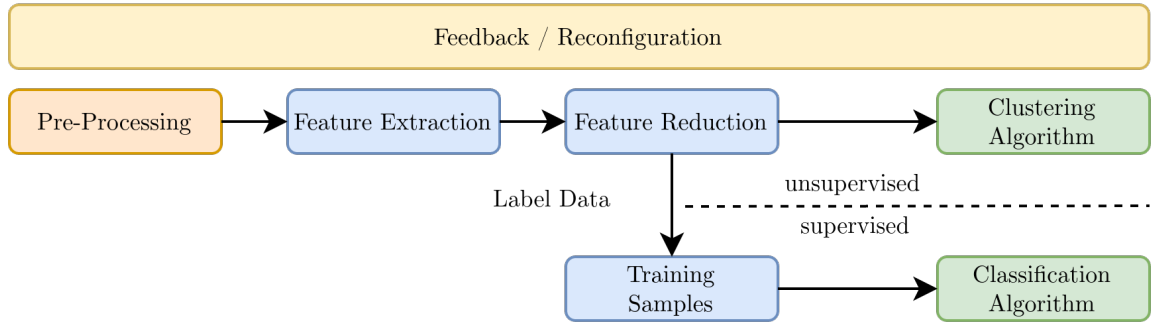


Figure 9: Typical workflow used in statistical pattern recognition (own illustration). It consists of data preparation (orange), preparation of the feature space (blue) and a statistical model (green). In the process of development, all components can be adapted (yellow) to achieve the desirable results. When building a framework for pattern recognition, all components underlie a constant adaptation and reconfiguration (yellow) achieve the desirable results. At the beginning, the data are pre-processed if necessary to a form suitable for machine processing. Features are then extracted and a feature space is built. Dependent on the data, a feature reduction may be performed to avoid the curse of dimensionality. If it is possible to label training samples, a supervised classification task can be performed. Else, an unsupervised clustering may be an option.

Figure 9 shows a typical workflow used in a statistical machine learning approach. In a first step the database is pre-processed to a form suitable for machine processing. Second, the properties that describe an object have to be measured and extracted. These selected properties, also described as *features*, form a *feature space* which is the basis for distinguishing and finding patterns (Beyerer, Richter and Nagel 2018). A feature reduction can be performed to remove redundant characteristics of the feature space (Géron 2020). The more dimensions a feature space has, the more data are necessary to avoid the danger of overfitting the model (Géron 2020). This phenomenon is called *the curse of dimensionality*. Since data are usually scarce or limited to a certain degree, the feature space needs to be adapted to the existing database. The feature space is the foundation on which the search of patterns is conducted. Generally, two main approaches exist in machine learning:

In supervised machine learning, the task is to assign objects to previously defined classes (patterns) based on similarity properties (Beyerer, Richter and Nagel 2018). Therefore, an algorithm is trained with labeled training samples to 'learn' and generalize the properties of the searched patterns (Géron 2020). A trained classification algorithm then can be used to detect and classify patterns it has 'learned' in unknown data. But again, it requires an a priori knowledge on patterns to label data and create a set of training samples (Sonka, Hlaváč and Boyle 2008). A formal definition is needed to collect samples and create a classification system of individual urban morphological patterns, which does not exist in the context of this master thesis.

This leaves the unsupervised approach in pattern recognition. It does not require training data and learns without supervision (Géron 2020). The goal is to group similar elements into clusters without prior knowledge on the resulting clusters. Thus, it allows a data exploration with a high impartiality and reduce possible biases. This method seems best suited for an explorative analysis to find urban morphological configurations. Yet, it requires a clear definition of the feature space on which the search of patterns should be conducted and a suitable choice of clustering algorithm.

Considering that there are many choices on the design of a clustering algorithm, is there one that performs *best*? Chen and Peter Ho (2008) answer this theoretical question with a clear *no*, since the result and thus the performance of an algorithm is always dependent on the data, the pre-processing and the selection of the feature space. To choose a reasonable model, it is essential to make assumptions on the data that are used. Some data can be described best by a linear model, while other data can be modeled better by artificial neural networks. This is known as the no free lunch (NFL) theorem which was formalized by Wolpert (1996). It states that all algorithms perform equally well "when averaged across all possible problems" (Wolpert and Macready 2005, p. 721). The only way to ensure the best model is to test all possible algorithms on the dataset (Géron 2020). Géron concludes that since this is not feasible, the best way is to make assumptions on the data and base thereon a reasonable design of classifiers. So far, this master thesis already made some theoretical assumptions on urban morphological patterns. For the methodological implementation of the framework, this will be continued with the data preparation and the choice of feature space (see figure 9).

To find urban morphological patterns, the search has to be conducted in a comparable framework. This concerns especially the design of a feature space. The database consists of cities with different sizes. This poses a major challenge, since patterns should have a similar extent to be comparable. A solution is to look at subsets of cities with a similar size, with the intent to recognize patterns as the spatial arrangement of the LCZs. Taubenböck, Debray et al. (2020) create their own feature space based on the relative coverage of the LCZs inside the city or inside concentric zones around a monocentric city center. The first method does not include any spatial relation and the latter only in a limited way (see chapter 2.5).

The LCZs can be regarded as image data, with the class representing a pixel value while the MUAs determine the boundary of the image. An image is already in a machine readable form. It can be broken down to a subset of smaller image parts. The creation of subsets with a comparable size allows the uniform search for patterns with an unsupervised clustering approach. Following, the similarity of cities can be described through the similarity of patterns that is found in their subsets.

“Capturing the complexity of urban form requires more than the characterization of street networks and a handful of other measurable characters.” (Fleischmann, Feliciotti and Kerr 2021, p. 2)

To conduct an explorative search of urban morphological patterns, a feature space that represents those informations needs to be developed. Since it is not possible to formalize a logic of urban patterns, the manual development of a suitable feature space seems unrealistic. Every feature selection would mean a formalization of those. To avoid this, a simple and pure assumption is made: An image itself is a pattern of spatially arranged LCZs and contains thus all information necessary. Hence, instead of constructing a sophisticated feature space, the search is conducted directly on the subsets of cities.

Having a basic idea of the database on which the search can be conducted, it is time to make assumptions on the clustering algorithm. Geographical data have often continuous properties without hard boundaries. Therefore it is questionable if homogeneous patterns of urban morphological configurations can be found. It is more likely that urban morphological patterns will have a typological overlap. Considering the LCZ database, a clustering algorithm should not perform a hard classification that assigns an image exclusively to a cluster (Jensen 2005). Hence, a fuzzy clustering algorithm that assigns a partial membership to different classes (patterns) would be better suited. It is still possible to associate an image with a specific class (based on the highest probability), yet a fuzzy membership contains information on a partial overlap with other classes. In this way, it is possible to achieve a better representation of the morphologic reality. Further, it allows a continuous similarity measurement when comparing the found patterns. With these general thoughts and assumptions on the methodological concept, the master thesis continues with an introduction to neuronal networks used in computer vision and takes a look at suitable machine learning methods.

4.1.2 Introduction to Deep Learning in Computer Vision

To ensure an understanding of the methods presented and used in this master thesis, this subchapter provides an introduction to the field of deep learning in computer vision. It starts with the basic elements of a neural network, the neurons, which can be connected to complex architectures. Based on this foundation, the concept of specialized state of the art networks for image processing are explained. They will play a key role in the methodology of this master thesis.

Artificial neural networks (ANNs) are machine learning models inspired from the architecture of the biological neurons in our brains. The initial idea came with the first architecture of an ANN developed by McCulloch and Pitts (1943). But it took several decades up to the 1990s, after a focus on other machine learning methods, for the idea to experience a revival in scientific research (Géron 2020). This is due to bigger amounts of available training data, development of new network architectures and algorithms, improvements of graphical processing units (GPUs) and the increase of computational power, which all play an important role in ANNs.

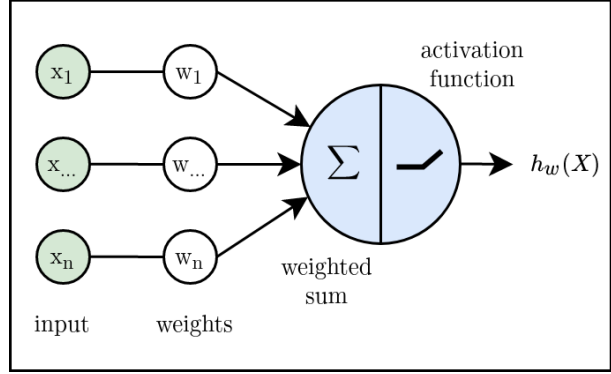


Figure 10: A single threshold logic unit (TLU) used in the perceptron developed by Rosenblatt (1958) (own illustration). The sum of the weighted input is given to an activation function (for example equation 1) and results in the output $h_w(X) = f(X^T w)$.

A simple architecture of an ANN is the perceptron invented by Rosenblatt (1958). A single unit, which imitates a neuron, is called a threshold logic unit (TLU) (see figure 10). It calculates the weighted sum $z = X^T w$ of a numerical input $X = \{x_1, \dots, x_n\}$ and applies an activation function $f(z) = f(X^T w)$ to the result (specific activation functions will be presented later). The numeric output of the perceptron is written as $h_w(X) = f(X^T w)$. The TLU is a single elementary processor of a network. Multiple TLUs can be combined by using the output of one as the input of another one (Sonka, Hlaváč and Boyle 2008). This idea is similar to the neural connections in our brain, which transmit electrical impulses. Figure 11 is a drawing from Ramón y Cajal of neurons in a human cerebral cortex. It shows the highly interconnected neuronal cells within more and less dense 'layered' structures, bringing up the idea of building similar artificial networks.

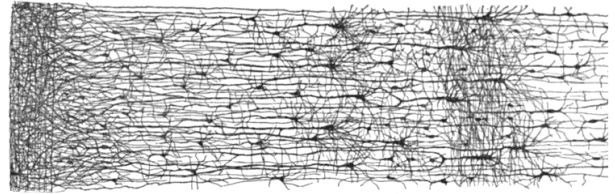


Figure 11: The postcentral gyrus, a part of the cerebral cortex, reprinted from Ramón y Cajal ([1899] 1988, p. 214). The biological neurons are highly connected with each other and form a biological neural network. A structure of more and less dense 'layers' can be observed. The image is a visualization where the inspiration of ANN architectures originates from. Similar to the biological version, layered structures of TLUs can be used to build complex networks.

Using multiple TLUs next to each other form a layer. If all TLUs are connected to all neurons of the *input layer*, it is called a *dense layer*. Usually, an additional bias-neuron b is added to each of the TLUs layers (Géron 2020). The output of a dense layer is $h_{W,b}(X) = f(XW + b)$. Stacking an input layer with one or more *hidden layers* of TLUs and a final TLU *output layer* (as shown in figure 12) forms a multilayer perceptron (MLP). Several activation functions exist to introduce non-linearity to the model. An activation function frequently used due to its fast calculation in the hidden layers of ANNs is the Rectified Linear Units (ReLU) function (see equation 1). If $z \geq 0$ the output is linear, else the output is 0 (rectified) (Beyerer, Richter and Nagel 2018).

$$f_{ReLU}(z) = \max(0, z) \quad (1)$$

In a MLP each layer is fully connected with the following layer. When neurons of one layer are connected only to the following layer and not to themselves, it is a *feed-forward* network (Beyerer, Richter and Nagel 2018). The input data of a network are represented in the *image space* I . It travels in one direction towards the output, which is called the *embedding space* E (Sonka, Hlaváč and Boyle 2008).

To train a network such as the MLP shown in figure 12, the weights are initialized first with random values. The goal of the training process is to adapt the weights of a network in such a way, that the input of the network produces a requested output. In a supervised way, this may be the already known labels of a given image. The tasks to train a network in an unsupervised and self-supervised way will be discussed later in section 4.1.3. Typically, the training data are divided into *mini-batches* to accelerate the training process.

But how is it possible to adapt the weights of a network, given an input and the desired output? For feed-forward networks, the back-propagation algorithm developed by Rumelhart, Hinton and Williams (1986) is used. In a first step, data are passed through the network to the output. For a mini-batch, each instance is fed to the input layer and passed through to the next layers (feed-forward). The intermediate results of each layer are kept. When using a MLP for classification tasks, the output layer has a neuron for each class c in $C = \{c_1, \dots, c_{|C|}\}$. The values of the output layer in the embedding space E are called *logits*. It is a $|C|$ -dimensional vector z . The logits are passed to a softmax activation function (see equation 2) to ensure that the probabilities of belonging to one or to another class are in the range $[0, 1]$ and sum up to 1 (Géron 2020).

$$f_{softmax}(z_i) = \frac{\exp(z_i)}{\sum_{j=1}^{|C|} \exp(z_j)} \quad \text{for } i = 1, \dots, |C|. \quad (2)$$

The output layer is then compared with the desirable result and an error is measured with a *loss function* L . In the second step, the back-propagation is performed. The gradient of the

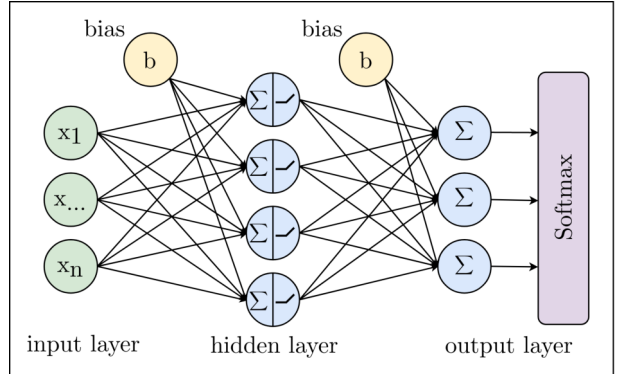


Figure 12: Multilayer perceptron (MLP) with one hidden layer and a softmax activation function on the output layer (own illustration). The input layer (green) passes the input data to the following TLU layer (blue). Each TLU of a layer has individual weights for the input. The added bias (yellow) is shared across the layer. An activation function is applied and the output is used as the input for the next TLU layer. For clustering, a special activation function like the softmax function can be applied on the last layer.

error is calculated to see how the weights and the bias of the previous layer need to be adapted to minimize the error (and thus get the desired output). Going backwards, this is performed for all previous layers based on the stored intermediate results (back propagation). Having processed the whole mini-batch, the gradient descent is performed in a *gradient-step*. When all mini-batches have run through the network, the model was trained for one *epoch*. A model is commonly trained for multiple epochs. To allow the calculation of gradients, the activation functions are ideally continuous functions.

Convolutional neural networks (CNNs) are specialized architectures to work with multi-dimensional image data in the field of computer vision. They perform well in tasks like object recognition, object detection and semantic segmentation. A CNN takes the entire pattern (an image) as an input (Beyerer, Richter and Nagel 2018). The information of each pixel is given to an input neuron of the network. Based on the previous concept of artificial neurons, two essential elements of a CNN are now introduced, the *convolutional layers* and the *pooling layers*. Together, they form the architecture of a CNN.

The neurons of a convolutional layer are not fully connected with the neurons of the input layer. Instead, the weighted sum of a small receptive field, again with a set of weights and a bias, is connected to a single neuron. The receptive field slides on the input layer, each time calculating the output of a new neuron (see figure 13). It functions as a (through the training process adaptable) filter applied to the input layers. Multiple input layers can be transformed with one receptive field to a single convolutional layer. Multiple receptive fields create, each with their set of individual weights, a convolutional layer. These layers can be stacked to one convolutional block, which again can serve as input data. The creation of a convolutional layer is controlled with three important parameters. The *kernel size* of the receptive field determines how many neurons from the input layer are connected to the new neuron. The *stride* determines the offset of the receptive field. The *padding* determines how many rows with which values are added around the input image. Since the size of the output layer is reduced from the input layer, a padding can be added to keep the size of the original input layer. A zero padding adds the value 0 around the input image, the receptive field now slides on this extended input and produces an output equal to the original extent of the input image (also referred to as 'same'-padding). A valid padding adds no values and simply ignores the input if the receptive field exceeds the borders.

A common problem with CNNs is the high number of parameters that need to be stored for training the network with a back-propagation. When using multiple receptive windows, stacking multiple convolutional layers to one block and building deep networks with multiple blocks, the number of parameters and amount of storage space grow very fast. To avoid this, pooling layers are used to reduce the number of parameters (neurons) in a network. Further, they have the capability to improve the tolerance for noise in the data and reduce the risk of overfitting (see the *curse of dimensionality* in chapter 4.1.1). A max-pooling layer, for example, sub-samples the input layer by passing only the maximum activation value of a small receptive window to the

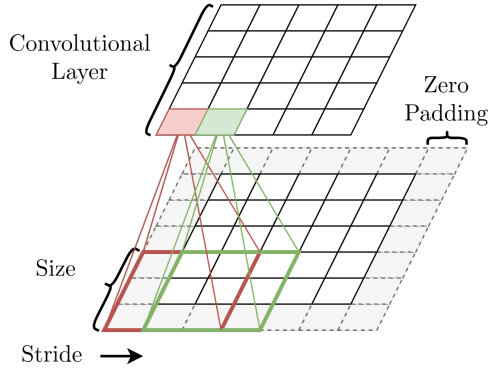


Figure 13: Example of creating a convolutional layer (own illustration). The receptive window with a kernel size 3x3 neurons slides on the input image with a stride of 1. Each time the output of a new neuron is calculated. To keep the original input extent (same-padding), a zero padding is added to the input layer.

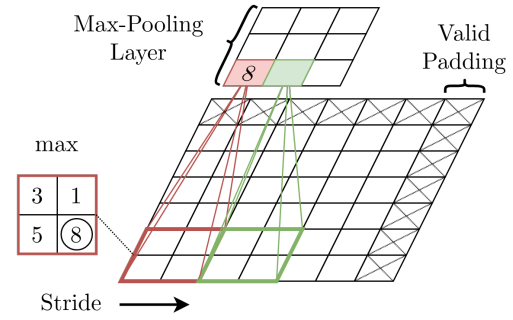


Figure 14: Example of a max-pooling layer (own illustration). The receptive windows has a kernel size of 2x2 and moves with a stride of 2 neurons. The maximum value of the kernel is passed to the next layer. Here, a valid padding is used, meaning some rows and columns are ignored if the perceptive field exceeds the border.

following layer (Beyerer, Richter and Nagel 2018). Average-pooling layers take the average of the values in a receptive window. The example of a pooling layer with a valid padding is shown in figure 14.

Now, building a CNN architecture includes typically convolution blocks with an increasing number of layers, followed by pooling layers to decrease the layer size. Deep networks can be built by stacking this combination multiple times. Figure 15 shows an exemplary architecture of a CNN capable of clustering images. Low level-features are learned in the first convolution blocks and aggregated to high-level features in the following blocks. A fully connected MLP appended to the end is used as a *cluster-head* to perform a clustering or classification of images.

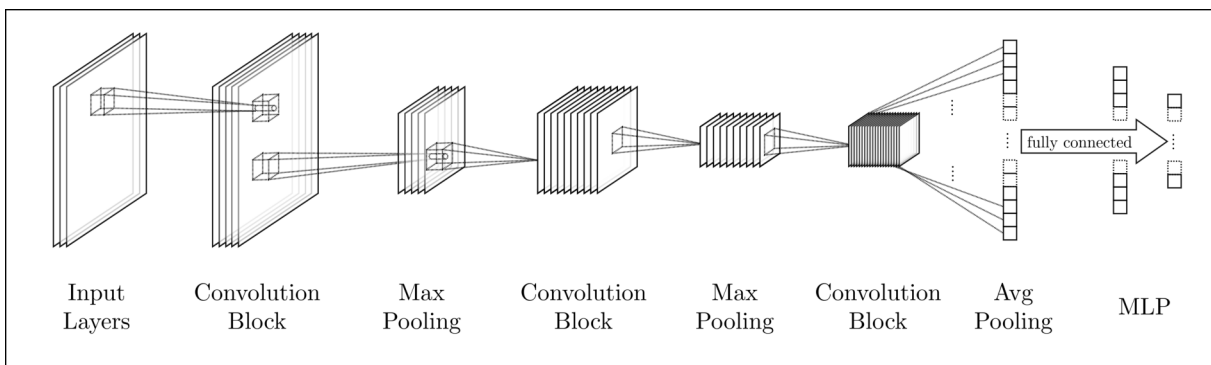


Figure 15: Example of a CNN architecture (own illustration). Beginning from the input layers, a block of convolutional layers is created. By doubling the number of kernels, the number of layers also doubled and passed through a ReLU activation function. The following max-pooling passes the maximum activation value in a kernel to the neurons of the next layer and thus reduces the layer size. Both steps can be performed multiple times, increasing the number of layers on one side while reducing the layer size on the other side. Towards the end, an average-pooling can be performed. It aggregates the average of each layer towards one output neuron. For clustering, a fully connected MLP (as shown in figure 12) can be added.

Building deeper networks increases the capability of learning more complex feature representations, yet they were difficult to train. The problems of vanishing or exploding gradients could be tackled by the normalized initialization of weights (Glorot and Bengio 2010) and batch normalization (BN) of intermediate layers (Ioffe and Szegedy 2015). Still, a degradation problem remained, meaning deeper networks performed worse than shallower networks (He, Zhang et al. 2015a). This could be resolved through *residual learning* introduced by He, Zhang et al. (2015a), now allowing to build and effectively train very deep networks. Their new residual neural network (ResNet) architecture is composed of multiple stacked residual units. Within a residual unit, the input x is mapped by multiple successive layers $g(x)$. Through a shortcut connection, the input is added to the mapping and results in the output $h(x) = g(x) + x$ (see figure 16a). An example for using this concept for convolutional layers is shown in figure 16b. If the dimensions are kept, the shortcut connection can directly be added to the output of the mapping $g(x)$. When changing the dimensions (for example doubling the number of layers from $d=64$ to $d=128$ while halving the size with a stride $s=2$) within the mapping $g(x)$, the dimensions of the shortcut connection are adapted with a 1×1 kernel and a stride of 2 as shown in figure 16c. Within the ResNet architecture, multiple residual units are stacked together, for example figure 16b can be used as the input for figure 16c.

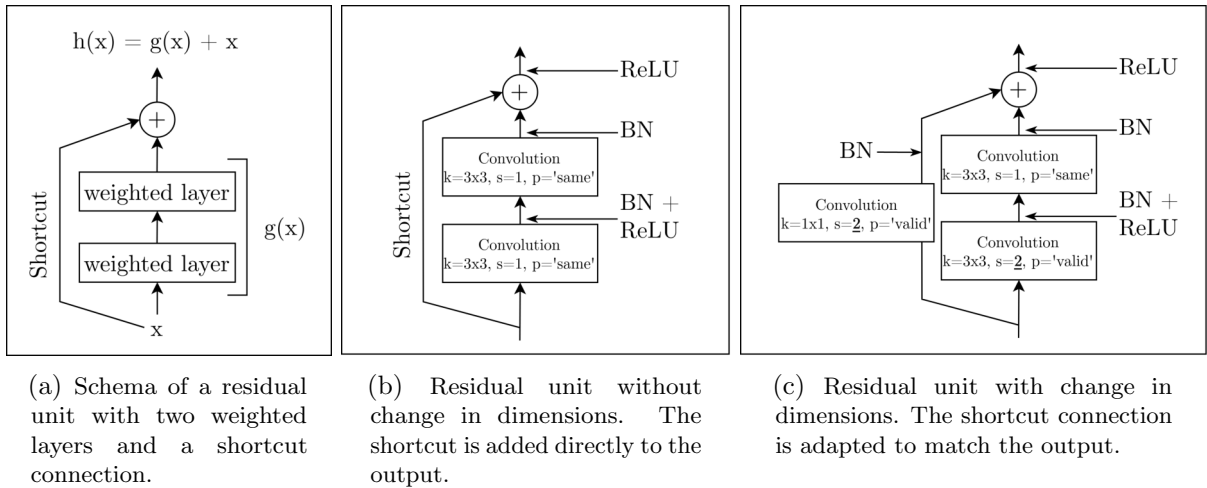


Figure 16: Concept of residual blocks (own illustration). A ResNet can be built by stacking multiple residual blocks, using the output of one block as the input for the following block.

Based on this short introduction on ANNs, the next subchapter continues with state of the art methods in unsupervised image clustering that are of interest for this master thesis.

4.1.3 Methods for Unsupervised Image Clustering

The unsupervised clustering of images is a major challenge in computer vision. The aim is to group a set of images $D = \{X_1, \dots, X_{|D|}\}$ into semantic similar classes $C = \{c_1, \dots, c_{|C|}\}$. A cluster algorithm determines the class membership of an image and assigns it to a class with similar traits (Park, Han et al. 2020). Since this challenge in computer vision is not new, there is a vast variety of fast improving algorithms that compete for comparability on image benchmark

datasets (cf. Papers with Code 2022). This subsection reviews a range from simple to current state of the art methods on unsupervised image clustering.

Reviewing the chapter ‘Conceptualization’ 4.1.1, each pixel of an image represents a value in a specific location. A pixel is therefore a feature of this image. Hence, it is possible to build a feature space out of the pixels of an image and cluster images based on this feature space. This very simple approach can easily detect identical images, but is not capable of recognizing semantic similarities. The method is vulnerable to slight offsets, rotation and differences of an image, since the location of a pixel value changes. It is assumed, that no two identical urban morphological patterns will be found. Instead, a class representing a type of urban morphology consists of a set with similar, but not identical images. Thus, a method that allows a certain degree of generalization is required to answer the research question.

Sequential approaches introduce generalization to a model by reducing the high-dimensional feature space of an image. It is possible to transform the feature space to a low-dimensional embedding space E or extract relevant features. This can be done through a principal component analysis (PCA) with a simple linear dimensionality reduction (cf. Ding and He 2004) or autoencoders which perform a parametric feature extraction (encoder) to an embedding space E and try to reconstruct the image space I (decoder) (cf. Baldi 2012; cf. Bourlard and Kamp 1988). Conventional density- or distance-based clustering methods can be applied on the embedding space E , but the extracted features of these models are not well suited for semantic clustering (Park, Han et al. 2020).

Recent advances in self-supervised learning within a CNN architecture allow more sophisticated feature extraction by solving pre-defined tasks, called *pretext tasks*. A pretext task τ is used to build an embedding space E and force the underlying network Φ to learn relevant semantic features of an image. The idea is to make transformations $T(X)$ of an image X agree with each other in the embedding space E (Chen, Kornblith et al. 2020). These representations are learned solely through the image itself, by adapting the weights Θ of a neural network Φ_Θ . Examples for pretext tasks are jigsaw puzzles (Noroozi and Favaro 2016), predicting image colors (Zhang, Isola and Efros 2016), rotation prediction (Gidaris, Singh and Komodakis 2018) or reconstructing the missing parts of an image (Pathak et al. 2016). In a second step, either a cluster-head is added and trained in a semi-supervised way (which is commonly used to compare the performance of the learned representations on a benchmark dataset) or an unsupervised clustering is performed directly on the embedding space of the pre-trained network. Contrastive learning methods combine these augmentations and achieve good results on benchmark datasets (Niu and Wang 2021). Especially the Simple Framework for Contrastive Learning of Visual Representations (SimCLR) developed by Chen, Kornblith et al. (2020) outperformed previous work and showed the capability of learning high-level semantic features in an unsupervised way.

End-to-end pipelines like Deep Cluster (Caron et al. 2018), Deep Embedded Clustering (Xie, Girshick and Farhadi 2016) and Deep Adaptive Image Clustering (Chang et al. 2017) use combined

approaches of feature learning and clustering. Features are learned through transformations tasks and pseudo-labels are assigned to the images. The weights of the underlying network are then updated by trying to predict these pseudo-labels. This means the learning of features and the clustering of images goes hand in hand. A constraint in learning meaningful representation in these end-to-end pipelines approaches is that they are sensitive to the network initialization and may focus only on low-level features (Gansbeke et al. 2020).

A two-step approach in unsupervised image classification is the framework Semantic Clustering by Adopting Nearest neighbors (SCAN) developed by Gansbeke et al. (2020). It was state of the art on four benchmark datasets in 2020 (cf. Papers with Code 2022). They combined the self-supervised representation learning and the end-to end approach in a two-step framework leveraging the achieved accuracy on the benchmark datasets CIFAR10 to 88.3%, CIFAR 100-20 to 50.7% and STL10 to 80.9% and outperforming previous methods by large margins (Gansbeke et al. 2020).

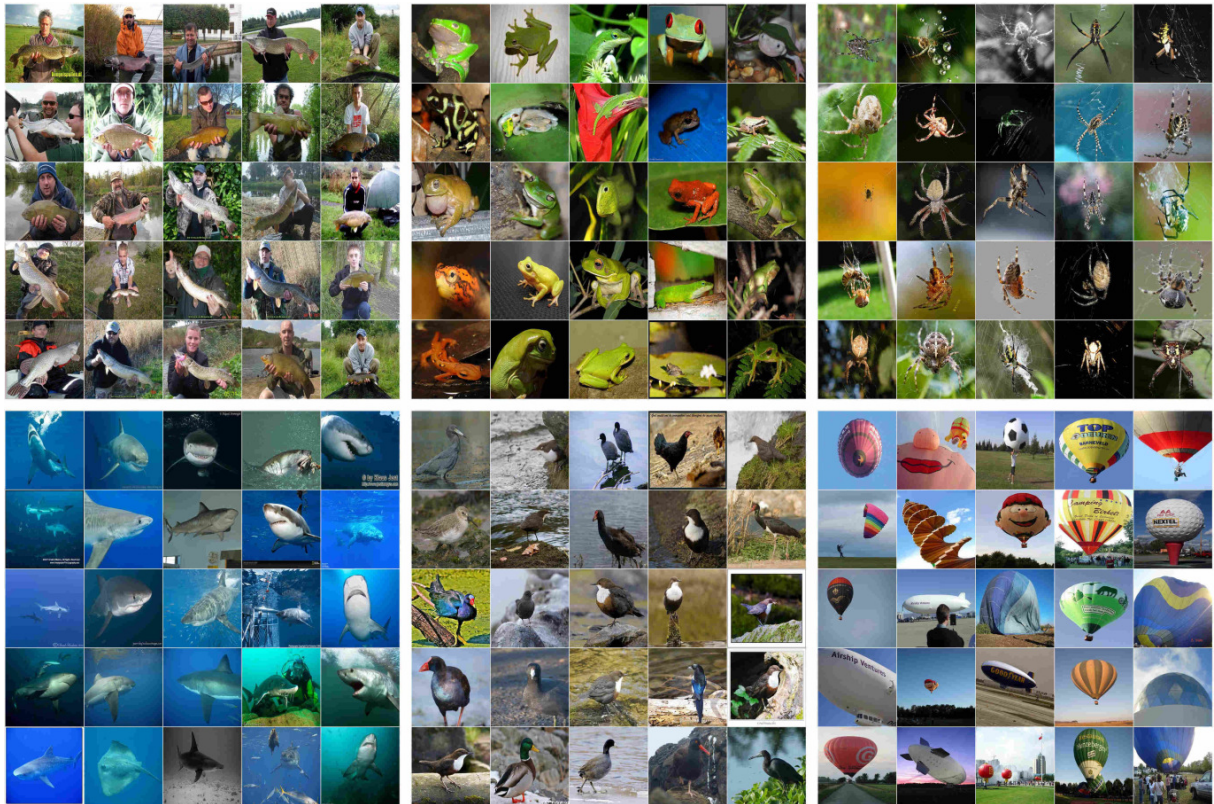


Figure 17: Example clusters achieved with SCAN on the ImageNet-1000 benchmark dataset, reprinted from Gansbeke et al. (2020, p. 23). It is a beautiful example of the algorithm capability to cluster semantically similar images without using labels.

Figure 17 shows a clustering example achieved with SCAN on the ImageNet-1000 dataset. These results are astonishing, considering they have not used any labels for the image clustering. The framework trains in a first step a CNN Φ_{Θ} by solving a self-supervised pretext task τ to learn semantically meaningful image representations. Any pretext task that minimizes the distance d between an image X and its transformation $T[X]$ in the embedding space E can be used (see equation 3).

$$\min_{\Theta} d(\Phi_{\Theta}(X), \Phi_{\Theta}(T[X])) \quad (3)$$

Gansbeke et al. (2020) adopt the SimCLR framework for benchmark datasets with up to 20 classes and the Momentum Contrast for Unsupervised Visual Representation Learning (MoCo) framework (He, Fan et al. 2019) for the ImageNet dataset with 1000 classes. They empirically found that in many cases the neighbors in the embedding space E of the pre-trained network Φ_{Θ} belong to the same semantic class. In a second step, they initialize a CNN Φ_{η} with the learned weights from Φ_{Θ} and train it to cluster an image X together with its k-nearest neighbors N_{X_i} . This two-step approach enables to retrieve high-level features a priori to the cluster learning (Gansbeke et al. 2020). Yet, the SCAN framework has some drawbacks. The k-nearest neighbors N_X contain noise, meaning not all nearest neighbors belong to the same semantic cluster. When using false samples in the early stage of training, the model can become overconfident and cause false predictions (Park, Han et al. 2020).

The remarkable results of Gansbeke et al. (2020) were even further improved through the Robust learning for Unsupervised Clustering (RUC) framework developed by Park, Han et al. (2020) which was state of the art on four benchmark datasets in 2021 (Papers with Code 2022). It is an additional module to existing clustering approaches (such as SCAN) developed to overcome some of their restrictions. RUC considers an already clustered dataset with pseudo-labels as *noisy* with miss-classified samples. During a re-training process, unclear samples are filtered via three possible sampling strategies: (i) a confidence-based strategy selects samples by a confidence threshold, (ii) a metric-based strategy trains an additional embedding network (for example SimCLR) to measure the credibility based on the k-nearest neighbors and (iii) a hybrid strategy combines both methods. The model then is re-trained only with the clean samples via robust learning techniques (Park, Han et al. 2020). The overconfident results of the previous model are softened and re-calibrated. The best accuracies were achieved through a combination of the SCAN and RUC framework, on the CIFAR-10 and STL-10 benchmark datasets with the confidence-based strategy and on CIFAR-20 with the hybrid-strategy.

Would it be possible to adapt the presented state of the art methods to answer the research questions of this master thesis? Finding spatial patterns of urban configurations in the database of the LCZs should be similar to the task of clustering semantic similar images of benchmark datasets. Yet, the LCZs pixel values represent categorical data, while the benchmark dataset images usually show a single object with gradient pixel values in a Red, Green and Blue channel. For example, one image of a white dog and one of a grey dog would still show the semantically similar object of a hypothetical class '*dog*'. This means, the exact pixel values of both images matter less than the actual semantic object in the image. Opposite to this, differences in the categorical pixel values of the LCZs matter more, since an image consisting of the LCZs class 1 and another image with class 2 show fundamentally different urban configurations (see figure 8). Hence, the presented methods cannot be adopted directly. However, it is assumed that adapting the state of the art methods to the characteristics of the LCZs dataset might produce acceptable

results. This assumption is made considering that there is no possible way of verifying the results of the explorative search on urban morphological patterns. Hence, the best performing state of the art methods might be the best choice to cluster unsupervised images and thus find intra-urban patterns and answer the research question of this master thesis.

In conclusion, this master thesis adapts the SCAN framework with the SimCLR implementation to find spatial urban morphological configurations in subsets of cities. Additionally, the RUC framework is implemented to correct miss-classified images and allow a more fuzzy class membership, compared to the overconfident results produced by SCAN. Concluding the theoretical background, the master thesis continues with the presentation of the data that will be used for this analysis.

4.2 Data Basis

In this subchapter the data basis used in this master thesis is presented. It consists of the LCZs to describe the urban morphology and the MUAs to determine the boundary of urban agglomerations. Essential for the usage of both datasets is the acquisition date from 2019. This ensures consistency in the delineation of the urban areas.

4.2.1 Local Climate Zones

The Local Climate Zones (LCZs) mapping developed and provided by Qiu, Schmitt and Zhu (2019) dates from 2019. It is an enhanced version of the database used by Taubenböck, Debray et al. (2020), consisting now of 1692 raster tiles covering urban areas around the world. The raster tiles are provided in their local WGS84 / UTM Zone projection. Some of the LCZ tiles are partially overlapping. The border of the raster tiles show occasionally stripes of miss-classified LCZs. The false data are likely an implication of processing the underlying remote sensing data. The size of the miss-classified stripe varies, but it does not exceed 3% of the complete raster tile size on each of the four edges. The overall accuracy of the LCZs mapping product is given with 86.7% (Qiu, Schmitt and Zhu 2019).

4.2.2 Morphological Urban Areas

The Morphological Urban Areas (MUAs) are developed and provided by Taubenböck, Weigand et al. (2019). The dataset consists of 1569 shapefiles delineating the boundary of the MUAs. The data are provided in the World Geodetic System 1984 (EPSG:4326). The MUAs delineate the boundaries of urban areas and partially exclude areas with water (as seen in figure 18). A full list of the MUAs is provided in table 7, found in the appendix.

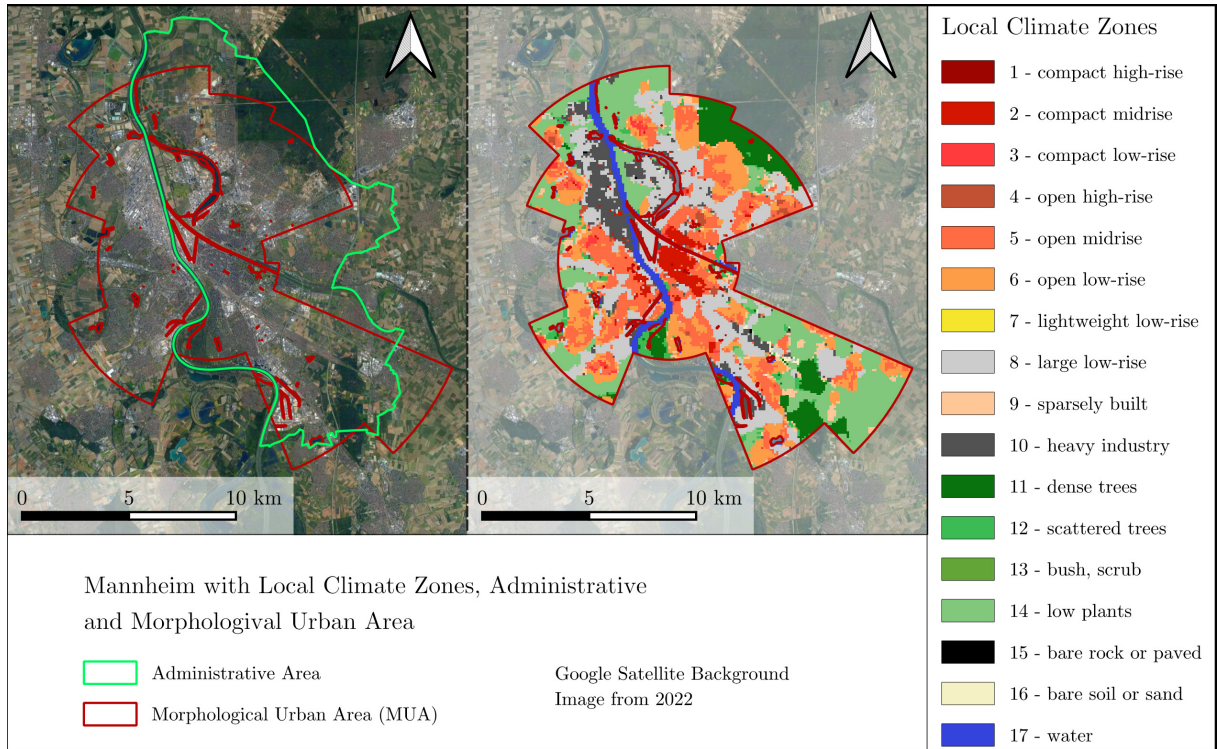


Figure 18: Mannheim in Germany with the Administrative Area, the Morphological Urban Area (MUA) and the Local Climate Zones (LCZs). It is a good example of an administrative area not representing the morphologic reality of urban areas. Partially included to the MUA of Mannheim is the city of Ludwigshafen, so a MUA can represent urban agglomerations of multiple cities. The MUAs provide a data-driven consistent delineation of urban areas around the world and allow thus a comparative approach of urban areas.

4.3 Methods

This subchapter presents the methods and analyses used to answer the research questions. First, the data are cleaned and prepared for the unsupervised image clustering (chapter 4.3.1). Second, the unsupervised image clustering is performed using an adaption of the SimCLR and SCAN framework (chapter 4.3.2). Third, the clustering results are corrected with the RUC framework (chapter 4.3.3). Based on the clustered images, which are supposed to represent patterns of urban morphological configurations, the similarity of cities is calculated (chapter 4.3.4). In a final step, the cities are clustered on the basis of the calculated similarity (chapter 4.3.5). Additionally, the methods developed by Taubenböck, Debray et al. (2020) are applied on the extended database of MUAs (chapter 4.3.6). Figure 19 shows the workflow of the components used in this explorative master thesis.

4.3.1 Data Preparation

To find patterns of urban morphological configurations and implement the proposed workflow (see chapter 4.1.3), the LCZs data have to be brought into a consistent form to feed them to a neural network. The data preparation consists of eliminating false data in the LCZs and extracting smaller subsets based on the delineation of the MUAs.

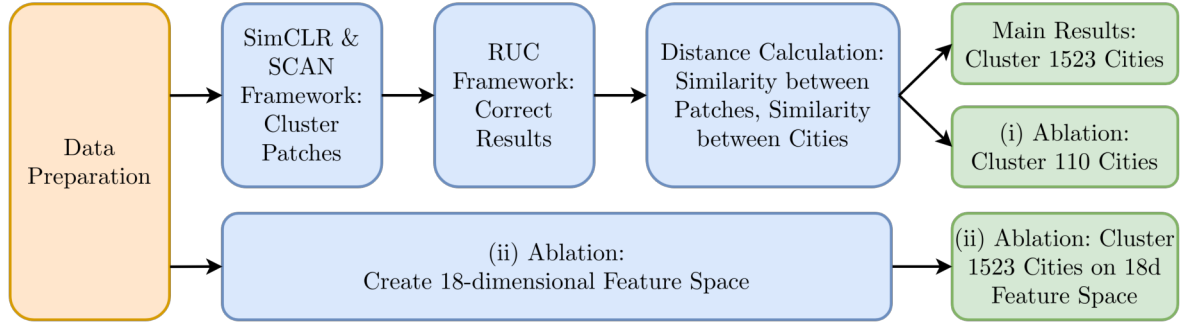


Figure 19: Methodological components of this master thesis. Starting with a data preparation, the SimCLR and SCAN framework are used to cluster patches. The results are corrected with the RUC framework. The similarity between cities is calculated based on the clustered patches and the cities are clustered. For the first ablation study, the number of cities used for clustering is reduced. For the second ablation study, a 18-dimensional feature space developed by Taubenböck, Debray et al. (2020) is derived for 1523 cities and used for clustering.

In a first step, the edges from the LCZ raster tiles are partially excluded from the analysis. The size of miss-classified data varies and cannot be determined unambiguously by its shape or values. Therefore, 3% of the raster tile size are cut off each edge. Figure 20 shows a LCZs raster tile of the urban area of Dongguan in China. On the left image, the original tile extent and miss-classified data on the image border can be seen. The right image shows the new raster tile extent after the cut off. This process reduced the raster tile image by a total of 6% on its x- and y-axis.

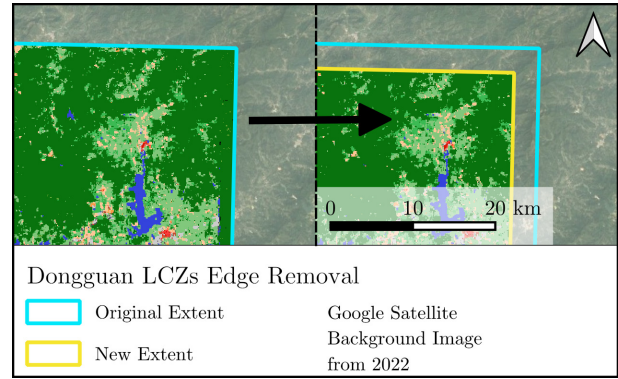


Figure 20: Edge removal on the LCZ raster tile of the urban region Dongguan in China. The raster tile contains a stripe with miss-classified data.

In a second step, the MUAs are pre-processed. The LCZ class '*water*' (LCZ-17) is so far partially excluded from the MUAs (see for example figure 18). This master thesis assumes that the class '*water*' has an influence on the urban morphological configurations and should therefore be included to a certain degree. To underline these assumptions with some examples, one might think about the location of '*heavy industry*' (LCZ-10) near rivers or coastlines for accessibility to waterway infrastructure, or '*open midrise*' (LCZ-6) buildings representing possible upper-class single-family houses near recreational areas such as lakes or parks. This brings in the argument to further extend the MUAs to include more of the '*non-built type*' (LCZ 11-17) surroundings. The MUAs are very restricted to the actual urban areas (LCZ 1-10), yet the '*non-built types*' might also have an influence on the urban configurations. Therefore, the MUAs are re-projected to a metric projection (EPSG:3857) and extended with a buffer size of 1000 meters. This buffer size is chosen quite arbitrarily, but under the consideration of including more of the surroundings while still keeping the focus on the urban areas.

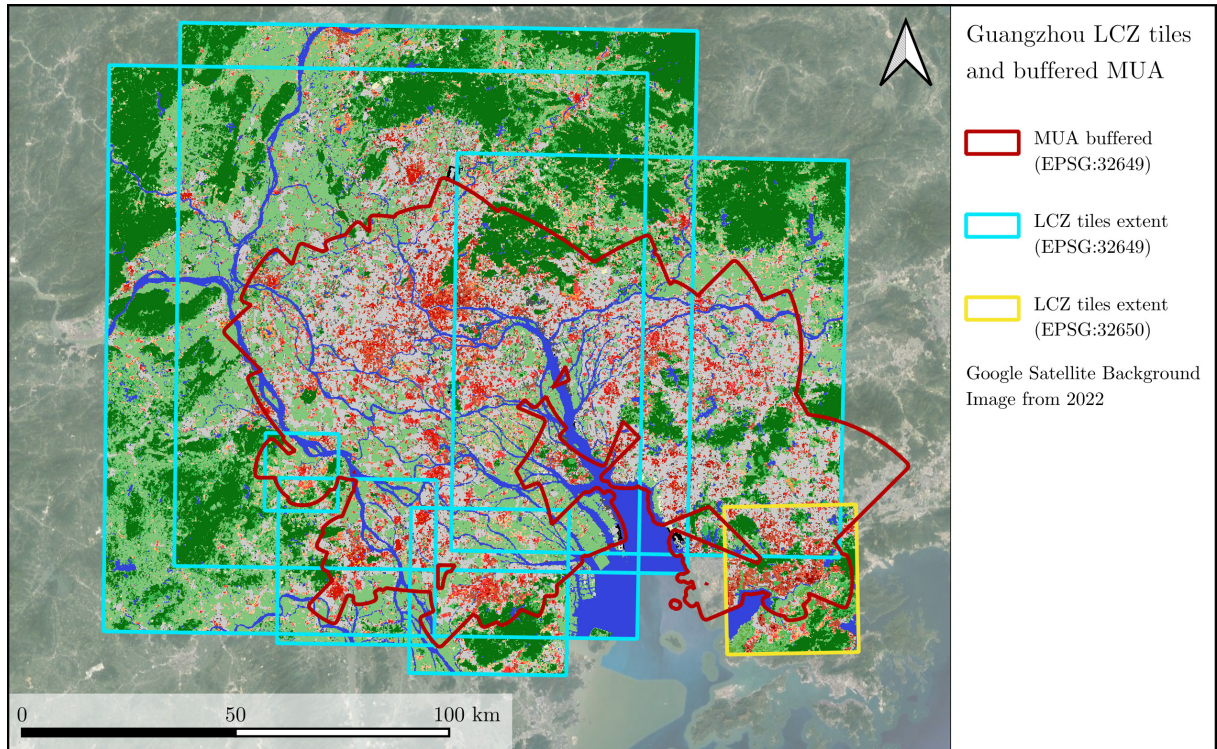


Figure 21: The agglomerative urban area of Guangzhou in China. The buffered MUA (red line) is covered by multiple LCZ tiles in different projections (cyan line in EPSG:32679; yellow line in EPSG:32650). The LCZs data do not cover completely the buffered MUA, missing data are found in the bottom right corner.

In a third step, the LCZ raster tiles are adapted to the MUAs. Both datasets do not align well with each other. Multiple LCZ tiles cover (partially) the MUAs, but also multiple MUAs are inside a LCZ raster tile. The LCZ raster tiles are sometimes overlapping and have different projections (relative to their location). As an example, the buffered MUA of Guangzhou in China is visualized in figure 21 with the available LCZ raster tiles and their projections. The goal of this step is to prepare a consistent LCZ coverage for each buffered MUA. Hence, the available LCZ raster tiles have to be merged. For each LCZ tile, the MUAs are reprojected to the same projection and the overlapping area is calculated. This is done to create a hierarchy for merging the LCZ tiles. The biggest overlapping raster for each MUA is chosen to be the basis for merging and reprojection. Descending from the second largest to the smallest overlapping raster, they are reprojected if necessary and gradually merged to the basis raster. A nearest neighbor interpolation is used to ensure the consistency of the categorical data. If parts are overlapping, the LCZs from the biggest raster are kept, while suppressing the data from the smaller raster tile. When all LCZ raster tiles are merged, the resulting file is clipped with the corresponding reprojected MUA. The result of the consistent LCZs in the MUA of Guangzhou is shown in figure 22. There is not always a full coverage of the MUAs with the LCZs. A full overview of the MUA size, the corresponding LCZs size and the resulting percental coverage of the MUAs is provided in table 7, found in the appendix.

In a fourth step, the image subsets that will be used in the unsupervised clustering are created. For each of the pre-processed LCZs linked to a MUA, subsets with a size of 32×32 pixels and an

offset of 10 pixels are extracted, beginning with the top-left corner. Hereby, subsets are only accepted if the class 'water' (LCZ-17) does not cover an area greater than or equal to 85% of the image subset to maintain a focus on the urban morphological configurations. Further, subsets are rejected if they are not fully covered by LCZ class values.

The size of the subsets is chosen based on two considerations. A subset size of $32 * 32$ pixels with a pixel length of $100\text{m} * 100\text{m}$ represents an area of 10.24km^2 . This can be seen as a balanced medium scale in which a search on urban morphological configurations can be conducted. A subset too small would consider not enough of the urban surroundings in the search of urban configurations and generalize the patterns too much. A subset too large would generalize not enough and look for too specific urban configurations. This size is assumed to allow a balanced approach for the search of patterns and the comparability of cities. Moreover, a direct implementation of the standard ResNet-18 network as used in the framework of Gansbeke et al. (2020) is possible. Larger subset sizes would need an additional pooling layer or the use of larger networks and hence more computational power.

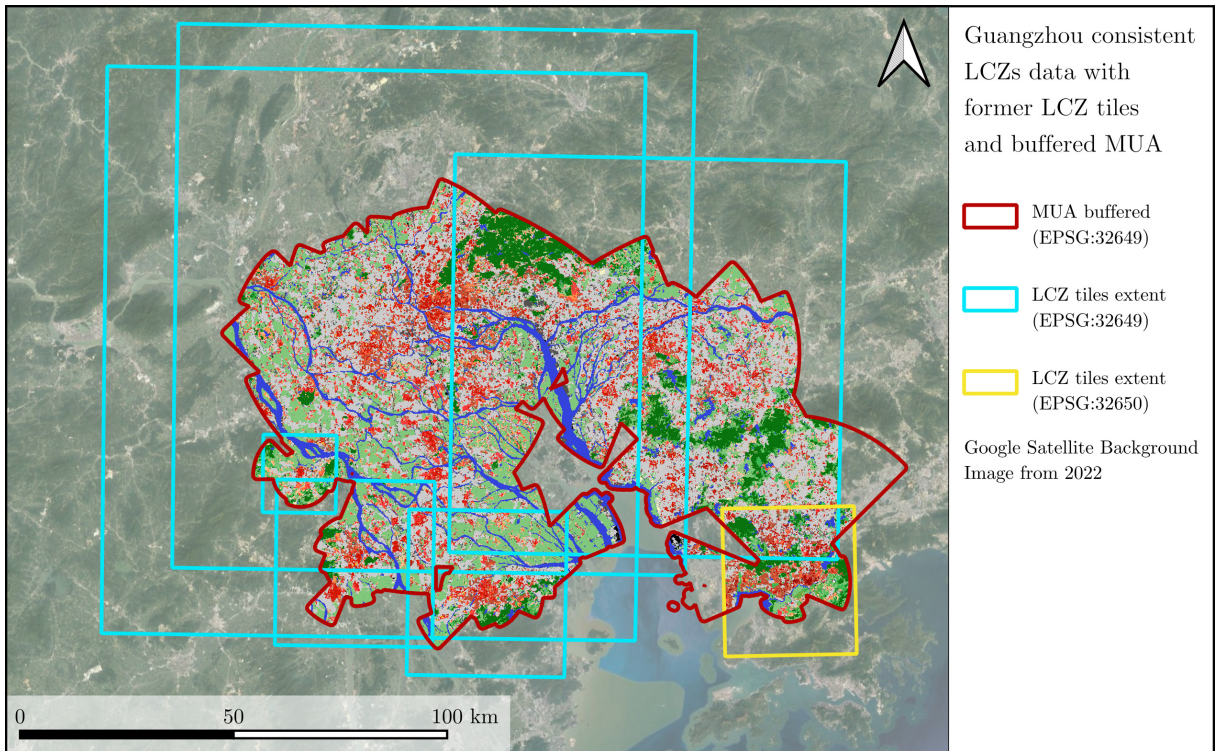


Figure 22: The agglomerative urban area of Guangzhou in China with consistent LCZs data. The LCZs are re-projected if necessary (in this example the yellow LCZ tile from EPSG:32650 to EPSG:32649) and gradually merged, beginning with the largest overlapping tile. The merged LCZs are then clipped to the extend of the buffered MUA.

The subsets are overlapping by approximately $2/3$ (due to the 10 pixel offset during the extraction process). This allows a more continuous representations of cities and search of patterns. In result, a total of $P = 316,536$ image subsets for 1523 out of the 1569 MUAs are extracted. 46 cities do not fulfill the subset size or the above mentioned restrictions and have in consequence no subsets extracted. Those 46 cities are excluded from the analysis. Let $D = x_1, \dots, x_P$ denote

the datasets with the image subsets x in this work. The subsets of cities are henceforth referred to as *patches*. An overview of the number of extracted patches for each city is found in table 7 in the appendix. Figure 23 shows some exemplary patches, randomly chosen from the large dataset. In the following subchapter, the search of urban morphological patterns is conducted on these patches with an unsupervised clustering approach.

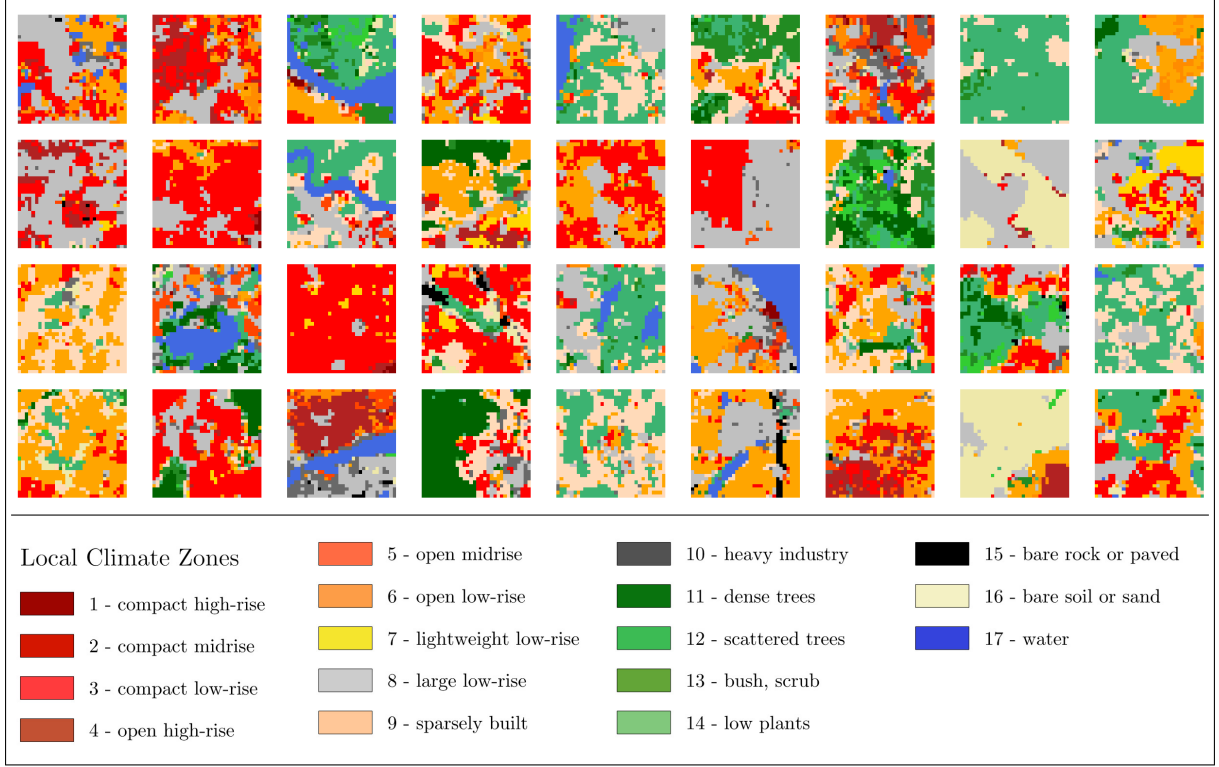


Figure 23: An exemplary random selection of extracted patches with $32 * 32$ pixels. They will be used in the following unsupervised clustering. The goal is to cluster patches with similar urban configurations into the same group and thus find patterns.

4.3.2 Semantic Clustering by Adopting Nearest Neighbors

The aim of this processing step is to cluster the extracted patches $x \in D$ into groups C_D . Patches that are semantically similar should be assigned to the same cluster, while patches that are considered as dissimilar should be assigned to different clusters. The challenge in grouping these patches results from the lack of a prior knowledge on the searched patterns. This master thesis aims to solve this by adapting Semantic Clustering by Adopting Nearest neighbors (SCAN), a state of the art method for unsupervised clustering. The framework consists of three processing steps:

- 1) Solving a pretext task with the SimCLR framework.
- 2) Clustering patches by adopting nearest neighbors.
- 3) Improving the cluster results with a self-labeling step.

Each of these steps makes relevant contributions towards the good results achieved in the work of Gansbeke et al. (2020). Yet, they need to be adapted to the characteristics of the LCZs data used in this master thesis (see chapter 4.1.3). Following the methodology of SCAN, adaptations will be mentioned along the way. The underlying network for all three steps of this framework is a ResNet-18 architecture (see figure 24). Larger networks (such as a ResNet-50 architecture) may achieve better results, but they need more processing power and computational time, both limited in the scope of this master thesis. The ResNet-18 is an 18-layered CNN built with residual blocks (see chapter 4.1.2). Let a *block* be a residual block with two convolutional layers as shown in figure 16. If the dimensions do not change within a block, the architecture of figure 16b is used, else the dimensions are adapted using the architecture of figure 16c. To handle the categorical data of the LCZs correctly within the network, the patches are one-hot encoded to 17 layers, each representing a binary image of one of the 17 LCZ classes. To feed the one-hot encoded patches with a dimensionality of $17 * 32 * 32$ to the ResNet-18, the first convolution layer of the network is adapted to receive a 17-layered input.

1) Solving a pretext task with the SimCLR framework.

The SCAN framework begins with solving a pretext task τ with the SimCLR framework. The goal is to learn semantically meaningful features and to mine nearest neighbors of a patch that will be used for clustering in the second step. SimCLR does this by maximizing the agreement between two augmented versions $T[X]$ of an image X with a Contrastive Loss L_{con} in the embedding space E (Chen, Kornblith et al. 2020). To start with, the architecture of the underlying network is explained.

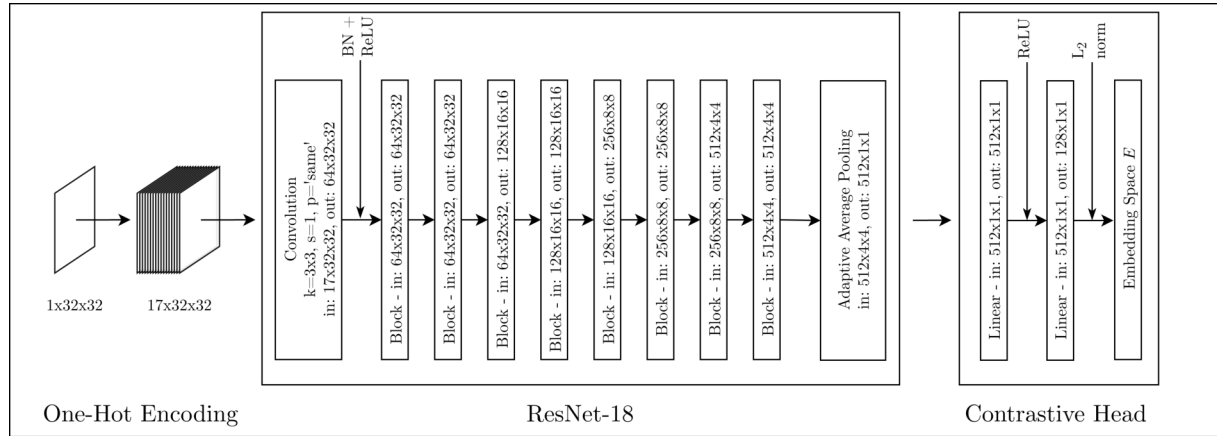


Figure 24: Network Φ_{Θ} built with a ResNet-18 $f(\cdot)$ and a Contrastive Head $g(\cdot)$ (own illustration). A patch X (or its augmentation $T[X]$) is one-hot encoded to 17 binary layers representing the 17 LCZs, before feeding it to the network. The input and output dimensions are shown in each step. The blocks of the ResNet-18 architecture correspond to the residual blocks presented in figure 16. The Contrastive Head is a fully connected MLP with one hidden layer.

A network Φ_{Θ} is built with a ResNet-18 $f(\cdot)$ and a Contrastive Head $g(\cdot)$. The Contrastive Head $g(\cdot)$ is a fully connected MLP with one hidden layer and a L_2 normalization. Figure 24 shows the architecture of the network. The weights of convolutional layers in $f(\cdot)$ are set with the He

initialization (cf. He, Zhang et al. 2015b). The last BN within each block is zero-initialized (cf. Goyal et al. 2017) while the remaining BNs are initialized with $weights = 1$ and $bias = 0$.

SimCLR consists of the major components shown in figure 25, following the work of Chen, Kornblith et al. (2020). For a patch x , two data augmentation operators are created from the adapted transformations T_{con} ($t_i \sim T_{con}$ and $t_j \sim T_{con}$) and applied to the patch, resulting in the pair \tilde{x}_i and \tilde{x}_j (an example of an original patch and possible resulting augmentations are shown in figure 55). The transformations T_{con} are adapted to the categorical nature of the LCZs data by discarding all color distortion operations while keeping, adapting and expanding the geometric distortion operators. The details of the transformations T_{con} used in this work are presented in table 1, found in the appendix. Both augmented versions \tilde{x}_i and \tilde{x}_j are one-hot encoded and passed through the ResNet-18 $f(\cdot)$. Thus, the representations

$h_{i,j} = f(\tilde{x}_{i,j}) = \text{ResNet-18}(\tilde{x}_{i,j})$ are obtained. They are passed further to the Contrastive Head $g(\cdot)$ (see figure 24) and result in the representations $z_{i,j}$ in the embedding space E . A mini-batch with B examples is randomly sampled and passed through these components, resulting in $2B$ augmentations \tilde{x}_k and thus representations z_k . The Contrastive Loss L_{con} (see equation 4) handles $\tilde{x}_{i,j}$ as a positive pair and the remaining $2(B - 1)$ as negative examples. Given \tilde{x}_i , the goal is to identify the corresponding \tilde{x}_j in $\{\tilde{x}_k\}_{k \neq i}$. With $\text{sim}(u, v) = u^T v / \|u\| \|v\|$ as the cosine similarity between u and v , the loss function L_{con} for a positive pair (i, j) is defined as

$$L_{con}(i, j) = -\log \frac{\exp(\text{sim}(z_i, z_j) / \kappa)}{\sum_{k=1}^{2B} \mathbb{1}_{[k \neq i]} \exp(\text{sim}(z_i, z_k) / \kappa)} \quad (4)$$

The temperature parameter κ regulates the training process of each epoch and the indicator function $\mathbb{1}_{[k \neq i]} \in \{0, 1\}$ evaluates if $k \neq i$. It is calculated for all positive pairs (i, j) and (j, i) of the mini-batch, resulting in the final loss. After processing one mini-batch, a training step is performed. Since SimCLR profits from large batch sizes, this work uses the largest possible mini-batch size of 256 samples (restrictions result from the GPU storage capacity available for this master thesis). The network Φ_Θ is trained for 500 epochs.

To recapitulate the SimCLR framework in one sentence: The loss function forces two augmented versions of one patch to be close to each other (*'agree with each other'*) in the embedding space E ,

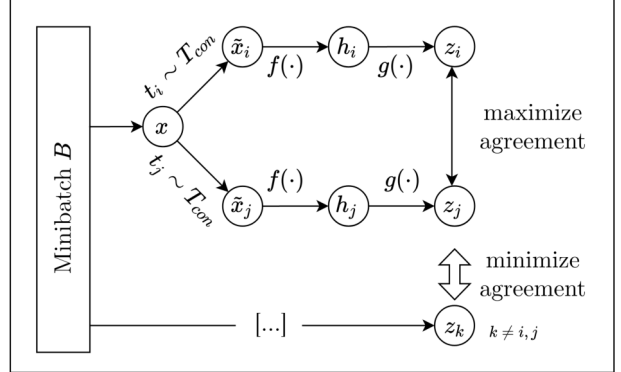


Figure 25: Major components of the SimCLR framework (own illustration). Two augmented versions $\tilde{x}_{i,j}$ are created for each sample x in a mini-batch B . The augmented versions are passed through the ResNet-18 $f(\cdot)$ and the Contrastive Head $g(\cdot)$ and result in the representations $z_{i,j}$ in the embedding space. The loss function L_{con} (see equation 4) maximizes the agreement of two augmented versions of the same patch, while minimizing the agreement with the other augmented versions of the other patches in the mini-batch.

while enforcing distance (*'disagree with each other'*) with the augmented versions of the remaining patches in the mini-batch. This allows the network $f(\cdot)$ to learn semantically meaningful features because the transformations of a patch show similar urban morphological configurations and thus force the network to look out for similarities in these versions, even though they are not identical. Most importantly, it is possible to feed the original (not augmented) patches to the trained network and mine the nearest neighbors of each patch. Since the network is trained to position similar patches close to each other in the embedding space E , the nearest neighbors of one patch are expected to share semantic similarities. Mining the k -nearest neighbors and clustering the patches will be the subject of the second step of the SCAN framework.

2) Clustering patches by adopting nearest neighbors.

In this step, the aim is to cluster the patches X of the dataset D with a new network Φ_η . The new network is built with a ResNet-18 $f(\cdot)$ and a Cluster Head $c(\cdot)$ (see figure 26). The previously learned weights Θ of the network Φ_Θ are used to initialize the new ResNet-18 $f(\cdot)$, while the former Contrastive Head $g(\cdot)$ is not used any further. The Cluster Head is a fully connected single layer with dimensions $|C| * 1 * 1$, where $|C|$ denotes the number of possible classes a patch can be assigned to.

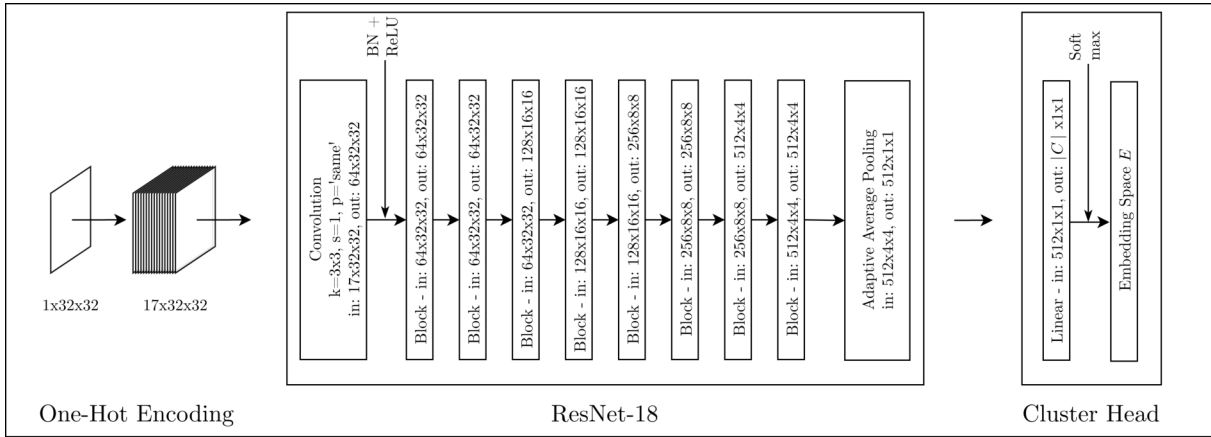


Figure 26: Network Φ_η built with a ResNet-18 $f(\cdot)$ and a Cluster Head $c(\cdot)$ (own illustration). A patch X (or its augmentation $T[X]$) is one-hot encoded to 17 binary layers representing the 17 LCZs, before feeding it to the network. The ResNet-18 $f(\cdot)$ is initialized with the weights Θ of the prior trained network Φ_Θ . With this initialization, the features learned by Φ_Θ are used in the new network. The Cluster Head is a fully connected single layer with the output dimensions $|C| * 1 * 1$. $|C|$ is referring to the number of classes.

Prior to this task, the network Φ_Θ was trained in a self-supervised manner by solving a pretext task τ with the SimCLR framework. The trained network is used to mine the K nearest neighbors for every patch $X_i \in D$ in the embedding space E . A set N_{X_i} is defined as the K nearest neighbors of a patch X_i . Gansbeke et al. (2020) empirically show that the nearest neighbors of an image tend to be in the same semantic class. Thus, they propose to adopt the nearest neighbors for their semantic clustering approach. Further they show that the results are insensitive to number of nearest neighbors K in the range $[5, 50]$. In this work the 20 nearest neighbors are mined for

each patch. Figure 27 shows an example of randomly selected patches (‘anchor image’) and their five nearest neighbors mined in the embedding space of Φ_Θ .

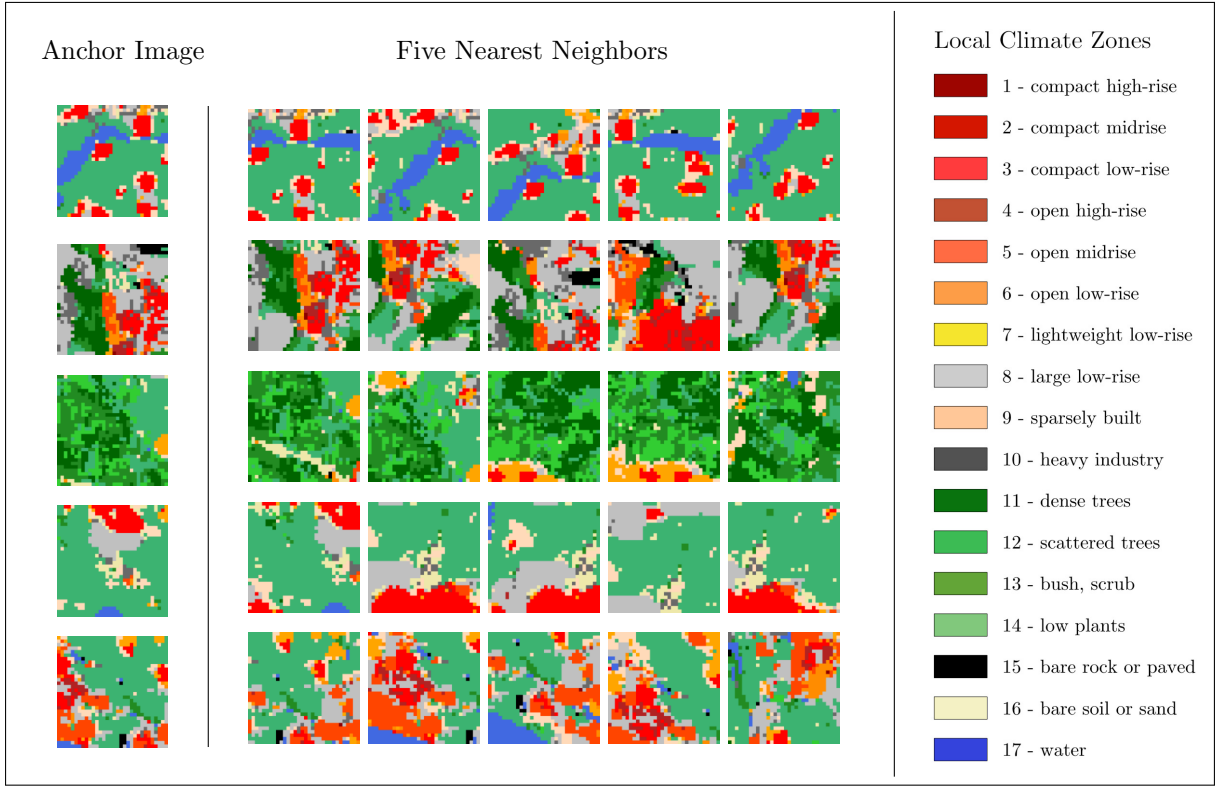


Figure 27: Exemplary patches (*anchor image*) and their five nearest neighbors. The patches can be considered as semantically similar. Some of the neighbors shown in this figure are partially overlapping with the anchor image. The K -nearest neighbors N_{X_i} for each patch are mined in the embedding space E of the trained network Φ_Θ .

Even though this master thesis has no metric to measure the similarity of these patches yet, a visual comparison discloses a high resemblance of the statistical and spatial LCZs patterns between an anchor image and its five nearest neighbors. These intermediate results motivate to continue with the SCAN framework.

To train the network Φ_η , a nearest neighbor $x_n \in N_{X_i}$ is randomly chosen for each sample x_i of a mini-batch (see figure 28). Both patches are transformed with two augmentation operators created from the adapted transformations T_{scan} ($t_i \sim T_{scan}$ and $t_n \sim T_{scan}$). Again, only geometric distortions are kept (for a detailed description of the adapted augmentations see table 2 in the

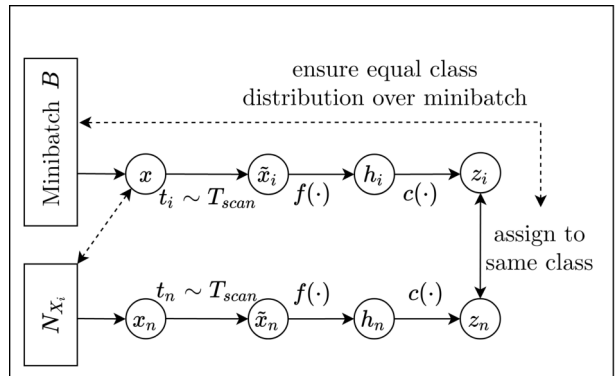


Figure 28: Clustering by adopting nearest neighbors (own illustration). For each sample $x \in B$ one of the 20 nearest neighbors N_{X_i} is randomly chosen. The image and its nearest neighbor are augmented with $t \sim T_{scan}$ and passed through the network $f(\cdot)$ and $c(\cdot)$. The loss function aims to maximize the agreement between the representations z_i and z_n , while maintaining an equal distribution of x assigned to the available classes.

appendix). An example of an original patch and possible augmentations are shown in figure 56. The augmented versions are passed through the network $f(\cdot)$ and $c(\cdot)$. The loss function L_{scan} (see equation 5) tries to assign the representations of a patch x and its nearest neighbor x_n to the same class, while maintaining an equal class distribution among the samples. The first term, the dot product operator $\langle \cdot \rangle$, is maximal when both patches are consistent (assigned to the same class) and confident (one-hot) of belonging to that class. The second term includes a cross entropy term for a uniform distributed assignment of the patches to the classes C (Gansbeke et al. 2020).

$$L_{scan} = -\frac{1}{|D|} \sum_{x \in D} \sum_{x_n \in N_X} \log \langle \Phi_\eta(x), \Phi_\eta(x_n) \rangle + \lambda \sum_{c \in C} \Phi_\eta'^c \log \Phi_\eta'^c \quad (5)$$

with $\Phi_\eta'^c = \frac{1}{|D|} \sum_{x \in D} \Phi_\eta^c(x)$

The number of classes has to be known in advance, which is a major impediment in unsupervised image clustering. Gansbeke et al. (2020) test their framework on benchmark datasets and simply take the already known number of classes. Further, they know that the samples of the benchmark datasets (such as STL-10 or CIFAR-10) are distributed equally among the classes. In the scope of this explorative master thesis, neither the number nor the distribution of different patterns representing urban morphological configurations are known. Hence, it is assumed that choosing a small number of classes might interfere with the search of patterns and generalize them too much. On the contrary, a large number of clusters is beneficial to the correct classification of urban patterns since it allows a more detailed distinction between the patterns. Even though there is a risk of over-clustering, it might be possible to merge similar clusters or re-assign patches to a smaller number of clusters in a later procedure. With these considerations, a number of 1024 classes is chosen. Since the loss function L_{scan} enforces a uniform distribution of a mini-batch over the given number of classes $|C|$, the largest batchsize for this network is chosen with 384 patches (again limited by the available hardware). The network is trained for 50 epochs, following the work of Gansbeke et al. (2020). The result of this second step is a clustering of all patches to the previously defined number of 1024 possible classes with a more or less uniform distribution. Patches are assigned to a class, based on their highest likelihood. The following third step in the SCAN framework aims to improve the obtained clustering results.

3) Improving the cluster results with a self-labeling step.

The network Φ_η was trained to assign semantically similar patches to one cluster with the help of the 20 nearest neighbors N_{X_i} mined for each sample in the first network Φ_Θ . Even though the nearest neighbors are likely to be semantically similar, it is inevitable that some samples with different characteristics (false positive samples) are among N_{X_i} . This introduces noise to the training process and clustering results of the network Φ_η . Gansbeke et al. (2020) developed a self-labeling step which re-trains the network Φ_η solely with highly confident samples. The confidence of a patch is described by the logits passed through the softmax function in the embedding space E of the network Φ_η (see chapter 4.1.2 and figure 26). Patches that have a high

confidence of belonging to a class c can be seen as *prototypes* representing this class. Re-training Φ_η with prototypes allows the network to correct itself by reclassifying missclassified patches and maintain the already correctly classified patches. Gansbeke et al. (2020) use a confidence threshold of $p = 99.9\%$ for selecting prototypes. This is not applicable in this work, since the patches do not reach such high confidence values for their classes. Figure 29 shows the sample confidence of exemplary classes. Each class contains only the patches that are assigned to it. Due to the variance among the classes, a fix threshold is not a good solution. Instead, a flexible, class dependent threshold is developed for this work.

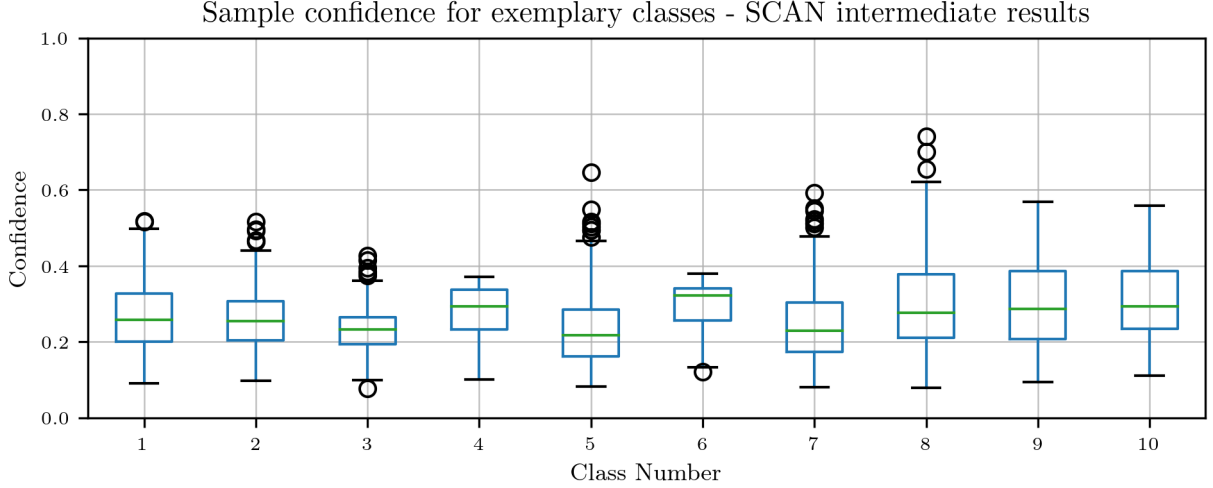


Figure 29: Boxplot of the sample confidence for exemplary classes (10 out of 1024 classes) resulting from the intermediate SCAN clustering. The samples have a mean confidence of 20-30% of belonging to the class they are assigned to. This is a result of over-clustering by choosing a high number of classes $|C|$. No samples exist with a confidence $p > 99.9\%$ as used in the framework of Gansbeke et al. (2020).

For each class $c \in C$, the threshold p_c is individually defined as $p_c = \max(\text{confidence}) - \sigma(\text{confidence})$ where σ denotes the standard deviation over the samples belonging to the class c . Further, the confidence p_x of a single patch x_i is not defined as the highest confidence $\max(z)$, but as the sum of the three highest class confidence values defined as $p_x = \sum \max3(z_i)$.

The network is re-trained by feeding each patch x (not augmented) of a mini-batch B to the network Φ_η and retrieving the target class c_t as well as the confidence value p_x (see figure 30). With the threshold value of the respective target class c_t it is checked if $p_x > p_{c_t}$. If the confidence is greater than the class dependent

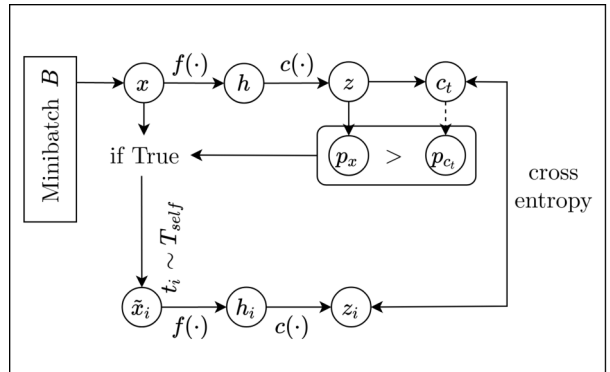


Figure 30: Re-training the network Φ_η with the self-labeling step (own illustration). Each patch $x \in B$ is passed through the network $f(\cdot)$ and $c(\cdot)$. The target class $c_t = \arg \max(z)$ is obtained to get the class specific threshold p_{c_t} . If the max3 threshold $p_x > p_{c_t}$, the patch is augmented and used for training. The aim of the loss function is to learn the target class c_t for the representation in z_i .

threshold (meaning the sample can be regarded as a prototype), it is used for training. Else it is deprecated for this epoch. The patches used for training are transformed with adapted augmentations operators $t_i \sim T_{self}$ to avoid overfitting (for details see table 3 in the appendix). Figure 57 shows some exemplary transformed patches. The augmented patches are passed through the network and a cross entropy loss function L_{CE}

$$L_{CE} = - \sum_{n=1}^C c_n \log(z_n) \quad (6)$$

is minimized to learn the target class c_t . The network is trained for 200 epochs. The batch size with 800 samples is again as large as possible (restrictions result from the available hardware).

This third step determines the final result of the SCAN framework. A network Φ_η was trained to cluster the patches into semantically similar classes. The dataset $D \leftarrow \{(x_i, y_i = \Phi_\eta(x_i)) | x \in D\}$ is updated to contain now the SCAN results y_i for each patch x_i . All in all, the patches are assigned to 138 different classes. The third step maintains the possible 1024 classes in the cluster head $c(\cdot)$, but is able to re-aggregate patches due to the new loss function. Figure 31 shows the patch confidence of some exemplary classes.

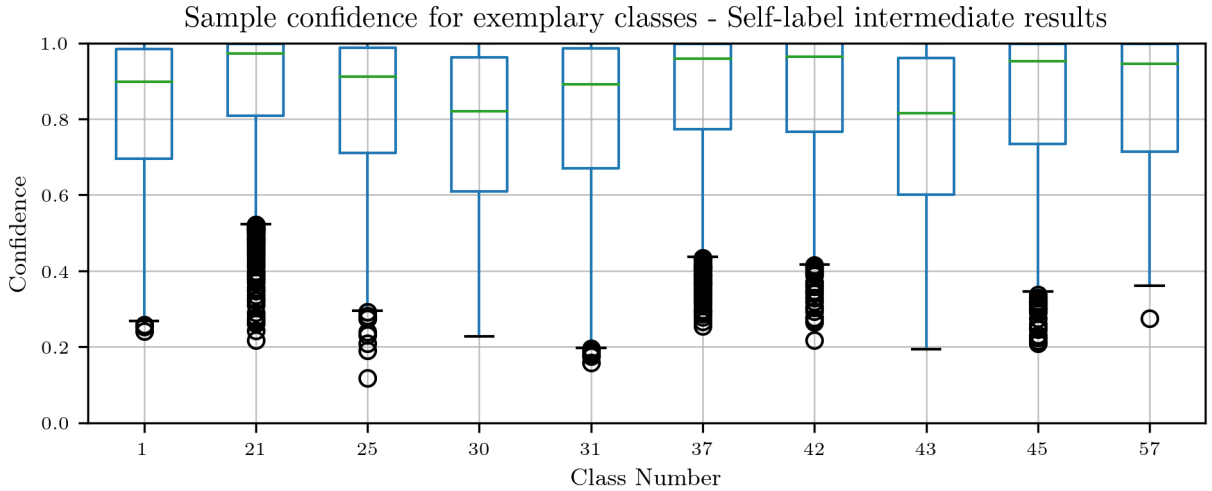


Figure 31: Boxplot of the sample confidence for exemplary classes (10 out of 138 classes) resulting from the intermediate self-label clustering.

Most of the patches show a high confidence of belonging to the class they have been assigned to. It visualizes the over-confident clustering criticized by Park, Han et al. (2020), which is also adverse for a direct patch comparison in this work (see chapter 4.1.3). This leads to the RUC framework to further improve the cluster results and smooth the sample confidence.

4.3.3 Robust learning for Unsupervised Clustering

The Robust learning for Unsupervised Clustering (RUC) framework developed by Park, Han et al. (2020) improves the existing model Φ_η by filtering unclear samples and re-training with

robust learning techniques. It is an additional module for already trained unsupervised networks. The RUC framework is composed of two components:

- 1) Extract clean samples with a confidence based strategy.
- 2) Re-train the model with the refined dataset.

The aim is to revise miss-classified samples and adjust the model confidence (Park, Han et al. 2020). This master thesis follows the RUC framework and adapts it to the needs of the LCZs data where necessary.

1) Extract clean samples with a confidence based strategy

The RUC framework begins with a confidence based sample strategy. The cluster results from the SCAN trained network Φ_η are considered as *pseudo-labels* y . The dataset $D = (x_i, y_i)_{i=1}^P$ in this framework contains for each image x_i the respective label $y_i = \Phi_\eta(x_i)$. The goal in this step is to divide the dataset $D = S \cup U$ in a set with clean samples $(x, y) \in S$ and unclean samples $u \in U$ where the labels are discarded. A sample $(x, y) \in D$ is assigned to the clean set S if the pseudo-label has a high confidence value $\max(y) > 99\%$. Otherwise the sample is assigned to U . The high threshold used by Park, Han et al. (2020) works well for the patches clustered with Φ_η , since each class has some very high confidence samples (see figure 31). The datasets S and U are used to refine the network in the second step.

2) Re-train the model with the refined dataset

In this step, two networks $\Phi_{\delta(1,2)}$ are initialized as a copy of the trained network Φ_η and re-trained via robust learning methods. For an epoch, both networks $\Phi_{\delta(1,2)}$ are iteratively trained (starting with $\Phi_{\delta(1)}$). In the following process the trained network (r) is referred to as $\Phi_{\delta(r)}$, its counter network (o) as $\Phi_{\delta(o)}$. Mini-batches B of an epoch are randomly sampled from $D = S \cup U$. The training process (shown in figure 32) consists of four major components: (i) Co-refinement, (ii) MixMatch, (iii) Label-smoothing and (iv) Co-refurbishing. The first three components contribute to the final loss function $L_{ruc(r)}$ and result in a training step for each mini-batch. The fourth component extracts clean samples at the end of each epoch for the following epoch.

(i) Co-refinement

The first component refines the labels within a mini-batch by using the predictions of both networks $\Phi_{\delta(r,o)}$. To describe the mechanism in simplified terms: For a patch x , the predictions of both networks are mixed and the result is sharpened again. This step intends to reduce overfitting and make the predictions more reliable (Park, Han et al. 2020). The component functions are as follows: The labels y of the clean samples $S = \{(x_b, y_b) : b \in (1, \dots, B)\}$ in the mini-batch B are co-refined to \bar{y} with the equations 7a and 7b, following the work of Li, Socher and Hoi (2020):

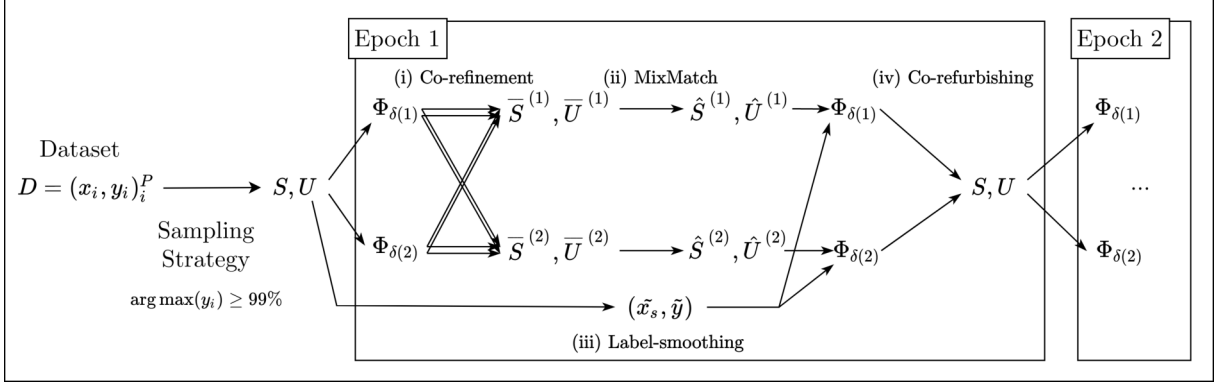


Figure 32: The major components of the RUC framework developed by Park, Han et al. (2020) (own illustration). The dataset D is split into a set of clean S and unclean U samples with a confidence based strategy. Two networks $\Phi_{\delta(1,2)}$ are trained within one epoch. Labels are (i) co-refined, the data are augmented with (ii) Mixmatch and a (iii) label-smoothing is performed. The final loss is calculated after these components and a training step is performed. After processing all mini-batches, the clean dataset S is updated via (iv) co-refurbishing for the following epoch.

$$\bar{y} = (1 - \omega^{(o)}) * y + \omega^{(o)} * \Phi_{\delta(o)}(x) \quad (7a)$$

$$\bar{y} = \text{Sharpen}(\bar{y}, \kappa) = \bar{y}^{c \frac{1}{\kappa}} / \sum_{c=1}^C \bar{y}^{c \frac{1}{\kappa}}, \text{ for } c = 1, 2, \dots, C \quad (7b)$$

with $\omega^{(o)}$ as the confidence value of x in the counter network $\Phi_{\delta(o)}$ and κ as a sharpening temperature parameter.

The unclean samples $U = \{(u_b) : b \in (1, \dots, B)\}$ in the mini-batch B are transformed with $M = 2$ weak augmentation operators ($t_w \sim T_{weak}$) resulting in \tilde{u}_m where $m = 1, \dots, M$ denotes the m -th augmentation (for details on T_{weak} see table 4 in the appendix). For a sample u , a pseudo-label \bar{q} is guessed with the help of both networks with the equations 8a and 8b:

$$\bar{q} = \frac{1}{2M} \sum_{m=1}^M (\Phi_{\delta(r)}(\tilde{u}_m) + \Phi_{\delta(o)}(\tilde{u}_m)) \quad (8a)$$

$$\bar{q} = \text{Sharpen}(\bar{q}, \kappa) = \bar{q}^{c \frac{1}{\kappa}} / \sum_{c=1}^C \bar{q}^{c \frac{1}{\kappa}}, \text{ for } c = 1, 2, \dots, C \quad (8b)$$

This co-refinement step results in the datasets $(x, \bar{y}) \in \bar{S}^{(r)}$ and $(u, \bar{q}) \in \bar{U}^{(r)}$ with refined labels.

(ii) MixMatch

The second component uses the MixMatch framework developed by Berthelot et al. (2019) to improve the noise resistance. Therefore patch x and its label y are mixed with the label from another randomly chosen patch. Let (x, y) denote a patch with its label from the combined co-refinement datasets $\bar{S}^{(r)} \cup \bar{U}^{(r)}$. A random data pair $(x_1, y_1) \wedge (x_2, y_2) \in \bar{S}^{(r)} \cup \bar{U}^{(r)}$ is augmented to (x', y') with the equations 9a- 9d:

$$\lambda \sim \text{Beta}(\alpha, \alpha) \quad (9a)$$

$$\lambda' = \max(\lambda, 1 - \lambda) \quad (9b)$$

$$x' = \lambda' x_1 + (1 - \lambda') x_2 \quad (9c)$$

$$y' = \lambda' y_1 + (1 - \lambda') y_2 \quad (9d)$$

where λ is a random sample from a *Beta*-distribution with the hyper-parameter $\alpha = 4$. The MixMatch component results in the augmented datasets $(\hat{x}, \hat{y}) \in \hat{S}^{(r)}$ and $(\hat{u}, \hat{q}) \in \hat{U}^{(r)}$. They are used to build two loss functions $L_{\hat{S}^{(r)}}$ defined in equation 10 and $L_{\hat{U}^{(r)}}$ defined in equation 11, both a part of the final loss function $L_{ruc(r)}$.

$$L_{\hat{S}^{(r)}} = \frac{1}{|\hat{S}^{(r)}|} \sum_{\hat{x}, \hat{y} \in \hat{S}^{(r)}} \sum_{n=1}^C \hat{y}_n \log(\Phi_{\delta(r)}(\hat{x}_n)) \quad (10)$$

$$L_{\hat{U}^{(r)}} = \frac{1}{|\hat{U}^{(r)}|} \sum_{\hat{u}, \hat{q} \in \hat{U}^{(r)}} \|\hat{q} - \Phi_{\delta(r)}(\hat{u})\|_2^2 \quad (11)$$

(iii) Label-smoothing

The third component regulates overconfident noise predictions of the model Park, Han et al. (2020). The clean samples $S = \{(x_b) : b \in (1, \dots, B)\}$ in the mini-batch B are transformed with a strong augmentation operator ($t_s \sim T_{strong}$) and result in \tilde{x}_s (for details on T_{strong} see table 5 in the appendix). Further, for all clean samples $(x, y) \in S$ uniform noise is injected into all classes C of the label y . The smoothed label \tilde{y} is defined with the following equation 12:

$$\tilde{y} = (1 - \epsilon) * y + \frac{\epsilon}{(|C| - 1)} * (1 - y) \quad (12)$$

with ϵ as a smoothing hyper-parameter and $|C|$ as the number of classes. If $\epsilon = 0$, then $\tilde{y} = y$ and if $\epsilon = 1$, then \tilde{y} is a uniform distribution over C . Park, Han et al. (2020) use in their exemplary code $\epsilon = 0.5$ which is adopted in this work. The strong augmentation \tilde{x}_s and its corresponding smoothed label \tilde{y} are used to build the loss function $L_{S_{strong}^{(r)}}$ (see equation 13) which is also a part of the final loss function $L_{ruc(r)}$.

$$L_{S_{strong}^{(r)}} = \frac{1}{|S^{(r)}|} \sum_{\tilde{x}_s, \tilde{y} \in S^{(r)}} \sum_{n=1}^C \tilde{y}_n \log(\Phi_{\delta(r)}(\tilde{x}_n)) \quad (13)$$

Thus, having the components (i)-(iii) processed for a mini-batch, the final loss function is defined as

$$L_{ruc(r)} = L_{S_{strong}^{(r)}} + L_{\hat{S}^{(r)}} + \lambda_U * L_{\hat{U}^{(r)}} \quad (14)$$

with a hyper-parameter λ_U to control the effect of $L_{\hat{U}(r)}$ (Park, Han et al. 2020). A training step is performed by adapting the weights δ of the network $\Phi_{\delta(r)}$ to minimize $L_{ruc(r)}$. After all mini-batches are processed, the trained network is referred to as $\Phi_{\delta(r')}$. The RUC framework continues with the last component to provide clean samples for the next epoch.

(iv) Co-refurbishing

With this component, the unclean samples U are revised and added to the clean samples X after training the network $\Phi_{\delta(r)}$ to $\Phi_{\delta(r')}$. Therefore unclean samples $u \in U$ are assigned to S if $\Phi_{\delta(r)}(u)$ and $\Phi_{\delta(r')}(u)$ predict the same label and the confidence value exceeds a threshold greater than or equal to 99% following equation 15:

$$S \leftarrow S \cup \{u \in U \mid \arg \max(\Phi_{\delta(r)}(u)) = \arg \max(\Phi_{\delta(r')}(u)) \wedge \max \Phi_{\delta(r')}(u) \geq 0.99\} \quad (15)$$

The four components (i-iv) are also conducted for $\Phi_{\delta(2)}$ as $\Phi_{\delta(r)}$ with $\Phi_{\delta(r')}$ now as the opposing network $\Phi_{\delta(o)}$. Concluding the training for both networks $\Phi_{\delta(1,2)}$, one epoch is successfully processed. The RUC framework is trained for 200 epochs with a mini-batch size of 110 samples. With the RUC framework uncertain samples are re-assigned to different clusters and the overconfidence of the results of the network is reduced. For each patch $x \in D$ a cluster result $y = \Phi_{\delta(1)}(x)$ is obtained. Thus the dataset is updated with $D \leftarrow \{(x, y = \Phi_{\delta(1)}(x)) \mid x \in D\}$. The cluster c to which a path is assigned can be obtained with $c = \arg \max(y)$. The patches are assigned to 138 different classes $C_D = c_1, \dots, c_{138}$, which is the same number of classes obtained by the SCAN framework. Figure 33 shows the confidence of patches assigned to their respective cluster. The confidence within each class is now better calibrated, compared to the confidence values obtained with the SCAN framework shown in figure 31.

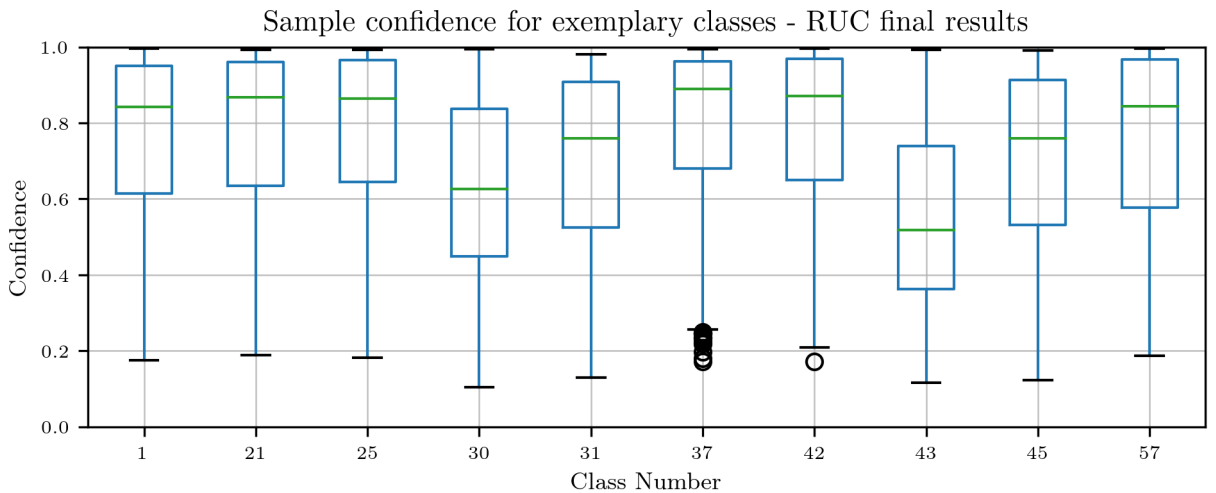


Figure 33: Boxplot of the sample confidence for exemplary classes (10 out of 138 classes) from the final RUC clustering result. The samples show a better calibration and less overconfident results compared to the SCAN results in figure 31. This work assumes that it is a better representation of the urban morphological reality and a good base for assessing the similarity of patches.

The clusters retrieved from the RUC framework denote the patterns of urban morphological configurations searched by the first research question. The found patterns build the foundation for the second research question, whose methodology is explained in the following subchapters 4.3.4-4.3.5. A detailed presentation of the cluster results obtained with RUC is provided in chapter 5.

4.3.4 Distance Calculation with Hungarian Matching

The second research questions aims to compare cities based on the found urban morphological patterns and examine the possible formation of geographic clusters. Therefore, a method to measure the similarity between cities has to be developed. This work uses the Kuhn–Munkres algorithm (Munkres 1957), also known as *Hungarian matching*, to calculate a similarity metric between the cities.

So far, this work has conducted a search of morphological patterns on the subsets of cities (referred to as patches in this work). Thus, a description y of each patch x was obtained, which represents not only knowledge on a specific type of cluster with $c = \arg \max(y)$, but also a fuzzy membership among all classes C (see chapter 4.1.1). Since each city is composed by a set of patches, the similarity of cities is measured by comparing the fuzzy membership y of their patches x . If the set of patches of one city is similar to the set of patches of another city, both cities can also be considered as similar.

The patches x are described by a 1024-dimensional vector y , thus y is the *feature space* of x . The 1024 dimensions represent the confidence values of the possible classes C as defined in the SCAN and RUC frameworks. Before starting with a comparison of patches, the feature space is reduced to remove irrelevant characteristics and decrease the computational time (see chapter 4.1.1). Since the patches are assigned to only 138 out of the possible 1024 classes, its is an indication that the remaining 886 classes did not play an important role for re-training the network and clustering the patches in RUC. Thus, a feature reduction is conducted by keeping only the relevant 138 dimensions of y .

The similarity between two patches can be described by the similarity of their representations y . Since y is a vector, the similarity of two patches x_1 and x_2 can be expressed by the Euclidean distance between y_1 and y_2 . The Euclidean distance is defined as

$$d_{1,2}^2 = \|y_1 - y_2\|^2 \quad (16)$$

with $d_{1,2}^2$ in range $[0, \sqrt{2}]$. If two patches are similar, the distance is close to 0. If two patches are different and hence very certain of belonging to two different classes c (and thus have two one-hot encoded vectors for different classes c), the distance is $\sqrt{2}$. This results from the softmax activation function where values of y are in range $[0, 1]$ (see chapter 4.1.2).

Let $V = \{v_1, \dots, v_{1523}\}$ denote a set of all 1523 MUAs used in this work and a MUA $v_i = \{y \mid (x, y) \in D \wedge \text{within}(v_i, x)\}$ denote a set of vectors y from the patches x within the MUA v_i . To compare the patches of one city v_i with the patches from another city v_j , a distance matrix $A_{i,j}$ is calculated in the equations 17- 18 with the Euclidean distance

$$pm_{m,n} = d_{m,n}^2 = \|a_m - b_n\|^2, \text{ for } a_1, \dots, a_m \in v_i, b_1, \dots, b_n \in v_j; \quad (17)$$

$$PM_{i,j} = (pm_{m,n}) \quad (18)$$

Having the distance matrix $PM_{i,j}$ with size $m \times n$ for two cities v_i and v_j , it is now possible to calculate a distance metric $d_{final}(v_i, v_j)$ between both cities via Hungarian matching. The Hungarian matching algorithm tries to assign each patch of a city to the patch of the other city by minimizing the overall costs $d_{total}(v_i, v_j)$ (see figure 34). The costs are in this case the distances between the patches, represented in the distance matrix $PM_{i,j} = (pm_{m,n})$. It is a linear assignment problem, where a patch is assigned to only one other patch. If $m \neq n$ it is an unbalanced assignment problem, meaning not all patches from a larger city can be assigned to a smaller city. The Hungarian matching algorithm determines the variables $x_{m,n}$ which denote the assignment of patches

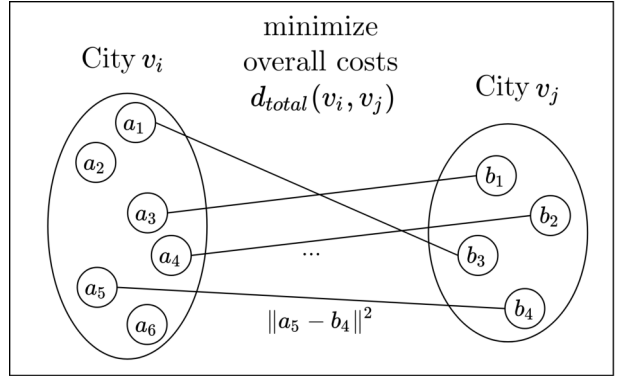


Figure 34: Distance calculation via Hungarian matching (own illustration). Each patch of one city is assigned to the patch of another city by minimizing the overall costs $d_{total}(v_i, v_j)$ with equations 19- 20. If $|v_i| > |v_j|$, not all patches are assigned.

$$x_{m,n} = \begin{cases} 1, & \text{if patch } m \text{ and } n \text{ are assigned to each other} \\ 0, & \text{otherwise} \end{cases} \quad (19)$$

by minimizing the overall costs $d_{total}(v_i, v_j)$

$$\begin{aligned} \text{minimize } d_{total}(v_i, v_j) &= \sum_{k=1}^m \sum_{l=1}^n x_{k,l} a_{k,l} \\ \text{under the conditions } \sum_{k=1}^m x_{k,l} &= 1 \text{ for } l = 1, \dots, n; \quad \sum_{l=1}^n x_{k,l} = 1 \text{ for } k = 1, \dots, m \end{aligned} \quad (20)$$

The costs d_{total} are dependent on the number of patches in the MUAs. The more patches can be coupled, the higher are the costs. To allow a comparison between pairs of small cities and pairs of large cities, the totals d_{total} are normalized by the number of assigned patches and result in the final costs d_{final} defined as

$$d_{final}(v_i, v_j) = \frac{d_{total}(v_i, v_j)}{\min(m, n)}, \text{ with } m = |v_i|, n = |v_j| \quad (21)$$

The final cost $d_{final}(v_i, v_j)$ is the metric value to describe the similarity of two cities v_i and v_j based on the similarity of their patches. With this mathematical foundation, the similarity between all cities $v \in V$ is calculated. The result is a distance matrix CM_V defined as

$$cm_{i,j} = d_{final}(v_i, v_j), \text{ for } i, j = 1, \dots, |V|; v \in V \quad (22)$$

$$CM_V = (cm_{i,j}) \quad (23)$$

The city distance matrix CM_V represents a similarity metric between all cities. It is the foundation on which the similarity of cities is further explored. To examine the formation of possible geographical clusters similar to the work of Taubenböck, Debray et al. (2020), the next subchapter continues with clustering similar cities.

4.3.5 Clustering Cities based on Urban Morphological Configurations

So far, this work searched patterns of similar urban morphological configurations and calculated thereon a metric to measure the similarity between cities with new explorative approaches. This work continues to group cities based on their similarity and examine the possible formation of geographic clusters with the baseline used by Taubenböck, Debray et al. (2020). They use the k-means algorithm for clustering cities, while the optimal number of clusters is determined with the gap statistic algorithm. However, having only the city distance matrix CM_V poses a problem for k-means clustering. The k-means algorithm implies Euclidean distances to minimize the squared deviations between data points and the centroids of k-clusters. Thus, the distance matrix has to be transformed to a n-dimensional Euclidean space.

Multi-dimensional scaling (MDS) embeds the elements of a distance matrix into a n-dimensional space E_{MDS} with the aim to preserve the distances between the elements (Borg and Groenen 2005). Let $d_{i,j}$ denote the original distance $d_{final}(v_i, v_j)$ between two cities calculated in equation 22. The *stress* value of a MDS gives a hint on how well the distance $d_{i,j}$ is represented by $\hat{d}_{i,j}$ in the transformed space. The stress value for transforming the city distance matrix via MDS is defined as

$$stress = \sum_{i \neq j=1, \dots, |V|} (d_{i,j} - \hat{d}_{i,j})^2 \quad (24)$$

The stress value is dependent on the number of dimensions n . The more dimensions are used by MDS to represent the elements in the embedding space E_{MDS} , the better the original distances of CM_V can be preserved. Figure 35 shows the MDS stress value for mapping CM_V with the use

of dimensions n in range $[1, 15]$. The stress value decreases with the use of more dimensions n .

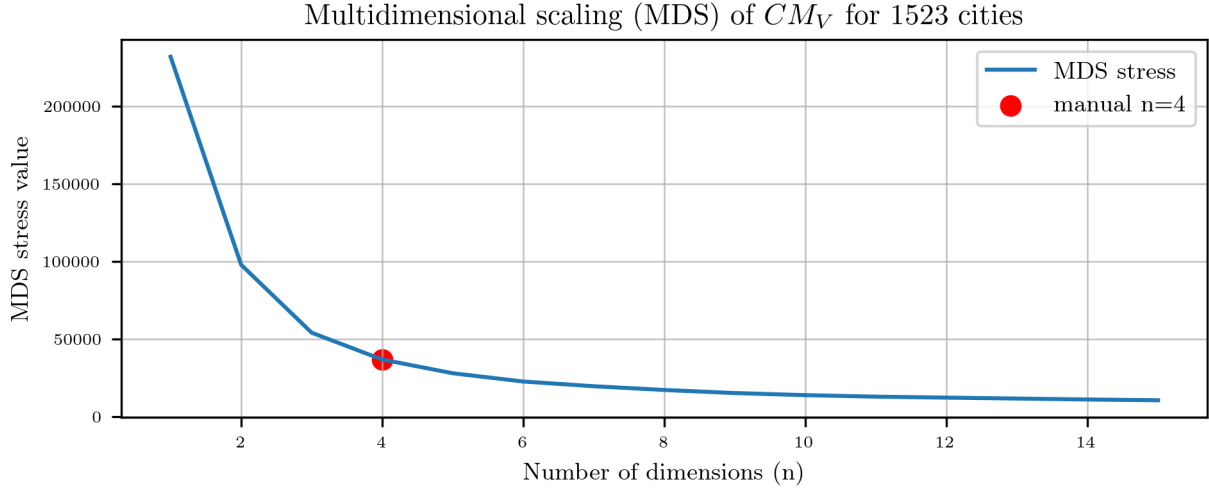


Figure 35: Multi-dimensional scaling (MDS) stress value for transforming the similarity of all 1523 cities represented in CM_V to n -dimensions in range $[1, 15]$. The stress value indicates how well the true distances in CM_V can be represented with the given number of dimensions. A lower stress value indicates a better representation. With the *elbow method* $n = 4$ dimensions are manually chosen.

On the other side, the reduction of the stress value decreases constantly with additional dimensions. This means, that additional dimensions contribute less and less to the correct representation of the original distance $d_{i,j}$. Since more dimensions constrain the clustering due to the *curse of dimensionality* (see chapter 4.1.1), it is not effective to use a very high number of dimensions. Thus, a number of dimensions has to be selected that adequately represents the true distances of CM_V while avoiding redundant dimensions. A method to determine an acceptable number is the manual selection of dimensions n with the *elbow method* (cf. Yuan and Yang 2019). In figure 35, $n = 4$ dimensions are selected, even though a sharp elbow is not identified unambiguously. Thus, CM_V is transformed to a 4-dimensional embedded space E_{MDS} . Let $HM_{1523} = \{\vec{v}_{i=1}^{|V|}\}$ denote the coordinates \vec{v} of CM_V in E_{MDS} for all cities $v \in V$. For the ablation study (see chapter 3), the same procedure is repeated with a reduced city distance matrix CM_{V110} , containing only the 110 cities used in the study of Taubenböck, Debray et al. (2020) (see figure 36).

Again, $n = 4$ dimensions are chosen manually. CM_{V110} is transformed via MDS and thus $HM_{110} = \{\vec{v}_{i=1}^{110}\}$ is obtained, representing the coordinates \vec{v} of CM_{V110} in E_{MDS} for all 110 cities of the reduced dataset.

Following the work of Taubenböck, Debray et al. (2020), the master thesis continues with the clustering of cities. As a prerequisite for using the k-means clustering algorithm, the similarity of cities was transformed with MDS to a representation of cities in a four-dimensional coordinate space. The k-means algorithm divides a dataset HM into k -clusters $C = \{c_{r=1}^k\}$ (cf. MacQueen et al. 1967). Each cluster in C is described by a centroid μ_r .

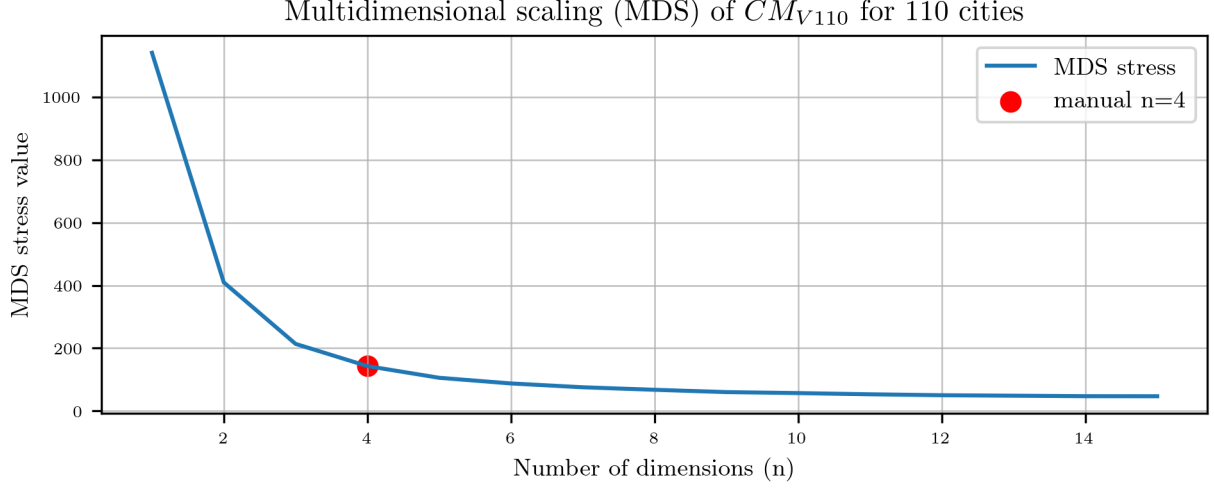


Figure 36: Multi-dimensional scaling (MDS) stress value for transforming the 110 cities used by Taubenböck, Debray et al. (2020) to n -dimensions in range $[1, 15]$. The true distances in CM_{V110} can be represented better with more dimensions (lower stress value). With the *elbow method* $n = 4$ dimensions are manually chosen.

The k-means algorithm works in the following way:

- (i) Initialize k-centroids μ_r with the k-means++ initialization (cf. Arthur and Vassilvitskii 2006).
- (ii) Assign elements $\vec{v} \in HM$ to clusters C , based on the smallest distance to the centroids μ_r .
- (iii) Re-calculate the centroids μ_r as the mean of elements assigned to the clusters C .

The steps (ii)-(iii) are repeated, until either 300 iterations are reached or the Frobenius norm of the difference between μ_r of two iterations falls below a threshold of $1e-4$ (meaning the centroids do not significantly move anymore). Thus, k-means aims to minimize

$$\sum_{i=1}^{|HM|} \min_{\mu_j \in K} (\|x_i - \mu_j\|^2) \quad (25)$$

Yet, the number of clusters k has to be known in advance. Taubenböck, Debray et al. (2020) use the gap statistic algorithm to determine the optimal number of clusters k^* . The gap statistic method developed by Tibshirani, Walther and Hastie (2001) measures with $Gap(k)$ the difference between the within-cluster dispersion W_k of the dataset HM and a random uniform distribution E^* with k clusters. Let $d_{i,j} = \|\vec{v}_i - \vec{v}_j\|^2$ denote the squared Euclidean distance between two cities $\vec{v}_{i,j} \in HM$ for $i, j = 1, \dots, |HM|$. Further, let $C = \{c_1, \dots, c_k\}$ be the k-means cluster result with k clusters and $c_r \in C$ denote the elements in cluster r with $n_r = |c_r|$. Following the work of Tibshirani, Walther and Hastie (2001), the sum of pairwise Euclidean distances for a cluster r is defined as

$$D_r = \sum_{i,j \in c_r} d_{i,j} \quad (26)$$

and the within-cluster variance over k clusters as

$$W(k) = \sum_{r=1}^k \frac{1}{2n_r} D_r \quad (27)$$

The gap metric $Gap(k)$ is defined as

$$Gap(k) = E^*\{\log(W(k))\} - \log(W(k)) \quad (28)$$

with $E^*\{\log(W(k))\}$ as the average of $B = 3$ random uniform distributions $\log(W^*(k))$ such that

$$E^*\{\log(W(k))\} = \frac{1}{B} \sum_b^B \log(W_b^*(k)) \quad (29)$$

The standard deviation $sd(k)$ is computed as

$$sd(k) = \left[\frac{1}{B} \sum_b^B \{\log(W_b^*(k)) - E^*\{\log(W(k))\}\}^2 \right]^{\frac{1}{2}} \quad (30)$$

and the metric $s(k)$ as

$$s(k) = sd(k) \sqrt{1 + \frac{1}{B}} \quad (31)$$

Different ways of finding the optimal number of k^* with these metrics of the gap statistic algorithm exist. Taubenböck, Debray et al. (2020) use the *firstSEmax* method to find k^* with

$$k^* = \arg \min_k \{k \in \mathbb{Z}^+ \wedge k \geq 2 \mid (Gap_{diff}(k) = Gap(k) - Gap(k+1) - s(k+1)) > 0\} \quad (32)$$

which is the smallest $k \geq 2$ that fulfills the condition of $Gap_{diff}(k) > 0$ in equation 32. The optimal value k^* can vary due to the random uniform distributions. Therefore the search of k^* is conducted 100 times with different seed keys. The final number of optimal clusters

$$k_{final}^* = \text{median}(k_{i=1}^{*100}) \quad (33)$$

is the median value of the 100 calculates of k^* .

With these foundations, the aim is to find k_{final}^* and cluster the datasets HM_{1523} and HM_{110} . Figure 37 shows the 100 calculations i of $Gap_{diff}(k)$ on HM_{1523} . The work of Taubenböck, Debray et al. (2020) determine k^* in the range $[1, 25]$, which is extended to the range $[1, 50]$ for the larger dataset HM_{1523} containing 1523 cities. The optimal number of cluster $k_{final}^* = 7$ is

the median value of the clusters k_i^* . Having found the optimal value, the 1523 cities in HM_{1523} are assigned to seven clusters with the k-means algorithm, resulting in the clusters $C_{HM_{1523}}$.

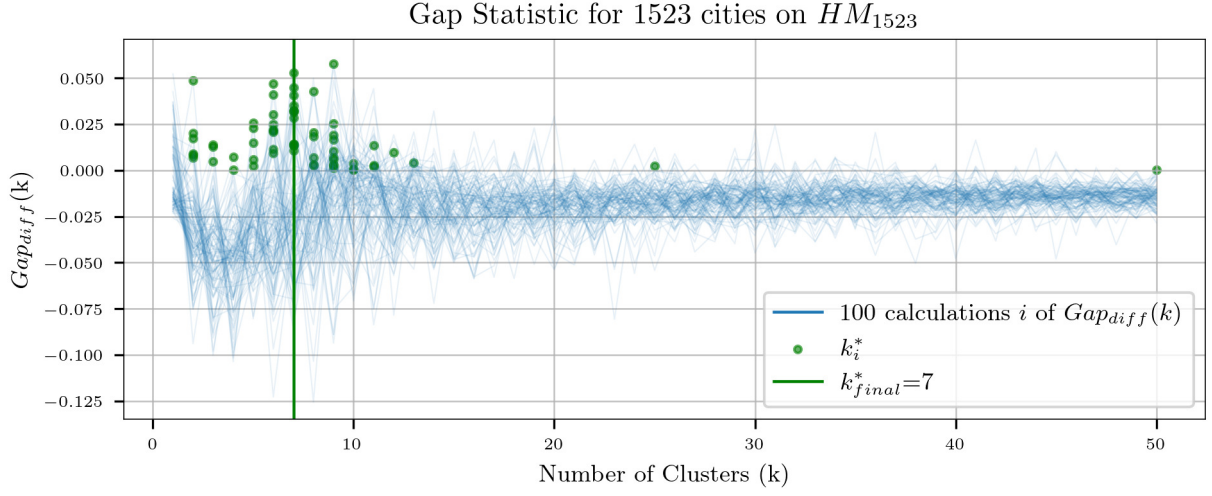


Figure 37: The $Gap_{diff}(k)$ value calculated 100 times on HM_{1523} with different key seeds following the work of Tibshirani, Walther and Hastie (2001). $Gap_{diff}(k)$ was calculated for k in range $[1, 50]$. The green dots show the optimal k_i^* found with equation 32. The final optimal cluster k_{final}^* is the median value of all found k_i^* (see equation 33).

The same procedure is performed for HM_{110} but with k in range $[1, 25]$ to allow a better comparison of the clustering result to the work of Taubenböck, Debray et al. (2020). The result is shown in figure 38. Also for HM_{110} the optimal number of clusters is $k_{final}^* = 7$. Hence, the 110 cities in HM_{110} are assigned with the k-means algorithm to seven clusters, resulting in $C_{HM_{110}}$.

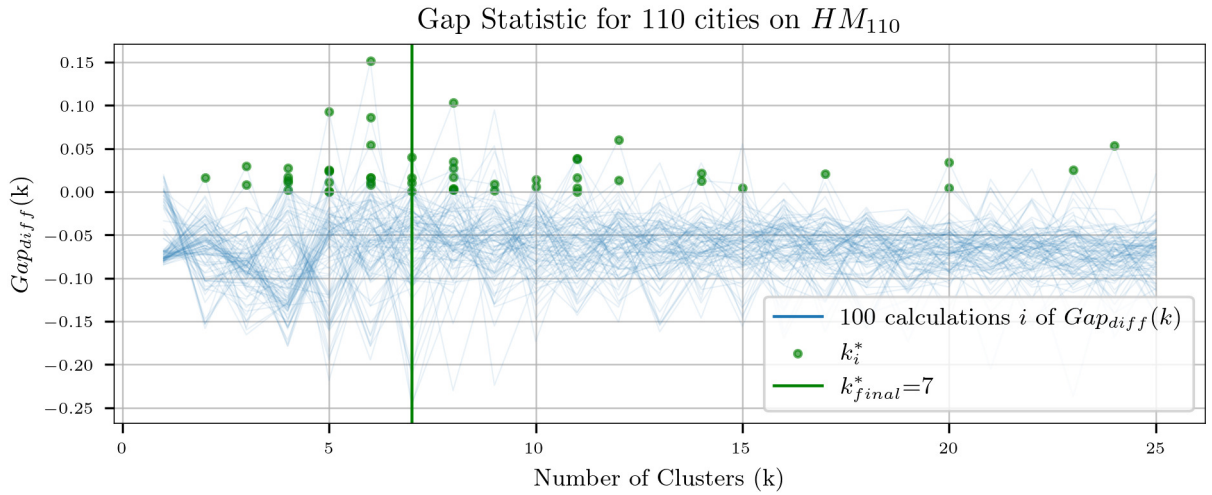


Figure 38: The $Gap_{diff}(k)$ calculated 100 times on HM_{110} for k in range $[1, 25]$. The green dots show the optimal k_i^* found with equation 32. The final optimal cluster k_{final}^* is the median value of all found k_i^* (see equation 33).

This final step completes the clustering of cities based on the similarity metric developed in the previous chapter 4.3.4. It also concludes the new explorative approaches developed in this master thesis. For the second ablation study (see chapter 3), the master thesis uses in the next subchapter the methodological baseline developed by Taubenböck, Debray et al. (2020) to extend their analysis to a larger database.

4.3.6 Clustering Cities with a 18-dimensional Feature Space

The second ablation study (see chapter 3) continues the work of Taubenböck, Debray et al. (2020) with a larger database. Therefore the 18-dimensional feature space used in their study is created additionally for the 1523 cities. Similar to the data preparation process (see chapter 4.3.1), the raw LCZs raster tiles are cleaned and merged, but clipped with the original MUAs which are not buffered. The first 17 dimensions represent the relative coverage of the 17 LCZs classes within each MUA. The 18th dimension represents the absolute area of LCZs within each MUA. The first 17 dimensions are centered by the median and normalized with the min-max normalization. The same procedure is performed for the 18th dimension, resulting in a 18-dimensional feature space for the 1523 cities F_{1523} .

To group similar cities together, their representations in F_{1523} are clustered with the k-means algorithm. Therefore, the optimal number of clusters is determined with the gap statistic algorithm, following the procedure presented in the previous chapter 4.3.5. The $Gap_{diff}(k)$ is calculated 100 times with different seed keys for k in range $[1, 50]$ (see figure 39).

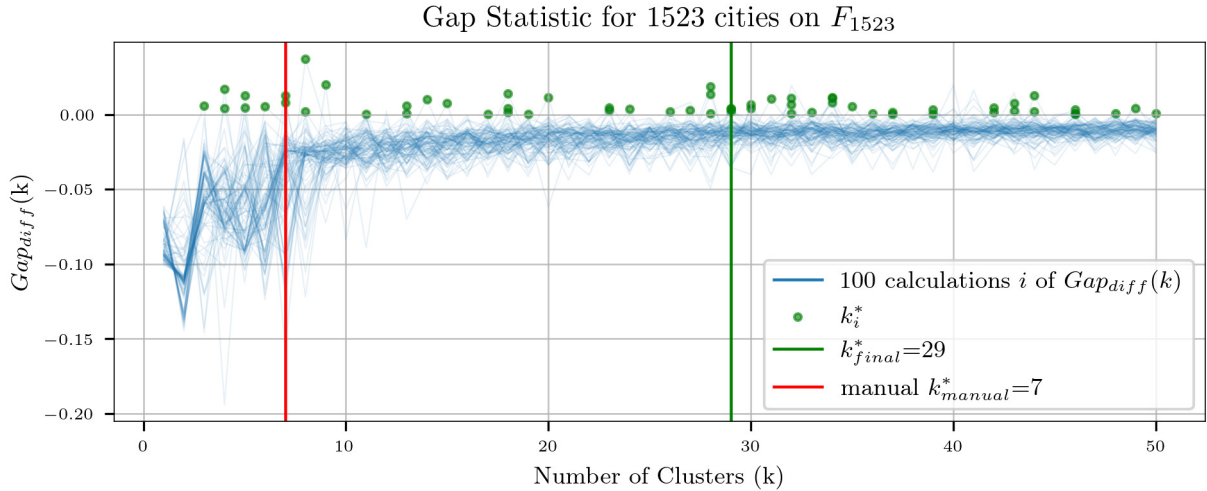


Figure 39: The $Gap_{diff}(k)$ on F_{1523} for k in range $[1, 50]$. $Gap_{diff}(k)$ is calculated 100 times with different seed keys (as presented in chapter 4.3.5). The green dots show the optimal k_i^* found with equation 32. The final optimal cluster $k_{final}^* = 29$ is the median value of all found k_i^* (see equation 33). The clusters k_i^* are spread more evenly over the range $[1, 50]$ and indicate no clear k_{final}^* , compared to the accumulated clusters found in $C_{HM_{1523}}$ and $C_{HM_{110}}$. Additionally F_{1523} is clustered with $k_{manual}^* = 7$ to allow a better visual comparison with $C_{HM_{1523}}$ and $C_{HM_{110}}$.

For F_{1523} the optimal number of clusters is $k_{final}^* = 29$. Therefore, the 1523 cities in F_{1523} are assigned to 29 clusters with the k-means algorithm, resulting in $C_{F_{1523}}$. To allow a better visual comparison with the seven clusters found in $C_{HM_{1523}}$, $C_{HM_{110}}$ and the seven clusters obtained by Taubenböck, Debray et al. (2020), the 1523 cities of F_{1523} are additionally assigned to seven clusters, resulting in $C_{F_{1523}}^{*7}$.

This concludes the methodology of this master thesis. The methods used in this work were presented in great detail, along with supplementary figures and tables in the appendix. The following chapter 5 presents the results obtained with this analysis to answer the research questions raised in chapter 3.

5 Results

This explorative work investigates the urban morphology on a global scale through the scope of new experimental approaches of unsupervised learning. The aim is to identify patterns of urban morphological configurations and develop a similarity metric of cities, based on these found patterns. Further, similar cities are grouped together to see if geographical clusters form. This chapter presents the result of the research question raised in chapter 3, beginning with the search of urban morphological patterns (chapter 5.1), followed by the clustering of cities (chapter 5.2) and the ablation studies on clustering cities (chapter 5.3).

5.1 Patterns of Intra-urban Morphological Configurations

Is it possible to find patterns of urban morphological configurations based on the LCZs across the world?

This work understands urban morphological patterns as re-occurring spatial arrangements of similar urban configurations. Since there is no prior knowledge for the formalization of urban patterns, the search was conducted with new unsupervised clustering approaches to group similar subsets (here *patches*) of cities (see chapter 4.3.1- 4.3.3). The patches were assigned to 138 different clusters $C_D = c_1, \dots, c_{138}$. This work assumes that the obtained clusters represent different patterns of urban morphological configurations. Each patch $x \in D$ is associated with a fuzzy membership function y which describes the likelihood of belonging to the clusters $c \in C_D$. A patch x can be assigned to a cluster with $c = \arg \max(y)$. Presenting the 138 clusters found in this work, along with the total amount of 316,536 patches assigned to them, is not practicable. Instead, a qualitative description is provided for some exemplary clusters.

Figure 40 shows patches assigned to a first exemplary cluster c_1 . The confidence value of belonging to this cluster is provided along with the patches. The cluster c_1 contains patches with a mixed variety of LCZ classes. The most dominant urban morphological components are LCZ-3 (*compact low-rise*) next to LCZ-6 (*open low-rise*) and LCZ-8 (*large low-rise*). These components come in different shapes and arrangements within the patches of this cluster. The size can reach from single scattered components covering only one pixel (100m * 100m) to medium sized components with almost 10 * 10 pixels (1km²). Some of the dominant non-built land coverage classes include LCZ-14 (*low plants*) and LCZ-16 (*bare soil or sand*). Further, some patches show the LCZ-17 (*water*) in form of a river or small lakes. The morphological impression of this exemplary class c_1 can be described as urban areas predominantly built with low-buildings in varying densities. In between the built-environment are islands of low vegetation or open spaces and sometimes small lakes or rivers. The patches themselves show a rather heterogeneous composition of the above mentioned components. Even among very confident samples (confidence > 90%) the compositions are ambiguous. Nevertheless, the overall picture of the patches assigned to this class looks consistent, even though a direct comparison of its patches would sometimes unveil different urban morphological configurations.

Another exemplary cluster c_2 is shown in figure 41. The patches of this cluster show large homogeneous areas of LCZ-3 (*compact low-rise*) and LCZ-8 (*large low-rise*), each covering up to 2/3 of the patch. Along with it, some medium sized components of LCZ-6 (*open low-rise*) can be found. Adjacent to the areas of LCZ-3 (*compact low-rise*), also higher built-up structures of LCZ-2 (*compact mid-rise*) and for very few patches even LCZ-1 (*compact high-rise*) can be found. The patches in c_2 denote truly urban areas with only little non-built LCZ classes among some patches. Some of the few non-built areas can be found with LCZ-12 (*scattered trees*), LCZ-13 (*bush-scrub*) and a very few examples with LCZ-17 (*water*). The exemplary class c_2 gives a morphological impression of more homogeneous urban areas (compared to cluster c_1), with mostly low-built buildings. Small clusters of mid-rise or high-rise buildings are among some patches. A subjective impression of the cluster c_2 suggests similar spatial forms and statistical distributions of the components shared across the patches. Thus, the patches among this cluster can be regarded as semantically similar, even though minor differences arise when comparing patches directly with each other.

The exemplary cluster c_3 shown in figure 42 contains patches with large homogeneous areas of LCZ-7 (*lightweight low-rise*) covering up to 1/2 of the patch. But also similar large areas of LCZ-8 (*large low-rise*) can be found. Further, components of LCZ-10 (*heavy industry*) and also LCZ-6 (*open low-rise*) are among some patches. Also scattered areas of LCZ-5 (*open midrise*) and LCZ-3 (*compact low-rise*) are shared across a few patches. Nevertheless, the non-built environment also covers up to 1/3 of a patch assigned to this class. The predominant non-built classes are LCZ-12 (*scattered trees*) and LCZ-13 (*bush, scrub*) also with some LCZ-11 (*dense-trees*). Rivers and lakes with LCZ-17 (*water*) can be found among some patches, often adjacent to other elements of the non-built environment. Similar to the other classes, no clear pattern among the confidence values can be observed with a visual comparison of the patches. The patches in this exemplary class c_3 look semantically similar to each other, even though the individual patches cover a broad variety of specific urban morphological configurations.

Figure 43 shows patches assigned to the exemplary cluster c_4 . The urban morphology among these patches is diverse and compartmentalized, similar to c_1 . Most dominant classes of the built environment are LCZ-8 (*large low-rise*) with a mixture of LCZ-5 (*open midrise*) and LCZ-6 (*open low-rise*). Further, small components of LCZ-4 (*open high-rise*) and LCZ-1 (*compact high-rise*) are found. The non-built environment is characterized by LCZ-14 (*low plants*) with small areas of LCZ-16 (*bare soil or sand*) as well as LCZ-17 (*water*) forming rivers.

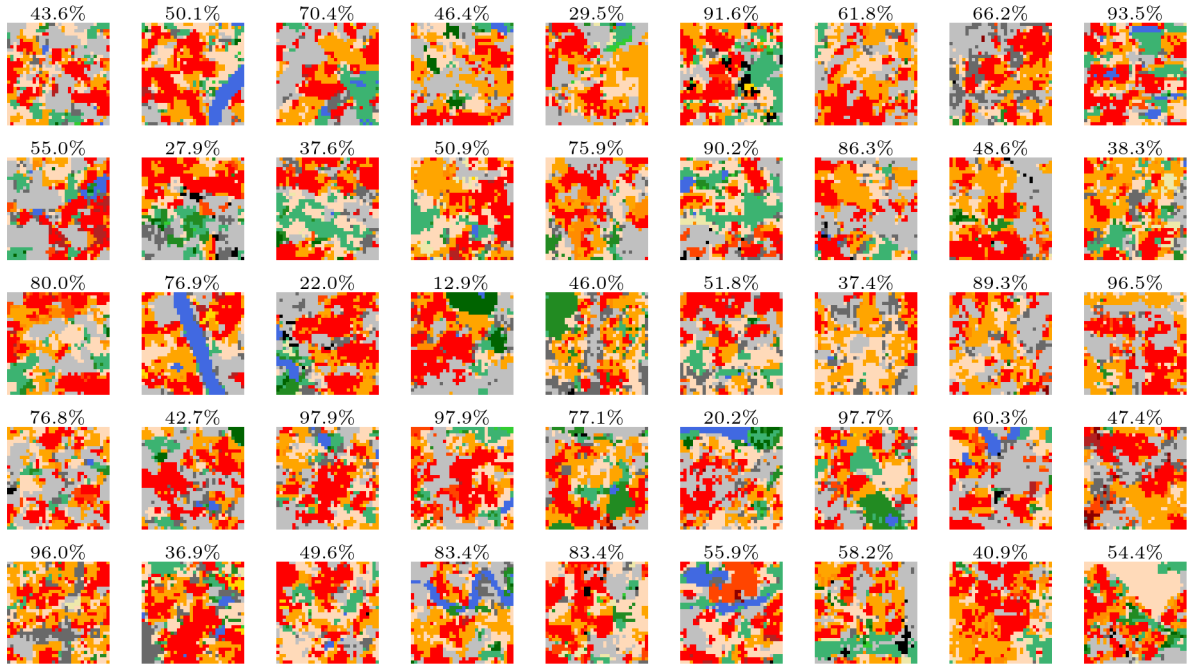


Figure 40: Exemplary cluster c_1 with 1043 patches in total. A random selection and their confidence values is shown.

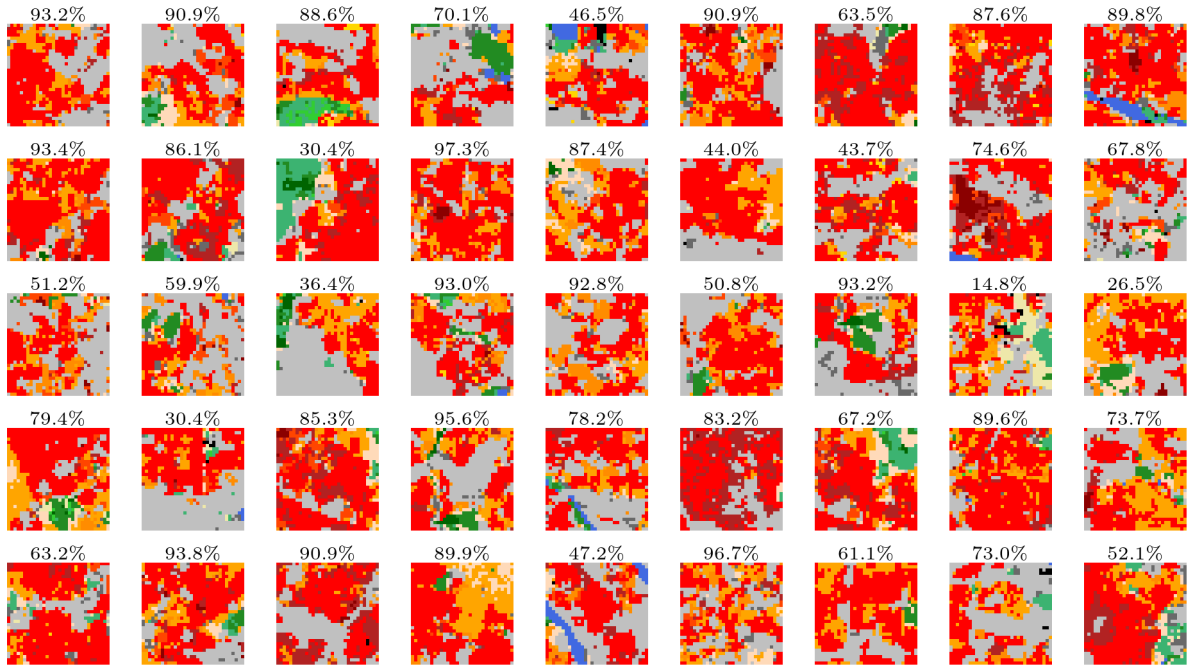
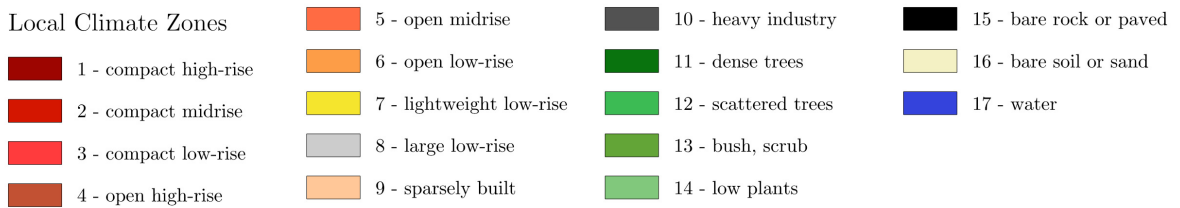


Figure 41: Exemplary cluster c_2 with 2618 patches in total. A random selection and their confidence values is shown.



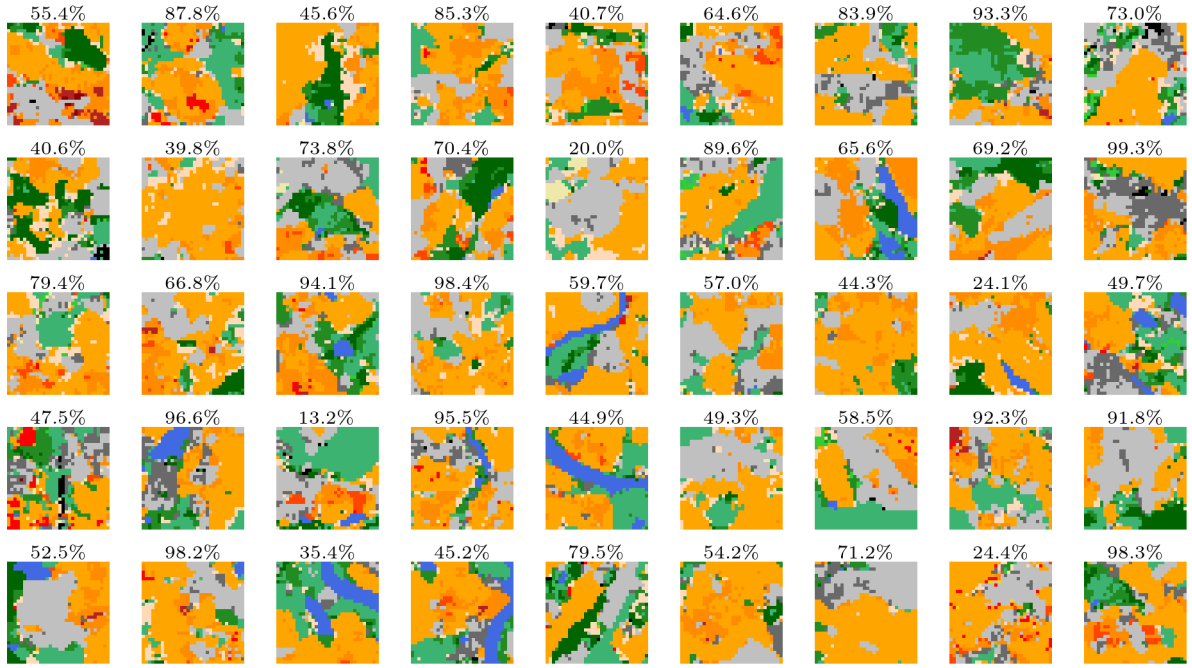


Figure 42: Exemplary cluster c_3 with 2476 patches in total. A random selection and their confidence values is shown.

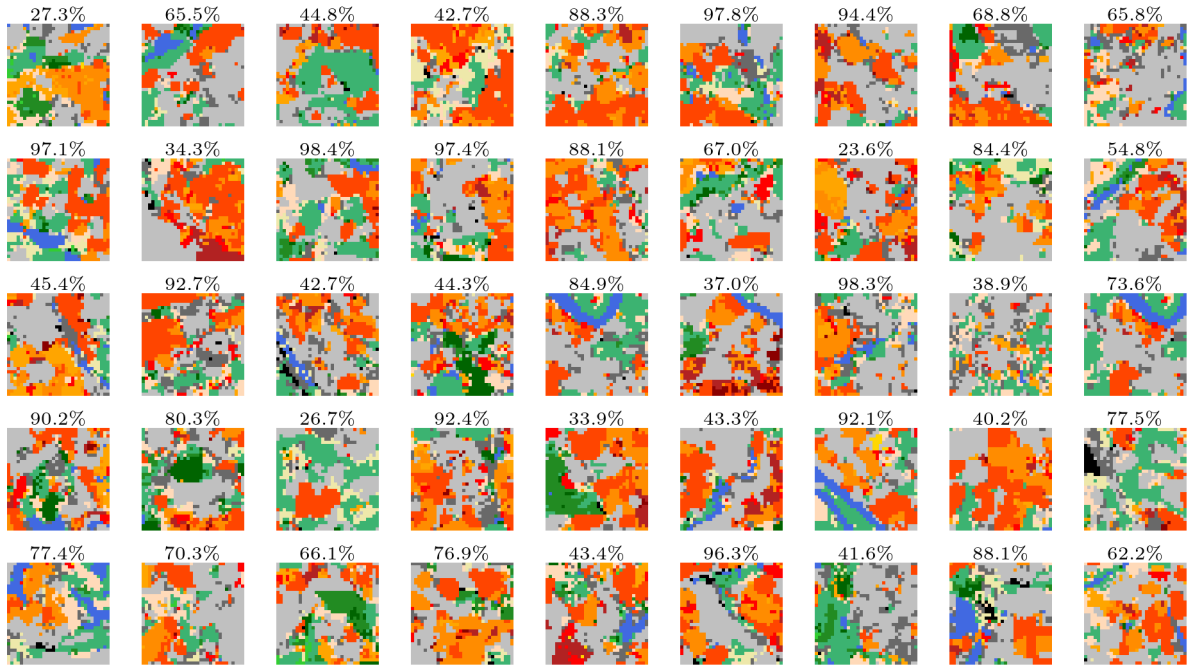
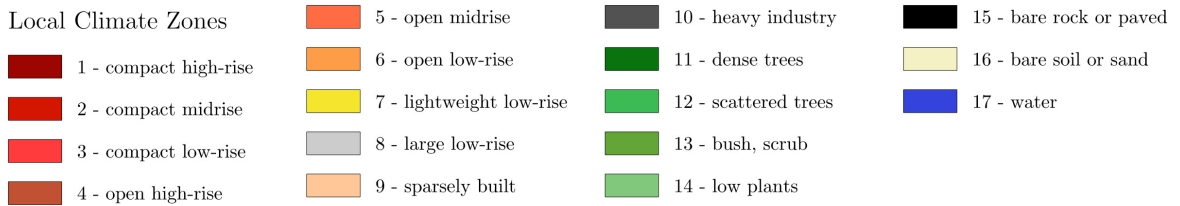


Figure 43: Exemplary cluster c_4 with 3305 patches in total. A random selection and their confidence values is shown.



More exemplary clusters are shown in the figures 58 - 65 in the appendix. The clusters were obtained without a supervised training process. This restricts the accuracy assessment to a qualitative visual comparison of the patches assigned to their clusters. It is impossible to validate whether the number of clusters or the patches assigned to it are correct. The patterns show continuous configurations of urban patterns, represented by a fuzzy class membership. Since there are no discrete classes of patterns defined in this work, the boundaries between the clusters are fluid. Therefore, the assessment of the results relies on a visual impression of similarity between the patches. The exemplary clusters in figures 40 - 65 show that patches assigned to one cluster are more similar to each other than to the patches assigned to a different cluster. The patches inside a cluster look consistent, even though there are of course some exceptions. In conclusion, the first research question can be answered with: *yes*, it is possible to find patterns of urban morphological configurations based on the LCZs across the world. A discussion of the results will follow in chapter 6. The work continues with the second research question.

5.2 Clustering Cities based on Urban Morphological Patterns

Do cities form geographical clusters when comparing them based on the found patterns?

To cluster cities, a metric to measure the similarity between cities was developed in this master thesis. The foundation of this measurement is the fuzzy class membership y for each patch x which describes the likelihood of belonging to the clusters $c \in C_D$ (see chapter 4.3.4 - 4.3.5). A pairwise distance measure $cm_{i,j}$ of cities was created by comparing their patches via Hungarian matching. Cities were thereon clustered with a k-means clustering algorithm. The optimal number of clusters was determined with the gap statistic algorithm, following the work of Taubenböck, Debray et al. (2020).

With this new explorative approach, the 1523 MUAs used in this master thesis were assigned to seven clusters $C_{HM_{1523}}$. An overview of the resulting seven clusters is shown in figure 44. Further, the cluster results are also provided in table 7. The cities assigned to the seven clusters form clusters of geographical regions. Even though not all geographical clusters have distinct and sharp spatial boundaries, their cities still form agglomeration in specific geographical locations. Since a visual recognition of 1523 individual cities assigned to seven clusters is tricky, the seven clusters are separately shown in the figures 45- 51. Each of the seven clusters is presented with a focus on their global geographical distribution. An example of cities assigned to the cluster is provided in the appendix with figures 66- 72.

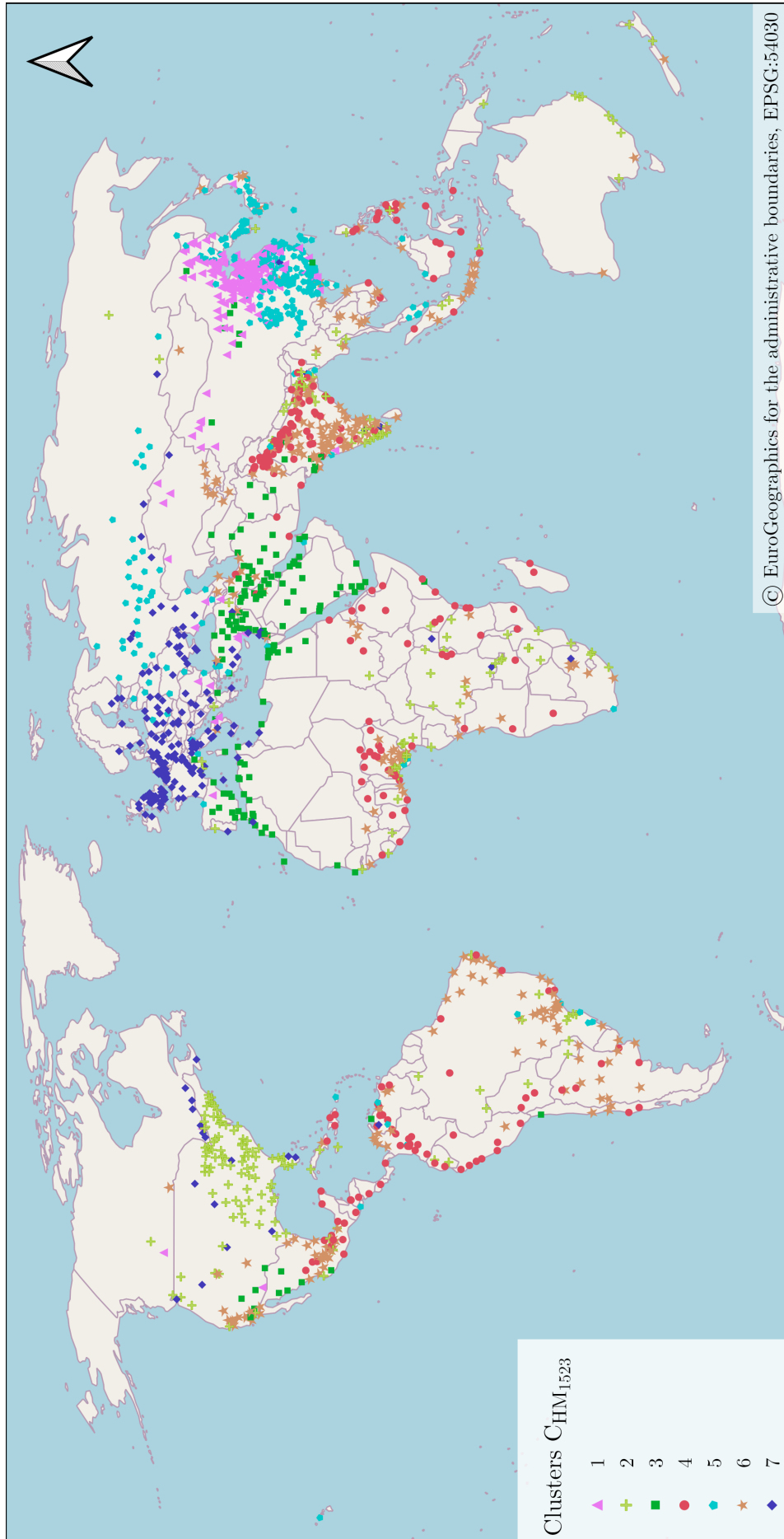


Figure 44: World map with an overview of the 1523 MUAs assigned to seven clusters $C_{HM_{1523}}$ by the k-means algorithm. In a first step, patterns of similar urban morphological configurations were searched within subsets of cities. Based on the found patterns, the similarity of cities was determined in a second step. The optimal number of clusters was found with the gap statistic method. The 1523 MUAs assigned to seven clusters $C_{HM_{1523}}$ form geographical regions.

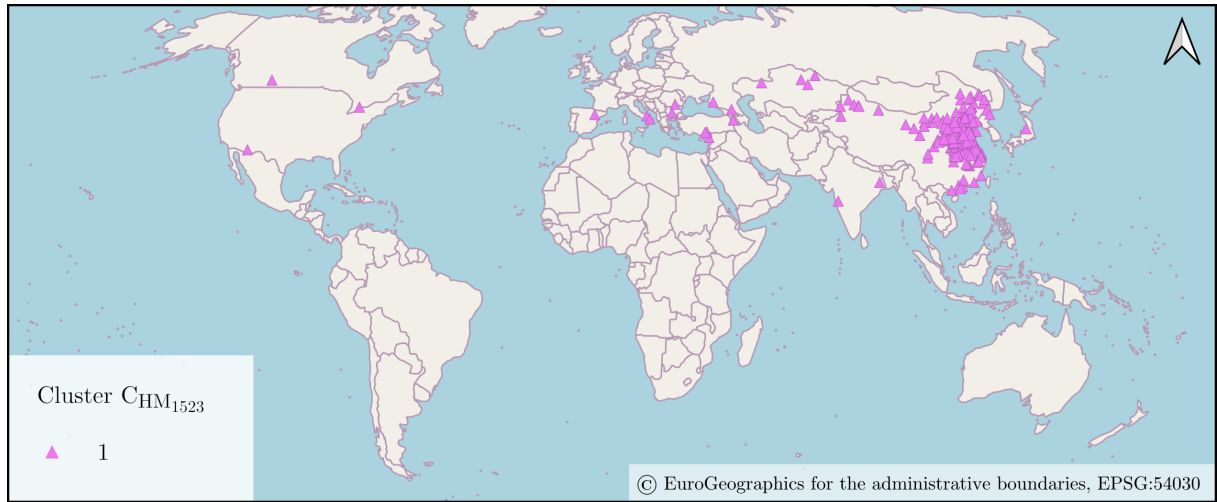


Figure 45: Cluster $C_{HM_{1523}}(1)$ with cities mostly based in Northeast China.

$C_{HM_{1523}}(1)$ - Northeast China

Figure 45 shows the cities assigned to the first cluster $C_{HM_{1523}}(1)$. A total of 212 cities belong to this cluster, of which 190 are located in China. This cluster is clearly dominated by eastern, northeastern and northern Chinese cities, around the metropolitan regions of Beijing, Shanghai and Zhengzhou. Some other Chinese cities are scattered near the southeastern coast of China and the northwestern region of Xinjiang in China. 18 cities of this clusters are found along Central Asia and Southern Europe. Three cities are located on the North American continent, two in India and one in Japan. Two exemplary cities assigned to this cluster are shown in figure 66 in the appendix.

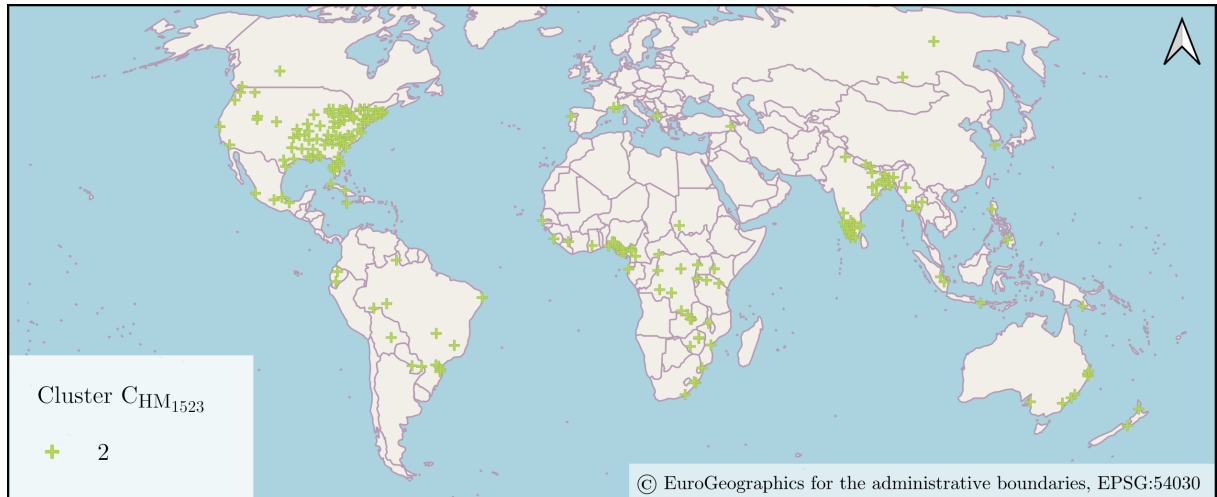


Figure 46: Cluster $C_{HM_{1523}}(2)$ with cities in the eastern United States, Nigeria, Southwest India and Bangladesh.

$C_{HM_{1523}}(2)$ - eastern United States of America, Nigeria, Southwest India, Bangladesh

The second cluster $C_{HM_{1523}}(2)$ with 215 cities is shown in figure 46. The cities assigned to this cluster are located more dispersed around the world, yet cities of this cluster form agglomerations

in specific geographic regions. One specific region within this cluster is the eastern half of the United States of America with 89 cities. Seven cities are located in Middle America and 15 spread across South America. Another small geographic region is formed by 16 cities located close to each other near the Nigerian coast. Other cities on the African continent are spread from the Central African region to the southeastern African region. In southern Asia, two distinct regions are found. 14 cities are located close to each other in the southwestern part of India. Another region with 12 cities is located in Northeast India and Bangladesh. Further cities are spread across Southeast Asia, the Australian West Coast and New Zealand. Four cities are found in Southern Europe, one in Turkey and two in Russia. Two exemplary cities assigned to this cluster are shown in figure 67 in the appendix.

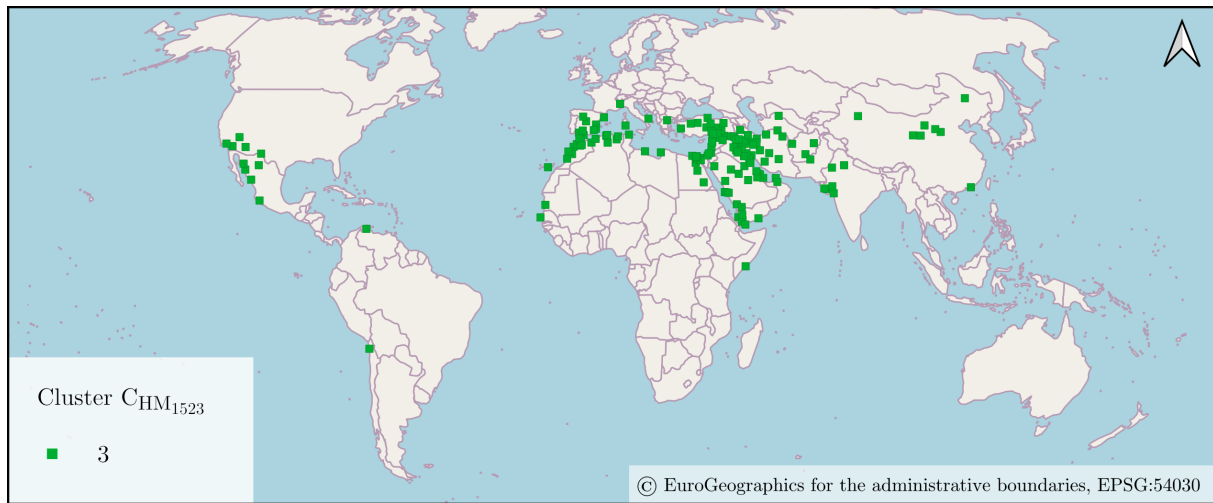


Figure 47: Cluster $C_{HM_{1523}}(3)$ with cities mostly in the Mediterranean region and Middle East.

$C_{HM_{1523}}(3)$ - Mediterranean region and Middle East

A distinctive geographic cluster is shown in figure 47 with $C_{HM_{1523}}(3)$. The cluster has 156 cities of which around 130 cities are located within the Mediterranean region, the Middle East and the Atlantic coast of Morocco. A smaller regional agglomeration with 10 cities is found in the region northwest of Mexico and southwest of the US. The remaining cities are mostly spread in China, India and South America. Two exemplary cities assigned to this cluster are shown in figure 68 in the appendix.

$C_{HM_{1523}}(4)$ - North India, Southeast Asia, East and West Africa, Middle America and northwestern South America

The 220 cities of cluster $C_{HM_{1523}}(4)$ are spread over multiple regions (see figure 48). The cities of this cluster form multiple agglomerations in specific parts of the world. One of the few dense agglomerations is the northern part of India, Pakistan and Bangladesh with around 60 cities. More cities are equally distributed over India. A less dense agglomeration is the Southeast Asian region including the Philippines, Indonesia and Malaysia. Also in the western and eastern region of Central Africa around 50 cities are assigned to this cluster. Further, 70 cities are located in

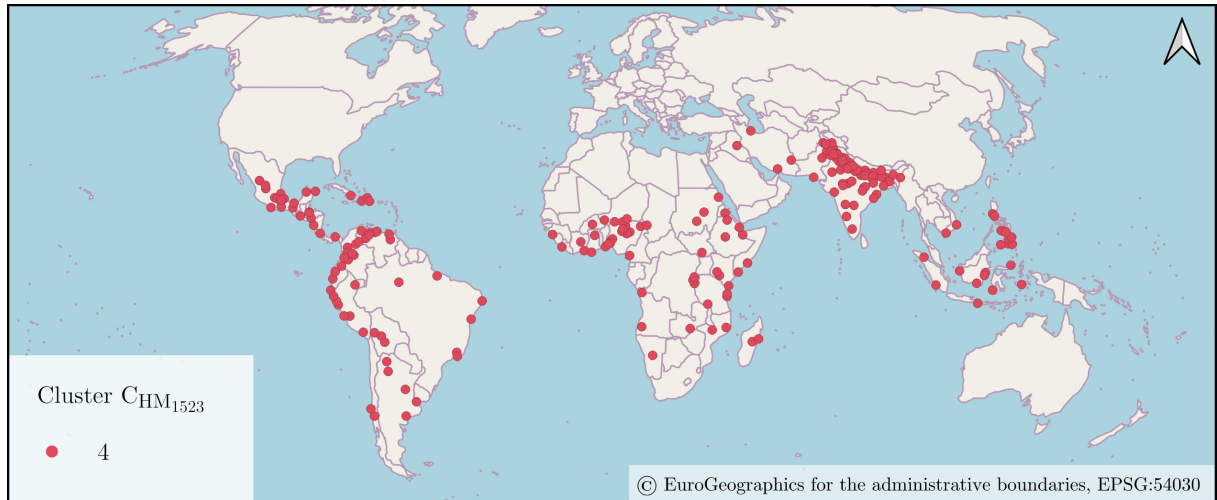


Figure 48: Cluster $CHM_{1523}(4)$ with cities in North India, Southeast Asia, East and West Africa, Middle America and northwestern South America.

Middle America and South America with a focus on the northwestern part of South America. Two exemplary cities assigned to this cluster are shown in figure 69 in the appendix.

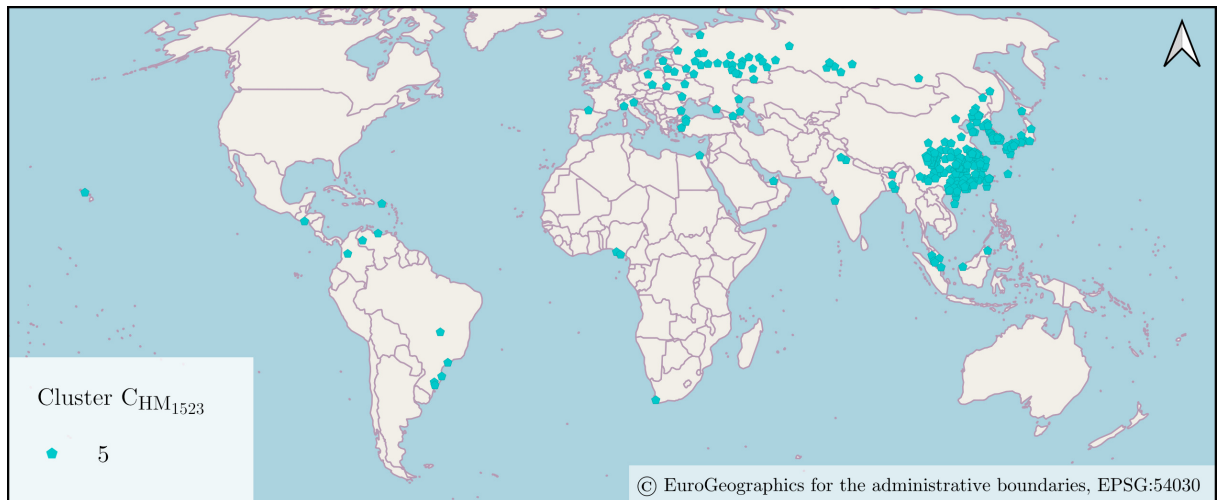


Figure 49: Cluster $CHM_{1523}(5)$ with cities in South China, Korea, Japan and the Eurasian region.

$CHM_{1523}(5)$ - South China, Korea, Japan and Eurasian region

$CHM_{1523}(5)$ shown in figure 49 has 263 cities assigned to it. The cluster has a very dominant agglomeration of cities in the southeastern region of China (151 cities), North and South Korea (16 cities) and Japan (12 cities). A second, more widespread agglomeration is the Eurasian region with 51 cities. The remaining cities are scattered around the world with five cities in the southwest of Brazil, five in Middle America, three in Africa and seven in Southeast Asia. Two exemplary cities assigned to this cluster are shown in figure 70 in the appendix.

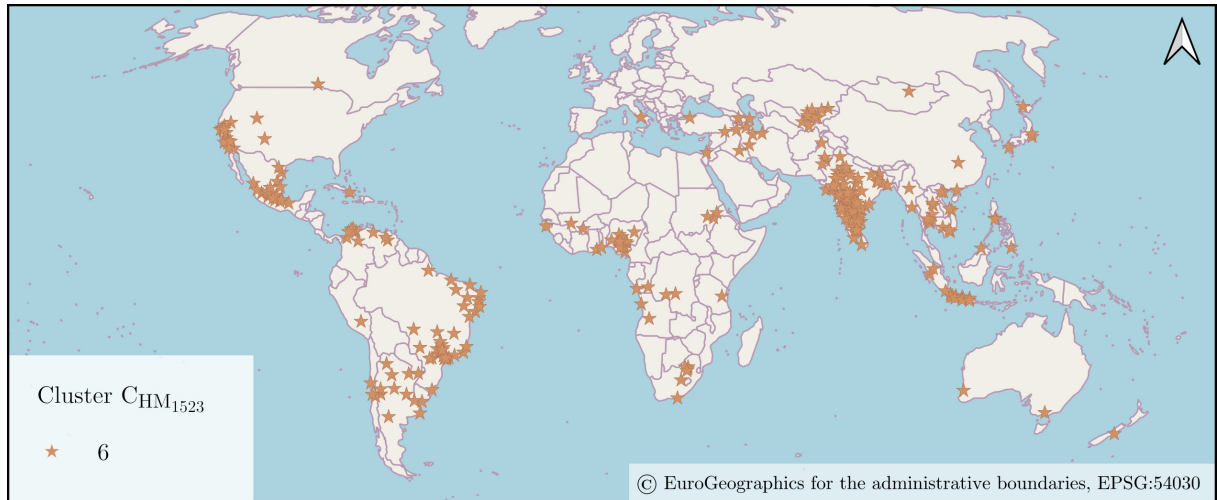


Figure 50: Cluster $C_{HM_{1523}}(6)$ with cities in India, California, Central Mexico, southeastern South America, West Africa and Southeast Asia.

$C_{HM_{1523}}(6)$ - India, California, Central Mexico, southeastern South America, West Africa, Southeast Asia

The cluster $C_{HM_{1523}}(6)$ with 263 cities is distributed more dispersed around the world. This cluster has many small and dense agglomerations in specific geographical regions. On the North American continent small agglomerations are in Central Mexico (19 cities) and California (14 cities). Continuing with South America, one agglomeration is in Colombia (nine cities) and two are in Brazil with 34 cities in an eastern and a southeastern cluster. 16 more cities are spread across the southern part of South America. Another dominant cluster are 64 cities spread across Central and South India. 33 more cities are spread across East Asia with a focus on the Indonesian island Java. Other regional clusters are found in Central Asia and the Middle East. 33 cities are spread across the African continent with a more dense agglomeration in Nigeria. Two exemplary cities assigned to this cluster are shown in figure 71 in the appendix.

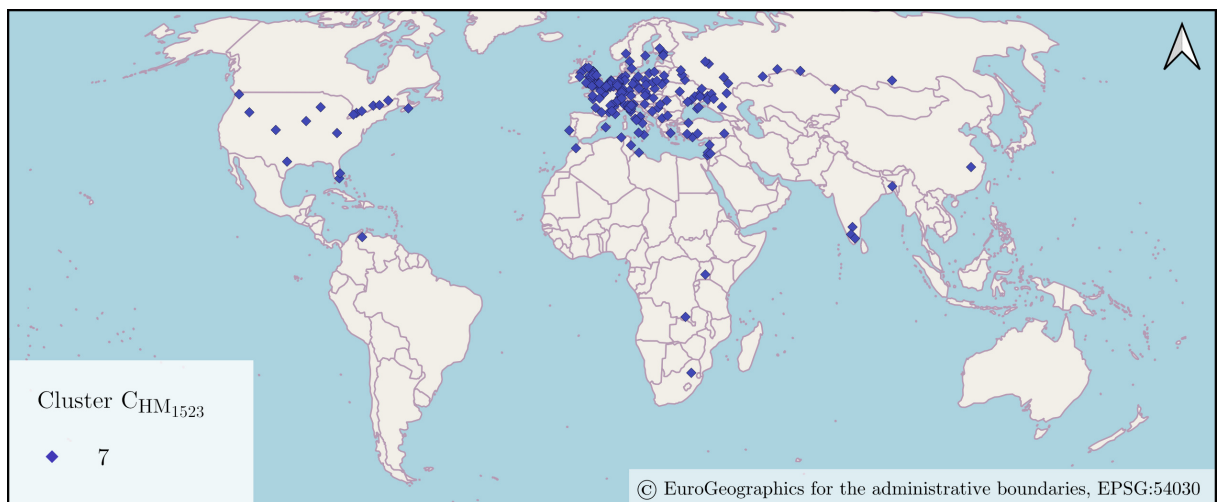


Figure 51: Cluster $C_{HM_{1523}}(7)$ with cities mainly located in the European, Mediterranean and West Russian region.

$C_{HM_{1523}}(7)$ - Europe, Mediterranean region and West Russia

The cluster $C_{HM_{1523}}(7)$ shown in figure 51 is strongly dominated by European and Mediterranean cities including the western region of Russia (165 out of 194 cities are within this area). Another 16 cities are spread across the US. The remaining few cities are scattered around the world. Two exemplary cities assigned to this cluster are shown in figure 72 in the appendix.

With the qualitative description of the seven clusters $C_{HM_{1523}}$ it is possible to give a definitive answer to the second research question: *Yes*, similar cities form geographical clusters when comparing them based on the found patterns of intra-urban morphological configurations. Some clusters form very distinct geographic regions, while others are a bit more ambiguous. Still, within the disperse clusters the cities form spatial agglomerations in specific geographical regions. The results will be discussed in chapter 6. To further examine the formation of geographic clusters, the work continues with the results obtained for the two ablation studies formulated within the research framework (see chapter 3).

5.3 Ablation Studies on Clustering Cities.

The first ablation aims to use the methodology developed in this work, but reduces the amount of cities to the 110 cities used by Taubenböck, Debray et al. (2020) (see chapter 4.3.5). The 110 cities were clustered with k-means and the optimal number of clusters determined with the gap statistic algorithm. In result, the 110 cities were assigned to seven clusters $C_{HM_{110}}$ shown in figure 52. Also the clusters obtained with the reduced dataset form partially distinctive geographical regions. The resulting seven clusters are shortly sketched:

$C_{HM_{110}}(1)$ - China and Eastern Europe (Eurasia)

The first cluster $C_{HM_{110}}(1)$ is mostly located in China and Eastern Europe.

$C_{HM_{110}}(2)$ - America, Africa, Asia

The second cluster $C_{HM_{110}}(2)$ is spread over North and South America, Central and South Africa, East Asia and Australia.

$C_{HM_{110}}(3)$ - Middle East

The third cluster $C_{HM_{110}}(3)$ includes cities mainly in the Middle East with three cities in the Indian/ Pakistani region and two in Africa.

$C_{HM_{110}}(4)$ - East Africa, Pakistan, Northwest India

The cluster $C_{HM_{110}}(4)$ is located mainly in East Africa and the region of Pakistan and Northwest India.

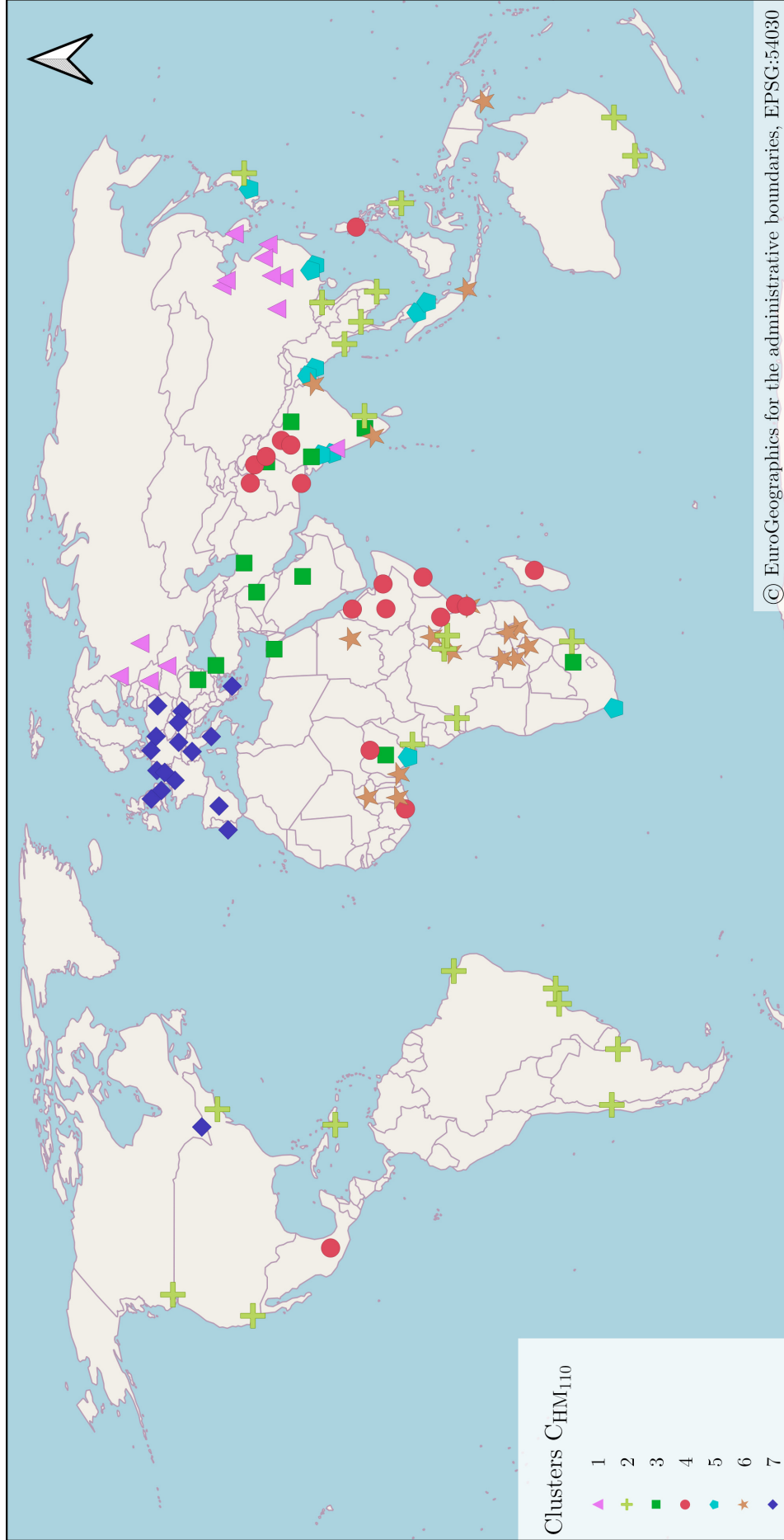


Figure 52: World map with an overview of the 110 MUAs assigned to seven clusters C_{HM110} by the k-means algorithm. It is the cluster result of the first ablation study. The similarity of cities was determined based on patterns of similar urban morphological configurations. The cluster result shows the formation of partially distinctive geographic regions.

$C_{HM_{110}}(5)$ - Asia

The cities assigned to Cluster $C_{HM_{110}}(5)$ spread more broadly in the Asian region. Two cities are located in Africa.

$C_{HM_{110}}(6)$ - Africa

The cluster $C_{HM_{110}}(6)$ contains cities located mainly in Africa. Two cities are found in India and two in Southeast Asia.

$C_{HM_{110}}(7)$ - Europe

The last cluster $C_{HM_{110}}(7)$ is a distinctive European cluster.

The second ablation aims to cluster the 1523 cities with the 18-dimensional feature space developed by Taubenböck, Debray et al. (2020). The feature space describes the statistical LCZ class distribution within the MUAs as well as the size of the MUAs. The cities are assigned with k-means to 29 clusters which were determined by the gap statistic algorithm (see chapter 4.3.6). Additionally, the 1523 cities were also assigned to only seven clusters $C_{F_{1523}}^{*7}$ to provide a better overview.

The 1523 cities assigned to the 29 individual clusters $C_{F_{1523}}$ are very difficult to distinguish (see figure 53). Thus, not all 29 clusters will be examined individually. Instead, this work confines itself to mention that similar cities form geographic regions when comparing them based on the feature space developed by Taubenböck, Debray et al. (2020). Yet, some mentionable exemplary clusters that can be distinguished on the map are $C_{F_{1523}}(19)$ in Great Britain, $C_{F_{1523}}(5)$ in the western part of the US or $C_{F_{1523}}(13)$ in Central Europe.

Figure 54 shows the 1523 cities assigned to the seven clusters $C_{F_{1523}}^{*7}$. The clusters form partially agglomerations in specific geographical regions, even though they are also spread widely around the world. Their geographical distributions are shortly sketched:

$C_{F_{1523}}^{*7}(1)$ - Eurasia with focus on Northeast and Southeast China

The first cluster has its cities located widespread in Eurasia with two dominant regions in China. One is located in the Northeast, the second is located in the southeast of China. Some other cities assigned to this cluster are found in Middle America.

$C_{F_{1523}}^{*7}(2)$ - Worldwide with focus on Nigeria, Northeast and South India, Southeast Asia

This cluster is very ambiguous. The cities assigned to it are widespread over the world, yet they form many local agglomerations in specific geographic regions. Geographical agglomerations are found in Nigeria, Pakistan, southern and northeastern India, as well as Southeast Asia. But also in Europe as well as on the North and South American continent, cities of this cluster are found.

$C_{F_{1523}}^{*7}(3)$ - Middle East

The cities assigned to this cluster form a distinctive area in the Middle East.

$C_{F_{1523}}^{*7}(4)$ - Worldwide with focus on India and Nigeria

The cluster is more widespread around the world, but has agglomerations mainly in India and Nigeria.

$C_{F_{1523}}^{*7}(5)$ - Eurasia with focus on Central East and Southeast China

Cities assigned to this cluster $C_{F_{1523}}^{*7}(5)$ are mainly located in Central East and Southeast China. Further, this cluster contains many cities in Europe and some located in Middle America.

$C_{F_{1523}}^{*7}(6)$ - America, Africa, Central and Southeast Asia

This cluster is widely spread across Middle and South America, the African continent as well as Central and Southeast Asia.

$C_{F_{1523}}^{*7}(7)$ - Europe

The cluster is dominated by two geographical regions, Europe and the eastern part of the US.

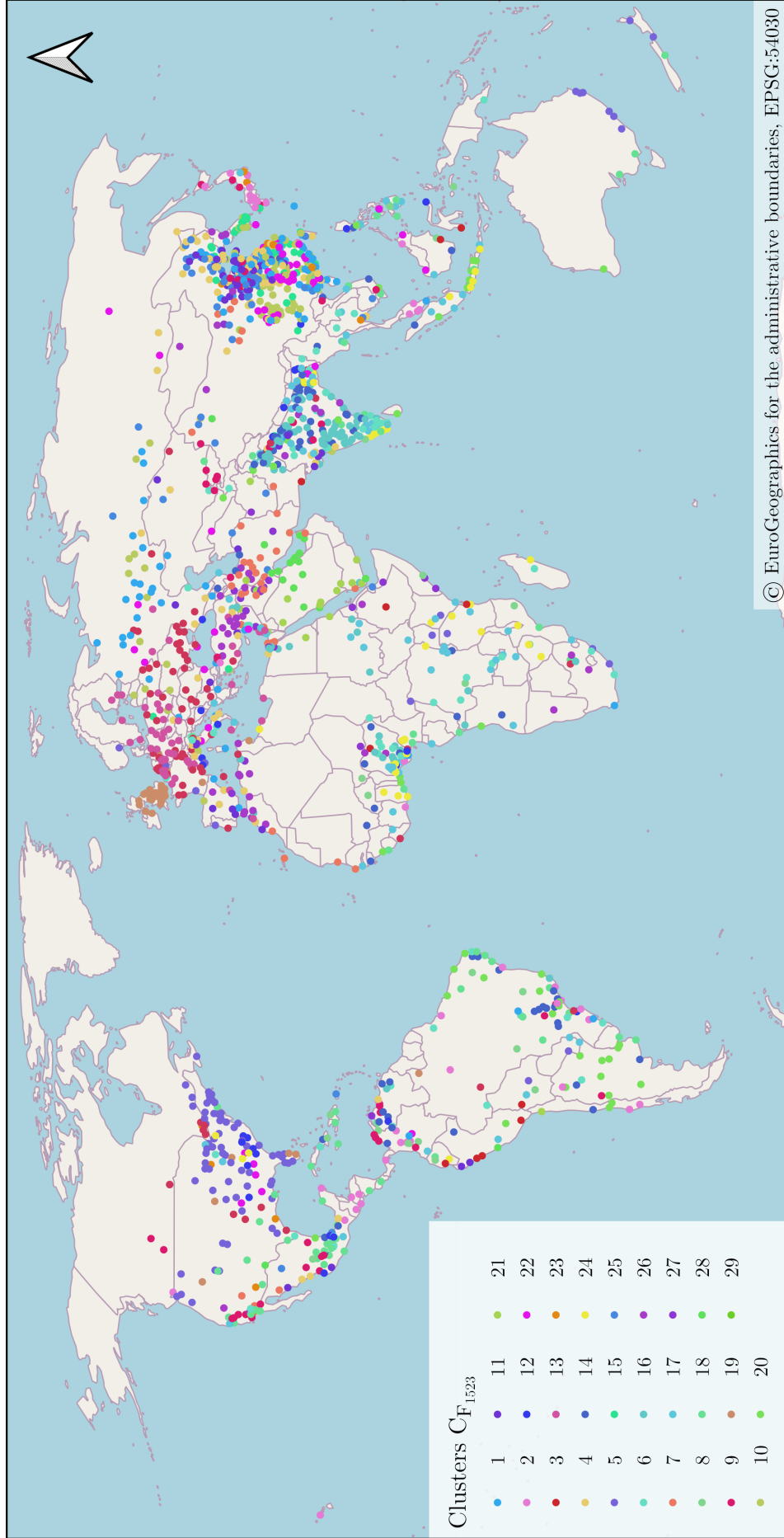


Figure 53: World map with an overview of the 1523 MUAs assigned to 29 clusters CF_{1523} by the k-means algorithm. It is the cluster result of the second ablation study. The cities were clustered on the 18-dimensional feature space. The high number of optimal clusters found impedes a profound qualitative description of the results. However, the formation of small clusters in partially distinctive geographic regions can be observed.

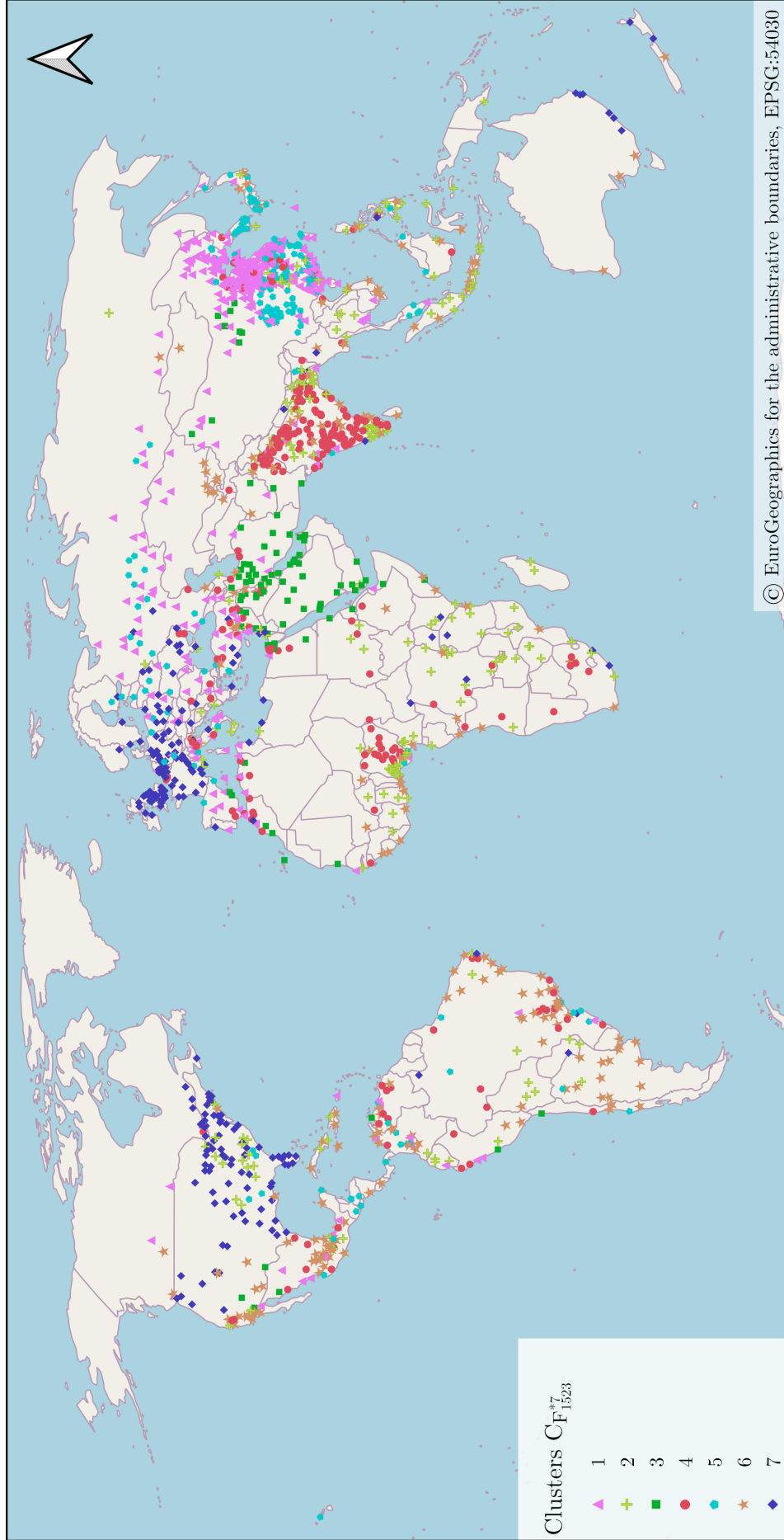


Figure 54: World map with an overview of the 1523 MUAs assigned to seven clusters $C_{F_{1523}}^{*7}$ by the k-means algorithm. The cities were clustered on the 18-dimensional feature space and thus belongs to the second ablation study. The number of seven clusters was chosen manually to allow a qualitative comparison of the results. The cluster result shows the formation of partially distinctive geographic regions.

6 Discussion

Each city is constituted by a unique composition of elements with their specific interrelations. The diversity of cities is reflected in a variety of functional, socio-economic and physical characteristics. Yet, similar elements and features are shared among multiple cities and form patterns that can be found across the world. The urban morphology can be regarded as a physical registration plate for all urban processes but also defines the human perception of and activity within a city. Multiple models were developed to examine the urban morphology and the formation of spatial patterns to uncover the rules that form and shape a city. However, they are often limited by a regional context, cultural bias and lack of comparable data (see chapter 2). Recent developments in remote sensing allowed the derivation of consistent datasets to describe the urban morphology and delineate city boundaries. Based on this empirical datasets, this work conducted in a first part an explorative search of intra-urban morphological configurations on a global scale with new unsupervised clustering approaches. The methodology and findings are discussed in the following chapter 6.1. In the second part of this work, the found patterns build the foundation to further group cities with similar urban morphological configurations through a comparative approach. The resulting city clusters and the methodology are discussed in chapter 6.2.

6.1 Understanding Patterns of Urban Morphological Configurations

The first research question aimed to find patterns of urban morphological configurations. In the context of this work, a pattern was defined as a re-occurring spatial arrangement of the urban morphological components. The urban morphology of a city is classified by the LCZs and the city boundaries delineated by the MUAs. The search of urban patterns was conducted on small subsets of LCZs data (*patches*) within the MUAs by adapting state of the art approaches of unsupervised clustering. In total, 138 clusters of urban morphological configurations were found (see chapter 5.1).

With a qualitative interpretation of the obtained results, the patches within the exemplary clusters (figures 40- 43 and figures 58- 65) can be considered as *semantically similar*, while the clusters themselves are *semantically different*. The patches within one cluster mostly share a similar statistical and spatial distribution of LCZs. In a patch, the individual pixels with the same LCZs classes aggregate to components on a larger scale. The size and shape of those components is also shared across the patches. This shows that the initial intention of clustering semantically similar patches and thus finding patterns was successful.

However, when visually comparing two patches of the same cluster, there can be major differences in some cases. That means, not all patches within a cluster have the same urban morphological configuration. This raises the following questions: Are some patches assigned to a *wrong* cluster and is there in consequence a *correct* cluster they should be assigned to instead? Or is the morphologic reality represented by the LCZs so diverse that the patches assigned to one cluster show inevitably differences? And is thus the assignment to the cluster still the best choice?

These questions are closely linked to the amount of generalization the clusters provide. With a very large number of individual clusters, one could expect each cluster to describe a more specific urban morphological configuration and in consequence include only a few patches that are very similar to each other. When reducing the overall amount of clusters, the patterns each cluster represents are introduced to a higher level of generalization and thus contain a larger amount of more diverse patches. It must be considered that the unsupervised clustering methods applied in this master thesis obtain for exemplary benchmark datasets accuracy performances of 90.3% on the CIFAR-10, 53.3% on CIFAR-20 and 86.7% on STL-10 (Park, Han et al. 2020). These accuracy measurements cannot be mapped to this work, yet they indicate that also these cluster results are likely to have miss-classified patches.

So what do these results suggest? Do the obtained clusters denote different *types* of urban morphological configurations? As stated in chapter 4.1.1, there is no prior knowledge on types of morphological configuration in the context of this master thesis. Subsequently, neither the obtained number of clusters nor the patches assigned to the clusters are quantitatively validated. Nevertheless, the results constitute a fundamental research baseline on urban morphological patterns and thus can build a foundation for a qualitative formalization and quantitative description of those. A similar motivated study of Fleischmann, Feliciotti, Romice et al. (2021) created a taxonomy of urban form on the level of individual buildings. Patterns are described with a hierarchical clustering based on morphometric characteristics. Could a similar idea be conducted with the results of this work? A hierarchical clustering of patches is possible but would require the derivation of a feature space which was omitted with these state of the art clustering approaches. Still, the obtained results may contain usable information to tackle these questions. The patches do not show distinguishable objects (such as the benchmark datasets) which would allow a clear distinction between the different clusters. Instead, the patches represent a continuous range of different LCZ classes and spatial arrangements. It may be possible, that some configurations are more frequently found among the cities than others. These configurations could denote a prototypical pattern *type*, while different *sub-types* describe patterns in between clusters which share configurations across multiple pattern *types*.

This leads to the next question: Does the confidence value of the patches not already denote such a system? Each patch is represented by a fuzzy membership function with a confidence value which describes the likelihood of belonging to each of the 138 found clusters. In theory, the patches assigned to a cluster with high confidence values could denote these prototypical patterns, while patches with low confidence values inevitably share configurations with other clusters. This also would mean that high confidence patches of one cluster have configurations similar to the pattern *type* represented by this cluster and are thus very similar to each other. Yet, in practice a first qualitative review of the patches and their confidence values could not unveil such regularities. All these open questions demand a more detailed analysis of the results, which goes beyond the scope of this master thesis. This might be a subject of interest for future studies.

Urban morphological patterns can be observed on multiple scales: for instance on the level of buildings, on the level of neighborhoods or whole cities. Searching patterns on the underlying LCZs data with a pixel size of $100\text{m} * 100\text{m}$ requires to regard the context of those basic elements and thus limits the scale. A patch size of $32 * 32$ pixels was chosen with an assumed balance between generalization and specification of patterns (see chapter 4.3.1). The findings of this study are thus limited only to this specific context and scale. However, the baseline developed in this master thesis can be adapted to other scales and datasets, to take either a closer look at more specific urban configurations or conduct a more generalized approach.

Throughout this master thesis, missing knowledge on the searched patterns had to be compensated by making assumptions on the characteristics of the data and on the adaption of the methodology. The results obtained in this work come along with a variety of uncertainties and limitations that need to be discussed. These include (i) the data basis, (ii) the data preparation, (iii) the pretext task, (iv) clustering with the SCAN framework and (v) the RUC framework.

(i) Data basis

The data basis consists of the externally provided datasets of LCZs and MUAs. Regarding the search of urban morphological patterns, especially the adequate representation of the morphological reality by the LCZ is of relevance. On one side, the representation of urban morphological characteristics through the LCZ classification scheme developed by Stewart and Oke (2012) comes with uncertainties. This is due to the level of generalization and the original development for climate studies. On the other side, the LCZ mapping itself comes with uncertainties. Qiu, Schmitt and Zhu (2019) give an overall accuracy of 86.7% for their LCZ mapping. Yet, the classification accuracy varies among different cities and geographical areas, but also among the specific LCZ classes different accuracies are achieved (cf. Qiu, Mou et al. 2019; cf. Qiu, Schmitt and Zhu 2019). This is partially due to the challenge of deriving three-dimensional knowledge (represented by the LCZ) from two-dimensional Sentinel-2 imagery. Other parameters are the training strategy, the number of selected samples in different cities and the aggregation of urban structures to spatial units of $100\text{m} * 100\text{m}$.

(ii) Data preparation

In the process of data preparation, multiple overlapping LCZ tiles are merged. The overlapping parts do not have an identical LCZ mapping, instead tiles that cover the same area have slight variations. This is resolved by keeping the mapping of the larger tile during the merging process (see chapter 4.3.1). In reality, it is not known if the LCZ mapping of one tile is preferable over the mapping of another tile. Further, some LCZ were re-projected in order to merge them together. To avoid inconsistencies with the categorical LCZs data along the process of merging and re-projecting, a nearest neighbor sampling was chosen. Still, minor divergences from the original LCZs data can arise during this process.

(iii) Pretext task with the SimCLR framework

A neural network Φ_{Θ} was trained to learn semantically meaningful representations by solving the pretext task τ with the SimCLR framework (see chapter 4.3.2). The framework was developed by Chen, Kornblith et al. (2020) to support object classification of images with multiple color channels.

Since the underlying LCZs of this work describe thematic categories and not color channels, this work assumes that the data augmentation parameters used in the SimCLR framework need to be adapted to the categorical LCZ pixel values (see chapter 4.1.3). The color distortions were removed and the geometric distortions enhanced by a *horizontal flip*. Chen, Kornblith et al. (2020) empirically show that the composition of augmentations, especially the combination of color and geometric distortions, is key for representation learning. They further show that images can be identified solely by their histogram of pixel intensities. A network might exploit these characteristics during the training process and thus focus on statistical features instead of learning generalized representations. This cannot be precluded with the use of LCZs data in this master thesis. Even though color distortions are completely excluded in this work, small color augmentations in a limited context could be helpful. Exploring their impact might be of interest for future studies.

Other parameters to consider are the batch size and depth of the network. Chen, Kornblith et al. (2020) achieve better results with larger batch sizes and deeper networks. In the context of this master thesis, both were limited due to the available hardware components. Even better results in this work may be achieved by leveraging both parameters. However, using the trained network Φ_{Θ} to mine k-nearest neighbors for the following task worked well and thus proved the functional principle of this concept (see figure 27).

(iv) Clustering with the SCAN framework

Also the implementation of the SCAN framework comes with several restrictions. Beginning with the augmentations, again all color distortions are discarded and the geometric distortions are partially adapted (see table 2 and 3). This concerns also the *cutout* parameter for which this work assumes that multiple smaller cutouts are better compared to one large cutout (1/4 of the original image). The overall urban configurations of the whole patch are considered to be important. These are assumed to be better maintained with smaller cutouts. One large cutout would force the network to focus on the remaining patch which could denote a different urban configuration without the cutout part. Yet, the specific effects on the network performance are also unknown here.

The SCAN framework expects that the number of clusters and the distribution of samples among the clusters are known in advance (see chapter 4.3.2). If the distribution is not known, an even distribution of the samples among given clusters is enforced by the second term of the loss function

L_{scan} (see equation 5). Gansbeke et al. (2020) achieve good results on benchmark datasets with this 'trick', yet they limit the approach for real-world applications where neither the number of classes nor the distribution of samples among the classes is known. This also poses a challenge for the search of urban morphological configurations. This work proposes a work-around by over-clustering the patches with a high number of clusters. For this task, the arbitrary number of 1024 classes was chosen. Even though the samples are automatically re-assigned during the self-labeling step to a smaller number of clusters, the effect of this parameter on the final results of this work is not known. Choosing a different number will likely lead to different results.

Further uncertainties arise from this adaption. For the clustering-step, the largest applicable batch size with the given hardware is 384. Having two similar patches within a mini-batch, the second part of the loss function L_{scan} inevitably punishes the assignment to the same cluster (since all 384 samples should be assigned evenly to the possible 1024 classes). Developing a better suited loss function for this task, which clusters similar patches on one side but still forces a dispersion over a given number of classes on the other side, was out of the scope of this master thesis. The uncertainty resulting from this dilemma could be reduced by leveraging the batch size to a number much larger than the given 1024 classes.

The chosen parameters also affect the self-labeling step of the SCAN framework. Due to the over-clustering, the patch confidence values turned out to be very low (see figure 29). Hence, the threshold for sampling high confidence patches proposed by Gansbeke et al. (2020) could not be used. Instead, a new sampling strategy was developed with a class dependent threshold. This certainly impacted the choice of prototypical patterns, the training process and consequently the resulting clusters.

(v) Refining results with the RUC framework

The intention of applying the additional RUC framework was to correct the overconfident results obtained with SCAN (see figure 29). The RUC framework improves the classification result on benchmark datasets by a few percental points (cf. Park, Han et al. 2020). It is uncertain how much the RUC framework will improve the accuracy of the results obtained with SCAN. Small changes will be difficult to notice with a visual comparison of the results. Yet, the framework succeeds in re-balancing the confidence values of the overconfident results obtained with SCAN (compare figure 29 with figure 33). This is assumed to be important for an adequate description of the morphologic reality and comparison of patches.

Despite all the uncertainties that come along the methodology, 138 clusters were obtained with a fully unsupervised approach. As discussed earlier, their validity cannot be verified and results may change with the use of different parameters. However, the cluster results prove the functional principle of the developed concept.

6.2 Formation of Geographic Clusters with similar Cities

The second research question aimed to group similar cities based on their morphological patterns and examine the formation of geographical clusters. Cities are delineated by the MUAs and their patches characterize urban morphological patterns with a fuzzy membership function. The fuzzy membership function of the patches was obtained with the previous results by clustering the patches (see chapter 4.3.1 - 4.3.3). In this work, a new similarity metric for cities was developed with a pairwise patch comparison (see chapter 4.3.4). Similar cities were grouped with a fully unsupervised baseline which was adopted from Taubenböck, Debray et al. (2020). The 1523 cities used in this work were assigned to seven clusters $C_{HM_{1523}}$ via the k-means algorithm (see chapter 4.3.5). Beginning with the main results, the spatial distribution of the clusters is discussed (1). Second, the ablation studies are set in context with the main results and the work of Taubenböck, Debray et al. (2020) (2). Lastly, uncertainties arising from the methodology are discussed (3).

1) Spatial distribution of the seven clusters $C_{HM_{1523}}$

The cities of each of the seven clusters are not dispersed randomly around the world. They form agglomerations in specific geographic regions (see figure 44). Some clusters are concentrated on one or two specific regions, such as $C_{HM_{1523}}(1)$ (China), $C_{HM_{1523}}(3)$ (Mediterranean and Middle East), $C_{HM_{1523}}(5)$ (Asia and Eurasia), $C_{HM_{1523}}(7)$ (Europe). Other clusters are spread over multiple continents but the cities assigned to them are still found in distinctive geographic regions on the world map, such as $C_{HM_{1523}}(2)$ (eastern United States, Nigeria, India) and $C_{HM_{1523}}(4)$ (Middle and South America, Africa, India) and $C_{HM_{1523}}(6)$ (India, Middle and South America). These results are interesting as they come along with many fundamental questions on urban morphology. Why do these spatial distributions appear within the clusters? May it be possible that these clusters denote a certain city *type*?

Different city *types* and models have been described and developed in the past (see chapter 2). These can be based on functional, socio-economical and morphological features. Yet, they are often limited by a regional context or by a cultural bias. The lack of comparable data impeded empirical approaches for a long time. Defining *types* of cities was often conducted with a cultural-genetic approach including the concept of *cultural areas*. Can the spatial patterns found in this master thesis be explained with this aspect? Does for example the *European cluster* $C_{HM_{1523}}(7)$ hints towards the theoretical concept of an *European city* as proposed by Hofmeister (1996)? It cannot be denied that cities which are spatially close to each other are often more similar compared to cities further away. But arguing with the concept of *cultural areas* is deliberately omitted in this work. It is considered that homogeneous cultural areas as proposed in the literature may not exist and thus are controversially discussed (see chapter 2.3). Instead, this work assumes that a geographical perspective on the spatial distribution of city clusters is more expedient.

To understand why the spatial distribution of clusters is linked to geographical regions, it is necessary to understand how these clusters form. Cities are clustered based on their intra-urban morphological configurations. In consequence, each cluster describes a set of cities with similar urban configurations. Comparing cities means automatically comparing a set of urban morphological patterns. The results of this master thesis suggest that the formation of urban morphological patterns are not random. Instead, similar patterns are likely to be found in cities that are spatially close to each other. So what does this mean in consequence? Urban morphological patterns emerge under the impact of factors linked to the geographic location. To understand the spatial distribution of city clusters requires thus finding and exploring these factors. This could be a promising tasks of future studies in urban morphology.

Even though the influencing factors have yet to be determined, seeing the clusters as a set of cities with similar urban configurations helps to understand the geographical distribution. The diverse distribution leads inevitably to the question why some clusters are restricted to geographic regions while others are dispersed more around the world. Having clusters of cities in distinctive regions such as *Europe* ($C_{HM_{1523}}(7)$, figure 51) or the *Middle East* ($C_{HM_{1523}}(3)$, figure 47) would correspond to a subjective impression of similarity many people have. Contrary to these clusters, what do the widely spread clusters such as $C_{HM_{1523}}(2)$ (see figure 46) suggest? Is a North American city similar to cities in Nigeria and Southwest India? Is a Californian city as found in cluster $C_{HM_{1523}}(6)$ (figure 50) characterized by urban morphological configurations that are likely to be found in Indian and South American cities but not in European cities? This is exactly what these results show. However, foremost it must be considered that these clusters denote not a similarity in terms of economic power, functions and cultural heritage. It is a similarity on the more abstract level of intra-urban morphological configurations described by the LCZs. Understanding how the formation and spatial arrangement of these LCZs, which denote urban morphological characteristics, are linked to all these diverse factors could result in a more profound understanding of cities.

The spatial distribution of the seven clusters $C_{HM_{1523}}$ obtained in this work may also be partially explained by the applied methodology. Especially the dispersed clusters $C_{HM_{1523}}(2)$, $C_{HM_{1523}}(4)$ and $C_{HM_{1523}}(6)$ indicate that more distinctive geographic regions might appear with the use of more clusters. This is underlined by the fact that even though these clusters are spread around the world they still form a number of small spatial agglomerations in distinct geographic regions. The number of seven clusters was determined via the gap statistic method. Multiple runs with different seed keys came to different cluster results (see figure 37). This shows that there is no definitive correct amount of clusters and the results are highly depend on the random initialization within the gap statistic method. Yet, the different gap statistic results are mainly grouped around the optimal seven cluster in a range of around two to twelve clusters. This means that even though the number of clusters is ambiguous their representation by $k_{final}^* = 7$ optimal clusters seems valid. Further, the number of seven clusters leaves enough room to examine specific clusters, while maintaining a high level of generalization. To examine the formation of geographic clusters in more detail, a higher number of clusters could be considered.

So do these seven clusters $C_{HM_{1523}}$ denote seven city *types*? An answer for this question would require a more detailed analysis of the characteristics of the found clusters. Yet, this work assumes that the boundary between the seven clusters is continuous since cities can be composed by a great variety of urban morphological patterns. However, it can be concluded that the cities among the seven clusters are likely to show similar patterns of urban morphological configurations.

2) Ablation studies

(i) Clustering 110 cities based on urban morphological patterns

The first ablation study is based on the distance calculation performed in the main study of this work, yet only 110 cities are used for clustering (see chapter 4.3.5). These 110 cities are the cities used in the study by Taubenböck, Debray et al. (2020) to allow a comparative approach. In result, seven clusters $C_{HM_{110}}$ are obtained (see figure 52). A comparison with the main findings of this work will unveil similarities and differences but do not contribute to thematic insights. If the same feature space is used but the number of samples to cluster is reduced the clusters found with k-means will consequently change. More interesting is a qualitative comparison to the work of Taubenböck, Debray et al. (2020). In their study they derive an 18-dimensional feature space describing the percental share of the 17 LCZ classes with the MUAs. The 18-th dimension describes the size of the MUAs. Based on this feature space, they also find seven clusters, denoted here as $C_{T_{110}}$.

Both cluster results $C_{T_{110}}$ and $C_{HM_{110}}$ are based on completely different feature spaces but their baseline for clustering is the same. Both find seven optimal clusters which simplifies a qualitative comparison. Surprisingly, the spatial distribution of both cluster results shows many similarities (see Taubenböck, Debray et al. 2020, p. 7). The first cluster $C_{T_{110}}(1)$ is described as mainly Asian and African and corresponds to $C_{HM_{110}}(4)$ which is also a mainly (East) African and Asian (Pakistani, Northwest Indian) cluster. The cluster $C_{T_{110}}(2)$ is described as mainly Asian and American and corresponds to $C_{HM_{110}}(2)$ which is an American, African and Asian cluster. Both are very similar with the exception that $C_{HM_{110}}(2)$ has cities included from the African continent, while $C_{T_{110}}(2)$ does not have any African cities. The third cluster $C_{T_{110}}(3)$ described as European corresponds to $C_{HM_{110}}(7)$ also including mainly European cities. Differences arise in the eastern part of Europe which is not included in $C_{HM_{110}}(7)$. Cluster $C_{T_{110}}(4)$ is described as eastern Asian and eastern African and may correspond to $C_{HM_{110}}(6)$ which has only African cities. The cluster $C_{T_{110}}(5)$ with Central African cities is not found in $C_{HM_{110}}$. Therefore $C_{HM_{110}}(1)$ describes a Chinese and Eastern European cluster, which is not found in $C_{T_{110}}$. $C_{T_{110}}(6)$ corresponds to $C_{HM_{110}}(3)$, both with cities in the Middle East. $C_{T_{110}}(7)$ is denoted to contain large cities but may correspond to $C_{T_{110}}(5)$ as a mainly Asian cluster.

The cluster results $C_{T_{110}}$ from Taubenböck, Debray et al. (2020) and $C_{HM_{110}}$ from this work are very similar. Six out of the seven clusters show a high overlapping. Differences occur due

to the different methodological approaches. Yet, the main findings of Taubenböck, Debray et al. (2020) can be partially confirmed with the new methodological approach developed in this work.

(ii) Clustering 1523 cities based on a 18-dimensional feature space

The second ablation study derives a 18-dimensional feature space for the 1523 cities used in this study with the methodology developed by Taubenböck, Debray et al. (2020) (see chapter 4.3.6). With the same baseline for clustering, an optimal number of 29 clusters $C_{F_{1523}}$ was found (see figure 53). To provide a generalized overview, the results were additionally clustered with only seven clusters $C_{F_{1523}}^{*7}$ (see figure 54).

The 29 clusters found with this ablation study were determined with the gap statistic method (see figure 39). The figure shows that multiple runs with different seed keys came to different cluster results. The gap statistic results range from 2 to 50 clusters and are more ambiguous than the gap statistic clusters obtained in the previous results (see for comparison figure 37 and figure 38). The gap statistic algorithm was limited to the range of $[1, 50]$ clusters. Using more clusters may have led to an even higher number of optimal clusters. In conclusion, the number of 29 optimal clusters $C_{F_{1523}}$ is very uncertain. However, the high number of clusters allows to examine the spatial distribution on a more detailed level. The clusters form spatial agglomerations, some in very distinctive regions, some more widely dispersed around the world. An extensive qualitative analysis of the 29 clusters would go beyond the scope of this master thesis. Further, a qualitative comparison with the previous results is impeded by the different number of clusters. Therefore, this work will analyze the 18-dimensional feature space clustered with seven clusters $C_{F_{1523}}^{*7}$ in comparison to the main findings of this work $C_{HM_{1523}}$.

The seven clusters $C_{F_{1523}}^{*7}$ show partially very clear formations of geographical agglomerations. Yet, they cannot be mapped one to one to the cluster results of $C_{HM_{1523}}$. The first cluster $C_{F_{1523}}^{*7}(1)$ is a mainly Chinese cluster, but also spread across Eurasia with some cities in Middle America. Cluster $C_{HM_{1523}}(1)$ is instead focused only on China. The second cluster $C_{F_{1523}}^{*7}(2)$ is very similar to $C_{HM_{1523}}(2)$ but has fewer cities in the eastern part of the US. The third cluster $C_{F_{1523}}^{*7}(3)$ is also very similar to $C_{HM_{1523}}(3)$, yet cities of the Mediterranean region are missing. Compared to $C_{HM_{1523}}(4)$, the spatial distribution of the fourth cluster $C_{F_{1523}}^{*7}(4)$ is on one side more dispersed on the world map but has on the other side a strong focus on the whole Indian country. Cluster $C_{F_{1523}}^{*7}(5)$ and $C_{HM_{1523}}(5)$ are both very similar and have only a few minor differences. Also $C_{F_{1523}}^{*7}(6)$ and $C_{HM_{1523}}(6)$ are similar, yet $C_{F_{1523}}^{*7}(5)$ is missing Indian cities which were assigned to $C_{F_{1523}}^{*7}(4)$ instead. The last cluster $C_{F_{1523}}^{*7}(7)$ shows the same European cluster as $C_{HM_{1523}}(7)$ with additional cities in the eastern part of the US. These are mainly the cities that were missing in $C_{F_{1523}}^{*7}(2)$.

In conclusion, both methods find very similar clusters. Some differences emerge from specific geographic regions shifting from one cluster to another cluster. The results which were obtained with the methodology developed by Taubenböck, Debray et al. (2020) applied to a larger dataset

can be confirmed with the findings obtained with the methodology developed in this work. Not only do similar cities form geographic clusters, but also the geographic clusters are mostly congruent. Why do the obtained results show similar findings despite fundamentally different methodological approaches? Both approaches use the same baseline for clustering, but have a different feature space. To tackle this question, the feature space on which the clustering of cities is performed needs to be examined.

F_{1523} is a feature space composed of the statistical distribution of the LCZ classes. Further, it includes the MUA size as an additional feature. In F_{1523} , the class LCZ-17 (*water*) is partially excluded from the original MUAs and thus less present compared to the patches derived within the buffered MUAs. The feature space HM_{1523} describes the similarity of cities based on their patches assigned to urban morphological *patterns*. The similarity of cities is embedded in a four-dimensional space via MDS. Similar patches have of course also a similar statistical distribution of the LCZ classes, yet their spatial arrangement plays a key role. However, it is unclear how the patch clustering is affected by the deprecation of the color augmentations (see chapter 6.1). It is not ensured that the networks used in this work do not exploit the statistical features and thus bias the representation learning and clustering of the patches (see chapter 6.1). Thus, both feature spaces may come to similar results since both include characteristics of the statistical distribution of the LCZs. A further possibility is that the differences in both feature spaces do not play an important role in clustering and thus similar clusters emerge. This includes the size of a city as well as the class LCZ-17 (*water*).

3) Uncertainties and limitations arising from the methodology

The development of a similarity metric for cities and the clustering of cities come with limitations and uncertainties which will be discussed in the following part. These include (i) the data basis, followed by (ii) the data preparation, (iii) distance calculation and (iv) the clustering of cities.

(i) Data basis

The data basis for clustering cities comes along with all the uncertainties mentioned in chapter 6.1. Still, further aspects play a role in the clustering of cities and thus need to be mentioned. The clustering of cities is closely linked to the clustering of patches and thus the underlying LCZ data. When finding similar cities in distinctive geographic regions, the similarity of those cities must be consequently reflected in the LCZs. However, the mapping of the LCZ itself may be affected by a hidden geographical bias. This begins with the selection of training samples limited to specific cities, continues with the interpretation based on expert knowledge and accumulates to different accuracy values for different regions in the world.

The MUAs are relevant for a comparative approach of cities, since they delineate their boundaries. They provide a consistent data driven approach and thus a comparable framework for cities. However, the underlying sectoral monocentric city model does not always reflect the morphologic

reality of cities (see chapter 2). This affects which urban morphological configurations are present within the MUAs. Further, MUAs are not fully covered by the LCZ classification (see chapter 4.3.1 and table 7). Hence, the urban morphological configurations of urban agglomerations cannot be fully represented by the available LCZs. This also limits the use of the MUAs size as an additional feature in F_{1523} . In consequence, the benefits of a comparable framework with the MUAs cannot be fully exploited in this work.

(ii) Data Preparation

Further limitations come along with the data preparation. Miss-classification of the LCZs was tackled with an edge-removal of 3% on each side of the LCZ raster tiles (see chapter 4.3.1. This reduced the available coverage of the LCZs data. The definition of patches with a squared shape and size of $32 * 32$ pixels limits the comparison of cities based on these patches to this spatial context. Additionally, the patches have to be within the MUAs. To relax the spatial limitations, include more LCZ data and enhance the available information to the surrounding context, the MUA were increased with a 1000m buffer. This buffer is independent on the MUAs size and thus has a relatively higher impact on small MUAs compared to large MUAs.

(iii) Distance calculation

The distance calculation developed in this work is based on a patch comparison of cities via Hungarian matching (see chapter 4.3.4). Each patch is described with a fuzzy membership functions which is dependent on the confidence balancing among the classes and underlies uncertainties as discussed in chapter 6.1. However, the confidence values among the patches seem to represent adequately the morphological diversity of patches (see figure 33) and the patch clusters look consistent (see chapter 5.1).

The similarity measurement of cities by a distance calculation with Hungarian matching is limited in several ways. The Hungarian algorithm tries to assign each patch of a city to a patch of another city by minimizing the overall Euclidean distance (*costs*) between the patches. Thus, the costs describe how similar a set of patches of one city is to a set of patches of another city. The costs are normalized by dividing the overall costs by the number of matched patches. Yet, the Hungarian algorithm performs solely a balanced assignment. If one city has more patches compared to another city, the redundant patches are not considered in the distance calculation. Differences in the city sizes introduce a bias towards a higher similarity. This becomes apparent with the following example: comparing a city v_a with one patch to a city v_b with 100 patches will assign the patch of v_a to the best matching patch of v_b and deprecate the other 99 patches. Having a high number of patches in v_b will increase the likelihood of having a patch which is similar to the patch in v_a . In consequence, the cities v_a and v_b will be considered as similar, even though the remaining 99 patches might show fundamentally different urban morphological configurations. Future studies can avoid this by including the distances of all patches within two sets of cities and normalizing the costs by the number of patches of both cities.

The Hungarian algorithm does not consider the spatial arrangement of patches within cities. The spatial relation of the LCZs is only considered within the patches and thus limited to this context. Considering the spatial relation of urban morphological patterns could provide interesting insights and lead to a better understanding of urban morphological configurations in cities.

(iv) Clustering of cities

To cluster cities based on the similarity metric developed within this work, the distances were transformed to a four-dimensional coordinate space via MDS. Embedding the distances to a fixed number of dimensions results in a loss of the correct distance representation, shown with the stress value in figures 35 and 36. The number of dimensions also affects the k-means algorithm due to the curse of dimensionality and are thus kept to a low number which is determined with the *elbow method* (see chapter 4.1.1 and 4.3.4). However, using a different number of dimensions in MDS may lead to different clustering results.

The clustering baseline with the k-means algorithm was adopted from Taubenböck, Debray et al. (2020). The k-means clustering algorithm results in Voronoi cells with a linear separation of the clusters. This might be sub-optimal in delineating the true cluster boundaries. The number of clusters is determined with the gap statistic method. As discussed earlier, the optimal number of clusters is ambiguous. However, it is a valid approach for a fully unsupervised clustering without prior knowledge on the searched clusters.

7 Conclusion and Outlook

Cities are complex systems with a unique composition of many different constituents and their specific relationships. Each city has a unique history and is formed by a range of functional, social, economical and environmental factors. However, similar elements and patterns can be observed among cities, especially in the physical environment. Exploring urban morphological configurations has been a subject of interest in many studies to gain a profound understanding of cities and to contribute towards an improvement of urban planning and the quality of living. Recent developments in EO and machine learning promoted the creation of new consistent datasets that allow the exploration of urban morphology with quantitative approaches on a global scale. Global mapping products provide a description of urban morphological components like the LCZ classification (cf. Qiu, Schmitt and Zhu 2019) or delineate urban areas in a comparable way like the MUAs (cf. Taubenböck, Weigand et al. 2019). A first study conducted by Taubenböck, Debray et al. (2020) hints towards different urban morphological configurations which are linked to geographic regions, yet the study was limited mostly to statistical characteristics in their methodological approaches.

This master thesis aimed to overcome previous shortcomings and contribute to the exploration of urban morphological patterns. Patterns of re-occurring morphological configurations were searched on subsets of cities without prior knowledge in a fully unsupervised manner. With the adaptation of new state of the art methods for unsupervised image clustering, 138 clusters of different urban morphological configurations were found. The clusters show subsets of cities with similar statistical distributions and spatial arrangements of their morphologic components. The found patterns describe urban morphological configurations on the aggregated level of the LCZs and are a first step towards a global typification of urban morphological patterns. Even though the results come with uncertainties, they denote a foundation for a qualitative formalization and quantitative description of urban morphological patterns which is missing to this day.

The found patterns were further used to compare the urban morphological configurations between cities with the aim to explore the formation of geographical regions. Therefore this work developed a similarity metric between cities by comparing their subsets, which belong to the previously found urban morphological patterns. The similarity metric was derived for 1523 cities worldwide and similar cities were grouped with a fully unsupervised clustering baseline. Using the k-means algorithm and gap statistic method seven clusters were statistically found. Grouping cities to seven clusters led to a partially emergence of geographic regions with similar cities. Some clusters are very congruent to distinctive geographic regions, while other clusters are ambiguous and spread more widely around the world. Cities are compared by their set of patterns, hence similar cities show similar urban morphological patterns. The formation of city clusters as geographic regions means in consequence that the formation of patterns with similar urban morphological configurations is linked to the geographical location. These findings provide a baseline to determine relevant factors that lead to the formation of urban morphological patterns and thus contributes towards a comprehensive understanding on the emergence of urban morphological elements.

Two ablation studies were conducted to compare the outcome of different methodological approaches. Therefore (i) the number of used cities is changed to 110 cities for a comparison with the results obtained by Taubenböck, Debray et al. (2020) and (ii) the methodology developed by Taubenböck, Debray et al. (2020) is applied on 1523 cities for a comparison with the results of this work. The ablation studies show that both methodological approaches come to similar results, on a large dataset with 1523 cities as well as on a small dataset with the 110 cities used by Taubenböck, Debray et al. (2020). Despite small differences, the results confirm that cities with similar urban morphological configurations form groups linked to partially distinctive geographic regions.

The findings of this work motivate to continue the exploration of urban morphological patterns and their link to the geographic location. To overcome the limitations of this study, an extension of the methodological implementation could be considered. For example the spatial arrangement of urban morphological patterns in cities could be included in a similarity metric. The results obtained in this master thesis raise multiple new questions. Why do cities in similar geographic locations have similar patterns of urban morphological configurations? What are the factors that lead to the formation of these patterns and how are these factors linked to the geographic location? The knowledge on space obtained with this work allows to examine how different functional, social, economical, physical and historical factors relate to the city clusters and thus might explain the formation of spatial patterns. Finding these answers will likely lead to a profound understanding of urban morphological processes and may be a subject of interest for future studies. Furthermore, the finding of urban morphological patterns in this work can contribute to the development of a comprehensive morphological city model. Based on these results, future studies might be able to create generalized prototypical patterns with the use of generative adversarial networks. This can be continued to create generalized models of prototypical cities, composed by prototypical patterns, for the seven clusters found in this work. In addition, the generation of those models could be linked to the factors responsible for the formation of patterns and linked to the geographical context. Thus, the master thesis provides a foundation for many new approaches in the exploration of urban morphology and contributes to a better understanding of cities.

References

- Adolphson, M. (2009). “Estimating a Polycentric Urban Structure. Case Study: Urban Changes in the Stockholm Region 1991–2004”. eng. In: *Journal of urban planning and development* 135.1, pp. 19–30. ISSN: 0733-9488 and 1943-5444 and 1943-5444.
- Alexander, C. (2015). *A City is Not a Tree: 50th Anniversary Edition*. Ed. by M. W. Mehaffy. Sustasis Press/Off The Common Books. ISBN: 978-0-9893469-7-9. URL: <http://www.sustasis.net/ACINAT-LR.pdf>.
- Alexander, C. (1988). “A city is not a tree”. In: *Design after modernism: Beyond the object*. Ed. by J. Thackara. London: Thames and Hudson, pp. 67–84.
- Alonso, W. (1964). *Location and land use. toward a general theory of of land rent*. eng. Publications of the Joint Center for Urban Studies of the Massachusetts Institute of Technology and Harvard University. Cambridge, Mass.: Harvard Univ. Pr., XI, 204 pages.
- Arthur, D. and Vassilvitskii, S. (2006). *k-means++: The Advantages of Careful Seeding*. Technical Report 2006-13. Stanford InfoLab. URL: <http://ilpubs.stanford.edu:8090/778/>.
- Baldi, P. (2012). “Autoencoders, Unsupervised Learning, and Deep Architectures.” In: *ICML Unsupervised and Transfer Learning*. Ed. by I. Guyon, G. Dror, V. Lemaire, G. W. Taylor and D. L. Silver. Vol. 27. JMLR Proceedings. JMLR.org, pp. 37–50. URL: <http://dblp.uni-trier.de/db/journals/jmlr/jmlrp27.html#Baldi12>.
- Barthelemy, M. (2016). *The Structure and Dynamics of Cities*. Cambridge University Press. ISBN: 9781107109179. URL: <https://books.google.de/books?id=PaDQDQAAQBAJ>.
- Bechtel, B., Alexander, P. J., Böhner, J., Ching, J., Conrad, O., Feddema, J., Mills, G., See, L. and Stewart, I. (2015). “Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities”. In: *ISPRS International Journal of Geo-Information* 4.1, pp. 199–219. ISSN: 2220-9964. DOI: 10.3390/ijgi4010199. URL: <https://www.mdpi.com/2220-9964/4/1/199>.
- Bechtel, B., Foley, M., Mills, G., Ching, J., See, L., Alexander, P., O’Connor, M., Albuquerque, T., Andrade, M., Brovelli, M., Das, D., Fonte, C., Petit, G., Hanif, U., Jiménez, J., Lackner, S., Liu, W., Perera, N., Rosni, N. A. and Gál, T. (2015). “CENSUS of Cities: LCZ Classification of Cities (Level 0) – Workflow and Initial Results from Various Cities”. In: *9th International Conference on Urban Climate*. Toulouse, France. URL: https://publicatio.bibl.u-szeged.hu/5873/1/GD2_2_2891238_a_u.pdf.
- Berghauser-Pont, M. and Haupt, P. (2010). *Spacematrix: Space, Density and Urban Form*. NAI Publishers. ISBN: 9789056627423. URL: <https://books.google.de/books?id=xwTZQgAACAAJ>.
- Berghauser-Pont, M. and Olsson, J. (2017). “Typology based on three density variables central to Spacematrix using cluster analysis”. In: *24th ISUF International Conference. Book of Papers*. Editorial Universitat Politècnica de València, pp. 1337–1348. DOI: 10.4995/ISUF2017.2017.5319.
- Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A. and Raffel, C. A. (2019). “MixMatch: A Holistic Approach to Semi-Supervised Learning”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox and R. Garnett. Vol. 32. Curran Associates, Inc. URL: <https://proceedings.neurips.cc/paper/2019/file/1cd138d0499a68f4bb72bee04bbec2d7-Paper.pdf>.

- Bettencourt, L. M. A. (2013). "The Origins of Scaling in Cities". In: *Science* 340.6139, pp. 1438–1441. DOI: 10.1126/science.1235823. URL: <https://www.science.org/doi/abs/10.1126/science.1235823>.
- Beyerer, J., Richter, M. and Nagel, M. (2018). *Pattern recognition. introduction, features, classifiers and principles*. eng. De Gruyter graduate. Berlin ; Boston: De Gruyter, XXI, 283 pages. ISBN: 978-3-11-053793-2 and 3-11-053793-1. URL: http://www.degruyter.com/search?f_0=isbnissn&q_0=9783110537932&searchTitles=true.
- Borg, I. and Groenen, P. (2005). *Modern Multidimensional Scaling: Theory and Applications*. 1st ed. Springer Series in Statistics. Springer, New York, NY. ISBN: 978-1-4757-2713-5. DOI: 10.1007/978-1-4757-2711-1.
- Bourlard, H. and Kamp, Y. (1988). "Auto-association by multilayer perceptrons and singular value decomposition". In: *Biological cybernetics* 59.4-5, pp. 291–294. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.453.846&rep=rep1&type=pdf>.
- Cai, J., Huang, B. and Song, Y. (2017). "Using multi-source geospatial big data to identify the structure of polycentric cities". In: *Remote Sensing of Environment* 202. Big Remotely Sensed Data: tools, applications and experiences, pp. 210–221. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2017.06.039>. URL: <https://www.sciencedirect.com/science/article/pii/S0034425717302985>.
- Caniggia, G. and Maffei, G. L. (1979). *Composizione architettonica e tipologia edilizia*. Marsilio Venezia.
- Caron, M., Bojanowski, P., Joulin, A. and Douze, M. (2018). "Deep Clustering for Unsupervised Learning of Visual Features". In: *CoRR* abs/1807.05520. URL: <http://arxiv.org/abs/1807.05520>.
- Castex, J., Depaule, J.-C. and Panerai, P. (1977). *Formes urbaines : de l'îlot à la barre*. français. ASPECTS DE L'URBANISME. Dunod. Paris; Centre d'études et de recherches architecturales (C.E.R.A) - École nationale supérieure d'architecture. Paris, 230 pages. ISBN: 2040101985. URL: <https://side.developpement-durable.gouv.fr/Default/doc/SYRACUSE/87524/formes-urbaines-de-l-ilot-a-la-barre>.
- Chang, J., Wang, L., Meng, G., Xiang, S. and Pan, C. (2017). "Deep Adaptive Image Clustering". In: *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 5880–5888. DOI: 10.1109/ICCV.2017.626.
- Chen, C. H. and Peter Ho, P.-G. (2008). "Statistical pattern recognition in remote sensing". In: *Pattern Recognition* 41.9, pp. 2731–2741. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2008.04.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0031320308001647>.
- Chen, T., Kornblith, S., Norouzi, M. and Hinton, G. E. (2020). "A Simple Framework for Contrastive Learning of Visual Representations". In: *CoRR* abs/2002.05709. URL: <https://arxiv.org/abs/2002.05709>.
- Ching, J., Mills, G., Bechtel, B., See, L., Feddema, J., Wang, X., Ren, C., Brousse, O., Martilli, A., Neophytou, M., Mouzourides, P., Stewart, I., Hanna, A., Ng, E., Foley, M., Alexander, P., Aliaga, D., Niyogi, D., Shreevastava, A., Bhalachandran, P., Masson, V., Hidalgo, J., Fung, J., Andrade, M., Baklanov, A., Dai, W., Milcinski, G., Demuzere, M., Brunsell, N., Pesaresi,

- M., Miao, S., Mu, Q., Chen, F. and Theeuwes, N. (2018). “WUDAPT: An Urban Weather, Climate, and Environmental Modeling Infrastructure for the Anthropocene”. In: *Bulletin of the American Meteorological Society* 99.9, pp. 1907–1924. DOI: 10.1175/BAMS-D-16-0236.1. URL: <https://journals.ametsoc.org/view/journals/bams/99/9/bams-d-16-0236.1.xml>.
- Christaller, W. (1933). *Die zentralen Orte in Süddeutschland. eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen.* ger. Jena: Fischer, 331 pages.
- Conzen, M. (1960). “Alnwick, Northumberland: A Study in Town-Plan Analysis”. In: *Transactions and Papers (Institute of British Geographers)* 27, pp. iii–122. ISSN: 14784017. URL: <http://www.jstor.org/stable/621094>.
- Conzen, M. (1981). *The urban landscape. historical development and management; papers.* eng. Ed. by J. W. R. Whitehand. Special publication / Institute of British Geographers. 13. London [u.a.]: Acad. Pr., VII, 166 pages. ISBN: 0-12-747020-4 and 978-0-12-747020-7.
- Cubuk, E. D., Zoph, B., Shlens, J. and Le, Q. V. (2020). “Randaugment: Practical Automated Data Augmentation With a Reduced Search Space”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- D’Acci, L. (2019). “On Urban Morphology and Mathematics”. In: *The Mathematics of Urban Morphology*. Ed. by L. D’Acci. Cham: Springer International Publishing, pp. 1–18. ISBN: 978-3-030-12381-9. DOI: 10.1007/978-3-030-12381-9_1. URL: https://doi.org/10.1007/978-3-030-12381-9_1.
- Dibble, J., Prelorndjos, A., Romice, O., Zanella, M., Strano, E., Pagel, M. and Porta, S. (2019). “On the origin of spaces: Morphometric foundations of urban form evolution”. In: *Environment and Planning B: Urban Analytics and City Science* 46.4, pp. 707–730. DOI: 10.1177/2399808317725075. URL: <https://doi.org/10.1177/2399808317725075>.
- Ding, C. and He, X. (2004). “K-means clustering via principal component analysis”. In: *ICML ’04: Proceedings of the twenty-first international conference on Machine learning*. Banff, Alberta, Canada: ACM, p. 29. ISBN: 1-58113-828-5. DOI: <http://doi.acm.org/10.1145/1015330.1015408>.
- Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A. and 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* 10.6, pp. 1617–1621. DOI: 10.1109/LGRS.2013.2272953.
- Faßmann, H. (2009). *Allgemeine Stadtgeographie.* ger. 2nd ed. Das geographische Seminar. Braunschweig: Westermann, 256 pages. ISBN: 978-3-14-160364-4.
- Fleischmann, M., Feliciotti, A. and Kerr, W. (2021). “Evolution of Urban Patterns: Urban Morphology as an Open Reproducible Data Science”. In: *Geographical Analysis*. ISSN: 0016-7363. DOI: 10.1111/gean.12302. URL: <https://doi.org/10.1111/gean.12302>.
- Fleischmann, M., Feliciotti, A., Romice, O. and Porta, S. (2021). “Methodological foundation of a numerical taxonomy of urban form”. In: *Environment and Planning B: Urban Analytics and City Science*. ISSN: 2399-8083. DOI: 10.1177/23998083211059835. URL: <https://doi.org/10.1177/23998083211059835>.

- Fritz, J. (1894). *Deutsche Stadtanlagen*. Beilage zum Programm Nr. 520 des Lyseums zu Strassburg i. Elsass. Heitz. URL: https://books.google.de/books?id=Op0%5C_AAAAYAAJ.
- Gaebe, W. (2004). *Urbane Räume. 61 Tabellen*. ger. UTB ; 2511 : Geographie. Stuttgart: Ulmer, 352 pages. ISBN: 3-8252-2511-9 and 3-8001-2760-1 and 978-3-8252-2511-7 and 978-3-8001-2760-3.
- Gamba, P. (2013). “Human Settlements: A Global Challenge for EO Data Processing and Interpretation”. In: *Proceedings of the IEEE* 101.3, pp. 570–581. ISSN: 1558-2256. DOI: 10.1109/JPROC.2012.2189089.
- Gansbeke, W. V., Vandenhende, S., Georgoulis, S., Proesmans, M. and Gool, L. V. (2020). “SCAN: Learning To Classify Images Without Labels”. In: *CoRR* abs/2005.12320. URL: <https://arxiv.org/abs/2005.12320>.
- Gebhardt, H., Glaser, R., Radtke, U., Reuber, P. and Vött, A., eds. (2020). *Geographie. physische Geographie und Humangeographie*. ger. 3rd ed. Lehrbuch. Berlin ; [Heidelberg]: Springer, XXI, 1272 pages. ISBN: 978-3-662-58378-4 and 3-662-58378-X. URL: <https://www.springer.com/de/book/9783662583784>.
- Gebhardt, H. and Reuber, P. (2020). “Der Wandel ländlicher Räume und die Forschungsaufgaben der Geographie”. ger. In: *Geographie: Physische Geographie und Humangeographie*. Ed. by H. Gebhardt, R. Glaser, U. Radtke, P. Reuber and A. Vött. 3rd ed. Lehrbuch. Berlin ; [Heidelberg]: Springer Spektrum, pp. 904–906. ISBN: 978-3-662-58378-4 and 3-662-58378-X. URL: <https://www.springer.com/de/book/9783662583784>.
- Géron, A. (2020). *Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. ger. 2nd ed. Heidelberg: O’Reilly, XXVII, 822 pages. ISBN: 978-3-96009-124-0 and 3-96009-124-9.
- Gidaris, S., Singh, P. and Komodakis, N. (2018). “Unsupervised Representation Learning by Predicting Image Rotations”. In: *CoRR* abs/1803.07728. URL: <http://arxiv.org/abs/1803.07728>.
- Glorot, X. and Bengio, Y. (2010). “Understanding the difficulty of training deep feedforward neural networks”. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Ed. by Y. W. Teh and M. Titterton. Vol. 9. Proceedings of Machine Learning Research. Chia Laguna Resort, Sardinia, Italy: PMLR, pp. 249–256. URL: <https://proceedings.mlr.press/v9/glorot10a.html>.
- Goethe, J. W. v. ([1817] 1954). *Morphologische Hefte*. ger. Ed. by R. Matthaei, W. Troll and L. Wolf. Weimar: Böhlau, XI, 389 pages.
- Goethe, J. W. v. ([1817] 1995). *Von 1816 bis 1824, Ergänzungen und Erläuterungen*. ger. Ed. by D. Kuhn and W. v. Engelhardt. Weimar: Böhlau, XXXIV, 978, XVI pages. ISBN: 3-7400-0953-5 and 978-3-7400-0953-3.
- Goh, S., Choi, M. Y., Lee, K. and Kim, K.-m. (2016). “How complexity emerges in urban systems: Theory of urban morphology”. In: *Phys. Rev. E* 93 (5). DOI: 10.1103/PhysRevE.93.052309. URL: <https://link.aps.org/doi/10.1103/PhysRevE.93.052309>.
- Goyal, P., Dollár, P., Girshick, R. B., Noordhuis, P., Wesolowski, L., Kyrola, A., Tulloch, A., Jia, Y. and He, K. (2017). “Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour”. In: *CoRR* abs/1706.02677. URL: <http://arxiv.org/abs/1706.02677>.

- Grassnick, M. and Hofrichter, H., eds. (1982). *Stadtbaugeschichte von der Antike bis zur Neuzeit*. ger. Materialien zur Baugeschichte 4. Braunschweig: Vieweg, 106, 205 pages. ISBN: 3-528-08684-X and 978-3-528-08684-8.
- Harris, C. D. and Ullman, E. L. (1945). "The Nature of Cities". In: *The Annals of the American Academy of Political and Social Science* 242.1, pp. 7–17. ISSN: 0002-7162 and 1552-3349. DOI: 10.1177/000271624524200103. URL: <https://doi.org/10.1177/000271624524200103>.
- Hassinger, H. (1916). *Kunsthistorischer Atlas der K. K. Reichshaupt- und Residenzstadt Wien und Verzeichnis der erhaltenswerten historischen, Kunst- und Naturdenkmale des Wiener Stadtbildes*. ger. Österreichische Kunsttopographie. 15. Wien: Schroll in Komm., p. 304.
- Hatt, P. (1946). "The Concept of Natural Area". In: *American Sociological Review* 11.4, pp. 423–427. ISSN: 00031224. URL: <http://www.jstor.org/stable/2087337>.
- He, K., Fan, H., Wu, Y., Xie, S. and Girshick, R. B. (2019). "Momentum Contrast for Unsupervised Visual Representation Learning". In: *CoRR* abs/1911.05722. URL: <http://arxiv.org/abs/1911.05722>.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015a). "Deep Residual Learning for Image Recognition". In: *CoRR* abs/1512.03385. URL: <http://arxiv.org/abs/1512.03385>.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015b). "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification". In: *CoRR* abs/1502.01852. URL: <http://arxiv.org/abs/1502.01852>.
- Heineberg, H. (2017). *Stadtgeographie*. ger. 5th ed. Grundriss Allgemeine Geographie. Paderborn: Ferdinand Schöningh, 504 pages. ISBN: 3-8252-4708-2 and 978-3-8252-4708-9.
- Heineberg, H. (2020). "Stadtstrukturmodelle und die innere Gliederung der Stadt". ger. In: *Geographie: Physische Geographie und Humangeographie*. Ed. by H. Gebhardt, R. Glaser, U. Radtke, P. Reuber and A. Vött. 3rd ed. Lehrbuch. Berlin ; [Heidelberg]: Springer Spektrum, pp. 849–859. ISBN: 978-3-662-58378-4 and 3-662-58378-X. URL: <https://www.springer.com/de/book/9783662583784>.
- Hillier, B. (1996). *Space is the machine. a configurational theory of architecture*. eng. 1. publ. Cambridge: Cambridge Univ. Press, XII, 463 pages. ISBN: 0-521-56039-X and 978-0-521-56039-9.
- Hofmeister, B. (1996). *Die Stadtstruktur. Ihre Ausprägung in den verschiedenen Kulturräumen der Erde*. ger. 3., überarb. Aufl. Erträge der Forschung. Literaturangaben. Darmstadt: Wiss. Buchges., [Abt. Verl.], VI, 194 pages. ISBN: 3-534-12998-9 and 978-3-534-12998-0.
- Howard, E. ([1902] 1944). *Garden cities of to-morrow*. eng. London: Faber, 168 pages.
- Hoyt, H. (1939). *The Structure and Growth of Residential Neighborhoods in American Cities*. U.S. Government Printing Office. URL: <https://books.google.de/books?id=VtjZdGS0WhgC>.
- Hu, J., Ghamisi, P. and Zhu, X. X. (2018). "Feature Extraction and Selection of Sentinel-1 Dual-Pol Data for Global-Scale Local Climate Zone Classification". In: *ISPRS International Journal of Geo-Information* 7.9. ISSN: 2220-9964. DOI: 10.3390/ijgi7090379. URL: <https://www.mdpi.com/2220-9964/7/9/379>.
- Hunt, A. and Watkiss, P. (2011). "Climate change impacts and adaptation in cities: a review of the literature". In: *Climatic Change* 104.1, pp. 13–49. ISSN: 1573-1480. DOI: 10.1007/s10584-010-9975-6. URL: <https://doi.org/10.1007/s10584-010-9975-6>.

- Huntington, S. P. (1993). "The Clash of Civilizations?" In: *Foreign Affairs* 72.3, pp. 22–49. ISSN: 00157120. URL: <http://www.jstor.org/stable/20045621>.
- Hurd, R. M. (1903). *Principles of city land values*. eng. RLIN, CTRG96-B2528 ; Reproduction of original from Harvard Law School Library. New York: Record and Guide, viii, 159 pages. URL: <http://nl.sub.uni-goettingen.de/id/20002116800?origin=/collection/nlh-mml>.
- Ioffe, S. and Szegedy, C. (2015). "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *CoRR* abs/1502.03167. URL: <http://arxiv.org/abs/1502.03167>.
- Jensen, J. R. (2005). *Introductory digital image processing. a remote sensing perspective*. eng. 3rd ed. Prentice Hall series in geographic information science. Includes bibliographical references. Upper Saddle River, NJ: Prentice Hall, XVI, 526 pages. ISBN: 0-13-145361-0 and 978-0-13-145361-6.
- Kolb, A. (1962). "Die Geographie und die Kulturerdteile". In: *Hermann von Wissmann-Festschrift. Tübingen*. Ed. by A. Leidlmair, pp. 42–49.
- Kropf, K. (2014). "Ambiguity in the definition of built form". In: *Urban Morphology* 18, pp. 41–57. URL: https://www.urbanform.org/online_unlimited/pdf2014/201418_41.pdf.
- Kulish, M., Richards, A. and Gillitzer, C. (2012). "Urban Structure and Housing Prices: Some Evidence from Australian Cities*". In: *Economic Record* 88.282, pp. 303–322. DOI: <https://doi.org/10.1111/j.1475-4932.2012.00829.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-4932.2012.00829.x>.
- Le Corbusier (1924). *Urbanisme*. fre. 7. éd. Collection de "L'esprit nouveau". Paris: Crès, V, 284 pages.
- Lehner, A. and Blaschke, T. (2019). "A Generic Classification Scheme for Urban Structure Types". In: *Remote Sensing* 11.2. ISSN: 2072-4292. DOI: 10.3390/rs11020173. URL: <https://www.mdpi.com/2072-4292/11/2/173>.
- Li, J., Socher, R. and Hoi, S. C. H. (2020). "DivideMix: Learning with Noisy Labels as Semi-supervised Learning". In: *CoRR* abs/2002.07394. URL: <https://arxiv.org/abs/2002.07394>.
- Lösch, A. (1940). *Die räumliche Ordnung der Wirtschaft*. ger. Jena: G.Fischer, VIII, 380 pages.
- MacQueen, J. et al. (1967). "Some methods for classification and analysis of multivariate observations". In: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. Vol. 1. 14. Oakland, CA, USA, pp. 281–297.
- Marcus, L., Berghauser Pont, M., Stavroulaki, G. and Bobkova, E. (2017). "Location-based density and differentiation-adding attraction variables to space syntax". In: *24th ISUF International Conference. Book of Papers*. Editorial Universitat Politècnica de València, pp. 1379–1389. DOI: 10.4995/ISUF2017.2017.5706.
- McCulloch, W. S. and Pitts, W. (1943). "A logical calculus of the ideas immanent in nervous activity". In: *The bulletin of mathematical biophysics* 5.4, pp. 115–133. ISSN: 1522-9602. DOI: 10.1007/BF02478259. URL: <https://doi.org/10.1007/BF02478259>.
- Mills, E. S. (1967). "An Aggregative Model of Resource Allocation in a Metropolitan Area". In: *The American Economic Review* 57.2, pp. 197–210. ISSN: 00028282. URL: <http://www.jstor.org/stable/1821621>.

- Munkres, J. (1957). “Algorithms for the assignment and transportation problems”. In: *Journal of the society for industrial and applied mathematics* 5.1, pp. 32–38.
- Muratori, S. (1959). “Studi per una operante storia urbana di Venezia”. In: *Palladio* 1959, pp. 1–113.
- Muth, R. F. (1969). *Cities and housing. the spatial pattern of urban residential land use.* eng. Graduate School of Business, Univ. of Chicago : Serie 3 : Studies in business and society. Chicago: Univ. of Chicago Press, XXII, 355 pages.
- Nissen, H. J. (1988). *The early history of the Ancient Near East. 9000 - 2000 B.C.* eng. Hier auch weitere unveränderte Drucke. Chicago [u.a.]: Univ. of Chicago Pr., XIV, 215 pages. ISBN: 0-226-58656-1 and 0-226-58658-8 and 978-0-226-58656-4 and 978-0-226-58658-8.
- Niu, C. and Wang, G. (2021). “SPICE: Semantic Pseudo-labeling for Image Clustering”. In: *CoRR* abs/2103.09382. URL: <https://arxiv.org/abs/2103.09382>.
- Noroozi, M. and Favaro, P. (2016). “Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles”. In: *CoRR* abs/1603.09246. URL: <http://arxiv.org/abs/1603.09246>.
- Oliveira, V. (2016). *Urban morphology. an introduction to the study of the physical form of cities.* eng. The urban book series. Switzerland: Springer, XXIII, 192 pages. ISBN: 978-3-319-32081-6.
- Oliveira, V. (2019). “Urban Forms, Agents, and Processes of Change”. In: *The Mathematics of Urban Morphology*. Ed. by L. D’Acci. Cham: Springer International Publishing, pp. 529–535. ISBN: 978-3-030-12381-9. DOI: 10.1007/978-3-030-12381-9_28. URL: https://doi.org/10.1007/978-3-030-12381-9_28.
- Papers with Code (2022). *Unsupervised Image Classification*. Custom data acquired via website (accessed 1 January 2022). URL: <https://paperswithcode.com/task/unsupervised-image-classification>.
- Park, R. E., Burgess, E. W. and McKenzie, R. D. ([1925] 1967). *The city.* eng. 4.impr. The heritage of sociology. Chicago, Ill. [u.a.]: Univ. of Chicago Pr., X, 239 pages.
- Park, S., Han, S., Kim, S., Kim, D., Park, S., Hong, S. and Cha, M. (2020). “Improving Unsupervised Image Clustering With Robust Learning”. In: *CoRR* abs/2012.11150. URL: <https://arxiv.org/abs/2012.11150>.
- Passarge, S., ed. (1930). *Stadtlandschaften der Erde.* ger. Hamburg: Friederichsen, de Gruyter, VI, 154 pages.
- Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T. and Efros, A. A. (2016). “Context Encoders: Feature Learning by Inpainting”. In: *CoRR* abs/1604.07379. URL: <http://arxiv.org/abs/1604.07379>.
- Qiu, C., Mou, L., Schmitt, M. and Zhu, X. X. (2019). “Local climate zone-based urban land cover classification from multi-seasonal Sentinel-2 images with a recurrent residual network”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 154, pp. 151–162. ISSN: 0924-2716. DOI: <https://doi.org/10.1016/j.isprsjprs.2019.05.004>. URL: <https://www.sciencedirect.com/science/article/pii/S0924271619301261>.
- Qiu, C., Schmitt, M. and Zhu, X. X. (2019). “Fusing Multi-Seasonal Sentinel-2 Images with Residual Convolutional Neural Networks for Local Climate Zone-Derived Urban Land Cover Classification”. In: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5037–5040. DOI: 10.1109/IGARSS.2019.8898223.

- Quinn, J. A. (1940). "The Burgess Zonal Hypothesis and its Critics". In: *American Sociological Review* 5.2, pp. 210–218. ISSN: 00031224. DOI: 10.2307/2083636. URL: <http://www.jstor.org/stable/2083636>.
- Ramón y Cajal, S. ([1899] 1988). *Cajal on the Cerebral cortex : an annotated translation of the complete writings. Studies on the Human Cerebral Cortex II: Structure of the Motor Cortex of Man and Higher Mamals*. eng. Ed. by J. DeFelipe. 2nd ed. 1. New York: Oxford Univ. Pr. ISBN: 0-19-505280-3.
- Reuber, P. and Gebhardt, H. (2020). "Wissenschaftliches Arbeiten in der Geographie". ger. In: *Geographie: Physische Geographie und Humangeographie*. Ed. by H. Gebhardt, R. Glaser, U. Radtke, P. Reuber and A. Vött. 3rd ed. Lehrbuch. Berlin ; [Heidelberg]: Springer Spektrum, pp. 75–85. ISBN: 978-3-662-58378-4 and 3-662-58378-X. URL: <https://www.springer.com/de/book/9783662583784>.
- Rosenblatt, F. (1958). "The perceptron: A probabilistic model for information storage and organization in the brain." In: *Psychological Review* 65.6, pp. 386–408. DOI: 10.1037/h0042519. URL: <https://doi.org/10.1037/h0042519>.
- Rubenstein, J. M. (1999). *The cultural landscape. an introduction to human geography*. eng. 6. ed. Includes bibliographical references and indexes. Upper Saddle River, NJ: Prentice Hall, XIII, 543 pages. ISBN: 0-13-079778-2 and 978-0-13-079778-0.
- Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986). "Learning representations by back-propagating errors". In: *Nature* 323.6088, pp. 533–536. ISSN: 1476-4687. DOI: 10.1038/323533a0. URL: <https://doi.org/10.1038/323533a0>.
- Sadeghi, G. and Li, B. (2019). "Urban Morphology: Comparative Study of Different Schools of Thought". In: *Current Urban Studies* 07, pp. 562–572. DOI: 10.4236/cus.2019.74029.
- Schlüter, O. (1899). "Über den Grundriss der Städte". In: *Zeitschrift der Gesellschaft für Erdkunde* 34.6, pp. 446–461. URL: <https://ia802205.us.archive.org/9/items/berdengrundriss00schlgoog/berdengrundriss00schlgoog.pdf>.
- Schneider, A., Friedl, M. A. and Potere, D. (2010). "Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'". In: *Remote Sensing of Environment* 114.8, pp. 1733–1746. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2010.03.003>. URL: <https://www.sciencedirect.com/science/article/pii/S003442571000091X>.
- Seto, K. C. and Fragkias, M. (2005). "Quantifying Spatiotemporal Patterns of Urban Land-use Change in Four Cities of China with Time Series Landscape Metrics". In: *Landscape Ecology* 20.7, pp. 871–888. ISSN: 1572-9761. DOI: 10.1007/s10980-005-5238-8. URL: <https://doi.org/10.1007/s10980-005-5238-8>.
- Shevky, E. and Bell, W. (1955). *Social area analysis; theory, illustrative application and computational procedures*. Stanford: Stanford University Press.
- Shevky, E. and Williams, M. (1949). *The social areas of Los Angeles, analysis and typology*. Berkley: John Randolph Haynes and Dora Haynes Foundation.
- Sjoberg, G. (1960). *The preindustrial city. past and present*. eng. Glencoe, Ill.: Free Press, XII, 353 pages.

- Small, C. and Sousa, D. (2016). “Humans on Earth: Global extents of anthropogenic land cover from remote sensing”. In: *Anthropocene* 14, pp. 1–33. ISSN: 2213-3054. DOI: <https://doi.org/10.1016/j.ancene.2016.04.003>. URL: <https://www.sciencedirect.com/science/article/pii/S2213305416300339>.
- Sonka, M., Hlaváč, V. and Boyle, R. (2008). *Image processing, analysis, and machine vision*. eng. 3. ed. International student edition. Includes bibliographical references and index. Stamford, Conn.: Cengage Learning, XXV, 829 pages. ISBN: 978-0-495-24438-7 and 0-495-24438-4.
- Stähle, A., Marcus, L. and Karlström, A. (2005). “Place Syntax - Geographic Accessibility with Axial Lines in GIS”. In: *5th International Space Syntax Symposium*, pp. 131–144.
- Stewart, I. D. and Oke, T. R. (2012). “Local Climate Zones for Urban Temperature Studies”. In: *Bulletin of the American Meteorological Society* 93.12, pp. 1879–1900. DOI: 10.1175/BAMS-D-11-00019.1. URL: <https://journals.ametsoc.org/view/journals/bams/93/12/bams-d-11-00019.1.xml>.
- Sullivan, L. H. (1896). “The tall office building artistically considered”. In: *Lippincotts Magazine*. URL: https://ocw.mit.edu/courses/architecture/4-205-analysis-of-contemporary-architecture-fall-2009/readings/MIT4_205F09_Sullivan.pdf.
- Taubenböck, H., Debray, H., Qiu, C., Schmitt, M., Wang, Y. and Zhu, X. (2020). “Seven city types representing morphologic configurations of cities across the globe”. In: *Cities* 105. ISSN: 0264-2751. DOI: <https://doi.org/10.1016/j.cities.2020.102814>. URL: <https://www.sciencedirect.com/science/article/pii/S0264275120301347>.
- Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A. and Dech, S. (2012). “Monitoring urbanization in mega cities from space”. In: *Remote Sensing of Environment* 117. Remote Sensing of Urban Environments, pp. 162–176. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2011.09.015>. URL: <https://www.sciencedirect.com/science/article/pii/S0034425711003427>.
- Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J. and Dech, S. (2019). “A new ranking of the world’s largest cities—Do administrative units obscure morphological realities?” In: *Remote Sensing of Environment* 232. ISSN: 0034-4257. DOI: <https://doi.org/10.1016/j.rse.2019.111353>. URL: <https://www.sciencedirect.com/science/article/pii/S0034425719303724>.
- Taubenböck, H., Wurm, M., Geiß, C., Dech, S. and Siedentop, S. (2019). “Urbanization between compactness and dispersion: designing a spatial model for measuring 2D binary settlement landscape configurations”. In: *International Journal of Digital Earth* 12.6, pp. 679–698. DOI: 10.1080/17538947.2018.1474957. URL: <https://doi.org/10.1080/17538947.2018.1474957>.
- Thünen, J. H. v. (2014). *Der isolierte Staat*. Salzwasser-Verlag GmbH. ISBN: 9783863832247. URL: <https://books.google.de/books?id=AWfVAgAAQBAJ>.
- Tibshirani, R., Walther, G. and Hastie, T. (2001). “Estimating the number of clusters in a data set via the gap statistic”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.2, pp. 411–423. DOI: <https://doi.org/10.1111/1467-9868.00293>. URL: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9868.00293>.

- United Nations, Department of Economic and Social Affairs, Population Division (2015). *World Urbanization Prospects: The 2014 Revision*. URL: <https://population.un.org/wup/Publications/Files/WUP2014-Report.pdf>.
- United Nations, Department of Economic and Social Affairs, Population Division (2018). *World Urbanization Prospects: The 2018 Revision*. Custom data acquired via website (accessed 1 January 2022). URL: <https://population.un.org/wup/DataQuery/>.
- Vance, J. E. (1971). “Land Assignment in the Precapitalist, Capitalist, and Postcapitalist City”. In: *Economic Geography* 47.2, pp. 101–120. ISSN: 00130095, 19448287. DOI: 10.2307/143040. URL: <http://www.jstor.org/stable/143040>.
- Weisdorf, J. L. (2005). “From Foraging To Farming: Explaining The Neolithic Revolution”. In: *Journal of Economic Surveys* 19.4, pp. 561–586. DOI: <https://doi.org/10.1111/j.0950-0804.2005.00259.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0950-0804.2005.00259.x>.
- Wheaton, W. C. (1974). “A comparative static analysis of urban spatial structure”. In: *Journal of Economic Theory* 9.2, pp. 223–237. ISSN: 0022-0531. DOI: [https://doi.org/10.1016/0022-0531\(74\)90068-4](https://doi.org/10.1016/0022-0531(74)90068-4). URL: <https://www.sciencedirect.com/science/article/pii/0022053174900684>.
- Whitehand, J. W. (2001). “British urban morphology: the Conzenion tradition”. In: *Urban Morphology* 5.2, pp. 103–109. ISSN: 1027-4278. URL: <https://urbanmorphology.org/pdf/whitehand2001.pdf>.
- Wirth, L. (1938). “Urbanism as a Way of Life”. In: *American Journal of Sociology* 44.1, pp. 1–24. ISSN: 00029602, 15375390. URL: <http://www.jstor.org/stable/2768119>.
- Wolpert, D. H. and Macready, W. (2005). “Coevolutionary free lunches”. In: *IEEE Transactions on Evolutionary Computation* 9.6, pp. 721–735. DOI: 10.1109/TEVC.2005.856205.
- Wolpert, D. H. (1996). “The Lack of A Priori Distinctions Between Learning Algorithms”. In: *Neural Computation* 8.7, pp. 1341–1390. ISSN: 0899-7667. DOI: 10.1162/neco.1996.8.7.1341. URL: <https://doi.org/10.1162/neco.1996.8.7.1341>.
- World Meteorological Organization (2022). *OSCAR: Observing Systems Capability Analysis and Review Tool*. Custom data acquired via website (accessed 1 January 2022). URL: <https://space.oscar.wmo.int/satellites>.
- Wright, R. P. (2010). *The ancient Indus. urbanism, economy, and society*. eng. 1. publ. Case studies in early societies. Includes bibliographical references and index. Cambridge [u.a.]: Cambridge Univ. Press, XIX, 396 pages. ISBN: 978-0-521-57652-9 and 978-0-521-57219-4 and 0-521-57219-3 and 0-521-57652-0.
- Xie, J., Girshick, R. B. and Farhadi, A. (2016). “Unsupervised Deep Embedding for Clustering Analysis”. In: *CoRR* abs/1511.06335v2. URL: <http://arxiv.org/abs/1511.06335v2>.
- Yokoya, N., Ghamisi, P., Xia, J., Sukhanov, S., Heremans, R., Tankoyeu, I., Bechtel, B., Le Saux, B., Moser, G. and Tuia, D. (2018). “Open Data for Global Multimodal Land Use Classification: Outcome of the 2017 IEEE GRSS Data Fusion Contest”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11.5, pp. 1363–1377. DOI: 10.1109/JSTARS.2018.2799698.

- Yuan, C. and Yang, H. (2019). “Research on K-Value Selection Method of K-Means Clustering Algorithm”. In: *J* 2.2, pp. 226–235. ISSN: 2571-8800. DOI: 10.3390/j2020016. URL: <https://www.mdpi.com/2571-8800/2/2/16>.
- Zhang, R., Isola, P. and Efros, A. A. (2016). “Colorful Image Colorization”. In: *Computer Vision – ECCV 2016*. Ed. by B. Leibe, J. Matas, N. Sebe and M. Welling. Cham: Springer International Publishing, pp. 649–666. ISBN: 978-3-319-46487-9.
- Zonneveld, W. (2005). “Polycentric City”. eng. In: *Encyclopedia of the city*. Ed. by R. W. Caves. 1. publ. Enth. Literaturangaben und Index. London ; New York, NY: Routledge, Taylor & Francis Group, p. 355. ISBN: 0-415-25225-3 and 978-0-415-25225-6.

Appendix

Table 1: Transformations used by Gansbeke et al. (2020) for the SimCLR pretext task and the adapted transformation used in this work for the SimCLR pretext task. Geometric transformations are kept, the *Random Resized Crop* is scaled to 75-100% and an additional *Vertical Flip* is added. The color distortions are deprecated due to the categorical nature of the LCZs data.

Transformation	T_{con} Original	T_{con} Adaptation
Random Resized Crop	Size: 32, Scale: [0.2-1.0]	Size: 32, Scale: [0.75-1.0]
Color Jitter	Probability: 80%	-
Random Grayscale	Probability: 20%	-
Normalization	<i>data dependent</i>	-
Horizontal Flip	Probability: 50%	Probability: 50%
Vertical Flip	-	Probability: 50%

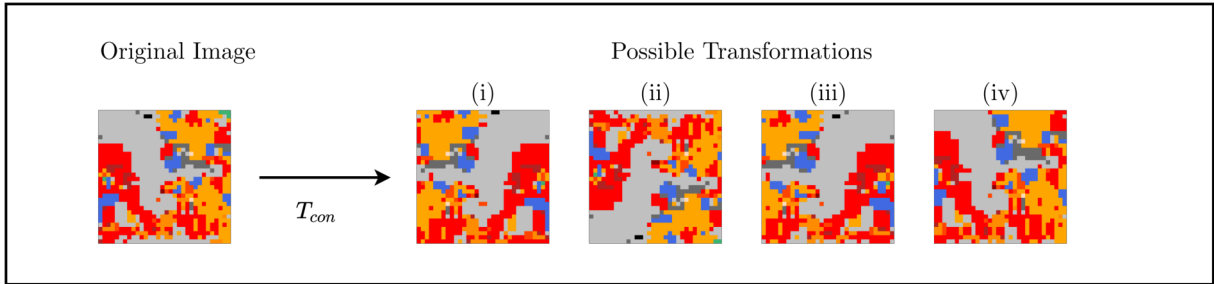


Figure 55: Four exemplary transformations of a patch with T_{con} (see table 1).

Table 2: Transformations used by Gansbeke et al. (2020) for the clustering task and the adapted transformation used in this work. Four random augmentations from the set of *Randaugment* are used (see table 6). The *Cutout* parameter with $16 * 16$ pixels is changed to 32 holes with $2 * 2$ pixels. The original *Cutout* parameter with $16 * 16$ is assumed to be beneficial for images with a single object (such as the benchmark datasets), but less for the LCZs data. Considering that a large cutout forces the network to focus on the remaining characteristics of an object shown in an image, this work assumes that this is not applicable on the LCZ data. Here, the overall distribution is more relevant, hence the parameter is changed to smaller and scattered cutouts (32 holes with size $2 * 2$) of the LCZs data.

Transformation	T_{scan} Original	T_{scan} Adaptation
Horizontal Flip	Probability: 50%	Probability: 50%
R_{SCAN} Adaptation (table 6)	no. augmentations: 4	no. augmentations: 4
Normalization	<i>data dependent</i>	-
Cutout	no. holes: 1, size: $16 * 16$	no. holes: 32, size: $2 * 2$

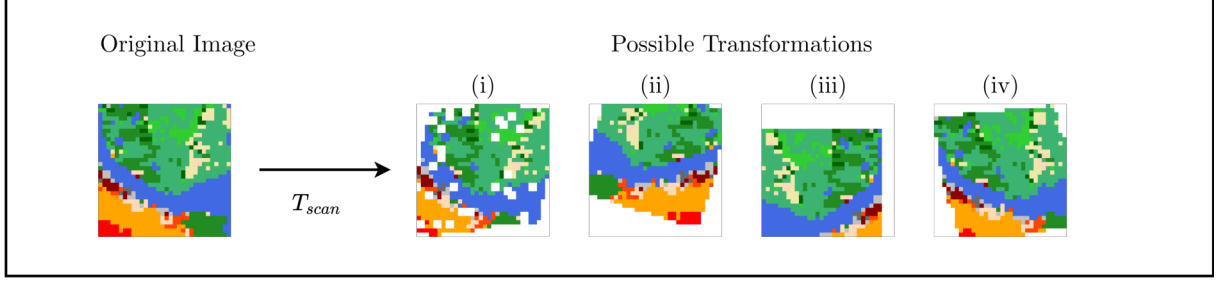


Figure 56: Four exemplary transformations of a patch with T_{scan} (see table 2).

Table 3: Transformations used by Gansbeke et al. (2020) for the self-labeling task and the adapted transformation used in this work. Four augmentations from *Randaugment* are randomly chosen (see table 6). The *Cutout* parameter with $16 * 16$ pixels is changed to four holes with $2 * 2$ pixels. This work expects that urban morphological patterns have no distinct boundaries and the transition between different types is fluent. To support the clustering of patches, the amount of generalization is reduced by using only four cutout holes.

Transformation	T_{self} Original	T_{self} Adaptation
Horizontal Flip	Probability: 50%	Probability: 50%
R_{SCAN} Adaptation (table 6)	no. augmentations: 4	no. augmentations: 4
Normalization	<i>data dependent</i>	-
Cutout	no. holes: 1, size: $16 * 16$	no. holes: 4, size: $2 * 2$

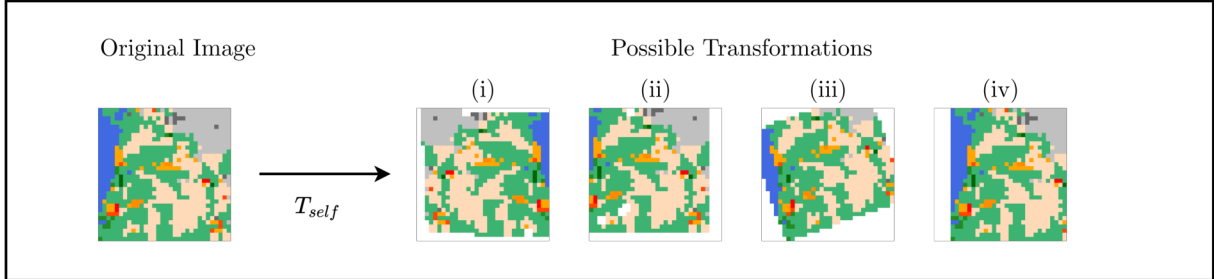


Figure 57: Four exemplary transformations of a patch with T_{self} (see table 3).

Table 4: The weak augmentations used in the RUC framework developed by Park, Han et al. (2020) and the adapted transformation for this work. An additional *Vertical Flip* is added.

Transformation	T_{weak} Original	T_{weak} Adaptation
Random Resized Crop	Size: 32, Scale: [0.2-1.0]	Size: 32, Scale: [0.75-1.0]
Normalization	<i>data dependent</i>	-
Horizontal Flip	Probability: 50%	Probability: 50%
Vertical Flip	-	Probability: 50%

Table 5: The strong augmentations used in the RUC framework developed by Park, Han et al. (2020) and the adapted transformation for this work. Again, the *Cutout* is changed to smaller, scattered cutouts.

Transformation	T_{strong} Original	T_{strong} Adaptation
Random Resized Crop	Size: 32, Scale: [0.2-1.0]	Size: 32, Scale: [0.75-1.0]
Horizontal Flip	Probability: 50%	Probability: 50%
R_{RUC} Adaptation (table 6)	no. augmentations: 2	no. augmentations: 2
Normalization	<i>data dependent</i>	-
Cutout	no. holes: 1, size: 16 * 16	no. holes: 32, size: 2 * 2

Table 6: Randaugment developed by Cubuk et al. (2020) is a set of possible augmentations. They are used in the SCAN framework of Gansbeke et al. (2020) with parameters R_{SCAN} *Original* which are adapted to R_{SCAN} *Adaptation* for this master thesis. They are also used in the RUC framework of Park, Han et al. (2020) with parameters R_{RUC} *Original* which are adapted to R_{RUC} *Adaptation* in this work. From this set, a specific number of augmentations is randomly chosen in the tables 2, 3 and 5.. All color distortions are deprecated due to the categorical nature of the LCZ data. The geometric distortions are kept.

Transformation	R_{SCAN} Original	R_{SCAN} Adaptation	R_{RUC} Original	R_{RUC} Adaptation
Identity	Yes	Yes	Yes	Yes
Autocontrast	Yes	-	Yes	-
Equalize	Yes	-	Yes	-
Rotate	[-30, 30]	[-30, 30]	[-45, 45]	[-45, 45]
Solarize	[0, 256]	-	[0, 256]	-
Color	[0.05, 0.95]	-	[0.01, 0.99]	-
Contrast	[0.05, 0.95]	-	[0.01, 0.99]	-
Brightness	[0.05, 0.95]	-	[0.01, 0.99]	-
Sharpness	[0.05, 0.95]	-	[0.01, 0.99]	-
Shear X	[-0.1, 0.1]	[-0.1, 0.1]	[-0.3, 0.3]	[-0.3, 0.3]
Translation X	[-0.1, 0.1]	[-0.1, 0.1]	[-0.3, 0.3]	[-0.3, 0.3]
Translation Y	[-0.1, 0.1]	[-0.1, 0.1]	[-0.3, 0.3]	[-0.3, 0.3]
Posterize	[4, 8]	-	[1, 8]	-
Shear Y	[-0.1, 0.1]	[-0.1, 0.1]	[-0.3, 0.3]	[-0.3, 0.3]

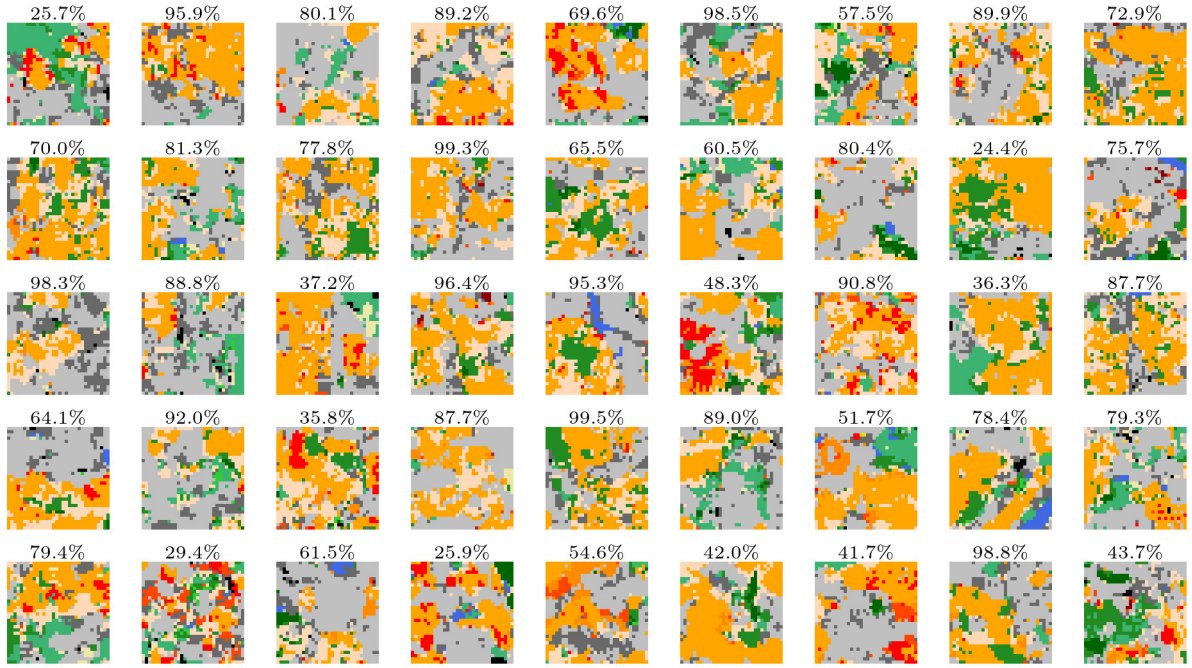


Figure 58: Exemplary cluster c_5 with 4121 patches in total. A random selection and their confidence values is shown.

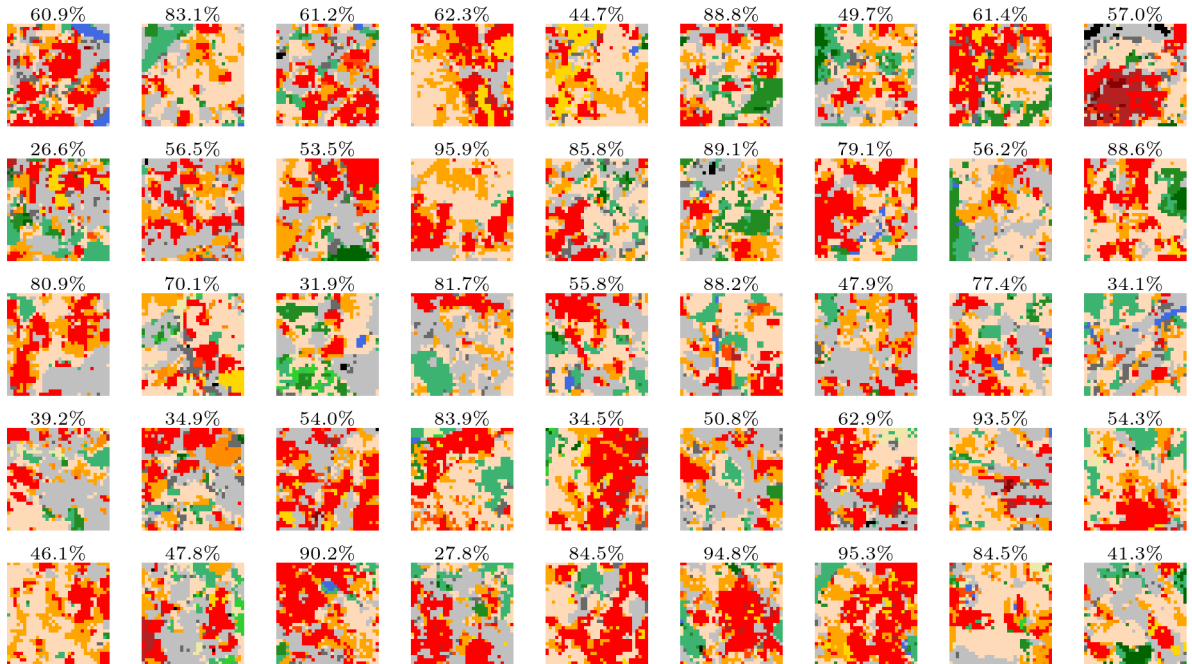


Figure 59: Exemplary cluster c_6 with 2069 patches in total. A random selection and their confidence values is shown.

Local Climate Zones		5 - open midrise	10 - heavy industry	15 - bare rock or paved
1 - compact high-rise	6 - open low-rise	11 - dense trees	16 - bare soil or sand	
2 - compact midrise	7 - lightweight low-rise	12 - scattered trees	17 - water	
3 - compact low-rise	8 - large low-rise	13 - bush, scrub		
4 - open high-rise	9 - sparsely built	14 - low plants		

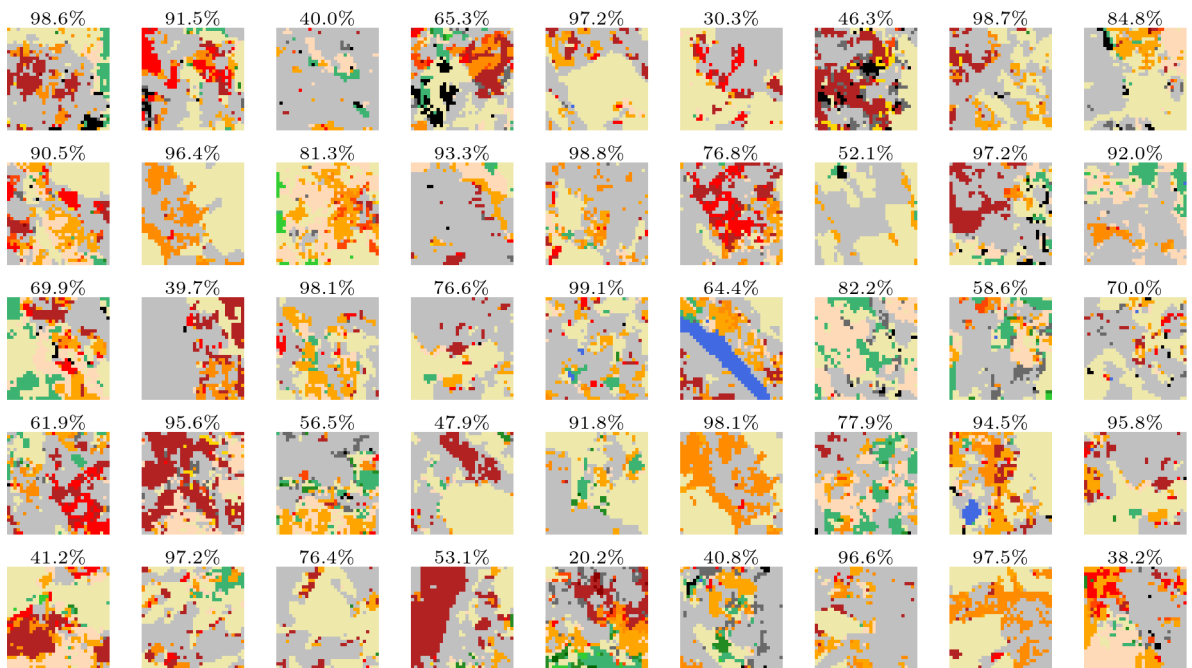


Figure 60: Exemplary cluster c_7 with 2032 patches in total. A random selection and their confidence values is shown.

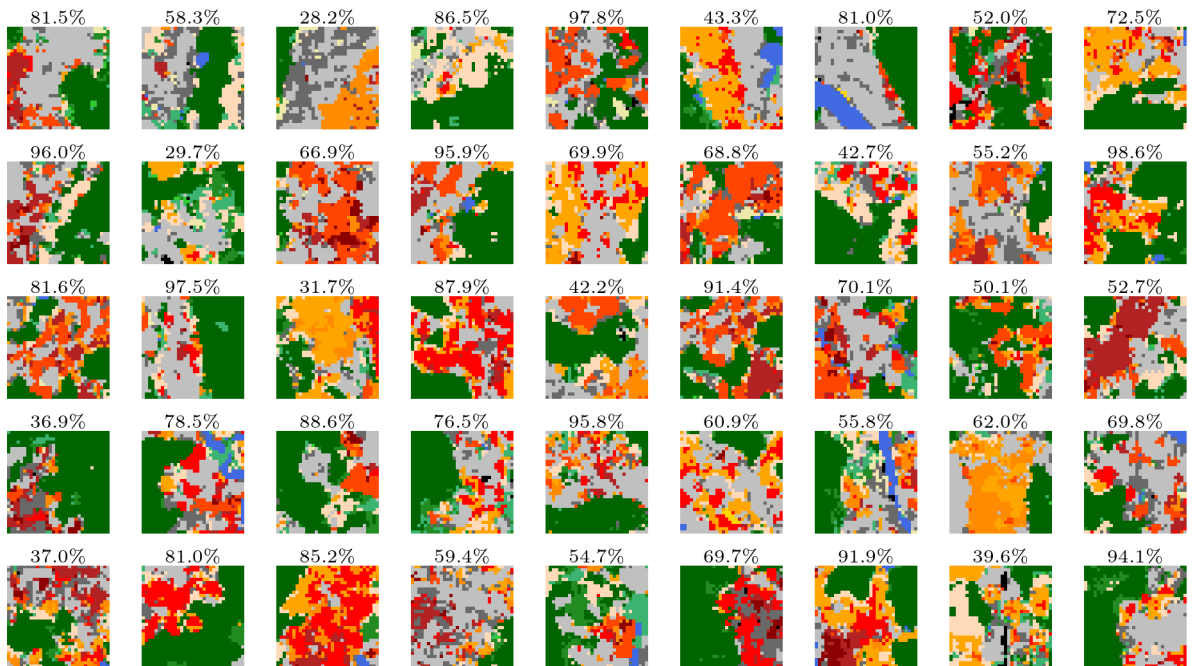
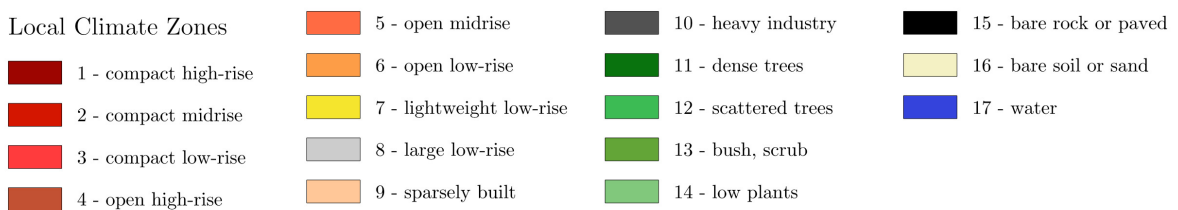


Figure 61: Exemplary cluster c_8 with 2307 patches in total. A random selection and their confidence values is shown.



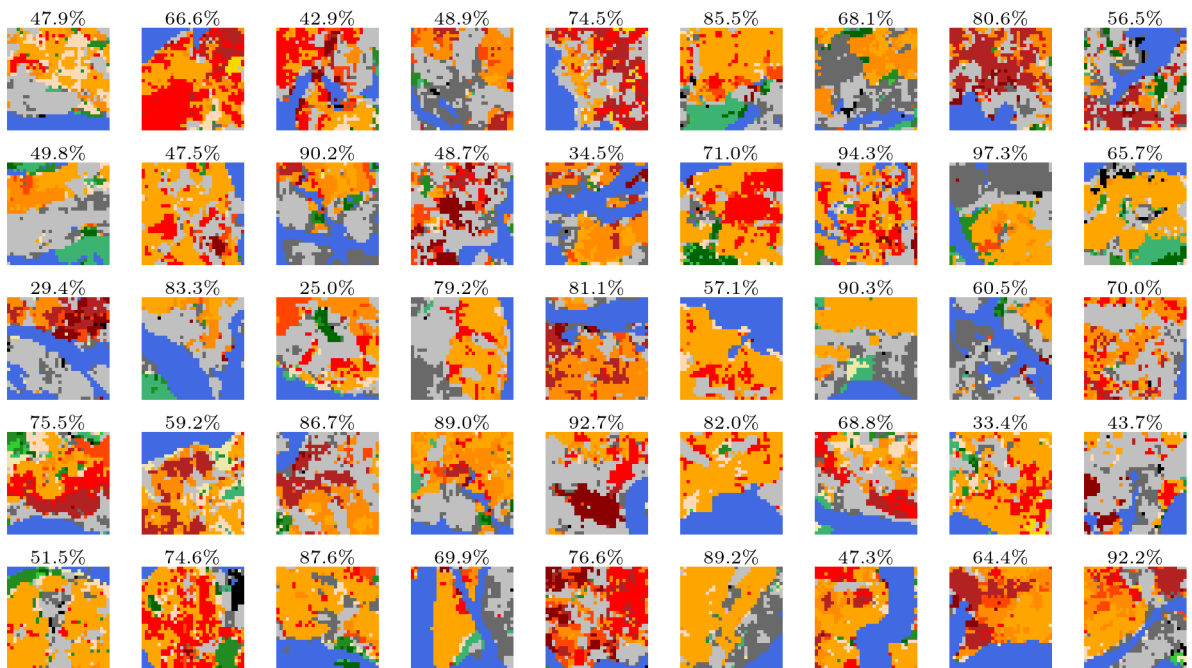


Figure 62: Exemplary cluster c_9 with 1180 patches in total. A random selection and their confidence values is shown.

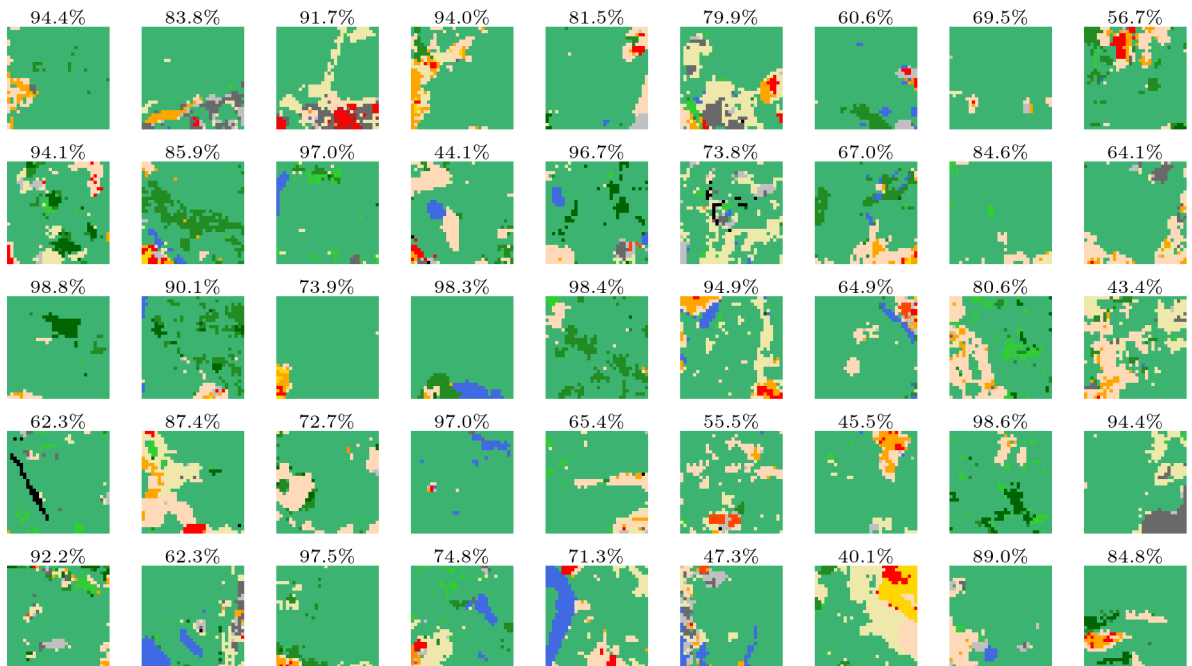
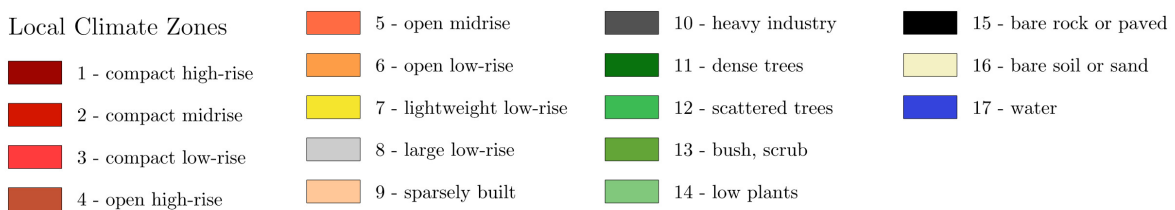


Figure 63: Exemplary cluster c_{10} with 2172 patches in total. A random selection and their confidence values is shown.



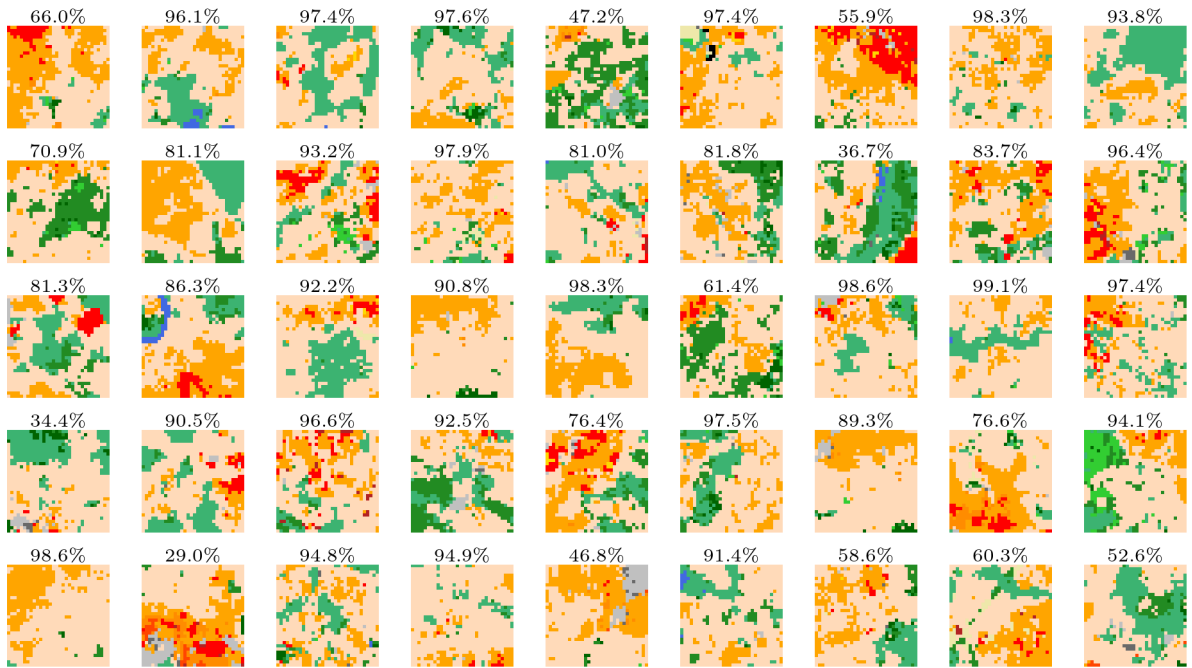


Figure 64: Exemplary cluster c_{11} with 2596 patches in total. A random selection and their confidence values is shown.

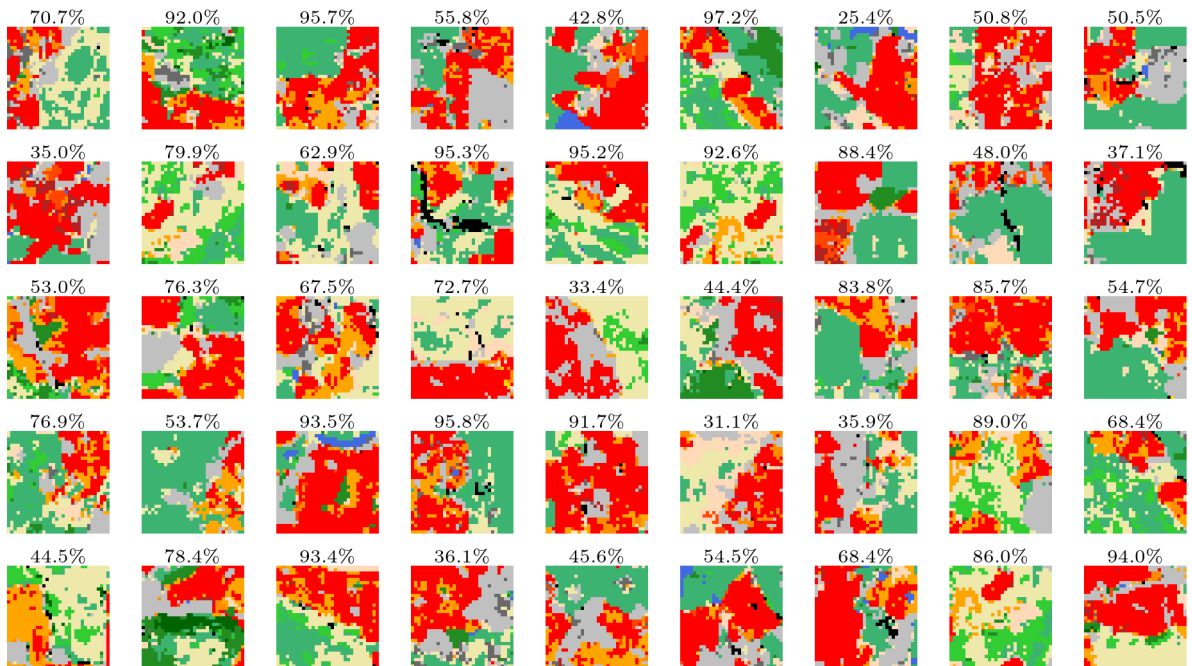
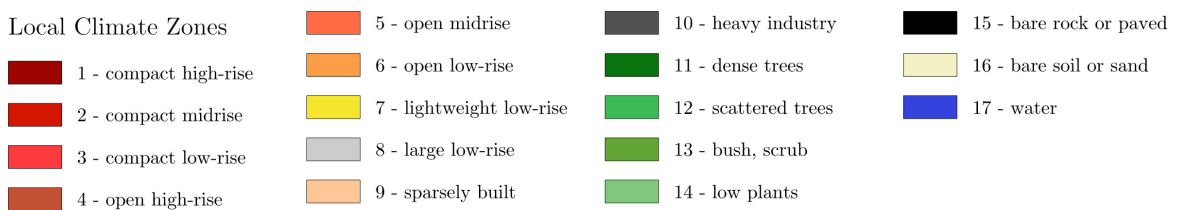


Figure 65: Exemplary cluster c_{12} with 1741 patches in total. A random selection and their confidence values is shown.



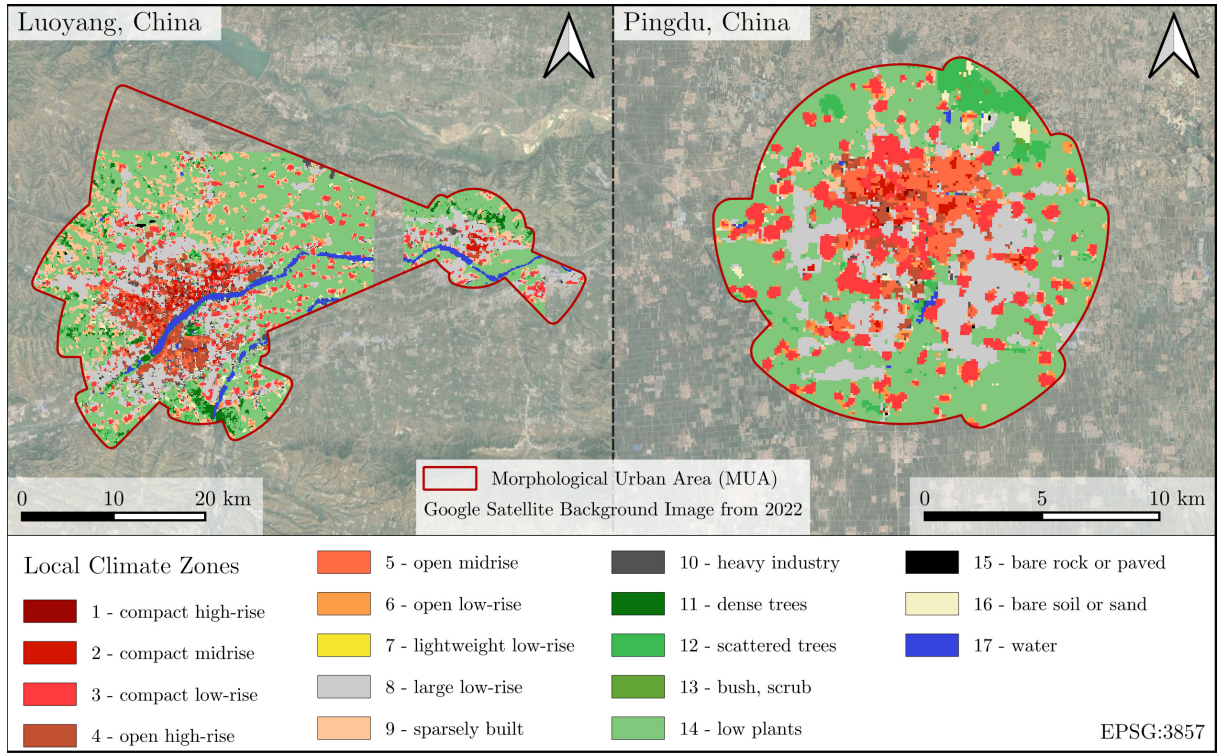


Figure 66: Two exemplary cities assigned to cluster $C_{HM_{1523}}(1)$.

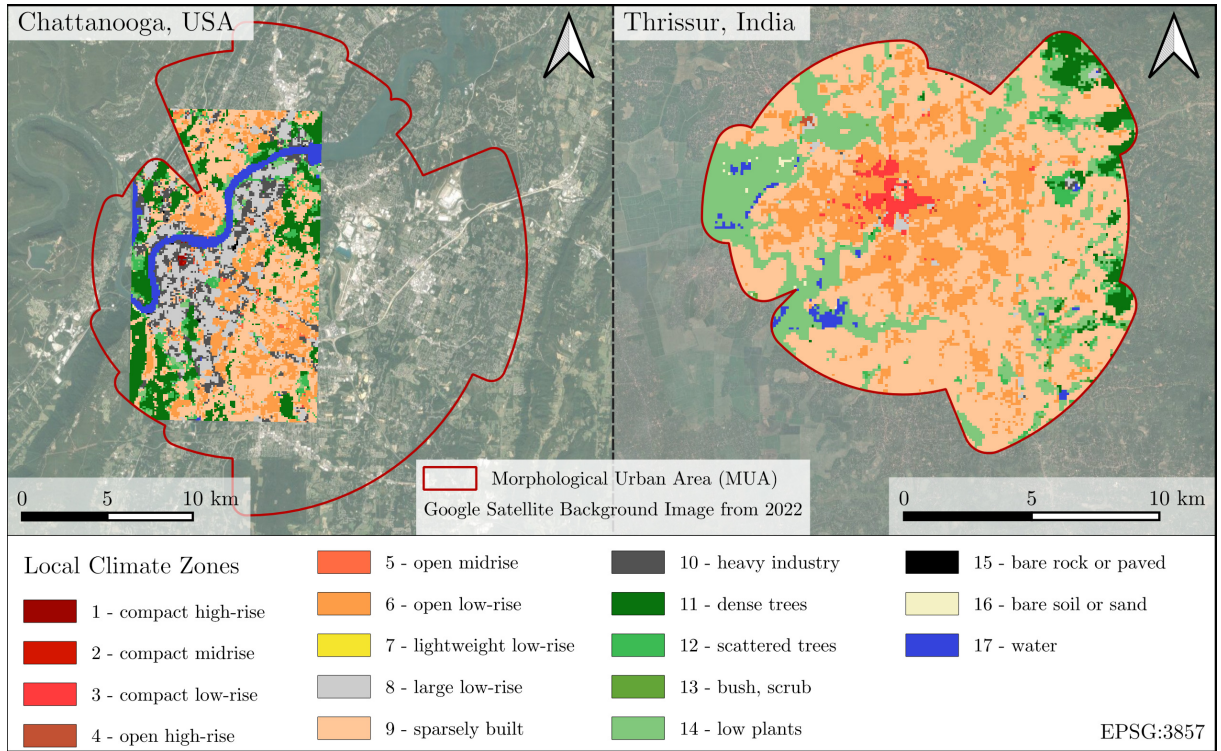


Figure 67: Two exemplary cities assigned to cluster $C_{HM_{1523}}(2)$.

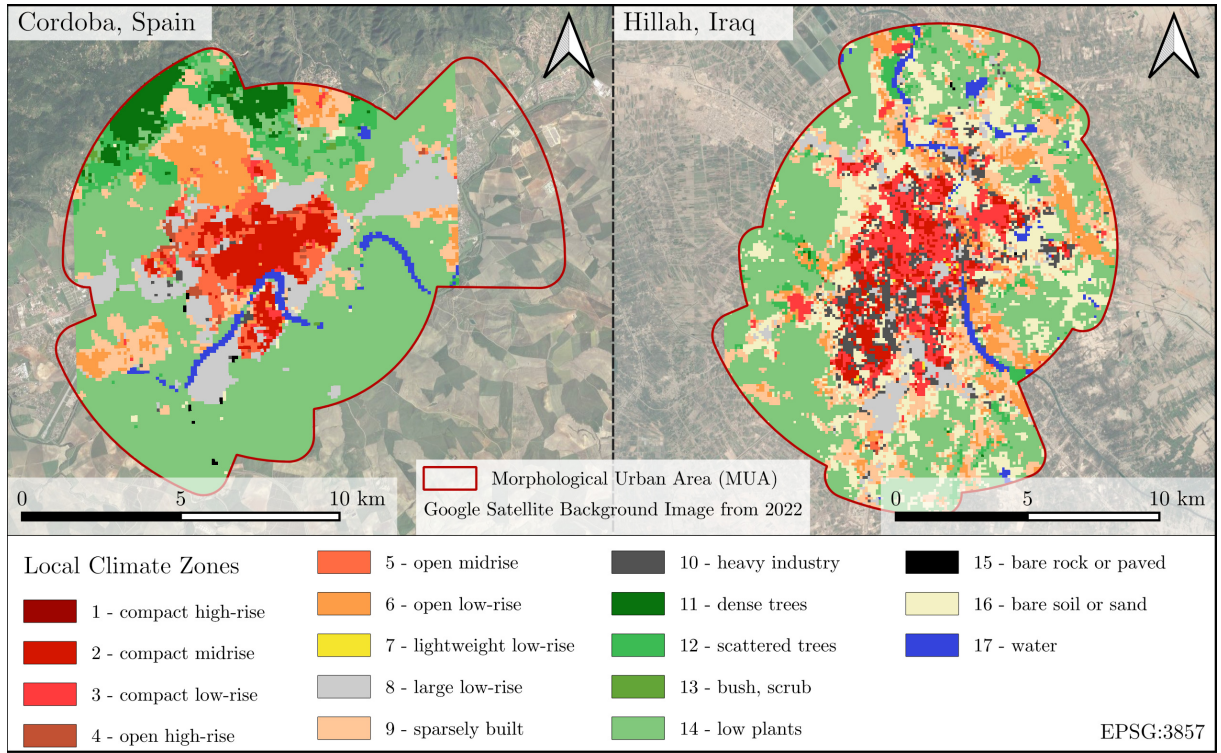


Figure 68: Two exemplary cities assigned to cluster $C_{HM_{1523}}(3)$.

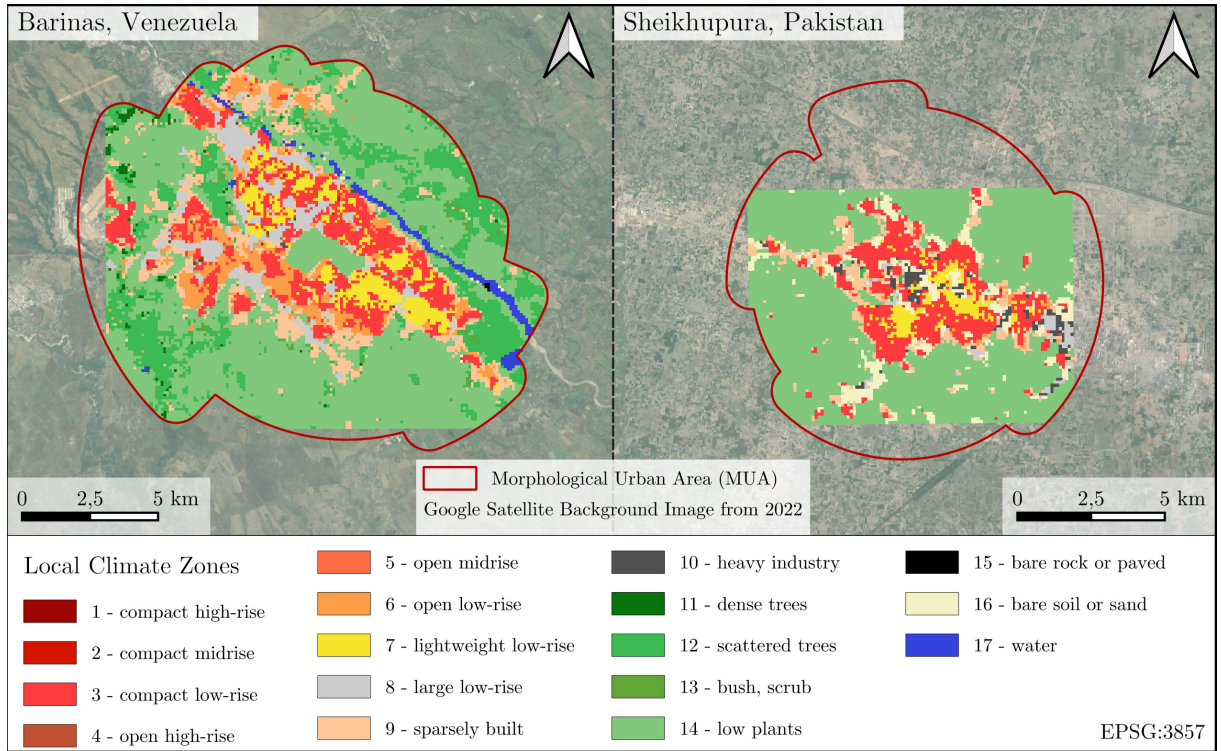


Figure 69: Two exemplary cities assigned to cluster $C_{HM_{1523}}(4)$.

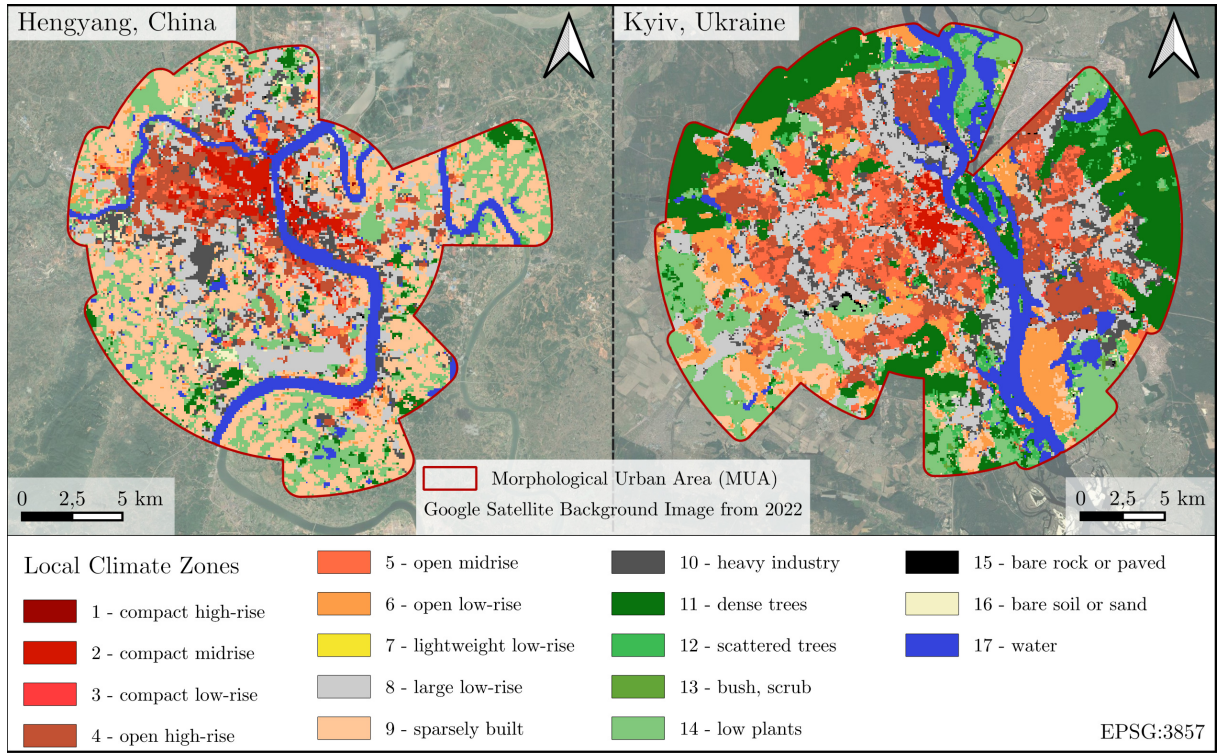


Figure 70: Two exemplary cities assigned to cluster $C_{HM_{1523}}(5)$.

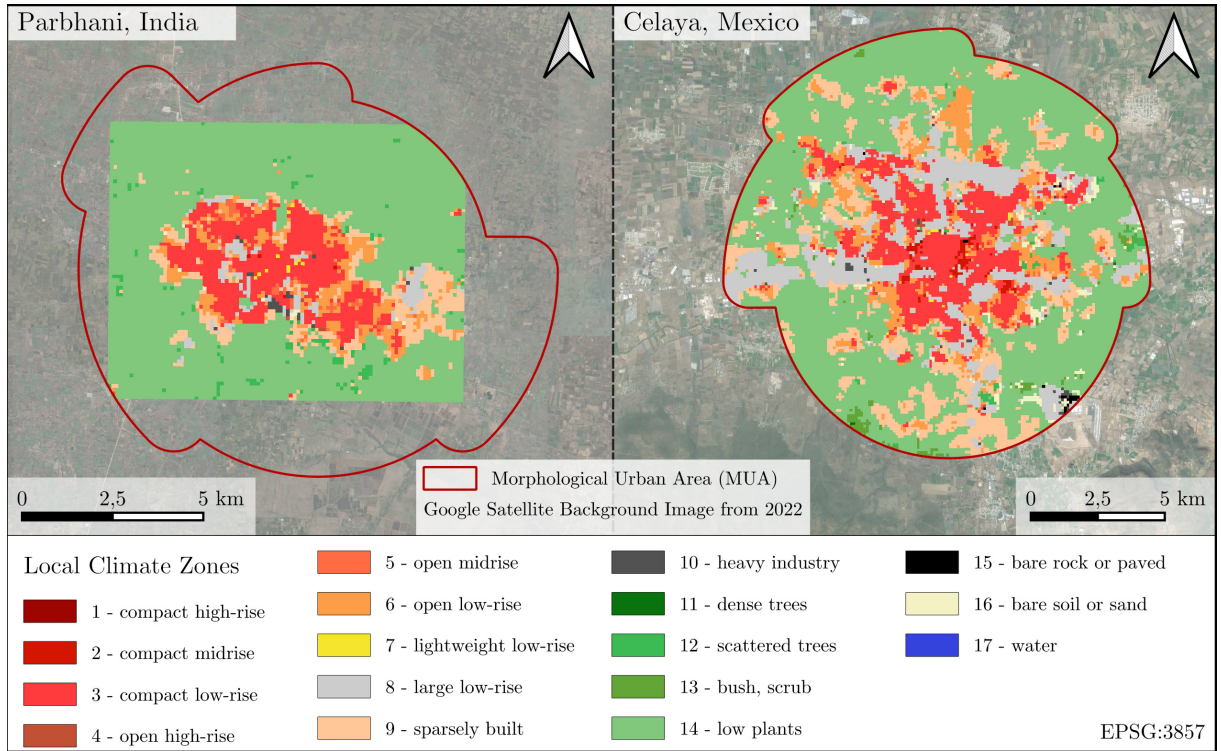


Figure 71: Two exemplary cities assigned to cluster $C_{HM_{1523}}(6)$.

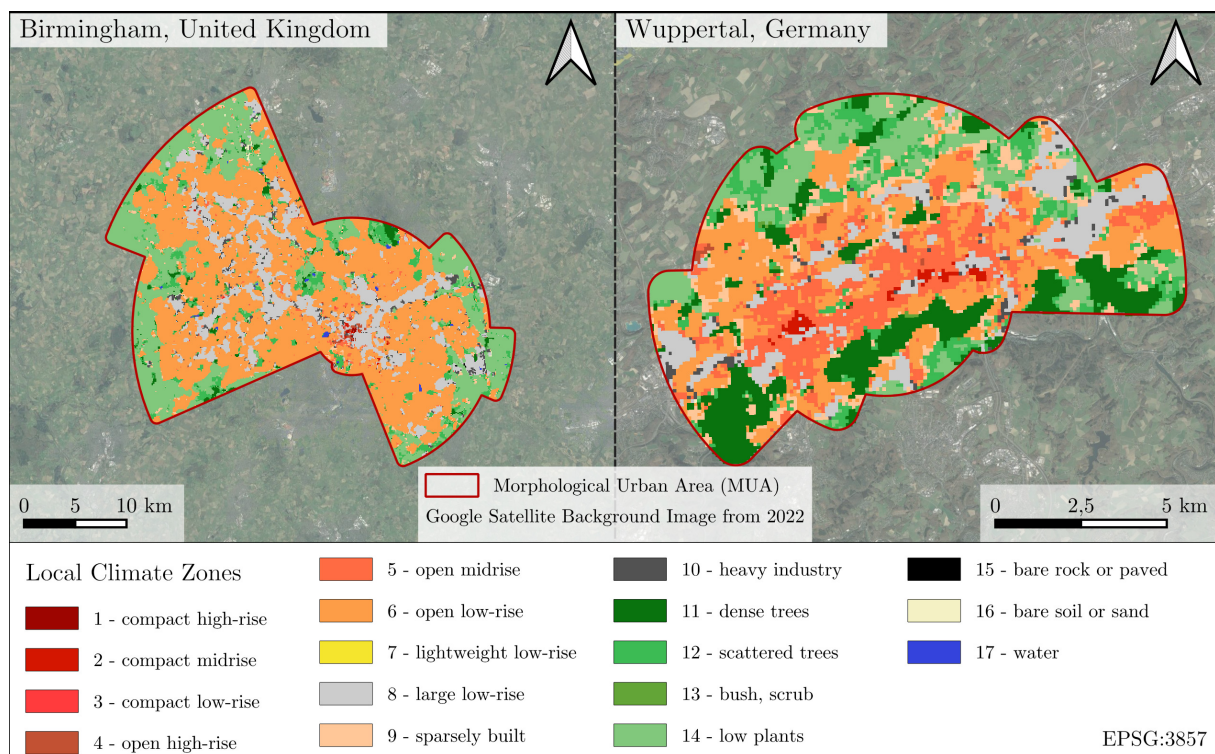


Figure 72: Two exemplary cities assigned to cluster $C_{HM_{1523}}(7)$.

Table 7: Overview of MUAs with their ID, the city names inside the MUAs, the MUA area, the LCZ area within the MUA, the percental coverage of the MUA with the LCZs, the number of extracted patches and the cluster results $C_{HM_{1523}}$ and the cluster results of the ablation studies $C_{HM_{110}}$, $C_{F_{1523}^*}$ and $C_{F_{1523}}$.

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}^*}$	$C_{F_{1523}}$
1	Winnipeg	Canada	553.61	543.35	98.15	364	6	-	1	1
2	Calabar	Nigeria	207.98	158.57	76.24	83	6	-	5	1
3	Anqiu	China	195.30	195.1	99.90	93	1	-	1	5
4	Maracay	Venezuela (Bolivarian Republic of)	449.46	278.24	61.91	140	4	-	1	10
5	Tel Aviv-Yafo (Tel Aviv-Jaffa)	Israel	920.68	269.58	29.28	158	7	-	1	14
6	Sanmenxia	China	162.47	107.52	66.18	49	5	-	1	26
7	Abidjan	Côte d'Ivoire	599.05	599.06	100.00	371	4	4	6	3
8	Abomey-Calavi Cotonou	Benin	640.77	296.61	46.29	136	4	-	6	21
9	Abuja	Nigeria	645.46	361.89	56.07	208	6	3	4	27
10	Abu Zaby (Abu Dhabi)	United Arab Emirates	930.94	244.26	26.24	122	5	-	3	29
11	Acapulco de Juárez	Mexico	262.42	151.86	57.87	52	4	-	6	19
12	Accra	Ghana	908.88	700.1	77.03	511	6	-	6	19
13	Adana	Turkey	337.80	337.7	99.97	175	1	-	4	12
14	Ad-Dammam	Saudi Arabia	776.17	663.49	85.48	390	3	-	3	29
15	Ad-Dawhah (Doha) Ad-Dawhah (Doha) Ar-Rayyan	Qatar	924.52	924.54	100.00	612	3	-	3	29
16	Addis Ababa	Ethiopia	566.74	566.41	99.94	376	4	4	6	4
17	Adelaide	Australia	1,008.29	1007.9	99.96	713	2	-	6	19
18	Ado-Ekiti	Nigeria	180.57	177.98	98.57	85	2	-	2	25
19	Agadir	Morocco	259.59	79.66	30.69	27	3	-	3	8
20	Agartala	India	133.64	133.47	99.88	58	7	-	7	1
21	Aguadilla-Isabela-San Sebastian	Puerto Rico	264.48	5.97	2.26	0	-	-	-	-
22	Aguascalientes	Mexico	314.38	314.12	99.92	179	4	-	6	19

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
23	Ahmadnagar	India	163.96	102.18	62.32	41	6	-	4	15
24	Ahvaz	Iran (Islamic Republic of)	445.38	444.43	99.79	293	3	-	1	26
25	Aizawl	India	95.97	12.48	13.00	0	-	-	-	-
26	Dubayy (Dubai) Sharjah Ajman	United Arab Emirates	1,291.71	782.53	60.58	556	3	-	3	29
27	Ajmer	India	186.23	183.2	98.37	89	6	-	2	7
28	Akesu	China	241.98	205.97	85.12	87	1	-	1	26
29	Cleveland Akron	United States of America	1,583.96	1242.36	78.43	828	2	-	7	6
30	Aktyubinsk	Kazakhstan	226.11	75.36	33.33	22	1	-	1	2
31	Al-Ain	United Arab Emirates	528.61	511.83	96.83	303	3	-	3	8
32	Albany	United States of America	352.93	194.81	55.20	99	2	-	7	6
33	Al-Basrah (Basra)	Iraq	279.66	276.66	98.93	148	3	-	3	29
34	Albuquerque	United States of America	856.55	856.47	99.99	604	6	-	6	19
35	Al-Fayyum	Egypt	84.05	61.24	72.86	16	3	-	4	17
36	Al Gadarif	Sudan	126.88	88.05	69.39	40	6	-	2	18
37	Al-Hasakah	Syrian Arab Republic	119.44	115.59	96.78	43	3	-	3	22
38	Al-Hudaydah	Yemen	138.97	120.32	86.58	52	3	-	3	8
39	Aligarh	India	158.69	158.49	99.88	76	4	-	4	15
40	Al-Iskandariyah (Alexandria) Kafr-ad-Dawwar	Egypt	773.55	205.25	26.53	62	3	-	1	5
41	Al-Ismailiyah	Egypt	158.40	122.13	77.10	55	3	-	3	8
42	Al-Kamishli	Syrian Arab Republic	128.16	123.08	96.04	54	6	-	4	17
43	Al-Khartum (Khartoum)	Sudan	1,213.55	1213.51	100.00	849	4	6	2	18
44	Allahabad	India	287.01	256.79	89.47	124	4	-	6	27

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}^*}$	$C_{F_{1523}}$
45	Allentown-Bethlehem	United States of America	117.53	117.51	99.98	36	2	-	7	6
46	Al-Madinah (Medina)	Saudi Arabia	533.55	531.73	99.66	337	3	-	3	29
47	Al-Manamah (Manama)	Bahrain	638.48	484.56	75.89	282	3	-	3	29
48	Almaty	Kazakhstan	597.27	597.07	99.97	394	6	-	6	19
49	Al-Mawsil (Mosul)	Iraq	373.16	353.33	94.68	223	3	-	3	28
50	Al-Mukalla	Yemen	162.08	34.01	20.98	8	3	-	3	22
51	Al Obeid (Al Ubayyid)	Sudan	237.17	236.25	99.61	134	4	-	2	18
52	Al-Qahirah (Cairo)	Egypt	1,388.72	1373.3	98.89	1043	3	3	3	28
53	Al-Raqqa	Syrian Arab Republic	108.98	107.19	98.35	44	3	-	4	27
54	Alwar	India	198.75	35.19	17.71	8	6	-	6	21
55	Amara	Iraq	170.56	155.34	91.07	80	3	-	3	8
56	Amman Zarqa Ar-Rusayfah	Jordan	978.44	976.16	99.77	684	3	-	3	8
57	Amritsar	India	239.04	238.26	99.67	132	4	-	4	15
58	Amsterdam	Netherlands	517.53	506.32	97.83	318	7	7	1	1
59	Anantapur	India	125.94	99.58	79.07	47	6	-	4	17
60	Anápolis	Brazil	254.81	254.34	99.81	131	6	-	6	7
61	Andizhan	Uzbekistan	250.10	249.84	99.90	130	6	-	6	10
62	Ankang	China	94.15	88.47	93.97	31	5	-	5	11
63	Ankara	Turkey	707.47	560.53	79.23	376	3	-	1	27
64	Ann Arbor	United States of America	143.06	142.95	99.92	42	2	-	7	6
65	Binzhou	China	381.44	275.77	72.30	157	1	-	1	12
66	Seoul Incheon Suweon Seongnam Goyang Bucheon Ansan Anyang Uijeongbu Siheung Gwangmyeong	Republic of Korea	3,188.16	3188.28	100.00	2525	5	1	5	16
67	Anshan Liaoyang	China	723.94	723.99	100.01	483	1	-	1	5
68	Anshun	China	103.39	83.82	81.07	34	5	-	5	11

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
69	Antalya	Turkey	304.70	288.22	94.59	133	7	-	7	14
70	Antananarivo	Madagascar	397.61	352.69	88.70	220	4	4	2	7
71	Antioch	United States of America	353.64	214.34	60.61	107	6	-	6	10
72	Anyang	China	294.41	293.71	99.76	172	1	-	1	5
73	Aracaju	Brazil	276.99	208.36	75.22	114	6	-	6	19
74	Ardabil	Iran (Islamic Republic of)	149.58	148.88	99.53	72	6	-	4	10
75	Arequipa	Peru	315.07	302.26	95.93	163	4	-	6	4
76	Arkhangelsk	Russian Federation	99.91	81.4	81.47	25	5	-	1	2
77	Armenia	Colombia	139.09	113.56	81.65	51	4	-	1	9
78	Ar-Riyadh (Riyadh)	Saudi Arabia	1,787.00	1786.95	100.00	1439	3	3	3	29
79	Arusha	United Republic of Tanzania	207.64	205.91	99.17	107	2	-	2	25
80	Asahikawa	Japan	226.90	226.85	99.98	123	6	-	1	10
81	Asheville	United States of America	207.25	21.73	10.49	2	2	-	7	23
82	Ashgabat	Turkmenistan	299.22	284.97	95.24	152	3	-	1	26
83	Asmara	Eritrea	144.82	137.76	95.13	69	4	4	3	28
84	As-Suways	Egypt	197.11	94.73	48.06	27	3	-	3	8
85	Astana	Kazakhstan	264.64	246.22	93.04	138	1	-	1	26
86	Astrakhan	Russian Federation	360.64	355.8	98.66	212	5	-	4	7
87	Aswan	Egypt	131.02	105.12	80.23	44	3	-	3	22
88	Asyut	Egypt	108.40	89.77	82.81	36	3	-	4	17
89	Atlanta	United States of America	487.58	284.55	58.36	142	2	-	2	23
90	Auckland	New Zealand	910.88	910.83	99.99	515	2	-	7	6
91	Augusta-Richmond County	United States of America	522.30	110.19	21.10	47	2	-	5	13

Continued on next page

ID	City Name(s)	Country	MUA Area (<i>km</i> ²)	LCZs Area (<i>km</i> ²)	LCZs Coverage (%)	No. Patches	<i>C</i> _{HM1523}	<i>C</i> _{HM110}	<i>C</i> _{<i>F</i>₁₅₂₃*7}	<i>C</i> _{<i>F</i>₁₅₂₃}
92	Aurangabad	India	270.05	220.6	81.69	113	6	-	4	15
93	Austin	United States of America	954.83	533.87	55.91	299	7	-	7	6
94	Az-Zaqazig	Egypt	104.46	86.4	82.71	35	3	-	4	17
95	Baaqoobah	Iraq	171.32	107.08	62.50	52	3	-	1	27
96	Baghdad	Iraq	974.49	974.33	99.98	722	4	3	6	21
97	Bago	Myanmar	119.88	106.94	89.20	46	2	-	4	17
98	Baguio City	Philippines	137.42	137.2	99.84	60	2	-	2	13
99	Bahampur	India	109.39	36.96	33.79	5	2	-	7	6
100	Bahawalpur	Pakistan	154.52	153.91	99.61	73	6	-	4	15
101	Baicheng	China	99.33	99.22	99.89	36	1	-	1	26
102	Baishan	China	95.56	91.61	95.86	34	5	-	5	16
103	Baixada Santista	Brazil	296.80	296.62	99.94	137	5	-	5	3
104	Baiyin	China	166.92	166.45	99.72	78	3	-	3	22
105	Bakersfield	United States of America	654.83	654.4	99.93	458	6	-	6	10
106	Baku Sumquayit	Azerbaijan	1,333.27	1030.12	77.26	676	6	-	2	18
107	Washington, D.C. Baltimore	United States of America	2,289.62	2289.2	99.98	1626	2	-	7	6
108	Bamako	Mali	497.40	492.47	99.01	325	6	-	6	15
109	Bamenda	Cameroon	160.02	159.95	99.96	77	2	-	2	25
110	Bandar Abbas	Iran (Islamic Republic of)	404.38	116.57	28.83	45	4	-	3	8
111	Bandung	Indonesia	534.07	533.83	99.95	353	6	-	6	21
112	Bangalore	India	863.66	863.67	100.00	659	7	3	6	19
113	Banghazi	Libya	361.92	162.01	44.76	73	3	-	7	14
114	Bangui	Central African Republic	217.09	209.75	96.62	103	2	-	7	6
115	Banjarmasin	Indonesia	200.96	200.86	99.95	105	4	-	4	15
116	Banjul	Gambia	248.88	189.54	76.16	89	2	-	2	18

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
117	Baoding	China	704.72	699.9	99.32	399	1	-	4	12
118	Baoji	China	315.25	10.75	3.41	0	-	-	-	-
119	Baofou	China	635.38	550.21	86.60	373	1	-	3	8
120	Barcelona	Spain	379.21	379.55	100.09	192	3	-	5	2
121	Bareilly	India	186.06	180.52	97.02	79	4	-	4	15
122	Bari	Italy	223.24	128.27	57.46	59	3	-	1	2
123	Barinas	Venezuela (Bolivarian Republic of)	186.67	178.04	95.38	96	4	-	4	15
124	Barisal	Bangladesh	111.51	93.21	83.59	46	2	-	2	25
125	Barletta	Italy	70.51	18.95	26.88	1	1	-	1	5
126	Barnaul	Russian Federation	236.16	4.16	1.76	0	-	-	-	-
127	Basel	Switzerland	217.46	214.2	98.50	106	7	-	7	1
128	Bathinda	India	142.19	134.51	94.60	60	4	-	4	15
129	Batman	Turkey	107.97	96.18	89.08	43	7	-	4	17
130	Batna	Algeria	117.69	115.68	98.30	50	3	-	3	28
131	Baton Rouge	United States of America	528.79	474.58	89.75	307	2	-	7	6
132	Bauchi	Nigeria	176.30	169.9	96.37	73	4	-	4	7
133	Bauru	Brazil	228.70	228.5	99.91	126	6	-	6	21
134	Bayrut (Beirut)	Lebanon	274.88	17.05	6.20	0	-	-	-	-
135	Be'er Sheva	Israel	171.69	131.26	76.45	56	3	-	3	22
136	Begusarai	India	199.61	68.29	34.21	25	4	-	2	18
137	Beijing Sanhe	China	2,667.88	2667.44	99.98	2075	1	1	1	24
138	Belém	Brazil	477.89	153.15	32.05	64	4	-	5	3
139	Belfast	United Kingdom	327.46	211.08	64.46	89	7	-	7	20
140	Belgorod	Russian Federation	183.93	4.05	2.20	0	-	-	-	-
141	Bellary	India	136.15	38.76	28.47	10	6	-	4	15
142	Bengbu	China	209.15	208.94	99.90	107	5	-	1	2
143	Bengkulu	Indonesia	149.34	81.83	54.80	22	4	-	2	18

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM1523}	C_{HM110}	C_{F1523}^*	C_{F1523}
144	Benin City	Nigeria	510.85	510.52	99.94	337	2	-	2	18
145	Benxi	China	153.31	153.28	99.98	74	5	-	5	16
146	Beograd (Belgrade)	Serbia	270.31	267.26	98.87	128	7	-	7	1
147	Berbera	Somalia	70.45	6.76	9.60	0	-	-	-	-
148	Milano (Milan) Bergamo Busto Arsizio Seregno Como	Italy	2,156.51	2156.28	99.99	1572	7	7	1	1
149	Berlin	Germany	699.60	699.48	99.98	491	7	7	7	14
150	Bern	Switzerland	146.13	146.19	100.04	53	7	-	7	1
151	Bhagalpur	India	124.85	84.89	67.99	34	4	-	4	15
152	Bhilwara	India	141.14	121.54	86.11	53	6	-	4	15
153	Mumbai (Bombay) Bhiwandi	India	1,310.37	849.28	64.81	451	5	5	5	11
154	Bhopal	India	359.46	338.16	94.08	190	4	-	4	15
155	Bhubaneswar	India	456.74	279.11	61.11	119	4	-	4	7
156	Bielefeld	Germany	112.77	112.73	99.97	36	7	-	7	20
157	Bijapur	India	146.20	93.83	64.18	41	6	-	4	17
158	Bijie	China	105.82	87.97	83.14	38	5	-	5	11
159	Bikaner	India	222.52	211.86	95.21	123	3	-	4	27
160	Bilaspur	India	137.42	118.18	86.00	55	4	-	4	7
161	Bilbao	Spain	261.67	47.33	18.09	17	5	-	5	11
162	Liverpool Birkenhead	United Kingdom	588.15	552.47	93.93	275	7	-	7	20
163	Birmingham (West Midlands)	United Kingdom	751.71	751.57	99.98	499	7	-	7	20
164	Bishkek	Kyrgyzstan	374.90	374.56	99.91	237	6	-	6	10
165	Bissau	Guinea-Bissau	142.74	130.24	91.24	54	6	-	4	15
166	Blantyre-Limbe	Malawi	380.29	335.73	88.28	205	4	6	2	25
167	Bloemfontein	South Africa	389.15	389.03	99.97	250	6	-	4	27
168	Boa Vista	Brazil	250.07	102.33	40.92	43	2	-	7	20
169	Bobo-Dioulasso	Burkina Faso	264.22	263.95	99.90	147	6	-	4	15
170	Jakarta Bogor	Indonesia	3,206.64	3206.76	100.00	2642	6	6	6	21
171	Bogra	Bangladesh	118.09	105.26	89.14	41	4	-	2	25

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
172	Boise City	United States of America	521.20	313.26	60.10	173	7	-	7	20
173	Cape Coral Bonita Springs-Naples	United States of America	2,550.43	601.83	23.60	385	2	-	7	6
174	Bonn	Germany	122.11	122.07	99.97	29	7	-	7	14
175	Bordeaux	France	364.52	364.38	99.96	214	7	-	7	6
176	Boston	United States of America	682.39	682.35	99.99	368	2	-	7	6
177	Botou	China	113.04	111.76	98.87	47	1	-	4	12
178	Bouake	Côte d'Ivoire	206.01	202.4	98.25	110	4	-	2	18
179	Bozhou	China	163.06	161.71	99.17	77	1	-	1	2
180	Brahmapur	India	118.51	92.69	78.21	39	4	-	4	17
181	Brasília	Brazil	892.91	126.68	14.19	57	5	-	1	2
182	Bratislava	Slovakia	197.07	176.41	89.52	64	7	-	1	2
183	Kinshasa Brazzaville	Democratic Republic of the Congo	927.47	927.44	100.00	631	6	2	6	15
184	Bremen	Germany	309.93	223.72	72.19	103	7	-	7	1
185	Brescia	Italy	328.89	238.34	72.47	101	7	-	1	5
186	Brest	Belarus	181.13	169.81	93.75	74	7	-	7	14
187	Brighton-Worthing- Littlehampton	United Kingdom	109.47	80.09	73.16	28	7	-	7	20
188	Brisbane	Australia	1,325.55	1325.58	100.00	953	2	-	7	6
189	Bristol	United Kingdom	310.53	310.24	99.91	178	7	-	7	20
190	Brno	Czech Republic	198.95	199.04	100.05	86	7	-	1	5
191	Bryansk	Russian Federation	294.54	270.65	91.89	139	5	-	5	14
192	Bucaramanga	Colombia	167.50	103.48	61.78	46	4	-	5	3
193	Bucuresti (Bucharest)	Romania	487.73	487.26	99.90	324	1	3	1	5
194	Budapest	Hungary	590.46	590.42	99.99	390	7	7	7	1

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM123}	C_{HM110}	C_{F1523}^*	C_{F1523}
195	Buffalo	United States of America	738.53	736.17	99.68	519	2	-	7	6
196	Bujumbura Uvira	Burundi	444.07	270.34	60.88	130	4	6	2	15
197	Bukavu	Democratic Republic of the Congo	127.18	126.31	99.31	33	4	-	2	18
198	Bulawayo	Zimbabwe	366.72	351.93	95.97	204	2	-	2	25
199	Bunia	Democratic Republic of the Congo	122.21	59.85	48.97	16	2	-	2	25
200	Buraydah	Saudi Arabia	357.20	235.13	65.83	130	3	-	3	29
201	Bur Sa'id	Egypt	117.71	99.43	84.47	27	5	-	3	8
202	Bydgoszcz	Poland	141.12	120.44	85.35	44	5	-	5	16
203	Cagliari	Italy	198.77	67.54	33.98	15	3	-	1	2
204	Calgary	Canada	875.14	875.11	100.00	637	1	-	6	10
205	Camaguey	Cuba	135.42	132.06	97.52	59	2	-	2	7
206	Campina Grande	Brazil	193.97	174.43	89.93	94	6	-	4	15
207	Campinas	Brazil	507.66	507.68	100.00	278	6	-	6	10
208	Campo Grande	Brazil	486.96	486.64	99.93	318	6	-	6	19
209	Campos dos Goytacazes	Brazil	198.62	190.94	96.14	85	6	-	4	17
210	Canberra	Australia	698.43	107.42	15.38	50	2	-	7	6
211	Cancún	Mexico	254.55	241.97	95.06	107	4	-	5	3
212	Can Tho	Viet Nam	139.28	139.26	99.99	58	4	-	6	19
213	Caracas	Venezuela (Bolivarian Republic of)	462.23	438.22	94.81	283	4	-	5	3
214	Caruaru	Brazil	169.34	139.54	82.40	66	6	-	4	15
215	Napoli (Naples) Caserta Nola	Italy	1,447.25	1182.61	81.71	785	6	-	2	7
216	Catania	Italy	465.85	432.95	92.94	260	7	-	2	18
217	Caxias Do Sul	Brazil	201.70	201.56	99.93	99	5	-	5	3
218	Cenxi	China	97.04	81.02	83.49	31	5	-	5	16
219	Chandrapur	India	154.13	96.27	62.46	44	6	-	4	17

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
220	Changchun	China	573.91	573.9	100.00	380	1	-	1	26
221	Changde	China	192.17	191.94	99.88	102	1	-	1	5
222	Changji	China	128.06	127.7	99.72	53	1	-	1	5
223	Changning	China	93.83	67.42	71.85	27	5	-	2	18
224	Changshu	China	433.16	433.21	100.01	240	1	-	1	5
225	Changyi	China	106.91	106.75	99.85	44	1	-	4	12
226	Changzhi	China	487.53	270.77	55.54	143	1	-	1	12
227	Chaohu	China	127.93	127.89	99.97	50	5	-	5	11
228	Chaoyang	China	139.60	120.85	86.57	42	1	-	1	26
229	Charleroi	Belgium	189.46	189.25	99.89	83	7	-	7	20
230	Charleston-North Charleston	United States of America	782.31	33.15	4.24	8	2	-	7	6
231	Charlotte	United States of America	366.24	192.84	52.65	83	2	-	7	6
232	Chattanooga	United States of America	498.49	188.4	37.79	104	2	-	2	13
233	Cheboksary	Russian Federation	155.31	107.5	69.22	39	5	-	5	23
234	Chelyabinsk	Russian Federation	377.29	377.18	99.97	180	5	-	1	2
235	Chennai (Madras)	India	910.92	910.61	99.97	665	6	2	6	19
236	Chenzhou	China	171.96	159.04	92.49	72	5	-	5	16
237	Cheonan	Republic of Korea	144.51	144.41	99.93	61	5	-	5	11
238	Cherepovets	Russian Federation	154.15	100.85	65.42	36	5	-	1	2
239	Cherthala	India	109.83	13.52	12.31	0	-	-	-	-
240	Chicago Round Lake Beach-McHenry-Grayslake	United States of America	3,591.29	3332.59	92.80	2687	2	-	7	24
241	Chifeng	China	251.73	241.23	95.83	116	1	-	1	5
242	Chihuahua	Mexico	415.22	397.43	95.71	230	3	-	1	26
243	Chimbote	Peru	169.23	87.09	51.46	21	4	-	1	4
244	Chita	Russian Federation	197.46	144.65	73.25	49	5	-	1	5

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}^Z$
245	Chittagong	Bangladesh	248.95	248.83	99.95	135	5	5	1	23
246	Chitungwiza	Zimbabwe	162.97	139.79	85.77	65	2	-	2	9
247	Chongqing	China	913.12	902.44	98.83	648	5	1	5	11
248	Christchurch	New Zealand	336.58	336.43	99.96	204	6	-	6	19
249	Chukyo M.M.A. (Nagoya)	Japan	2,926.69	2926.56	100.00	2321	5	-	6	24
250	Chuxiong	China	143.58	142.5	99.24	52	5	-	5	13
251	Chuzhou	China	205.90	203.9	99.03	96	5	-	1	2
252	Cincinnati	United States of America	655.72	284.75	43.43	154	2	-	7	6
253	Cirebon	Indonesia	337.98	157	46.45	63	6	-	6	21
254	Ciudad Bolivar	Venezuela (Bolivarian Republic of)	296.05	245.07	82.78	112	6	-	4	7
255	Ciudad de Guatemala (Guatemala City)	Guatemala	529.00	528.82	99.97	324	4	-	5	3
256	Ciudad del Este	Paraguay	419.57	418.89	99.84	248	2	-	2	18
257	Ciudad de México (Mexico City)	Mexico	2,530.12	2529.65	99.98	2058	4	4	6	19
258	Ciudad Guayana	Venezuela (Bolivarian Republic of)	535.04	238.37	44.55	101	4	-	6	15
259	Ciudad Obregón	Mexico	175.35	175.27	99.95	88	3	-	4	12
260	Ciudad Victoria	Mexico	191.73	191.56	99.91	101	6	-	4	10
261	Cluj-Napoca	Romania	121.39	121.25	99.88	51	7	-	1	1
262	Coatzacoalcos	Mexico	241.96	78.77	32.56	30	2	-	6	21
263	Cochabamba	Bolivia (Plurinational State of)	501.71	470.33	93.75	276	4	-	2	18
264	Coimbatore	India	451.79	434.15	96.10	275	7	-	2	7
265	Colima	Mexico	180.54	164.68	91.21	88	3	-	4	12
266	Colombo	Sri Lanka	725.92	77.22	10.64	22	6	-	6	21

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
267	Colorado Springs	United States of America	639.33	639.17	99.98	444	7	-	7	6
268	Columbia, South Carolina	United States of America	855.39	684.56	80.03	395	2	-	5	13
269	Columbus, Ohio	United States of America	877.07	876.62	99.95	532	2	-	7	6
270	Comilla	Bangladesh	147.27	24.44	16.60	2	5	-	2	7
271	Concepción	Chile	286.29	96.89	33.84	38	4	-	5	3
272	San Francisco-Oakland San Jose Concord	United States of America	3,067.36	2267.53	73.92	1469	6	-	6	19
273	Córdoba	Argentina	458.64	458.17	99.90	297	6	-	6	21
274	Corpus Christi	United States of America	523.37	383.98	73.37	201	2	-	7	1
275	Resistencia Corrientes	Argentina	350.55	319.26	91.08	148	6	-	6	21
276	Coventry-Bedworth	United Kingdom	218.92	218.76	99.93	86	7	-	7	20
277	Cuautla Morelos	Mexico	153.43	94.87	61.83	42	6	-	4	7
278	Cuenca	Ecuador	230.65	189.95	82.35	98	2	-	2	25
279	Cuiabá	Brazil	467.91	468.04	100.03	302	6	-	2	9
280	Culiacán	Mexico	303.59	303.43	99.95	179	3	-	1	5
281	Curitiba	Brazil	814.89	814.69	99.98	600	2	-	6	19
282	Cusco	Peru	147.45	75.48	51.19	30	6	-	6	21
283	Dafeng	China	157.31	148.48	94.39	70	1	-	1	5
284	Dakar	Senegal	365.45	148.51	40.64	51	3	-	1	8
285	Dali	China	166.99	118.61	71.03	39	5	-	5	16
286	Dalian	China	724.72	532.45	73.47	308	5	-	5	11
287	Dallas-Fort WorthDallas-Fort Worth Denton-Lewisville	United States of America	2,680.98	1519.89	56.69	819	2	-	7	19
288	Da Nang	Viet Nam	298.73	263.15	88.09	101	6	-	6	19
289	Dandong Sinuiju	China	249.82	168.19	67.32	81	5	-	1	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
290	Danyang	China	172.26	172.19	99.96	83	1	-	1	5
291	Daqing	China	548.17	92.26	16.83	38	3	-	1	26
292	Darbhangha	India	132.99	23.92	17.99	2	4	-	4	15
293	Dar-el-Beida (Casablanca)	Morocco	362.54	336.49	92.81	193	3	-	1	5
294	Dar es Salaam	United Republic of Tanzania	661.50	661.46	99.99	429	4	6	2	25
295	Datong	China	335.24	314.64	93.86	155	1	-	1	2
296	Davangere	India	130.44	77.79	59.64	34	4	-	4	17
297	Huangshi Daye	China	477.65	143.06	29.95	59	5	-	5	11
298	Dayton	United States of America	196.07	195.92	99.92	74	2	-	7	6
299	Daytona Beach-Port Orange	United States of America	434.89	266.35	61.25	153	2	-	7	6
300	Dazhou	China	120.63	22.95	19.03	4	5	-	5	11
301	Deir El-Zor(Deir ez-Zor)	Syrian Arab Republic	103.42	33.11	32.02	5	3	-	3	28
302	Delhi	India	1,326.41	1326.08	99.98	1000	3	4	1	2
303	Dengfeng	China	97.25	92.58	95.20	36	5	-	1	5
304	Dengzhou	China	122.71	103.59	84.42	48	1	-	1	5
305	Denizli	Turkey	165.55	164.65	99.46	71	7	-	7	1
306	Denver-Aurora	United States of America	1,897.60	1897	99.97	1520	2	-	7	6
307	Dera Ghazikhan	Pakistan	120.43	104.97	87.16	49	4	-	4	17
308	Des Moines	United States of America	480.55	333.49	69.40	174	2	-	7	6
309	Detroit Windsor	United States of America	3,356.52	3356.02	99.98	2837	2	-	7	6
310	Dezhou	China	252.57	252.45	99.95	122	1	-	1	26
311	Dhaka	Bangladesh	640.39	614	95.88	402	4	5	6	23
312	Dimashq (Damascus)	Syrian Arab Republic	499.18	499.01	99.97	315	3	-	3	28

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
313	Dindigul	India	118.31	58.84	49.73	19	6	-	4	17
314	Diwaniyah	Iraq	150.79	150.76	99.98	67	6	-	6	28
315	Diyarbakir	Turkey	127.89	102.59	80.22	36	3	-	1	5
316	Dnipropetrovsk	Ukraine	374.56	374.51	99.99	218	7	-	7	1
317	Donetsk Makeyevka	Ukraine	613.94	605.26	98.59	405	7	-	4	1
318	Dongtai	China	129.54	129.18	99.72	49	1	-	1	5
319	Yiwu Dongyang	China	546.52	540.5	98.90	353	5	-	5	11
320	Dortmund	Germany	110.50	110.38	99.89	28	7	-	7	14
321	Douala	Cameroon	372.33	372.22	99.97	191	4	2	6	19
322	Dresden	Germany	224.70	224.61	99.96	103	7	-	7	14
323	Dublin	Ireland	492.15	487.34	99.02	288	7	-	7	20
324	Durango	Mexico	253.96	235.29	92.65	132	4	-	4	10
325	Durgapur	India	314.34	54.67	17.39	16	1	-	4	1
326	Raleigh Durham	United States of America	1,128.03	293.19	25.99	144	2	-	2	13
327	Dushanbe	Tajikistan	389.98	389.62	99.91	227	6	-	4	7
328	East London	South Africa	204.47	86.25	42.18	30	2	-	7	6
329	Edmonton	Canada	679.00	672.04	98.97	420	2	-	1	10
330	Eindhoven	Netherlands	204.47	204.04	99.79	86	7	-	7	1
331	Elazig	Turkey	128.68	52.52	40.82	14	3	-	7	14
332	El Djelfa	Algeria	104.89	100.39	95.71	42	3	-	4	27
333	Eldoret	Kenya	168.12	100.91	60.02	43	2	-	2	25
334	El Tigre-San José de Guanipa	Venezuela (Bolivarian Republic of)	244.46	116.93	47.83	56	6	-	4	17
335	English Bazar	India	106.04	69.19	65.25	24	4	-	4	15
336	Enshi	China	115.34	77.8	67.45	25	5	-	5	11
337	Erduosi (Ordoss)	China	142.16	31.8	22.37	6	1	-	1	2
338	Erzurum	Turkey	111.40	72.72	65.28	23	3	-	1	5

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
339	Esfahan	Iran (Islamic Republic of)	823.60	571.37	69.38	382	3	-	1	8
340	Eskisehir	Turkey	211.25	184.74	87.45	86	3	-	4	27
341	Tehran/Tehran Karaj Eslamshahr Malard Qods of)	Iran (Islamic Republic of)	3,454.74	2478.17	71.73	1988	3	3	3	8
342	Ezhou	China	209.44	34.45	16.45	8	5	-	5	11
343	Faisalabad	Pakistan	338.27	338.18	99.97	210	4	3	6	28
344	Faloojah	Iraq	110.62	110.4	99.81	48	3	-	3	22
345	Farrukhabad	India	71.27	36	50.51	8	4	-	4	17
346	Fayetteville	United States of America	737.92	21.93	2.97	2	2	-	2	13
347	Fayetteville-Springdale	United States of America	453.57	167.29	36.88	80	2	-	2	23
348	Feicheng	China	107.50	107.38	99.89	40	1	-	1	12
349	Feira De Santana	Brazil	218.53	216.82	99.22	115	6	-	6	15
350	Firozabad	India	111.15	67.24	60.49	26	4	-	4	17
351	Flint	United States of America	303.00	286.02	94.40	158	2	-	2	18
352	Florence	Italy	466.91	199.65	42.76	81	7	-	2	23
353	Fortaleza	Brazil	608.97	608.73	99.96	414	6	2	6	21
354	Fort Wayne	United States of America	425.73	425.47	99.94	256	2	-	2	7
355	Franca	Brazil	237.08	236.92	99.93	127	6	-	4	15
356	Frankfurt am Main	Germany	468.35	367.72	78.51	175	7	-	7	14
357	Freetown	Sierra Leone	350.58	141.55	40.38	48	2	-	6	19
358	Fresno	United States of America	686.03	685.85	99.97	488	6	-	6	10
359	Fuan	China	101.49	53.96	53.17	16	5	-	5	16
360	Fuqing	China	1,421.18	239.51	16.85	105	5	-	5	3

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM123}	C_{HM110}	C_{F1523}^*	C_{F1523}
361	Fushun, Liaoning	China	347.08	298.13	85.90	117	5	-	1	2
362	Fuxin	China	268.96	235.09	87.41	111	1	-	1	26
363	Fuyang	China	226.04	225.76	99.87	115	1	-	1	26
364	Fuyang	China	210.43	210.21	99.90	95	5	-	5	16
365	Fuzhou, Fujian	China	574.90	574.47	99.93	384	5	-	5	11
366	Fuzhou, Jiangxi	China	178.08	158.63	89.08	63	1	-	1	2
367	Găncă	Azerbaijan	187.28	180.6	96.43	80	6	-	4	15
368	Ganzhou Nankang	China	910.25	539.7	59.29	329	5	-	2	23
369	Gaocheng	China	94.35	94.28	99.92	31	1	-	4	12
370	Gaoyou	China	134.41	134.2	99.85	58	1	-	1	5
371	Gaozhou	China	108.49	90.96	83.84	38	5	-	1	23
372	Gaya	India	117.52	99.37	84.56	41	4	-	4	17
373	Gaza (incl. Ash Shati Camp)	State of Palestine	221.66	155.24	70.04	67	6	-	2	18
374	Gboko	Nigeria	139.30	138.29	99.28	65	6	-	4	17
375	Gda?sk	Poland	271.46	209.2	77.06	85	7	-	1	1
376	Istanbul Gebze	Turkey	1,442.36	994.44	68.95	554	5	3	5	2
377	General Santos City	Philippines	173.35	121.99	70.37	41	6	2	2	18
378	Genève	Switzerland	127.41	120.24	94.37	23	7	-	7	14
379	Genova (Genoa)	Italy	111.53	54.7	49.04	14	5	-	5	11
380	Glasgow	United Kingdom	531.83	477.76	89.83	301	7	-	7	20
381	Goiânia	Brazil	862.06	861.84	99.97	630	2	-	6	19
382	Gold Coast	Australia	431.24	141.76	32.87	66	2	-	7	6
383	Gombe	Nigeria	168.46	167.43	99.39	85	6	-	4	17
384	Gomel	Belarus	192.89	192.81	99.96	84	7	-	1	1
385	Gongyi	China	140.67	133.01	94.56	59	1	-	1	2
386	Gongzhuling	China	93.65	93.5	99.84	32	1	-	1	12
387	Gorakhpur	India	203.76	200.84	98.57	91	4	-	2	18
388	Gorgan	Iran (Islamic Republic of)	116.59	116.3	99.75	46	3	-	4	12

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
389	Göteborg	Sweden	196.94	196.78	99.92	53	7	-	7	14
390	Grand Rapids	United States of America	501.74	501.44	99.94	313	2	-	7	6
391	Greensboro	United States of America	321.94	180.31	56.01	54	2	-	7	6
392	Greenville	United States of America	553.17	118.84	21.48	50	2	-	2	25
393	Grenoble	France	140.02	140.03	100.01	63	7	-	5	14
394	Grodno	Belarus	151.69	142.24	93.77	53	7	-	1	1
395	Guadalajara	Mexico	904.26	904.13	99.99	678	6	-	6	19
396	Guang'an	China	128.27	67.02	52.25	25	5	-	5	11
397	Guangyuan	China	165.40	104.5	63.18	45	5	-	5	11
398	Guayaquil	Ecuador	581.93	581.38	99.91	358	4	-	1	4
399	Guigang	China	169.34	169.34	100.00	78	5	-	1	23
400	Guilin	China	337.38	296.19	87.79	169	5	-	1	2
401	Guiping	China	117.32	36.66	31.25	4	5	-	5	11
402	Guiyang	China	404.10	346.48	85.74	167	5	-	5	11
403	Gujranwala	Pakistan	215.15	215.12	99.99	113	4	-	4	15
404	Gujrat	Pakistan	119.33	113.84	95.40	46	4	-	4	17
405	Gulbarga	India	164.66	153.97	93.51	77	6	-	4	17
406	Guntur	India	138.33	97.61	70.57	50	6	-	4	15
407	Guwahati (Gauhati)	India	318.68	292.59	91.81	179	5	-	5	13
408	Gwalior	India	228.92	228.94	100.01	121	6	-	4	17
409	Gwangju	Republic of Korea	360.99	214.12	59.32	117	5	-	5	11
410	Habra	India	118.03	108.93	92.29	42	2	-	2	7
411	Haerbin	China	395.77	395.62	99.96	246	1	-	1	26
412	Hafar al-Batin	Saudi Arabia	276.55	238.38	86.20	137	3	-	3	29
413	Haikou	China	527.07	309.61	58.74	165	5	-	1	2
414	Ha'il	Saudi Arabia	332.20	328.92	99.01	180	3	-	3	22

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM123}	C_{HM110}	C_{F1523}^*	C_{F1523}
415	Nantong Haimen	China	1,729.54	1101.64	63.70	705	1	-	1	23
416	Hai Phòng	Viet Nam	609.49	359.51	58.99	206	6	-	4	15
417	Halab (Aleppo)	Syrian Arab Republic	297.25	297.06	99.94	169	3	-	3	8
418	Halifax	Canada	261.08	69.61	26.66	19	7	-	7	6
419	Hamadan	Iran (Islamic Republic of)	185.40	184.86	99.71	98	3	-	3	27
420	Hamah	Syrian Arab Republic	133.45	132.02	98.93	59	3	-	2	7
421	Hamburg	Germany	290.97	290.85	99.96	154	7	7	7	14
422	Hamhung	Dem. People's Republic of Korea	147.83	92.86	62.82	34	5	-	4	27
423	Hami	China	170.01	147.24	86.61	64	1	-	1	27
424	Toronto Hamilton	Canada	2,640.55	2566.25	97.19	1898	1	7	1	24
425	Handan	China	415.09	358.07	86.26	195	1	-	1	10
426	Hannover	Germany	207.67	207.48	99.91	75	7	-	7	14
427	Hà Noi	Viet Nam	537.33	537	99.94	302	6	2	6	10
428	Hanzhong	China	208.51	169.94	81.50	78	1	-	1	23
429	Harare	Zimbabwe	680.35	678.66	99.75	453	2	6	2	25
430	Hargeysa	Somalia	156.73	156.56	99.89	79	4	4	3	28
431	Harrisburg	United States of America	404.24	385	95.24	204	2	-	7	6
432	Hartford	United States of America	560.01	452.4	80.78	242	2	-	7	6
433	Hat Yai	Thailand	167.68	6.76	4.03	0	-	-	-	-
434	Hebi	China	140.97	136.59	96.89	60	1	-	4	12
435	Hefa (Haifa)	Israel	217.10	120.37	55.44	58	7	-	7	14
436	Hefei	China	818.29	818.14	99.98	562	1	-	1	26
437	Hegang	China	240.84	175.65	72.93	83	1	-	1	5
438	Hejian	China	92.30	92.15	99.83	35	1	-	4	12
439	Helsinki	Finland	354.00	132.27	37.36	47	7	-	7	14

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}^7$
440	Hengshui	China	165.95	165.81	99.91	74	1	-	1	12
441	Hengyang	China	333.32	333.02	99.91	184	5	-	2	23
442	Herat	Afghanistan	154.24	154.12	99.92	73	3	-	6	28
443	Hermosillo	Mexico	356.87	356.87	100.00	211	3	-	3	8
444	Heyuan	China	138.29	137.15	99.17	58	5	-	5	11
445	Heze	China	257.38	230.99	89.75	119	1	-	1	26
446	Hezhou	China	165.61	36.92	22.29	8	5	-	1	5
447	Hims (Homs)	Syrian Arab Republic	158.63	158.56	99.96	79	3	-	4	7
448	Hiroshima	Japan	437.90	433.01	98.88	212	5	-	5	3
449	Hisar	India	137.99	89.37	64.77	38	4	-	4	12
450	Hohhot	China	487.58	487.33	99.95	307	1	-	1	26
451	Holguin	Cuba	116.79	77.15	66.06	28	6	-	6	15
452	Hong Kong	China, Hong Kong SAR	379.12	88.34	23.30	19	5	5	5	16
453	Honolulu	United States of America	737.74	288.27	39.07	144	5	-	5	3
454	Hosur	India	182.79	131.28	71.82	58	2	-	2	18
455	Houston The Woodlands	United States of America	3,543.34	3543.02	99.99	2874	2	-	7	24
456	Hsinchu	China	219.69	218.1	99.27	104	5	-	5	2
457	Huai'an	China	460.47	363.35	78.91	207	1	-	1	26
458	Huaibei	China	325.45	320.16	98.37	177	1	-	4	12
459	Huaihua	China	139.81	129.33	92.51	59	5	-	5	11
460	Huainan	China	473.75	256.66	54.18	118	1	-	1	12
461	Huambo	Angola	186.49	186.27	99.88	88	6	-	4	15
462	Huancayo	Peru	284.98	228.39	80.14	118	4	-	2	18
463	Huangshan	China	223.84	106.58	47.61	47	5	-	5	16
464	Hue	Viet Nam	148.75	125	84.04	60	6	-	2	18
465	Hufuf-Mubarraz	Saudi Arabia	665.48	515.96	77.53	346	3	-	3	29
466	Huizhou	China	297.94	254.48	85.41	132	5	-	1	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
467	Huntsville	United States of America	429.53	169.58	39.48	77	2	-	7	6
468	Huzhou	China	332.85	206.41	62.01	100	5	-	1	2
469	Hyderabad	India	1,002.51	1002.71	100.02	763	4	-	6	19
470	Hyderabad	Pakistan	187.50	5.98	3.19	0	-	-	-	-
471	Ibagué	Colombia	130.15	53.56	41.15	14	4	-	5	3
472	Ibb	Yemen	108.89	73.45	67.45	25	3	-	2	18
473	Ife	Nigeria	217.51	217.32	99.91	118	2	-	2	13
474	Imphal	India	177.23	177.06	99.90	85	4	-	4	7
475	Indianapolis	United States of America	999.39	997.67	99.83	714	2	-	2	18
476	Indio-Cathedral City-Palm Springs	United States of America	494.30	343.48	69.49	194	3	-	3	8
477	Ipoh	Malaysia	541.14	533.77	98.64	338	5	-	5	3
478	Iquique	Chile	132.61	52.94	39.92	9	3	-	3	22
479	Iquitos	Peru	131.95	77.83	58.99	32	4	-	4	15
480	Irbid	Jordan	216.18	214.82	99.37	113	7	-	2	18
481	Irbil (Erbil)	Iraq	349.52	287.08	82.14	174	3	-	3	8
482	Irkutsk	Russian Federation	316.35	292.17	92.36	153	7	-	1	5
483	Rawalpindi Islamabad	Pakistan	738.99	738.91	99.99	488	4	4	2	18
484	Ivanovo	Russian Federation	170.94	170.76	99.89	89	7	-	7	1
485	Izhevsk	Russian Federation	211.44	198.6	93.93	76	5	-	5	2
486	Izmir	Turkey	464.12	462.9	99.74	226	3	-	5	11
487	Izmit	Turkey	299.16	151.76	50.73	40	7	-	1	1
488	Jabalpur	India	248.71	248.76	100.02	130	6	-	4	15
489	Jackson, Mississippi	United States of America	326.68	90.32	27.65	30	2	-	5	23
490	Jacksonville, Florida	United States of America	908.39	895.03	98.53	598	2	-	7	6

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM1523}	C_{HM110}	C_{F1523}^*	C_{F1523}
491	Jaipur	India	540.92	532.71	98.48	328	4	4	2	18
492	Jalandhar	India	239.48	239.34	99.94	127	4	-	1	5
493	Jalgaon	India	138.32	138.4	100.06	61	6	-	4	17
494	Jalna	India	124.56	88.51	71.06	33	6	-	4	17
495	Jambi	Indonesia	269.65	247.67	91.85	142	2	-	2	18
496	Jammu	India	272.37	270.47	99.30	134	4	-	2	18
497	Jamshedpur	India	338.29	331.35	97.95	186	2	-	4	15
498	Jhang	Pakistan	131.14	92.79	70.76	40	4	-	4	17
499	Jiamusi	China	172.71	117.82	68.22	50	1	-	1	26
500	Ji'an, Jiangxi	China	149.32	127.37	85.30	60	5	-	1	23
501	Jianyang	China	83.14	34.19	41.12	5	5	-	5	11
502	Jiaxing	China	257.57	251.71	97.72	136	1	-	1	5
503	Jiddah	Saudi Arabia	825.97	825.35	99.92	583	3	-	3	29
504	Jilin	China	287.17	286.9	99.90	146	5	-	1	2
505	Ji'nan, Shandong Zhangqiu	China	1,253.68	1090.83	87.01	717	1	-	1	5
506	Jincheng	China	188.64	155.63	82.50	73	1	-	1	26
507	Jingdezhen	China	228.12	202.92	88.95	98	5	-	5	11
508	Jingmen	China	222.63	147.08	66.07	66	5	-	5	2
509	Quanzhou Shishi Jinjiang	China	1,751.51	1742.15	99.47	1268	5	-	1	2
510	Taiyuan, Shanxi Jinzhong	China	890.46	890.5	100.00	625	1	-	1	2
511	Jiujiang	China	307.46	295.66	96.16	152	5	-	5	11
512	Jixi, Heilongjiang	China	157.49	143.15	90.90	48	1	-	1	5
513	João Pessoa	Brazil	358.57	336.72	93.91	196	2	-	2	7
514	Jodhpur	India	338.94	302.56	89.27	180	4	-	2	7
515	Johannesburg	South Africa	1,797.30	1797.51	100.01	1361	6	3	6	19
516	Jos	Nigeria	354.48	273.82	77.24	153	4	-	4	7
517	Juazeiro Do Norte	Brazil	210.91	140.78	66.75	63	6	-	2	9
518	Juba	South Sudan	186.27	171.59	92.12	90	4	-	2	18
519	Jubayl	Saudi Arabia	484.66	34.6	7.14	4	3	-	3	29

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
520	Juiz De Fora	Brazil	237.13	205.69	86.74	106	4	-	6	19
521	Junagadh	India	132.66	89.63	67.56	35	6	-	4	17
522	Jurong	China	109.53	107.72	98.35	43	1	-	1	5
523	Kabul	Afghanistan	602.25	562.83	93.46	373	3	4	3	22
524	Kadapa	India	115.74	74.19	64.10	33	6	-	4	17
525	Kaduna	Nigeria	453.02	453.13	100.02	301	4	-	4	7
526	Kagoshima	Japan	280.39	239.73	85.50	121	5	-	5	3
527	Kaifeng	China	254.91	243.01	95.33	121	1	-	1	10
528	Kaili	China	115.67	90.9	78.59	41	5	-	5	16
529	Taishan Kaiping	China	456.78	363.53	79.59	180	1	-	4	12
530	RajahmundryRajahmundry	India	1,076.49	222.67	20.68	94	6	-	4	17
531	Kaliningrad, Kaliningrad Oblast	Russian Federation	208.93	208.66	99.87	114	7	-	7	14
532	Kaluga	Russian Federation	146.20	3.78	2.59	0	-	-	-	-
533	Kampala	Uganda	610.19	610.05	99.98	424	7	6	7	6
534	Kananga	Democratic Republic of the Congo	163.56	160.08	97.87	79	2	-	2	9
535	Kanazawa	Japan	377.89	336.12	88.95	166	5	-	6	10
536	Kandahar	Afghanistan	302.58	168	55.52	90	3	-	1	26
537	Kannur	India	140.87	83.96	59.60	29	2	-	2	25
538	Kano	Nigeria	447.18	447.04	99.97	284	4	4	6	4
539	Kanpur	India	390.67	384.23	98.35	217	4	-	1	10
540	Kansas City	United States of America	383.63	256.46	66.85	140	2	-	7	6
541	KaohsiungKaohsiung Tainan	China	2,438.44	907.08	37.20	653	5	-	1	5
542	Karachi	Pakistan	1,022.96	1022.63	99.97	738	4	4	3	4
543	Karaganda	Kazakhstan	165.03	158.25	95.89	69	1	-	1	5
544	Karbala	Iraq	186.45	186.23	99.88	88	3	-	3	28

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
545	Kassala	Sudan	249.72	179.32	71.81	99	6	-	4	17
546	Kathmandu	Nepal	270.77	270.6	99.94	162	2	-	2	18
547	Katowice	Poland	243.93	133.29	54.64	41	5	-	5	11
548	Katsina	Nigeria	174.18	172.74	99.17	88	4	-	4	27
549	Thiruvananthapuram Kollam Kayamkulam	India	2,152.70	581.52	27.01	354	2	-	2	25
550	Kayseri	Turkey	286.80	252.97	88.20	134	3	-	1	27
551	Kazan	Russian Federation	372.39	365.88	98.25	187	5	-	1	2
552	Keelung	China	150.42	144.06	95.77	48	5	-	5	16
553	Kelamayi	China	100.14	98.51	98.37	38	1	-	3	8
554	Kemerovo	Russian Federation	256.33	184.59	72.01	89	5	-	1	2
555	Kénitra	Morocco	165.57	99.83	60.29	40	3	-	4	27
556	Kerman	Iran (Islamic Republic of)	268.41	254.29	94.74	146	3	-	3	28
557	Kermanshah	Iran (Islamic Republic of)	241.37	211.08	87.45	107	3	-	3	27
558	Khamis Mushayt	Saudi Arabia	625.05	327.97	52.47	183	3	-	3	22
559	Kharkiv	Ukraine	410.04	270.52	65.97	161	7	-	7	14
560	Khorramabad	Iran (Islamic Republic of)	161.56	71.82	44.45	22	6	-	3	8
561	Khulna	Bangladesh	151.13	94.86	62.77	39	6	-	2	18
562	Kigali	Rwanda	270.79	236.99	87.52	138	2	2	2	18
563	Kikwit	Democratic Republic of the Congo	149.84	146.68	97.89	64	2	-	2	7
564	Kingston	Jamaica	335.12	277.38	82.77	127	2	-	6	19
565	Kinki M.M.A. (Osaka)	Japan	2,723.69	2723.87	100.01	2046	5	5	5	3
566	Kirkuk	Iraq	250.02	249.79	99.91	138	3	-	3	8
567	Kirov	Russian Federation	196.51	158.58	80.70	68	5	-	1	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
568	Kisangani	Democratic Republic of the Congo	186.14	172.9	92.89	85	2	-	2	18
569	Kismaayo	Somalia	77.19	26.3	34.07	2	4	-	6	28
570	Orlando Kissimmee	United States of America	2,159.08	2149.17	99.54	1653	2	-	7	6
571	Kitchener	Canada	561.92	407.49	72.52	258	7	-	4	1
572	Knoxville	United States of America	343.37	219.62	63.96	106	2	-	2	25
573	Kolwezi	Democratic Republic of the Congo	305.97	104.69	34.22	49	2	-	2	18
574	Konya	Turkey	293.26	293.13	99.95	173	7	-	7	14
575	Kota	India	240.24	240.06	99.93	120	6	-	4	15
576	Kota Bharu	Malaysia	277.22	6.76	2.44	0	-	-	-	-
577	Kota Kinabalu	Malaysia	252.09	208.42	82.68	89	6	-	6	3
578	Kottayam	India	224.70	200.27	89.13	96	2	-	2	25
579	Kraków (Cracow)	Poland	139.52	139.31	99.85	47	7	-	7	14
580	Krasnodar	Russian Federation	401.02	355.32	88.60	186	1	-	1	5
581	Krasnoyarsk	Russian Federation	354.19	148.48	41.92	59	5	-	1	26
582	Krivoi Rog	Ukraine	399.96	378.15	94.55	222	7	-	7	1
583	Kuala Lumpur	Malaysia	1,996.63	1996.84	100.01	1620	5	5	5	3
584	Kuala Terengganu	Malaysia	224.33	6.76	3.01	0	-	-	-	-
585	Kuantan	Malaysia	199.04	181.55	91.21	89	5	-	5	3
586	Kuerle	China	304.25	288.13	94.70	148	3	-	3	29
587	Kumamoto	Japan	451.10	451	99.98	271	6	-	6	10
588	Kumasi	Ghana	704.58	704.2	99.95	517	2	6	2	25
589	Kunming	China	691.49	691.38	99.98	455	5	-	5	11
590	Shanghai Kunshan Taicang	China	3,482.13	3481.59	99.98	2837	1	1	1	24
591	Kurgan	Russian Federation	152.88	149	97.46	68	7	-	1	5
592	Kurnool	India	145.43	119.51	82.18	57	6	-	4	17

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
593	Kursk	Russian Federation	266.34	4.05	1.52	0	-	-	-	-
594	Kut	Iraq	145.80	134.73	92.41	66	3	-	3	8
595	Kyiv (Kiev)	Ukraine	604.51	604.14	99.94	387	5	1	5	11
596	La Habana (Havana)	Cuba	322.05	317	98.43	178	2	-	6	19
597	Laibin	China	128.83	126.94	98.53	55	5	-	1	2
598	Laixi	China	181.09	180.84	99.86	81	1	-	4	12
599	Laiyang	China	203.14	175.99	86.64	93	1	-	4	12
600	Lampang	Thailand	145.16	145.04	99.92	68	2	-	2	7
601	Lancaster-Palmdale	United States of America	407.41	406.94	99.88	238	6	-	6	27
602	Lansing	United States of America	183.76	178.79	97.29	38	2	-	7	6
603	Lanzhou	China	479.46	99.13	20.68	38	1	-	3	8
604	La Paz	Bolivia (Plurinational State of)	441.51	430.21	97.44	263	4	-	6	4
605	La Plata	Argentina	263.17	263.08	99.97	143	4	-	6	21
606	Larkana	Pakistan	104.76	5.98	5.71	0	-	-	-	-
607	La Serena-Coquimbo	Chile	243.31	149.36	61.39	61	6	-	4	15
608	Las Palmas Gran Canaria	Spain	104.64	78.37	74.89	27	3	-	3	8
609	Las Vegas	United States of America	1,461.95	1461.71	99.98	1154	3	-	3	8
610	Lattakia	Syrian Arab Republic	102.30	84.29	82.40	20	1	-	1	2
611	Latur	India	138.14	96.23	69.66	44	6	-	4	17
612	Lausanne	Switzerland	193.12	99.36	51.45	23	7	-	7	14
613	Leicester	United Kingdom	213.98	213.63	99.84	105	7	-	7	20
614	Leipzig	Germany	149.71	149.67	99.97	49	7	-	7	14
615	Leyang	China	124.25	124.17	99.93	48	5	-	1	23
616	Leping	China	145.57	136.1	93.50	55	7	-	4	15

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
617	Lexington-Fayette	United States of America	360.54	360.3	99.93	231	7	-	7	20
618	Liaocheng	China	222.63	222.41	99.90	103	1	-	1	26
619	Liaoyuan	China	126.22	126.05	99.86	44	5	-	1	2
620	Liège	Belgium	244.01	243.84	99.93	118	7	-	7	6
621	Likasi	Democratic Republic of the Congo	153.74	128.42	83.53	58	7	-	4	7
622	Liling	China	141.00	135.87	96.36	65	5	-	2	23
623	Lilongwe	Malawi	453.36	420.75	92.81	279	2	6	2	18
624	Lin'an	China	120.57	79.49	65.93	27	5	-	5	16
625	Linfen	China	184.08	178.67	97.06	74	1	-	1	12
626	Linhai	China	220.18	186.17	84.56	71	5	-	5	16
627	Linqing	China	111.81	111.6	99.81	46	1	-	1	10
628	Linyi, Shandong	China	712.95	712.78	99.98	479	1	-	1	26
629	Linzhou	China	151.81	151.73	99.95	57	1	-	4	10
630	Lipa City	Philippines	95.31	39	40.92	10	6	-	2	18
631	Lipetsk	Russian Federation	247.37	4.16	1.68	0	-	-	-	-
632	Lisboa (Lisbon)	Portugal	873.66	781.6	89.46	407	7	7	1	1
633	Lishui, Zhejiang	China	108.12	77.17	71.37	26	5	-	5	16
634	Little Rock	United States of America	378.15	266.48	70.47	105	2	-	5	13
635	Lluan	China	226.24	180.91	79.96	82	1	-	1	2
636	Liupanshui	China	141.27	60.16	42.58	16	5	-	5	11
637	Liuzhou	China	425.94	425.68	99.94	267	5	-	1	26
638	Liyang	China	158.75	153.39	96.62	69	1	-	1	5
639	London	United Kingdom	1,205.57	1205.07	99.96	834	7	7	7	20
640	London	Canada	378.02	377.64	99.90	239	7	-	7	1
641	Londrina	Brazil	291.69	291.51	99.94	165	6	-	6	15
642	Xiamen Zhangzhou Longhai	China	1,772.09	822.93	46.44	396	5	-	1	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}^*}$	$C_{F_{1523}}$
643	Longyan	China	241.76	222.08	91.86	99	5	-	5	16
644	Los Angeles-Long Beach-Santa Ana Riverside-San Bernardino Mission Viejo	United States of America	6,681.00	6681.24	100.00	5534	6	2	6	24
645	Loudi	China	165.36	165.41	100.03	77	5	-	5	23
646	Louisville	United States of America	496.86	491.94	99.01	293	2	-	7	6
647	Loun	Cameroon	85.87	13.52	15.74	0	-	-	-	-
648	Luanda	Angola	845.30	565.43	66.89	347	6	-	6	21
649	Lubango	Angola	188.89	140.53	74.40	68	4	-	2	18
650	Lublin	Poland	157.69	155.72	98.75	71	7	-	1	1
651	Lubumbashi	Democratic Republic of the Congo	378.94	378.78	99.96	243	2	-	2	18
652	Lucknow	India	459.25	458.78	99.90	284	4	3	6	19
653	Ludhiana	India	336.28	336.21	99.98	209	4	-	4	10
654	Lugansk	Ukraine	202.97	202.85	99.94	111	7	-	7	1
655	Luoding	China	111.74	71.95	64.39	26	1	-	4	12
656	Luohe	China	194.93	194.75	99.91	98	1	-	1	26
657	Luoyang Yanshi	China	1,090.33	965.81	88.58	650	1	-	1	12
658	Lusaka	Zambia	514.17	513.71	99.91	350	4	6	2	18
659	Luzhou	China	216.41	94.13	43.50	40	5	-	5	11
660	Lvliang	China	109.98	91.28	83.00	30	3	-	3	8
661	Lvov	Ukraine	170.60	170.42	99.90	72	5	-	5	11
662	Lyon	France	307.66	307.54	99.96	168	7	-	7	14
663	Zhuhai Macao	China	463.01	360.52	77.86	206	5	-	1	2
664	Macapá	Brazil	189.86	119.38	62.88	49	6	-	4	7
665	Macheng	China	92.58	78.06	84.32	32	1	-	2	23

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
666	Madison	United States of America	267.39	119.54	44.71	20	2	-	7	6
667	Madrid	Spain	801.21	801.11	99.99	541	3	7	1	5
668	Madurai	India	234.25	234.25	100.00	133	7	-	4	7
669	Magnitogorsk	Russian Federation	300.68	281.11	93.49	135	7	-	1	1
670	Maiduguri	Nigeria	314.22	313.54	99.78	186	4	-	4	15
671	Makhachkala	Russian Federation	252.18	109.29	43.34	33	5	-	6	2
672	Makkah (Mecca)	Saudi Arabia	469.62	468.9	99.85	274	3	-	3	22
673	Makurdi	Nigeria	215.06	79.31	36.88	28	6	-	4	7
674	Malaga	Spain	220.47	130.36	59.13	47	3	-	1	2
675	Malatya	Turkey	118.63	113.74	95.88	45	3	-	4	27
676	Managua	Nicaragua	242.94	242.87	99.97	130	4	-	6	19
677	Manaus	Brazil	520.35	501.97	96.47	331	4	-	5	3
678	Manchester	United Kingdom	611.76	611.51	99.96	372	7	7	7	20
679	Mandalay	Myanmar	281.77	256.58	91.06	145	6	-	6	15
680	Mangalore	India	310.56	216.1	69.58	121	2	-	7	9
681	Manisa	Turkey	143.81	32.24	22.42	6	5	-	5	16
682	Manizales	Colombia	127.67	40.5	31.72	9	5	-	5	3
683	Mardan	Pakistan	288.32	81.61	28.31	31	4	-	2	18
684	Mar Del Plata	Argentina	198.94	198.84	99.95	95	6	-	6	19
685	Maringá	Brazil	280.54	280.35	99.93	170	6	-	4	15
686	Mariupol	Ukraine	313.05	95.01	30.35	28	7	-	4	1
687	Marrakech	Morocco	253.39	253.16	99.91	146	3	-	3	27
688	Mashhad	Iran (Islamic Republic of)	515.38	512.3	99.40	321	3	-	3	8
689	Masqat (Muscat)	Oman	309.91	5.94	1.92	0	-	-	-	-
690	Matadi	Democratic Republic of the Congo	127.05	87.03	68.50	38	4	-	4	17
691	Matamoros	Mexico	94.07	6.24	6.63	0	-	-	-	-

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
692	Mataram	Indonesia	179.08	177.42	99.07	88	2	-	2	18
693	Maturin	Venezuela (Bolivarian Republic of)	309.59	292.87	94.60	139	4	-	4	15
694	Maunath Bhanjan	India	95.53	47.98	50.22	12	4	-	4	15
695	Mawlamyine	Myanmar	143.09	25.48	17.81	4	2	-	2	18
696	Mazatlán	Mexico	200.76	200.29	99.77	89	6	-	1	5
697	Mbandaka	Democratic Republic of the Congo	126.87	78	61.48	36	2	-	4	7
698	Mbeya	United Republic of Tanzania	263.95	159.87	60.57	79	4	-	2	25
699	Mbuji-Mayi	Democratic Republic of the Congo	385.57	27.04	7.01	4	6	-	7	7
700	McAllen Reynosa	United States of America	1,064.04	914.03	85.90	614	6	-	4	7
701	Medan	Indonesia	715.86	711.21	99.35	440	4	-	6	15
702	Medellín	Colombia	378.92	303.15	80.00	168	4	-	5	3
703	Meerut	India	205.80	205.59	99.90	94	4	-	2	18
704	Meishan	China	164.34	155.2	94.44	75	1	-	1	23
705	Meizhou	China	197.05	196.99	99.97	101	5	-	5	23
706	Mekele	Ethiopia	124.54	109.49	87.91	38	4	-	4	27
707	Meknès	Morocco	196.04	173.72	88.61	81	3	-	4	17
708	Melbourne	Australia	2,152.82	2152.94	100.01	1698	6	2	6	19
709	Memphis	United States of America	947.24	264.65	27.94	143	2	-	7	6
710	Mendoza	Argentina	391.16	390.99	99.96	249	6	-	6	21
711	Merca	Somalia	61.01	22.3	36.55	1	3	-	3	28
712	Mersin	Turkey	404.68	176.38	43.58	71	1	-	1	5
713	Mexicali	Mexico	428.48	428.37	99.97	262	6	-	1	10

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
714	Miami	United States of America	2,494.07	1764.58	70.75	1355	7	-	7	20
715	Mianyang, Sichuan	China	360.76	263.8	73.12	148	5	-	5	23
716	Miluo	China	84.79	51.86	61.17	16	6	-	2	18
717	Milwaukee	United States of America	841.47	787.78	93.62	576	2	-	7	6
718	Minatitlán	Mexico	154.85	154.86	100.01	78	6	-	4	15
719	Minna	Nigeria	214.01	209.75	98.01	106	6	-	4	17
720	Minsk	Belarus	419.38	419.26	99.97	254	5	1	5	2
721	Misratah	Libya	328.40	280.81	85.51	157	3	-	7	14
722	Mobile	United States of America	325.36	130.68	40.16	39	2	-	7	6
723	Mogilev	Belarus	199.59	192.42	96.41	95	7	-	7	1
724	Mombasa	Kenya	270.19	264.47	97.88	113	4	4	2	18
725	Monclova	Mexico	227.91	6.21	2.72	0	-	-	-	-
726	Monrovia	Liberia	399.93	85.63	21.41	32	4	-	6	1
727	Monteria	Colombia	144.39	103.09	71.40	53	6	-	4	15
728	Monterrey	Mexico	1,148.43	5.98	0.52	0	-	-	-	-
729	Montes Claros	Brazil	179.83	178.95	99.51	88	6	-	6	9
730	Montevideo	Uruguay	396.33	378.72	95.56	215	6	-	6	19
731	Montpellier	France	322.42	320.7	99.47	135	7	-	7	1
732	Monywa	Myanmar	122.33	72.47	59.24	27	2	-	7	7
733	Moradabad	India	141.64	130.45	92.10	56	4	-	4	17
734	Morelia	Mexico	233.33	232.94	99.83	129	6	-	6	19
735	Morogoro	United Republic of Tanzania	191.88	158.63	82.67	74	6	-	2	18
736	Moskva (Moscow)	Russian Federation	947.79	947.31	99.95	643	5	1	5	11
737	Mudanjiang	China	181.54	180.93	99.66	80	1	-	1	5
738	Muenster (Westfalen)	Germany	137.16	136.9	99.81	48	7	-	7	1

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
739	Multan	Pakistan	284.26	284.01	99.91	170	6	-	6	21
740	Muqdisho (Mogadishu)	Somalia	228.97	216.14	94.40	107	4	4	6	9
741	Muzaffarpur	India	119.49	77.59	64.94	24	2	-	2	18
742	Mwanza	United Republic of Tanzania	230.46	95.27	41.34	42	2	2	7	6
743	Mymensingh	Bangladesh	106.14	101.77	95.88	45	2	-	4	15
744	Mysore	India	291.71	282.99	97.01	159	6	-	2	18
745	Nagasaki	Japan	275.86	108.69	39.40	42	5	-	5	16
746	Nagpur	India	378.24	358.55	94.80	224	6	-	6	10
747	Naha	Japan	535.82	317.8	59.31	161	5	-	1	2
748	Nairobi	Kenya	694.82	694.54	99.96	455	4	4	7	6
749	Najran	Saudi Arabia	441.42	140.76	31.89	67	3	-	3	22
750	Nakhon Ratchasima	Thailand	242.15	238.52	98.50	115	6	-	2	7
751	Nakuru	Kenya	184.72	95.03	51.45	41	4	-	2	18
752	Namangan	Uzbekistan	340.50	298.82	87.76	167	6	-	6	10
753	Nampula	Mozambique	172.77	171.9	99.49	87	4	-	2	9
754	Nanchong	China	230.10	188.72	82.02	85	5	-	5	11
755	Nanded Waghala	India	155.65	112.41	72.22	53	6	-	4	17
756	Nanjing, Jiangsu	China	1,179.56	1179.41	99.99	848	5	1	1	2
757	Nanning	China	617.65	597.98	96.82	390	5	-	1	26
758	Nantes	France	392.40	392.31	99.98	223	7	-	7	1
759	Nanyang, Henan	China	241.55	241.43	99.95	136	1	-	1	26
760	Nashik	India	329.86	319.3	96.80	182	6	-	4	7
761	Nashville-Davidson	United States of America	731.99	578.98	79.10	355	2	-	7	6
762	Nasiriyah	Iraq	168.80	132.79	78.67	66	3	-	3	28
763	Natal	Brazil	448.77	114.35	25.48	45	6	-	6	21
764	Navsari	India	119.30	106.11	88.95	44	4	-	4	17
765	Nawabshah	Pakistan	110.16	5.98	5.43	0	-	-	-	-

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
766	Nay Pyi Taw	Myanmar	702.04	12.96	1.85	0	-	-	-	-
767	N'Djaména	Chad	451.51	411.69	91.18	268	4	-	4	7
768	Ndola	Zambia	253.38	227.43	89.76	118	2	-	2	18
769	Neijiang	China	120.71	35.07	29.05	4	5	-	5	11
770	Neiva	Colombia	139.42	138.73	99.51	60	4	-	6	9
771	Nellore	India	155.23	128.18	82.57	52	6	-	4	17
772	Neuquén-Plottier-Cipolletti	Argentina	212.94	141.5	66.45	64	6	-	6	21
773	Newcastle and Lake Macquarie	Australia	435.42	234.06	53.76	116	2	-	7	6
774	Newcastle upon Tyne Sunderland	United Kingdom	457.34	453.09	99.07	284	7	-	7	20
775	New Orleans	United States of America	717.00	716.79	99.97	475	2	-	6	19
776	New York-Newark	United States of America	2,321.15	2320.56	99.97	1549	2	2	6	19
777	Niamey	Niger	343.91	342.91	99.71	203	4	-	6	9
778	Nice-Cannes	France	518.36	184.78	35.65	101	2	-	2	13
779	Niigata	Japan	349.15	216.45	61.99	111	1	-	1	10
780	Nikolaev	Ukraine	191.14	181.39	94.90	67	7	-	1	1
781	Ningbo	China	1,320.99	1175.03	88.95	885	5	-	1	5
782	Nizamabad	India	105.98	101.41	95.69	42	6	-	4	17
783	Nizhny Novgorod	Russian Federation	442.94	22.26	5.03	4	5	-	1	2
784	Nizhny Tagil	Russian Federation	196.47	146.11	74.37	61	5	-	5	11
785	Nottingham	United Kingdom	255.76	251.24	98.23	131	7	-	7	20
786	Nouakchott	Mauritania	291.86	286.61	98.20	166	3	-	3	8
787	Novokuznetsk	Russian Federation	130.92	101.74	77.71	37	5	-	5	11
788	Novosibirsk	Russian Federation	376.03	297.71	79.17	173	5	-	1	2
789	Nuevo Laredo	Mexico	421.08	399.6	94.90	241	6	-	6	19
790	Nuremberg	Germany	302.93	291.31	96.16	167	7	-	7	14

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
791	Nyala	Sudan	248.97	242.5	97.40	142	2	-	4	17
792	Nzérékoré	Guinea	123.98	109.71	88.49	47	2	-	2	18
793	Oaxaca de Juárez	Mexico	273.53	271.41	99.23	144	4	-	2	18
794	Łódź	Poland	213.13	212.63	99.77	95	7	-	7	14
795	Odesa	Ukraine	440.41	289.3	65.69	118	7	-	1	1
796	Ogden-Layton	United States of America	499.39	348.39	69.76	195	2	-	7	6
797	Okara	Pakistan	99.79	69.58	69.73	20	4	-	4	17
798	Okayama	Japan	1,678.11	559.86	33.36	356	5	-	5	3
799	Okene	Nigeria	191.43	111.76	58.38	45	6	-	4	17
800	Oklahoma City	United States of America	1,038.96	1038.85	99.99	725	2	-	7	1
801	Omaha	United States of America	664.20	490.76	73.89	280	7	-	7	20
802	Omsk	Russian Federation	384.52	301.67	78.45	161	7	-	1	2
803	Ondo	Nigeria	158.93	149.59	94.12	79	2	-	2	13
804	Orel	Russian Federation	151.68	3.78	2.49	0	-	-	-	-
805	Orenburg	Russian Federation	219.86	211.98	96.42	97	5	-	1	2
806	Orumiyeh	Iran (Islamic Republic of)	287.90	242.52	84.24	122	6	-	6	28
807	Oshawa	Canada	247.86	218.44	88.13	117	7	-	7	1
808	Oslo	Norway	233.24	201.6	86.44	80	7	-	7	6
809	Ottawa-Gatineau	Canada	487.77	487.85	100.02	308	7	-	7	6
810	Ouagadougou	Burkina Faso	605.04	604.77	99.95	430	4	6	2	9
811	Oujda	Morocco	165.41	165.4	100.00	84	3	-	4	27
812	Oxnard	United States of America	268.20	266.2	99.25	136	6	-	6	10
813	Pachuca de Soto	Mexico	294.59	186.45	63.29	104	6	-	6	19
814	Padang	Indonesia	230.19	168.58	73.24	76	6	-	6	21

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM123}	C_{HM110}	C_{F1523}^*	C_{F1523}
815	Palakkad	India	162.71	125.67	77.24	62	2	-	2	18
816	Palembang	Indonesia	526.06	40.56	7.71	10	2	-	2	25
817	Palermo	Italy	256.85	217.89	84.83	90	7	-	2	14
818	Palma	Spain	201.14	151.01	75.08	58	7	-	1	1
819	Palm Bay-Melbourne	United States of America	772.16	33.81	4.38	6	2	-	7	6
820	Panjin	China	367.35	324.21	88.26	167	1	-	1	12
821	Panzhihua	China	158.75	11.75	7.40	0	-	-	-	-
822	Parbhani	India	115.39	73.43	63.63	30	6	-	4	17
823	Paris	France	1,161.16	1160.87	99.98	886	7	7	7	14
824	Parna	Italy	118.11	117.9	99.82	46	7	-	4	12
825	Pasto	Colombia	121.81	99.16	81.41	41	4	-	2	18
826	Pavlodar	Kazakhstan	211.15	145.82	69.06	67	1	-	1	12
827	Pekan Baru	Indonesia	442.41	406.3	91.84	257	6	-	2	18
828	Pelotas	Brazil	161.80	148.44	91.74	70	6	-	4	7
829	Pensacola	United States of America	542.98	329.33	60.65	164	2	-	7	6
830	Penza	Russian Federation	233.08	227.99	97.82	112	7	-	7	14
831	Pereira	Colombia	207.95	118.59	57.03	54	4	-	6	9
832	Perm	Russian Federation	288.39	158.57	54.98	69	5	-	5	11
833	Pescara	Italy	114.20	106.18	92.98	21	7	-	7	23
834	Peshawar	Pakistan	341.15	288.68	84.62	154	6	-	6	21
835	Petrolina	Brazil	214.15	27.04	12.63	4	6	-	6	21
836	Phnum Pénh (Phnom Penh)	Cambodia	350.13	338.36	96.64	194	6	-	1	26
837	Phoenix-Mesa	United States of America	3,405.28	3405.18	100.00	2875	3	-	6	24
838	Pingdu	China	195.94	195.75	99.90	102	1	-	1	12
839	Pinghu	China	137.33	137.16	99.88	63	1	-	1	5
840	Pingxiang, Jiangxi	China	173.43	161.03	92.85	77	5	-	5	11

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
841	Piracicaba	Brazil	257.93	257.79	99.94	133	6	-	4	15
842	Pittsburgh	United States of America	232.70	232.79	100.04	108	2	-	7	6
843	Piura	Peru	174.06	167.24	96.08	76	4	-	4	12
844	Plovdiv	Bulgaria	138.57	138.5	99.95	66	1	-	4	12
845	Pohang	Republic of Korea	210.78	192.39	91.28	81	5	-	5	11
846	Pointe-Noire	Congo	255.28	253.09	99.14	133	6	-	6	19
847	Pokhara	Nepal	128.18	23.97	18.70	2	2	-	7	13
848	Ponta Grossa	Brazil	266.18	266.01	99.94	146	2	-	4	7
849	Pontianak	Indonesia	224.28	224.01	99.88	113	4	-	2	18
850	Port-au-Prince	Haiti	400.03	376.23	94.05	215	4	2	6	19
851	Port Elizabeth	South Africa	407.18	407.09	99.98	235	6	-	2	7
852	Portland	United States of America	961.67	961.66	100.00	718	2	-	7	6
853	Porto	Portugal	569.68	569.44	99.96	341	2	-	7	6
854	Pôrto Alegre	Brazil	763.91	23.07	3.02	2	5	-	1	2
855	Port St. Lucie	United States of America	892.01	625.31	70.10	410	7	-	7	20
856	Port Sudan (Bur Sudan)	Sudan	277.40	270.63	97.56	145	4	-	3	22
857	Posadas	Argentina	274.19	210.53	76.78	70	6	-	2	18
858	Poughkeepsie-Newburgh	United States of America	602.47	5.2	0.86	0	-	-	-	-
859	Poza Rica de Hidalgo	Mexico	179.72	175.39	97.59	90	4	-	6	15
860	Pozna?	Poland	270.13	267.6	99.06	104	7	-	7	1
861	Praha (Prague)	Czech Republic	286.41	286.32	99.97	157	7	-	7	14
862	Preston	United Kingdom	136.54	136.28	99.81	55	7	-	7	20
863	Pretoria Soshanguve	South Africa	1,461.36	1440.99	98.61	967	6	-	6	19
864	Providence	United States of America	331.50	331.3	99.94	173	2	-	7	6

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}^Z$
865	Provo-Orem	United States of America	529.57	213.15	40.25	100	2	-	7	6
866	Puducherry	India	133.85	123.54	92.30	50	2	-	6	7
867	Puebla Tlaxcala	Mexico	1,235.48	1017.4	82.35	713	4	-	6	19
868	Pulandian	China	98.81	75.84	76.76	33	1	-	1	26
869	Pune (Poona)	India	823.65	823.26	99.95	590	1	1	1	5
870	Punto Fijo	Venezuela (Bolivarian Republic of)	239.83	205.31	85.60	101	3	-	3	28
871	Purnia	India	105.17	70.31	66.85	20	6	-	4	15
872	Putian	China	905.84	364.24	40.21	209	1	-	5	23
873	Puyang	China	217.46	217.31	99.93	105	1	-	1	2
874	P'yongyang	Dem. People's Republic of Korea	379.59	379.46	99.96	228	5	-	1	2
875	Qacentina	Algeria	325.73	172.34	52.91	88	3	-	4	17
876	Qazvin	Iran (Islamic Republic of)	117.66	112.63	95.73	42	6	-	3	27
877	Qinhuangdao	China	273.81	273.66	99.95	147	1	-	1	26
878	Qinzhou	China	146.21	140.69	96.22	68	5	-	1	2
879	Qiqihaer	China	218.74	218.27	99.78	113	1	-	1	5
880	Qom	Iran (Islamic Republic of)	250.12	230.4	92.12	123	3	-	3	8
881	Québec	Canada	496.18	484.49	97.64	236	7	-	7	6
882	Quetta	Pakistan	266.61	265.31	99.51	140	3	-	3	8
883	Quito	Ecuador	836.97	549.65	65.67	371	2	-	2	19
884	Qujing	China	238.91	200.54	83.94	96	5	-	1	2
885	Quzhou	China	266.04	36.19	13.60	10	1	-	1	2
886	Rabat	Morocco	304.90	272.83	89.48	133	7	-	7	14
887	Rahim Yar Khan	Pakistan	127.01	127.02	100.01	55	6	-	4	17
888	Rajkot	India	218.77	217.57	99.45	120	3	-	1	5

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
889	Rajshahi	Bangladesh	118.68	82.66	69.65	35	2	-	2	25
890	Ramadi	Iraq	167.05	145	86.80	56	3	-	3	8
891	Rampur	India	104.64	86.19	82.36	33	4	-	4	15
892	Rangpur	Bangladesh	107.21	24.48	22.83	4	2	-	2	18
893	Rasht	Iran (Islamic Republic of)	157.29	157.14	99.90	73	4	-	6	15
894	Raurkela	India	203.50	109.53	53.82	51	2	-	4	1
895	Rayong	Thailand	351.15	187.8	53.48	93	6	-	1	5
896	Reading-Wokingham	United Kingdom	114.97	43.63	37.95	11	7	-	7	20
897	Recife	Brazil	545.19	463.95	85.10	293	4	-	7	19
898	Rennes	France	130.22	130.21	99.99	47	7	-	7	1
899	Reno	United States of America	524.73	524.5	99.96	340	6	-	7	27
900	Renqiu	China	184.74	184.52	99.88	74	1	-	1	5
901	Richmond	United States of America	282.69	266.35	94.22	94	2	-	7	6
902	Riga	Latvia	204.54	186.77	91.31	76	5	-	5	11
903	Rio Branco	Brazil	237.03	198.06	83.56	108	2	-	4	7
904	Rio de Janeiro	Brazil	1,510.37	277	18.34	110	4	2	6	3
905	Rizhao	China	642.88	377.42	58.71	183	1	-	1	2
906	Rochester	United States of America	439.28	434.74	98.97	265	2	-	7	6
907	Rockford	United States of America	348.54	282.79	81.14	162	2	-	7	6
908	Rohtak	India	139.92	114.72	81.99	54	4	-	4	17
909	Roma (Rome)	Italy	840.11	834.69	99.36	559	7	7	1	1
910	Rongcheng	China	219.73	124.01	56.44	52	1	-	1	12
911	Rosario	Argentina	353.39	320.13	90.59	179	6	-	6	21

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
912	Rostov-na-Donu (Rostov-on-Don)	Russian Federation	369.58	369.26	99.91	222	7	-	1	1
913	Rotterdam s-Gravenhage (The Hague)	Netherlands	1,250.75	1067.45	85.34	733	7	-	7	1
914	Rouen	France	197.11	197.11	100.00	94	7	-	7	14
915	WenzhouWenzhou Ruian Yueqing	China	1,725.28	969.16	56.17	612	5	-	5	11
916	Rustenburg	South Africa	192.53	191.77	99.61	98	6	-	4	1
917	Ryazan	Russian Federation	238.01	3.9	1.64	0	-	-	-	-
918	Sacramento	United States of America	1,487.46	1353.95	91.02	1046	6	-	6	19
919	Safakis	Tunisia	326.19	311.73	95.57	170	7	-	7	20
920	Safi	Morocco	97.98	73.67	75.19	20	3	-	4	12
921	Sagar	India	97.94	83.93	85.70	34	4	-	4	17
922	Saint-Étienne	France	194.16	168.72	86.90	64	7	-	7	1
923	Sakarya	Turkey	158.46	158.34	99.92	71	6	-	4	10
924	Salem	India	199.17	199.04	99.94	99	6	-	2	18
925	Salerno	Italy	83.28	33.02	39.65	3	7	-	5	13
926	Salta	Argentina	257.25	185.72	72.19	89	6	-	6	21
927	Saltillo	Mexico	299.49	6.21	2.07	0	-	-	-	-
928	Salt Lake City	United States of America	949.88	889.38	93.63	651	6	-	6	19
929	Samara	Russian Federation	406.40	132.41	32.58	36	5	-	1	2
930	Samarinda	Indonesia	221.87	221.66	99.90	117	4	-	2	9
931	Samarkand	Uzbekistan	297.59	297.4	99.94	179	6	-	6	21
932	Samsun	Turkey	95.94	93.47	97.43	24	3	-	1	23
933	Sanandaj	Iran (Islamic Republic of)	166.79	135.5	81.24	66	3	-	3	22

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
934	San Antonio	United States of America	1,411.41	1410.28	99.92	1095	2	-	7	6
935	San Diego Tijuana	United States of America	1,841.38	1764.54	95.83	1271	6	-	6	19
936	Sangali	India	172.80	90.87	52.59	39	6	-	4	7
937	San Juan	Puerto Rico	619.29	31.97	5.16	2	5	-	1	26
938	San Juan	Argentina	264.13	263.72	99.85	148	6	-	6	10
939	Sankt Peterburg (Saint Petersburg)	Russian Federation	580.14	580.1	99.99	365	5	1	1	2
940	Sanliurfa	Turkey	157.94	144.68	91.60	51	3	-	1	27
941	San Miguel de Tucumán	Argentina	326.65	326.41	99.93	193	4	-	6	15
942	Sanning	China	125.30	24.44	19.50	4	5	-	5	16
943	San Salvador	El Salvador	300.50	281.17	93.57	157	5	-	5	3
944	San Salvador de Jujuy	Argentina	191.56	152.7	79.71	70	4	-	5	3
945	Santa Clarita	United States of America	258.65	258.46	99.93	143	3	-	6	19
946	Santa Cruz	Bolivia (Plurinational State of)	457.13	456.99	99.97	307	2	-	2	6
947	Santa Fe	Argentina	210.37	196.64	93.47	71	4	-	6	21
948	Santa Rosa	United States of America	362.49	336.25	92.76	196	2	-	2	18
949	Santiago	Chile	1,091.12	1090.81	99.97	824	6	2	6	21
950	Santiago	Dominican Republic	228.38	228.35	99.99	123	4	-	2	9
951	Santiago de Cuba	Cuba	134.19	125.56	93.57	58	4	-	2	9
952	Santiago Del Estero	Argentina	233.78	228.25	97.64	108	6	-	2	7
953	Santipur	India	93.55	36.71	39.24	8	2	-	2	18
954	Santo Domingo	Dominican Republic	464.30	459.48	98.96	275	4	-	6	19
955	Santo Domingo	Ecuador	157.51	144.98	92.04	75	4	-	2	9
956	Sanya	China	236.49	212.05	89.66	79	5	-	1	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
957	São José do Rio Preto	Brazil	245.29	245.14	99.94	142	6	-	6	10
958	São José dos Campos	Brazil	395.13	179.86	45.52	86	6	-	6	15
959	São Paulo	Brazil	2,521.03	2520.81	99.99	2061	6	2	6	21
960	Sapporo	Japan	515.87	515.82	99.99	326	5	-	5	3
961	Sarajevo	Bosnia and Herzegovina	188.02	143.03	76.07	66	7	-	2	13
962	Sarasota-Bradenton	United States of America	878.88	830.51	94.50	507	2	-	7	6
963	Saratov	Russian Federation	354.01	279.25	78.88	85	7	-	1	2
964	Sargodha	Pakistan	144.84	141.39	97.62	62	4	-	4	15
965	Sari	Iran (Islamic Republic of)	121.43	120.89	99.56	53	6	-	4	15
966	Satna	India	152.36	114.4	75.09	58	6	-	4	17
967	Scranton	United States of America	226.43	157.29	69.46	72	2	-	7	6
968	Seattle	United States of America	1,391.90	949.85	68.24	555	2	-	7	6
969	Sekondi Takoradi	Ghana	200.16	33.9	16.94	4	4	-	2	25
970	Semarang	Indonesia	414.93	414.81	99.97	251	6	-	6	21
971	Semipalatinsk	Kazakhstan	189.18	4.16	2.20	0	-	-	-	-
972	Sendai	Japan	613.13	612.92	99.97	396	5	-	5	3
973	Seremban	Malaysia	330.23	328.9	99.60	191	5	-	5	3
974	Sevastopol	Ukraine	166.37	157.29	94.54	50	7	-	2	23
975	Sevilla	Spain	442.45	410.22	92.72	252	3	-	4	12
976	Shahjahanpur	India	138.11	55.86	40.45	9	4	-	4	17
977	Shangqiu	China	369.68	285.83	77.32	154	1	-	1	26
978	Shangrao	China	210.35	174.69	83.05	89	5	-	5	23
979	Shaoguan	China	314.18	195.58	62.25	76	5	-	5	11
980	Shaoyang	China	206.77	179.99	87.05	83	5	-	2	23
981	Sheffield	United Kingdom	354.86	299.38	84.36	173	7	-	7	20

Continued on next page

ID	City Name(s)	Country	MUA Area (<i>km</i> ²)	LCZs Area (<i>km</i> ²)	LCZs Coverage (%)	No. Patches	<i>C</i> _{HM1523}	<i>C</i> _{HM110}	<i>C</i> _{<i>F</i>₁₅₂₃⁷}	
982	Shengzhou	China	176.37	175.99	99.79	81	5	-	5	23
983	Shenyang	China	917.54	917.29	99.97	672	1	-	1	2
984	Shihezi	China	200.42	181.19	90.40	96	1	-	1	26
985	Shijiazhuang	China	883.02	882.87	99.98	591	1	-	1	26
986	Shillong	India	123.41	112.49	91.15	48	4	-	5	13
987	Shimkent	Kazakhstan	399.68	382.65	95.74	241	6	-	6	7
988	Shimoga	India	124.74	87.52	70.16	34	6	-	4	17
989	Shiraz	Iran (Islamic Republic of)	391.96	391.66	99.92	234	3	-	3	8
990	Shiyan	China	219.11	170.55	77.84	83	5	-	5	16
991	Shizuishan	China	188.29	77.39	41.10	24	3	-	3	22
992	Shizuoka-Hamamatsu	M.M.A. Japan	531.96	277.86	52.23	140	5	-	5	3
993	Shreveport	United States of America	456.23	428.29	93.88	274	2	-	7	1
994	Shuangyashan	China	174.26	70.87	40.67	24	5	-	1	5
995	Shuozhou	China	151.37	151.14	99.85	66	1	-	1	12
996	Sialkot	Pakistan	207.30	188.93	91.14	93	4	-	4	15
997	Sihui	China	122.36	122.32	99.97	46	5	-	1	5
998	Siliguri	India	197.10	196.71	99.80	94	4	-	2	18
999	Simferopol	Ukraine	188.17	188.01	99.91	81	7	-	7	1
1000	Sivas	Turkey	104.69	71.48	68.27	19	3	-	4	27
1001	Skopje	TFYR Macedonia	222.04	199.68	89.93	103	7	-	4	1
1002	Smolensk	Russian Federation	127.39	125.15	98.24	56	5	-	2	23
1003	Sochi	Russian Federation	91.41	89.23	97.61	19	5	-	5	13
1004	Sofia	Bulgaria	343.75	343.65	99.97	205	7	-	1	1
1005	Sokoto	Nigeria	229.52	228.86	99.71	124	4	-	4	15
1006	Solapur	India	185.19	158.86	85.78	79	4	-	4	17
1007	Songyuan	China	69.32	12.24	17.66	0	-	-	-	-

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1008	Southampton/Portsmouth (South Hampshire)	United Kingdom	443.58	287.31	64.77	126	7	-	7	20
1009	Spokane	United States of America	448.47	330.84	73.77	188	2	-	7	6
1010	Springfield, Massachusett, Connecticut	United States of America	297.88	294.92	99.01	158	2	-	7	6
1011	Springfield, Missouri	United States of America	356.49	356.42	99.98	213	2	-	2	6
1012	Srinagar	India	240.20	239.99	99.91	112	4	-	6	19
1013	Stavropol	Russian Federation	282.92	282.79	99.95	148	7	-	5	1
1014	St. Catharines-Niagara	Canada	181.30	181.26	99.98	85	2	-	7	1
1015	St. Louis	United States of America	384.39	341.71	88.90	208	2	-	7	6
1016	Stockholm	Sweden	278.85	212.29	76.13	48	7	-	7	14
1017	Stoke-on-Trent (The Potteries)	United Kingdom	198.60	198.34	99.87	91	7	-	7	20
1018	Strasbourg	France	242.77	242.35	99.83	132	7	-	7	1
1019	Stuttgart	Germany	220.44	220.33	99.95	79	7	-	7	1
1020	Sucre	Bolivia (Plurinational State of)	128.80	94.53	73.40	40	4	-	2	9
1021	Suihua	China	85.17	85.02	99.82	29	1	-	1	26
1022	Suining, Sichuan	China	151.77	117.72	77.57	51	5	-	5	11
1023	Suizhou	China	127.20	127.21	100.01	51	5	-	5	23
1024	Sukabumi	Indonesia	186.39	163.03	87.47	75	6	-	2	25
1025	Sukkur	Pakistan	127.85	5.98	4.68	0	-	-	-	-
1026	Sulaimaniya	Iraq	240.14	220.53	91.83	115	3	-	6	27
1027	Sunshine Coast	Australia	289.19	45.33	15.67	11	2	-	7	6
1028	Suqian	China	611.91	414.89	67.80	234	1	-	1	5
1029	Surakarta	Indonesia	286.06	285.88	99.94	156	6	-	6	19

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1030	Surat	India	628.54	491.26	78.16	304	3	5	4	12
1031	Surgut	Russian Federation	110.89	96.25	86.80	37	5	-	1	26
1032	Suzhou, Anhui	China	154.48	153.08	99.10	75	1	-	1	5
1033	Sydney	Australia	1,638.04	1390.59	84.89	1019	2	2	7	6
1034	Sylhet	Bangladesh	129.15	127.36	98.61	56	2	-	2	18
1035	Syracuse	United States of America	266.74	244.13	91.52	109	2	-	7	6
1036	Szczecin	Poland	141.36	138.99	98.32	49	7	-	5	14
1037	Tabriz	Iran (Islamic Republic of)	257.52	252.19	97.93	130	3	-	3	8
1038	Tabuk	Saudi Arabia	259.42	258.83	99.77	135	3	-	3	22
1039	Taichung	China	1,403.08	1058.76	75.46	792	5	-	1	26
1040	Taif	Saudi Arabia	275.38	258.69	93.94	138	3	-	3	22
1041	Taipei	China	575.27	575.23	99.99	365	5	-	5	11
1042	Taizhou, Jiangsu	China	506.83	335.62	66.22	177	1	-	1	5
1043	Taizhou, Zhejiang Wenling	China	1,327.51	1148.42	86.51	832	5	-	1	2
1044	Ta'izz	Yemen	159.76	135.07	84.54	62	3	-	3	22
1045	Tallinn	Estonia	155.90	84.35	54.10	8	7	-	5	14
1046	Tamale	Ghana	182.56	177.07	96.99	93	4	-	2	25
1047	Tampa-St. Petersburg	United States of America	1,449.20	1134.24	78.27	721	2	-	7	6
1048	Tampere	Finland	121.29	78.16	64.44	12	7	-	5	14
1049	Tampico	Mexico	338.72	133.8	39.50	51	6	-	6	21
1050	Tangshan, Hebei	China	1,019.81	990.44	97.12	603	1	-	4	12
1051	Tarabulus (Tripoli)	Libya	868.48	693.64	79.87	499	7	-	2	25
1052	Taraz	Kazakhstan	204.33	204.11	99.89	110	6	-	6	10
1053	Tartus	Syrian Arab Republic	89.26	65	72.82	13	7	-	1	23
1054	Tashauz	Turkmenistan	131.37	125.81	95.77	55	3	-	1	23
1055	Tashkent	Uzbekistan	799.12	798.88	99.97	573	6	-	6	10

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1056	Tasikmalaya	Indonesia	165.28	165.17	99.93	77	6	-	2	25
1057	Tbilisi	Georgia	380.47	95.68	25.15	41	5	-	1	2
1058	Teeside (Middlesbrough)	United Kingdom	259.04	140.91	54.40	60	7	-	7	20
1059	Tegucigalpa	Honduras	249.82	249.67	99.94	143	4	-	5	3
1060	Tehuacán	Mexico	255.16	151.37	59.32	84	6	-	6	19
1061	Temecula-Murrieta	United States of America	326.31	286.67	87.85	152	2	-	6	19
1062	Temuco	Chile	145.29	94.75	65.22	38	4	-	6	3
1063	Tengzhou	China	205.82	205.54	99.86	102	1	-	1	26
1064	Tepic	Mexico	189.09	159.19	84.19	82	6	-	6	10
1065	Teresina	Brazil	375.10	374.85	99.93	238	6	-	6	19
1066	Tétouan	Morocco	177.80	79.38	44.65	31	3	-	4	27
1067	Thessaloniki	Greece	273.85	242.89	88.70	90	3	-	1	2
1068	Thoothukkudi (Tuticorin)	India	129.61	120.54	93.00	48	2	-	4	7
1069	Tianjin	China	1,351.56	1351.49	99.99	1018	1	1	1	26
1070	Tianmen	China	107.59	99.84	92.80	43	1	-	1	5
1071	Tianshui	China	224.51	78.16	34.81	30	5	-	1	2
1072	Tieling	China	139.10	119.46	85.88	46	1	-	1	5
1073	Timisoara	Romania	153.24	153.15	99.94	68	7	-	4	1
1074	Tiranë (Tirana)	Albania	178.32	173.27	97.17	89	2	-	2	18
1075	Tirunelveli	India	268.84	39.27	14.61	10	6	-	4	17
1076	Toamasina	Madagascar	140.32	24.89	17.74	4	4	-	2	25
1077	Tokyo	Japan	5,088.01	5087.43	99.99	4291	6	2	6	24
1078	Toledo	United States of America	594.84	519.57	87.35	333	2	-	7	6
1079	Tomsk	Russian Federation	156.94	144.16	91.86	64	5	-	5	11
1080	Tongchuan	China	172.75	104.58	60.54	49	1	-	1	5
1081	Tonghua	China	168.02	81.01	48.21	30	5	-	5	16
1082	Tongliao	China	219.08	218.84	99.89	105	1	-	1	5

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1083	Tongling	China	244.23	244.03	99.92	124	5	-	1	2
1084	Tongxiang	China	191.91	191.82	99.96	88	1	-	1	5
1085	Torino (Turin)	Italy	336.60	330.26	98.12	164	3	-	1	5
1086	Toulon	France	262.54	145.34	55.36	49	7	-	7	6
1087	Toulouse	France	382.69	382.22	99.88	208	7	-	7	20
1088	Tours	France	170.81	170.56	99.85	55	7	-	7	1
1089	Trujillo	Peru	228.50	226.97	99.33	118	4	-	1	4
1090	Tshikapa	Democratic Republic of the Congo	146.11	114.39	78.29	49	6	-	4	17
1091	Tucson	United States of America	1,026.13	1025.72	99.96	706	1	-	6	19
1092	Tula	Russian Federation	250.31	3.78	1.51	0	-	-	-	-
1093	Tulsa	United States of America	723.32	721.98	99.81	504	2	-	7	6
1094	Tumkur	India	144.36	92.66	64.19	34	2	-	4	7
1095	Tunis	Tunisia	566.38	561.1	99.07	292	3	-	1	27
1096	Tuxtla Gutierrez	Mexico	261.23	254.08	97.26	136	4	-	6	3
1097	Tver	Russian Federation	156.80	156.75	99.97	71	5	-	1	23
1098	Tyumen	Russian Federation	239.48	228.02	95.21	118	5	-	1	2
1099	Überaba	Brazil	206.97	206.92	99.98	108	6	-	4	15
1100	Überlândia	Brazil	329.05	328.81	99.93	196	6	-	6	15
1101	Udaipur	India	232.46	206.86	88.99	117	6	-	2	18
1102	Udon Thani	Thailand	163.82	158.23	96.59	71	6	-	2	18
1103	Ufa	Russian Federation	260.40	130.87	50.26	52	5	-	5	11
1104	Ulaanbaatar	Mongolia	343.23	319.81	93.18	186	6	-	6	27
1105	Ulan-Ude	Russian Federation	196.60	154.98	78.83	71	2	-	6	23
1106	Ulsan	Republic of Korea	324.78	305.72	94.13	161	5	-	5	16
1107	Ulyanovsk	Russian Federation	244.66	81.8	33.43	24	5	-	5	11
1108	Ürümqi (Wulumqi)	China	630.04	576.3	91.47	364	1	-	1	26

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1109	Ust-Kamenogorsk	Kazakhstan	178.85	130.97	73.23	48	7	-	1	1
1110	Utsunomiya	Japan	666.81	666.63	99.97	386	6	-	2	18
1111	Valencia	Venezuela (Bolivarian Republic of)	585.47	499.9	85.38	307	6	-	6	10
1112	Valencia	Spain	499.82	469.06	93.85	298	3	-	1	5
1113	Valenciennes	France	153.09	141.43	92.39	62	7	-	7	1
1114	Valladolid	Spain	173.46	172.34	99.35	81	3	-	1	12
1115	Valledupar	Colombia	149.42	143.52	96.05	65	6	-	6	9
1116	Valparaíso	Chile	223.46	21.77	9.74	2	6	-	6	21
1117	Vancouver	Canada	953.96	644.75	67.59	364	2	2	6	3
1118	Varanasi (Benares)	India	271.69	271.39	99.89	141	4	-	2	18
1119	Varna	Bulgaria	146.65	125.81	85.79	39	5	-	5	23
1120	Veracruz	Mexico	243.20	138.13	56.80	61	4	-	1	26
1121	Vereeniging	South Africa	397.49	300.1	75.50	155	6	-	4	17
1122	Verona	Italy	204.65	157.23	76.83	61	7	-	1	5
1123	Victorville-Hesperia-Apple Valley	United States of America	646.28	126.09	19.51	22	6	-	2	7
1124	Vientiane	Lao People's Democratic Republic	415.85	325.19	78.20	160	6	-	2	7
1125	Vijayawada	India	173.61	157.54	90.74	69	6	-	4	15
1126	Villahermosa	Mexico	232.38	155.85	67.07	86	4	-	1	5
1127	Villavicencio	Colombia	135.36	108.58	80.22	48	4	-	6	23
1128	Vilnius	Lithuania	168.17	168.07	99.94	77	5	-	5	11
1129	Vinnitsa	Ukraine	167.41	155.2	92.71	80	7	-	7	1
1130	Virginia Beach	United States of America	1,075.85	486.61	45.23	285	2	-	7	6
1131	Visakhapatnam	India	345.27	113.76	32.95	49	6	-	6	27
1132	Vitebsk	Belarus	157.30	139.46	88.66	60	7	-	7	1
1133	Vladimir	Russian Federation	141.39	136.88	96.81	53	5	-	1	1

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1134	Volgograd Volzhsky	Russian Federation	644.99	592.83	91.91	307	7	-	1	12
1135	Vologda	Russian Federation	120.98	109.37	90.40	33	5	-	1	2
1136	Volta Redonda	Brazil	227.54	181.63	79.82	88	6	-	4	15
1137	Voronezh	Russian Federation	318.09	4.16	1.31	0	-	-	-	-
1138	Wad Medani	Sudan	279.07	226.48	81.16	125	6	-	4	17
1139	Wafangdian	China	168.31	132.59	78.78	51	1	-	1	5
1140	Wah	Pakistan	231.03	103.16	44.65	50	4	-	4	15
1141	Wahran (Oran)	Algeria	275.03	120.55	43.83	60	3	-	1	5
1142	Warri	Nigeria	293.20	285.62	97.41	150	5	-	5	3
1143	Warszawa (Warsaw)	Poland	421.61	421.23	99.91	262	7	7	7	14
1144	Weihai	China	330.80	120.63	36.47	31	5	-	5	11
1145	Weinan	China	166.62	119.1	71.48	51	1	-	1	5
1146	West Yorkshire	United Kingdom	340.12	205.01	60.28	94	7	-	7	20
1147	Wichita	United States of America	563.08	551.55	97.95	369	2	-	7	1
1148	Wien (Vienna)	Austria	421.84	421.61	99.95	216	7	7	7	14
1149	Windhoek	Namibia	273.70	260.17	95.06	139	4	-	4	27
1150	Winston-Salem	United States of America	417.41	321.78	77.09	212	2	-	2	13
1151	Witbank	South Africa	300.92	116.99	38.88	55	7	-	4	27
1152	Worcester	United States of America	85.32	85.12	99.76	24	2	-	7	6
1153	Wroc'aw	Poland	170.61	170.63	100.01	77	7	-	7	14
1154	Wuhai	China	518.93	145.73	28.08	57	1	-	3	28
1155	Wuwei	China	111.76	111.39	99.66	45	1	-	1	26
1156	Wuzhou	China	154.11	87.47	56.76	37	5	-	5	11
1157	Xalapa	Mexico	188.55	188.32	99.88	89	2	-	2	13
1158	Xiangcheng	China	210.44	101.18	48.08	39	1	-	1	5
1159	Xiangyang	China	359.32	108.02	30.06	53	5	-	5	11

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1160	Xianning	China	186.07	103.41	55.57	44	5	-	5	11
1161	Xi'an, Shaanxi Xianyang, Shaanxi	China	1,411.39	1405.99	99.62	1090	1	-	1	26
1162	Xiantao	China	147.75	147.59	99.89	65	1	-	1	5
1163	Xinghua	China	162.13	160.11	98.76	77	1	-	1	5
1164	Xingning	China	121.20	119.41	98.52	46	5	-	2	23
1165	Xingtai	China	397.90	340.72	85.63	195	1	-	1	12
1166	Zhengzhou Xingyang	China	959.82	887.65	92.48	590	1	-	1	26
1167	Xingyi, Guizhou	China	139.88	85.59	61.19	36	5	-	5	11
1168	Xining	China	322.10	258.23	80.17	107	3	-	3	8
1169	Xinni	China	117.35	117.17	99.85	51	1	-	1	15
1170	Xintai	China	183.81	127.43	69.33	59	1	-	1	26
1171	Xinyang	China	184.47	184.1	99.80	97	5	-	5	11
1172	Xinyi	China	93.53	84.03	89.84	33	5	-	5	9
1173	Xinyu	China	233.58	229.15	98.10	120	5	-	5	11
1174	Xinzheng	China	156.08	108.64	69.61	49	1	-	1	12
1175	Xinzhou	China	115.61	115.6	99.99	43	1	-	1	12
1176	Xuancheng	China	124.54	98.1	78.77	40	5	-	1	2
1177	Xuzhou	China	398.45	398.24	99.95	224	1	-	1	26
1178	Yakutsk	Russian Federation	152.33	148.43	97.44	57	2	-	2	23
1179	Yanan	China	142.18	5.4	3.80	0	-	-	-	-
1180	Yancheng, Jiangsu	China	364.31	357.22	98.05	216	1	-	1	5
1181	Yangon	Myanmar	629.44	589.49	93.65	402	6	2	6	19
1182	Yangquan	China	216.23	200.29	92.63	87	1	-	1	27
1183	Yangzhou	China	482.19	451.49	93.63	258	1	-	1	2
1184	Yanji	China	142.13	141.33	99.43	62	1	-	1	5
1185	Yantai	China	699.00	268.76	38.45	122	5	-	1	2
1186	Yaoundé	Cameroon	389.05	388.83	99.94	246	2	-	2	19
1187	Yaroslavl	Russian Federation	179.80	176.51	98.17	71	7	-	5	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1188	Yazd	Iran (Islamic Republic of)	296.25	296.29	100.01	167	3	-	3	29
1189	Yekaterinburg	Russian Federation	327.45	327.11	99.90	164	5	-	5	11
1190	Yerevan	Armenia	336.04	335.99	99.98	201	1	-	1	27
1191	Yibin	China	182.60	117.71	64.47	42	5	-	5	11
1192	Yichang	China	259.21	242.73	93.64	104	5	-	5	11
1193	Yichun, Heilongjiang	China	167.50	17.68	10.55	0	-	-	-	-
1194	Yichun, Jiangxi	China	187.13	143.84	76.87	65	5	-	5	11
1195	Yinchuan	China	451.57	367.67	81.42	181	1	-	1	26
1196	Yingde	China	118.85	24.48	20.60	4	1	-	1	2
1197	Yining	China	199.69	199.66	99.99	96	1	-	1	10
1198	Yogyakarta	Indonesia	324.00	323.93	99.98	195	6	-	2	25
1199	Yongcheng	China	158.55	134.9	85.08	44	1	-	4	12
1200	Yongin	Republic of Korea	89.40	89.28	99.86	20	5	-	5	16
1201	Yongzhou	China	161.79	95.65	59.12	42	5	-	1	23
1202	Youngstown	United States of America	456.32	310.77	68.10	155	2	-	2	25
1203	Yulin, Guangxi	China	238.98	238.76	99.91	126	5	-	1	23
1204	Yulin, Shaanxi	China	180.14	177.38	98.47	86	3	-	3	8
1205	Yushu	China	75.74	61.94	81.78	18	1	-	1	12
1206	Yuxi	China	207.83	193.02	92.88	85	5	-	1	2
1207	Zacatecas	Mexico	206.69	85.63	41.43	31	4	-	6	9
1208	Zagreb	Croatia	245.74	237.93	96.82	120	7	-	7	1
1209	Zahedan	Iran (Islamic Republic of)	194.36	194.15	99.89	106	4	-	3	8
1210	Zanjan	Iran (Islamic Republic of)	162.58	128.59	79.09	57	3	-	3	8
1211	Zaoyang	China	124.73	88.45	70.92	37	1	-	1	5
1212	Zaozhuang	China	224.86	210.13	93.45	103	1	-	1	5

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1213	Zaporizhzhya	Ukraine	292.22	270.19	92.46	146	7	-	7	1
1214	Zaragoza	Spain	209.86	190.5	90.77	77	1	-	1	27
1215	Zaria	Nigeria	259.32	212.05	81.77	112	4	-	4	17
1216	Zengcheng	China	135.51	135.34	99.88	62	5	-	5	11
1217	Zhangjiakou	China	416.78	196.2	47.07	79	1	-	1	5
1218	Zhangye	China	96.82	96.6	99.77	38	1	-	1	5
1219	Zhaodong	China	87.48	55.69	63.66	15	1	-	1	12
1220	Zhenjiang, Jiangsu	China	274.91	263.73	95.93	126	5	-	1	2
1221	Zhongxiang	China	117.55	102.25	86.98	47	1	-	1	5
1222	Zhoukou	China	201.88	201.72	99.92	88	1	-	1	26
1223	Zhoushan	China	289.24	99.58	34.43	34	5	-	5	11
1224	Zhuji	China	305.45	272.15	89.10	145	5	-	5	11
1225	Zhumadian	China	292.83	204.76	69.93	96	1	-	1	26
1226	Zibo	China	762.87	732.06	95.96	451	1	-	1	26
1227	Zigong	China	223.41	183.56	82.16	96	5	-	5	23
1228	Zinder	Niger	128.76	98.37	76.40	48	4	-	2	28
1229	Ziyang	China	125.14	103.14	82.42	42	5	-	5	23
1230	Zunyi	China	188.04	162.66	86.50	79	5	-	5	16
1231	Minneapolis-St. Paul	United States of America	645.73	645.35	99.94	332	7	-	7	6
1232	Xiaogan	China	57.89	57.73	99.73	14	1	-	1	2
1233	Merida	Venezuela (Bolivarian Republic of)	145.32	39.78	27.37	8	5	-	5	13
1234	San José	Costa Rica	605.91	603.92	99.67	380	4	-	6	19
1235	Cordoba	Spain	145.66	125.22	85.97	50	3	-	4	12
1236	Jining, Shandong	China	214.56	204.61	95.36	109	1	-	1	26
1237	Douai-Lens	France	115.87	115.93	100.05	39	7	-	4	1
1238	Zhaoqing	China	237.49	100.6	42.36	25	5	-	5	11
1239	Qingyuan	China	213.98	148.52	69.41	57	5	-	1	2

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1240	Guangzhou, Guangdong Shenzhen Dongguan Foshan Zhongshan Jiangmen Heshan	China	10,123.17	9792.53	96.73	8214	5	5	1	24
1241	Donggang	China	133.18	128.56	96.53	41	1	-	1	12
1242	Djibouti	Djibouti	135.28	30.75	22.73	3	4	-	1	26
1243	Dhule	India	152.60	149.42	97.91	71	6	-	4	17
1244	Dewas	India	143.03	24.48	17.12	4	6	-	4	15
1245	Ujjain	India	126.02	81.81	64.92	31	4	-	4	15
1246	Cuernavaca	Mexico	481.84	391.99	81.35	257	2	-	2	7
1247	San Cristóbal	Venezuela (Bolivarian Republic of)	192.03	142.93	74.43	66	4	-	6	7
1248	Cotabato	Philippines	108.22	27.04	24.99	4	4	-	2	18
1249	Córdoba	Mexico	139.69	112.08	80.23	52	6	-	6	18
1250	Conakry	Guinea	340.87	157.14	46.10	48	4	-	6	21
1251	Haining	China	172.55	156.69	90.81	75	1	-	1	5
1252	Chongjin	Dem. People's Republic of Korea	118.80	69.76	58.72	24	1	-	1	27
1253	Choloma	Honduras	126.65	54.14	42.75	6	4	-	5	9
1254	Chi'in?u	Republic of Moldova	230.59	225.6	97.84	127	5	-	1	1
1255	Dujiangyan	China	165.65	154.76	93.43	67	5	-	5	23
1256	Chiclayo	Peru	178.55	176.36	98.78	91	4	-	4	12
1257	Cheongju	Republic of Korea	210.81	187.4	88.90	88	5	-	5	11
1258	Deyang	China	224.99	224.93	99.97	120	1	-	1	5
1259	Puning	China	712.85	252.94	35.48	132	5	-	5	11
1260	Daejeon	Republic of Korea	323.81	266.98	82.45	165	5	-	5	16
1261	Chengde	China	74.26	68.94	92.83	19	5	-	5	16
1262	Shantou Chaozhou Jieyang	China	1,930.58	1601.66	82.96	1081	1	-	1	2
1263	Maracaibo	Venezuela (Bolivarian Republic of)	589.90	507.58	86.05	298	4	-	6	19

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM123}	C_{HM110}	C_{F1523}^*	C_{F1523}
1264	Suzhou, Jiangsu Wuxi, Jiangsu Changzhou, Jiangsu Jiangyin Zhangjiagang Jingjiang	China	6,142.75	6086.07	99.08	4798	5	-	1	24
1265	Yixing	China	287.27	108.8	37.87	51	5	-	1	2
1266	Cabimas Lagunillas	Venezuela (Bolivarian Republic of)	442.14	197.6	44.69	88	7	-	4	15
1267	Denpasar	Indonesia	339.27	323.36	95.31	140	4	-	2	25
1268	Hardwar	India	385.80	22.79	5.91	3	5	-	5	16
1269	Davao City	Philippines	262.50	159.92	60.92	63	4	-	6	19
1270	Yingkou	China	253.44	253.39	99.98	130	1	-	1	26
1271	Dashiqiao	China	198.69	192.54	96.91	67	1	-	1	5
1272	Haicheng	China	144.28	139.89	96.96	52	1	-	1	12
1273	Pingdingshan, Henan	China	238.21	232.08	97.43	110	1	-	1	2
1274	Xinyi	China	143.08	142.9	99.87	63	1	-	1	5
1275	Puerto Vallarta	Mexico	118.50	65.3	55.10	18	2	-	5	13
1276	Qitalhe	China	230.41	46.5	20.18	6	1	-	1	26
1277	Taixing	China	203.14	195.56	96.27	88	1	-	1	5
1278	Changwon	Republic of Korea	265.55	184.9	69.63	84	5	-	5	16
1279	Liuyang	China	135.68	103.45	76.25	47	5	-	5	16
1280	Xiangtan, Hunan	China	419.32	404.55	96.48	245	5	-	1	23
1281	Xuchang	China	197.33	197.26	99.96	103	1	-	1	26
1282	Patiala	India	196.83	196.07	99.62	102	6	-	4	17
1283	Celaya	Mexico	184.37	184.17	99.89	97	6	-	4	15
1284	Querétaro	Mexico	411.05	410.82	99.94	238	6	-	6	19
1285	Cardiff	United Kingdom	219.79	217.59	99.00	89	7	-	7	20
1286	Cape Town	South Africa	855.88	60.85	7.11	20	5	5	6	2
1287	Cangzhou	China	190.27	190.14	99.93	93	1	-	1	10
1288	Cagayan de Oro City	Philippines	146.12	141.23	96.66	57	4	-	2	9

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM1523}	C_{HM110}	C_{F1523}^*	C_{F1523}
1289	Butuan	Philippines	102.84	26.52	25.79	4	4	-	2	18
1290	Gimhae	Republic of Korea	653.03	460.62	70.54	233	5	-	5	16
1291	Gumi	Republic of Korea	202.10	103.19	51.06	38	5	-	5	16
1292	Bursa	Turkey	299.72	142.22	47.45	59	5	-	6	3
1293	Buenaventura	Colombia	161.12	13.38	8.30	0	-	-	-	-
1294	Bridgeport-Stamford New Haven	United States of America	667.78	591.14	88.52	279	2	-	2	6
1295	Modena	Italy	130.28	127.17	97.61	56	7	-	4	12
1296	Blumenau	Brazil	267.81	105.67	39.46	49	2	-	7	13
1297	Bologna	Italy	198.32	185.8	93.69	86	7	-	1	5
1298	Blida	Algeria	158.24	157.88	99.77	70	3	-	4	27
1299	Bihar Sharif	India	115.44	36.57	31.68	6	4	-	4	17
1300	Vungtau	Viet Nam	120.87	53.62	44.36	11	6	-	6	21
1301	Bhubaneswar	India	242.75	91.72	37.78	40	2	-	2	7
1302	B��thune	France	70.80	69.16	97.68	22	7	-	4	7
1303	Hubli-Dharwad	India	375.70	99.73	26.55	46	2	-	4	17
1304	Beira	Mozambique	149.02	51.3	34.42	14	2	-	6	19
1305	Beihai	China	194.94	190.12	97.53	73	1	-	1	2
1306	Bazhong	China	126.05	22.95	18.21	4	5	-	5	23
1307	Bayannaoer	China	143.43	143.36	99.95	63	1	-	1	27
1308	Batangas City	Philippines	164.71	80.52	48.88	24	6	-	2	18
1309	Zamboanga City	Philippines	164.16	88.68	54.02	25	4	-	2	19
1310	Santa Marta	Colombia	137.37	135.36	98.54	48	6	-	6	10
1311	Bardhaman	India	134.64	132.3	98.26	60	6	-	4	17
1312	Bokaro Steel City	India	306.52	6.24	2.04	0	-	-	-	-
1313	Dhanbad	India	252.39	104.33	41.34	50	6	-	4	15
1314	Cumana	Venezuela (Bolivarian Republic of)	195.38	6.76	3.46	0	-	-	-	-
1315	Bandar Lampung	Indonesia	321.93	307.61	95.55	172	6	-	2	18

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1316	Bahia Blanca	Argentina	175.55	173.56	98.87	77	4	-	6	15
1317	Mbouda	Cameroon	106.49	51.56	48.42	14	2	-	2	25
1318	Bacolod	Philippines	193.11	160.88	83.31	80	4	-	4	7
1319	Marseille-Aix-en-Provence	France	82.14	82	99.83	27	2	-	2	13
1320	Asansol	India	280.50	94.96	33.85	40	6	-	4	15
1321	Arak	Iran (Islamic Republic of)	168.03	158.05	94.06	76	3	-	3	8
1322	Gent	Belgium	186.86	186.66	99.89	86	7	-	7	14
1323	Bruxelles-Brussel	Belgium	260.43	260.41	99.99	137	7	7	7	14
1324	Antofagasta	Chile	129.98	28.38	21.83	0	-	-	-	-
1325	Qingzhou	China	227.41	227.26	99.93	117	1	-	1	26
1326	Shouguang	China	1,169.68	1089.29	93.13	680	1	-	1	12
1327	Dongying	China	600.13	165.36	27.55	66	1	-	1	2
1328	Tongcheng	China	131.70	102.62	77.92	42	1	-	1	5
1329	Annaba	Algeria	170.56	63.02	36.95	11	7	-	1	1
1330	Angeles City	Philippines	166.51	166.46	99.97	84	4	-	4	15
1331	Ambon	Indonesia	114.09	86.47	75.79	25	4	-	2	9
1332	Taian, Shandong	China	383.79	352.47	91.84	198	5	-	1	2
1333	Qufu	China	115.10	111.97	97.28	46	1	-	1	10
1334	Yongkang	China	395.24	241.25	61.04	110	5	-	1	2
1335	Jingzhou, Hubei	China	321.77	292.31	90.85	140	1	-	1	5
1336	Jhansi	India	190.23	186.19	97.88	96	6	-	4	17
1337	Jerusalem	Israel	374.92	339.99	90.68	194	3	-	7	14
1338	Jeonju	Republic of Korea	190.87	177.6	93.05	87	5	-	5	11
1339	Jeju	Republic of Korea	149.23	120.82	80.96	46	2	-	2	23
1340	Jamnagar	India	174.54	168.78	96.70	79	3	-	4	12
1341	Iligan	Philippines	95.44	67.34	70.56	20	2	-	5	9
1342	Irapuato	Mexico	166.25	166.15	99.94	82	6	-	1	10
1343	Kollhapur	India	191.11	174.87	91.50	79	6	-	4	17

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1344	Oyo	Nigeria	186.35	185.93	99.78	93	4	-	2	18
1345	Hillah	Iraq	197.38	196.69	99.65	97	3	-	4	27
1346	Jinzhou	China	201.13	201.07	99.97	102	1	-	1	5
1347	Huludao	China	238.09	174.76	73.40	77	1	-	1	5
1348	Hanchuan	China	103.26	66.58	64.47	22	1	-	4	12
1349	Guarenas-Guatire	Venezuela (Bolivarian Republic of)	193.77	96.22	49.66	40	5	-	1	5
1350	Goma	Democratic Republic of the Congo	164.15	154.61	94.19	65	4	-	6	18
1351	Kahramanmaras	Turkey	117.85	95.01	80.62	42	3	-	6	27
1352	Zhucheng	China	272.67	272.58	99.97	138	1	-	1	10
1353	Gaomi	China	201.38	201.35	99.98	106	1	-	1	10
1354	Fengcheng	China	178.90	125.26	70.01	51	1	-	4	12
1355	Fangchenggang	China	93.54	67.55	72.21	24	5	-	5	11
1356	Erode	India	166.76	166.54	99.87	78	2	-	2	18
1357	Okpogho	Nigeria	131.40	6.76	5.14	0	-	-	-	-
1358	Edinburgh	United Kingdom	188.68	172.08	91.20	86	7	-	7	20
1359	Durg-Bhilainagar	India	292.22	137.1	46.92	63	6	-	4	17
1360	Durban	South Africa	933.91	343.49	36.78	225	2	-	7	6
1361	Lima	Peru	1,101.52	1081.07	98.14	740	4	-	3	29
1362	Lianyungang	China	253.94	241.39	95.06	123	1	-	1	5
1363	Libreville	Gabon	300.63	292.07	97.15	146	2	-	2	18
1364	Leshan	China	242.48	110.05	45.39	53	5	-	5	11
1365	Latina	Italy	85.09	84.95	99.83	29	7	-	4	7
1366	León de los Aldamas	Mexico	404.82	404.8	99.99	249	6	-	6	10
1367	Langfang	China	198.60	198.62	100.01	83	1	-	1	26
1368	Lancaster	United States of America	142.49	122.87	86.23	45	2	-	7	6
1369	Longkou	China	443.54	241.16	54.37	124	1	-	4	12

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	C_{HM1523}	C_{HM110}	C_{F1523}^*	C_{F1523}
1370	Zhaoyuan	China	136.46	130.23	95.43	60	1	-	1	12
1371	Laizhou	China	300.42	197.42	65.71	80	1	-	1	12
1372	Kuching	Malaysia	305.25	301.3	98.71	172	5	-	5	23
1373	Malappuram	India	658.08	6.76	1.03	0	-	-	-	-
1374	Korba	India	337.51	24.99	7.40	2	4	-	4	27
1375	Thrissur	India	196.84	196.89	100.02	96	2	-	2	25
1376	Kitwe	Zambia	365.69	209.71	57.35	115	2	6	2	7
1377	Kitakyushu-Fukuoka M.M.A.	Japan	802.96	778.61	96.97	460	5	-	5	3
1378	Kingston upon Hull	United Kingdom	184.00	169.25	91.99	73	7	-	7	20
1379	Khabarovsk	Russian Federation	270.47	253.07	93.57	123	5	-	1	5
1380	Mannheim	Germany	218.45	218.5	100.02	83	7	-	1	14
1381	Karimnagar	India	124.76	37.44	30.01	5	6	-	4	17
1382	Sorocaba	Brazil	311.76	311.43	99.89	181	6	-	6	21
1383	Joinville	Brazil	375.86	375.58	99.93	225	2	-	5	1
1384	Thanjavur	India	179.25	25.95	14.48	4	6	-	4	17
1385	Tirupati	India	125.73	26	20.68	4	6	-	4	15
1386	Tolyatti	Russian Federation	450.47	48.19	10.70	15	5	-	1	2
1387	Van	Turkey	159.84	159.79	99.97	75	2	-	2	18
1388	Victoria	Canada	270.16	174.92	64.75	75	7	-	7	6
1389	Vladikavkaz	Russian Federation	173.21	172.14	99.38	78	1	-	1	5
1390	Newport	United Kingdom	143.41	127.86	89.16	40	7	-	7	20
1391	Nha Trang	Viet Nam	243.38	118.97	48.88	51	4	-	6	15
1392	Nnewi	Nigeria	351.63	264.8	75.31	121	2	-	2	25
1393	Owerri	Nigeria	227.09	227.02	99.97	123	2	-	2	18
1394	Umuahia	Nigeria	262.17	102.83	39.22	49	2	-	2	13
1395	Oita	Japan	306.75	259.73	84.67	151	5	-	5	3
1396	Perth	Australia	1,211.33	1211.37	100.00	890	6	-	6	21
1397	Ribeirão Preto	Brazil	349.26	349.18	99.98	215	6	-	4	15
1398	Saharanpur	India	128.83	111.83	86.80	52	4	-	4	15

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1399	Yamunanagar	India	198.69	79.33	39.93	23	4	-	4	17
1400	Salvador	Brazil	538.73	186.6	34.64	94	4	-	6	3
1401	Sandakan	Malaysia	274.58	29.88	10.88	4	5	-	5	23
1402	Wuhu, Anhui	China	322.72	293.63	90.99	158	5	-	1	2
1403	Chizhou	China	117.51	95.83	81.55	38	1	-	4	12
1404	Jiaozuo	China	298.42	290.88	97.47	148	1	-	1	5
1405	Muzaffarnagar	India	109.13	35.7	32.71	5	4	-	4	15
1406	Makassar (Ujung Pandang)	Indonesia	357.15	238.93	66.90	129	4	-	6	4
1407	Shanwei	China	102.82	65.55	63.75	17	3	-	1	2
1408	Siping	China	160.06	160.05	100.00	69	1	-	1	26
1409	Southend-On-Sea	United Kingdom	266.44	50.12	18.81	9	7	-	7	20
1410	Swansea	United Kingdom	118.51	54.62	46.09	16	7	-	7	20
1411	Tanger	Morocco	206.38	132.26	64.09	59	3	-	1	27
1412	Treviso	Italy	328.65	162.15	49.34	52	7	-	4	7
1413	Utrecht	Netherlands	247.87	193.62	78.11	90	7	-	7	1
1414	Barcelona-Puerto La Cruz	Venezuela (Bolivarian Republic of)	296.31	6.76	2.28	0	-	-	-	-
1415	Bafoussam	Cameroon	176.84	152.03	85.97	86	2	-	2	25
1416	Barranquilla	Colombia	316.21	311.45	98.50	178	6	-	1	10
1417	Akure	Nigeria	287.54	287.35	99.93	158	2	-	2	25
1418	Port Harcourt	Nigeria	471.01	421.89	89.57	270	5	5	1	2
1419	Lagos Ikorodu	Nigeria	2,089.88	600.32	28.73	336	2	6	2	18
1420	Barquisimeto	Venezuela (Bolivarian Republic of)	314.17	290.61	92.50	144	4	-	6	10
1421	Abakaliki	Nigeria	154.23	119.6	77.55	52	6	-	4	17
1422	Ilorin	Nigeria	330.49	326.61	98.83	200	4	-	2	18
1423	Iloilo City	Philippines	188.39	75.56	40.11	25	4	-	7	19
1424	Kolkata (Calcutta)	India	1,201.10	1,200.73	99.97	882	4	6	2	25

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1425	Basilan City (including City of Isabela)	Philippines	80.28	6.76	8.42	0	-	-	-	-
1426	Singapore Johor Bahru	Singapore	1,905.60	1681.45	88.24	1273	5	5	1	2
1427	Belgaum	India	185.41	163.86	88.38	83	6	-	4	17
1428	Thành Pho Ho Chí Minh (Ho Chi Minh City)ThÃ nh Pho Ho ChÃ Minh (Ho Chi Minh City) Bien H	Viet Nam	1,487.06	1487.53	100.03	1181	6	2	6	10
1429	Patna	India	290.38	288.2	99.25	140	6	-	4	15
1430	Manila	Philippines	1,538.15	1263.56	82.15	901	4	4	6	21
1431	Ahmadabad	India	524.54	524.26	99.95	360	3	3	1	5
1432	Amravati	India	146.29	142.9	97.68	66	6	-	4	15
1433	Ranchi	India	214.40	214.32	99.96	108	4	-	2	18
1434	Avignon	France	199.64	126.08	63.15	43	7	-	7	1
1435	Antwerpen	Belgium	313.78	313.78	100.00	143	7	-	7	14
1436	Alicante	Spain	214.57	148.08	69.01	54	3	-	3	14
1437	Anqing	China	207.17	201.76	97.39	108	5	-	1	2
1438	Tanta	Egypt	122.59	121.39	99.02	54	3	-	4	17
1439	Agra	India	318.86	317.84	99.68	196	4	-	4	15
1440	Toluca de Lerdo	Mexico	397.45	396.98	99.88	243	6	-	6	19
1441	Orizaba	Mexico	147.49	145.42	98.60	63	4	-	5	13
1442	Hangzhou Shaoxing Cixi Yuyao Shangyu	China	4,799.01	4417.68	92.05	3417	1	-	1	24
1443	San Pedro Sula	Honduras	193.83	191.95	99.03	92	4	-	5	3
1444	Changge	China	135.60	134.11	98.90	48	1	-	1	12
1445	Changsha	China	909.09	905.27	99.58	641	5	1	5	23
1446	Zhuzhou	China	291.86	291.78	99.97	160	5	-	5	11
1447	Chengdu	China	1,368.34	1368.28	100.00	1062	5	-	1	2
1448	Reggio Emilia	Italy	100.58	100.28	99.70	38	7	-	4	12

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1449	Chandigarh	India	440.04	407.97	92.71	253	5	-	6	1
1450	Cali	Colombia	320.00	319.79	99.93	193	4	-	6	10
1451	Daegu	Republic of Korea	300.15	300.23	100.03	173	5	-	5	16
1452	Cúcuta	Colombia	322.86	279.81	86.67	166	6	-	6	15
1453	Pietermaritzburg	South Africa	354.95	294.07	82.85	165	2	-	2	7
1454	Enugu	Nigeria	260.00	259.78	99.92	145	6	-	4	15
1455	Ibadan	Nigeria	735.26	734.84	99.94	512	4	-	2	25
1456	Najaf	Iraq	250.33	216.94	86.66	120	3	-	3	8
1457	Gaziantep	Turkey	234.64	234.59	99.98	125	3	-	6	27
1458	Lille	France	397.62	381.51	95.95	222	7	-	7	14
1459	Ichalakaranji	India	142.64	99.55	69.79	47	6	-	4	17
1460	Wuhan	China	503.00	502.85	99.97	243	5	1	1	2
1461	Qingdao Jimo Jiaozhou	China	2,320.77	1842.16	79.38	1239	1	-	1	5
1462	Dehradun	India	223.14	223.19	100.02	120	2	-	2	13
1463	Roorkee	India	109.07	68.4	62.71	32	4	-	4	15
1464	Malegaon	India	123.89	91.97	74.23	42	6	-	4	17
1465	Indore	India	307.52	300.62	97.76	180	4	-	6	10
1466	Raipur	India	324.05	242.96	74.98	136	4	-	4	27
1467	Tiruppur	India	254.86	184.17	72.26	94	4	-	2	7
1468	Nanchang	China	785.82	785.72	99.99	549	5	-	1	2
1469	Gaoan	China	123.62	67.11	54.29	25	1	-	4	5
1470	Vellore	India	213.03	200.73	94.23	102	6	-	4	17
1471	Onitsha	Nigeria	644.51	310.48	48.17	141	6	-	6	7
1472	Panipat	India	157.96	157.74	99.86	77	4	-	4	17
1473	Stockton	United States of America	518.10	481.72	92.98	302	6	-	4	10
1474	Zhanjiang	China	369.44	317.52	85.95	149	5	-	1	5
1475	Surabaya	Indonesia	1,028.02	889.71	86.55	642	6	-	6	21
1476	Ma'anshan	China	442.03	339.92	76.90	199	5	-	5	11

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1477	Kochi (Cochin)	India	569.98	104.49	18.33	38	2	-	6	19
1478	Kozhikode (Calicut)	India	240.00	224.44	93.52	101	2	6	2	25
1479	Pizhou	China	158.66	66.61	41.98	25	1	-	1	2
1480	Jiyuan	China	215.76	215.84	100.04	97	1	-	1	12
1481	Yiyang, Hunan	China	198.88	198.88	100.00	92	5	-	2	23
1482	Xinxiang	China	394.30	327.79	83.13	171	1	-	1	26
1483	Qianjiang	China	96.17	93.97	97.72	34	1	-	4	12
1484	Jinhua	China	326.88	310.03	94.85	165	1	-	1	23
1485	Tiruchirappalli	India	200.09	149.9	74.92	86	6	-	4	7
1486	Laiwu	China	520.16	401.74	77.23	240	1	-	1	12
1487	Jundiaí	Brazil	271.33	271.02	99.88	137	6	-	6	3
1488	Warangal	India	226.92	37.96	16.73	8	6	-	6	21
1489	Karlsruhe	Germany	196.89	159.46	80.99	68	7	-	5	14
1490	Lahore	Pakistan	639.00	638.78	99.97	432	4	4	6	4
1491	Port Moresby	Papua New Guinea	238.12	141.89	59.59	61	2	6	2	7
1492	Bournemouth/Poole	United Kingdom	301.99	161.81	53.58	68	7	-	7	20
1493	Yuncheng	China	179.82	173.83	96.67	82	1	-	1	26
1494	Qinzhou	China	146.21	140.69	96.22	68	5	-	1	2
1495	Cebu City Lapu-Lapu City Mandaue City	Philippines	513.94	444.74	86.54	181	4	-	2	23
1496	Buenos Aires	Argentina	2,235.05	2171.8	97.17	1772	6	2	6	21
1497	Bhavnagar	India	164.02	141.41	86.22	75	6	-	4	12
1498	Birmingham	United States of America	228.32	228.47	100.06	80	2	-	2	23
1499	Cartagena	Colombia	252.40	190.83	75.61	61	6	-	6	10
1500	Taranto	Italy	200.99	83.79	41.69	12	1	-	1	5
1501	Venezia	Italy	182.62	27.42	15.02	3	5	-	6	19
1502	Naberezhnye Tselny	Russian Federation	282.33	139.66	49.47	66	5	-	1	2
1503	Lokoja	Nigeria	135.57	77.53	57.19	35	6	-	4	17

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}^*}$	$C_{F_{1523}}$
1504	Pôrto Velho	Brazil	225.60	165.2	73.23	86	2	-	4	1
1505	Winnipeg	Canada	553.61	543.35	98.15	364	6	-	1	1
1506	Duesseldorf	Germany	284.42	284.29	99.95	119	7	-	7	14
1507	Wuppertal	Germany	102.59	102.5	99.91	33	7	-	7	14
1508	Mérida	Mexico	423.02	422.87	99.96	280	4	-	6	19
1509	Zhuanghe	China	142.94	114.15	79.86	42	1	-	1	5
1510	Sheikhupura	Pakistan	119.26	83.45	69.97	36	4	-	4	17
1511	Kasur	Pakistan	103.71	65.66	63.31	26	4	-	4	17
1512	Padova	Italy	236.38	236.4	100.01	88	7	-	7	1
1513	Vicenza	Italy	194.46	107.77	55.42	33	7	-	4	7
1514	Philadelphia Trenton	United States of America	1,225.06	1224.95	99.99	796	2	-	7	6
1515	Duisburg Essen Bochum	Germany	1,107.30	1088.34	98.29	778	7	-	7	14
1516	Aba	Nigeria	264.09	264	99.97	155	2	-	2	25
1517	Abeokuta	Nigeria	278.71	278.33	99.86	160	4	-	2	25
1518	Acarigua-Aruare	Venezuela (Bolivarian Republic of)	193.91	181.28	93.49	94	4	-	4	15
1519	Adan (Aden)	Yemen	225.65	86.45	38.31	21	3	-	3	29
1520	Uyo	Nigeria	322.04	27.04	8.40	4	2	-	2	13
1521	Mathura	India	173.41	151.29	87.24	69	6	-	4	12
1522	Al Kuwayt (Kuwait City)	Kuwait	620.93	521.48	83.98	321	3	-	3	29
1523	Matsuyama	Japan	295.34	295.14	99.93	131	5	-	5	3
1524	Modesto	United States of America	487.34	443.99	91.11	244	6	-	4	10
1525	Maputo Matola	Mozambique	729.49	396.64	54.37	201	2	2	2	18
1526	Wuchuan	China	123.29	122.4	99.28	53	6	-	1	15
1527	Maoming	China	201.40	193.32	95.99	91	5	-	2	23
1528	Manado	Indonesia	180.39	168.45	93.38	66	4	-	2	13
1529	Malang	Indonesia	227.41	227.45	100.02	128	6	-	6	21

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{123}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1530	Ranipet	India	167.28	13.26	7.93	0	-	-	-	-
1531	Balikpapan	Indonesia	219.71	66.77	30.39	16	4	-	6	4
1532	Vadodara	India	304.51	304.39	99.96	180	6	-	1	5
1533	Zanzibar	United Republic of Tanzania	216.58	30.47	14.07	1	4	4	6	4
1534	Yueyang	China	230.73	230.74	100.01	107	5	-	5	23
1535	Yuanjiang	China	125.39	103.08	82.21	28	5	-	2	23
1536	Yangjiang	China	222.63	214.98	96.56	107	1	-	1	5
1537	Wonsan	Dem. People's Republic of Korea	82.04	65.41	79.73	13	5	-	1	27
1538	Wellington	New Zealand	117.25	75.66	64.53	18	2	-	7	6
1539	Vladivostok	Russian Federation	161.20	30.42	18.87	0	-	-	-	-
1540	Anand	India	312.39	91.38	29.25	41	6	-	2	18
1541	Akola	India	167.79	37.2	22.17	8	6	-	4	17
1542	Oshogbo	Nigeria	407.49	215.97	53.00	125	2	-	2	18
1543	Ogbomosh	Nigeria	177.72	142.07	79.94	70	6	-	2	18
1544	Al-Mansurah	Egypt	117.50	116.93	99.52	47	3	-	4	12
1545	Al-Mahallah al-Kubra	Egypt	110.02	106.15	96.48	44	3	-	4	17
1546	Murcia	Spain	411.35	385.46	93.71	232	3	-	3	28
1547	Chon Buri	Thailand	395.78	301.86	76.27	146	6	-	1	26
1548	Krung Thep (Bangkok) Samut Thep (Bangkok) Samut Prakan Nonhaburi)	Thailand	4,106.91	4101.18	99.86	3535	6	2	1	24
1549	Bogotá	Colombia	696.98	695.88	99.84	485	4	-	1	2
1550	Belo Horizonte Vale do Aco	Brazil	834.52	834.42	99.99	604	2	-	6	21
1551	Montréal	Canada	1,330.27	1309.72	98.46	918	7	-	7	6
1552	Sana'a'	Yemen	447.18	406.76	90.96	247	3	-	3	22
1553	Athínai (Athens)	Greece	773.67	771.45	99.71	486	7	7	7	14
1554	El Djazair (Algiers)	Algeria	749.85	462.16	61.63	279	3	-	1	27

Continued on next page

ID	City Name(s)	Country	MUA Area (km^2)	LCZs Area (km^2)	LCZs Coverage (%)	No. Patches	$C_{HM_{1523}}$	$C_{HM_{110}}$	$C_{F_{1523}}^*$	$C_{F_{1523}}$
1555	Asunción	Paraguay	567.61	472.31	83.21	289	2	-	7	6
1556	Ciudad Juárez El Paso	Mexico	1,324.74	1324.63	99.99	964	3	-	3	29
1557	Ciudad de Panamá (Panama City)	Panama	363.93	279.99	76.93	145	4	-	5	3
1558	Grande Vitória	Brazil	602.69	232.38	38.56	104	6	-	6	21
1559	Grande São Luís	Brazil	516.56	358.32	69.37	199	6	-	6	21
1560	München (Munich)	Germany	442.38	442.43	100.01	253	7	7	7	14
1561	Torreón	Mexico	518.93	6.24	1.20	0	-	-	-	-
1562	Maceió	Brazil	244.68	100.7	41.16	29	6	-	6	19
1563	Zürich (Zurich)	Switzerland	785.16	318.16	40.52	133	7	-	7	1
1564	København (Copenhagen)	Denmark	418.80	354.57	84.66	206	7	-	7	14
1565	Florianópolis	Brazil	406.28	61.7	15.19	12	5	-	5	3
1566	Fès	Morocco	211.47	202.41	95.71	102	3	-	4	27
1567	San Luis Potosí	Mexico	384.91	382.84	99.46	222	6	-	6	19
1568	Köln (Cologne)	Germany	139.86	139.83	99.98	55	7	-	5	14
1569	Lomé	Togo	440.93	297.16	67.39	155	6	-	6	21