Remote Sensing of Environment 232 (2019) 111353

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A new ranking of the world's largest cities—Do administrative units obscure morphological realities?

H. Taubenböck[a,b,*], M. Weigand[a], T. Esch[a], J. Staab[a], M. Wurm[a], J. Mast[a], S. Dech[a,b]

[a] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, Germany
[b] Institute for Geography and Geology, Julius-Maximilians-Universität Würzburg, Würzburg 97074, Germany

ARTICLE INFO

Edited by Emilio Chuvieco

Keywords:
City size
Urban agglomeration
Rank-size distribution
Remote sensing
Global urban footprint
Urban morphology

ABSTRACT

With 37 million inhabitants, Tokyo is the world's largest city in UN statistics. With this work we call this ranking into question. Usually, global city rankings are based on nationally collected population figures, which rely on administrative units. Sprawling urban growth, however, leads to morphological city extents that may surpass conventional administrative units. In order to detect spatial discrepancies between the physical and the administrative city, we present a methodology for delimiting Morphological Urban Areas (MUAs). We understand MUAs as a territorially contiguous settlement area that can be distinguished from low-density peripheral and rural hinterlands. We design a settlement index composed of three indicators (settlement area, settlement area proportion and density within the settlements) describing a gradient of built-up density from the urban center to the periphery applying a sectoral monocentric city model. We assume that the urban-rural transition can be defined along this gradient. With it, we re-territorialize the conventional administrative units. Our data basis are recent mapping products derived from multi-sensoral Earth observation (EO) data – namely the Global Urban Footprint (GUF) and the GUF Density (GUF-DenS) – providing globally consistent knowledge about settlement locations and densities. For the re-territorialized MUAs we calculate population numbers using WorldPop data. Overall, we cover the 1692 cities with > 300,000 inhabitants on our planet. In our results we compare the consistently re-territorialized MUAs and the administrative units as well as their related population figures. We find the MUA in the Pearl River Delta the largest morphologically contiguous urban agglomeration in the world with a calculated population of 42.6 million. Tokyo, in this new list ranked number 2, loses its top position. In rank-size distributions we present the resulting deviations from previous city rankings. Although many MUAs outperform administrative units by area, we find that, contrary to what we assumed, in most cases MUAs are considerably smaller than administrative units. Only in Europe we find MUAs largely outweighing administrative units in extent.

1. Introduction

"Tokyo is the world's largest city with an agglomeration of 37 million inhabitants, followed by Delhi with 29 million, Shanghai with 26 million, and Mexico City and São Paulo, each with around 22 million inhabitants. Today, Cairo, Mumbai, Beijing and Dhaka all have close to 20 million inhabitants" (UN, 2018). This quote from the 2018 World Urbanization Prospects publication is a clear statement about the ranking of the largest cities in the world. This paper, however, calls this ranking into question.

The UN ranking is based on population numbers relying on administrative space units. Naturally, these spatial units are crucial as they determine a clear-cut city's boundary to establish jurisdictional competence of its municipal government (Parr, 2007). In a world, however, which is experiencing highly dynamic transformation processes through global urbanization (e.g. Angel et al., 2011; Taubenböck et al., 2012), conventional, administrative spatial units represent reality less and less. Scholars describe a continuous transformation from formerly compact, concentrated land use patterns into spatially extended and morphologically less clear-cut urban configurations (e.g. Anas et al., 1998; Batty et al., 2004; Siedentop and Fina, 2010; Small et al., 2011; Taubenböck et al., 2019). Terms such as 'peri-urban', 'urban fringe', 'leapfrog development', 'suburb', among many others describe a complex, indistinct, irregular, and dynamic transitional zone which is more a urban-rural continuum than a clear-cut boundary.

The implications, however, of the common administrative space units are critical: Beyond the political control competence, the question of city size is of fundamental statistical significance. Size is of relevance

* Corresponding author at: German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, Germany.
E-mail address: hannes.taubenboeck@dlr.de (H. Taubenböck).

https://doi.org/10.1016/j.rse.2019.111353
Received 25 February 2019; Received in revised form 22 July 2019; Accepted 29 July 2019

with respect to the ranking and hierarchical ordering of cities (Auerbach, 1913; Zipf, 1941; Eeckhout, 2004), to the relation to such variables as per capita income, unemployment rate, inequality, among many others (Parr, 2007) or to aggregated numbers such as total population or gross domestic product (UN, 2018; Openshaw, 1983). Without a comparable and meaningful unit of space, these numbers lead to distorted patterns of explanation, e.g. in multivariate analysis.

Let us take Tokyo as an example: The city of Tokyo has, according to official figures, 9.5 million inhabitants on an administrative area of 622 km^2. These numbers would rank Tokyo at number 31 in the list of the largest cities in the world (UN, 2018). The metropolitan region (more or less the continuous built urban area), however, is listed with 37 million inhabitants at 13,572 km^2, making it officially the largest city in the world. In global statistics, the latter number is commonly used and compared to other cities such as Guangzhou in China, which is reflected in this statistic with 12 million inhabitants, ranking it at 23. However, the continuous built urban area in Guangzhou is, as it has been shown by scholars, significantly larger (e.g. Florida et al., 2008; Taubenböck et al., 2014). Neighboring cities such as Foshan, Dongguan or Shenzhen have physically merged together with Guangzhou but make the list by themselves. So, Guangzhou is treated individually here. Surprisingly, these inconsistencies are largely hidden in the statistics; they are tacitly accepted and barely questioned.

In consequence, we argue a fundamental challenge for any empirical investigation is to define comparable spatial units in order to provide a realistic statistics of city sizes or above mentioned downstream statistical indicators. Critics claim that much of existing urban research studies can be contested due to a data-driven approach and a rather arbitrary use of geographical boundaries (e.g. Lechner et al., 2013; Riitters et al., 1995; Berry and Okulicz-Kozaryn, 2012; Masucci et al., 2015; Taubenböck et al., 2016). In a non-academic sense, the size of the city contributes to its global perception. The opening quotation *"Tokyo is the world's largest city"* assigns this city to a certain global city network and implies, triggers, attracts and advertises economic and political spin-offs in terms of, for example, investments by international companies, attraction to high-skilled labor force or presence of a broad skilled labor force (e.g. Weber, 2001). Rankings have become a fundamental element in the contest of globalization.

In this paper we question whether and to what extent today's morphological urban areas (MUAs) differ from conventional administrative spatial units. To do this, we derive MUAs from globally consistent geodata on settlement structures derived from remote sensing data. In other words, we investigate whether the current ranking of city sizes based on the United Nations' statistics are strongly influenced by the artificial spatial units of administrative boundaries compared to the built-up extent of cities. But perhaps more importantly, we propose a new ranking of the largest cities in the world based on consistently calculated MUAs and related population figures.

The remainder of this paper is organized as follows. Section 2 provides more background and introduces the conceptual foundation. In Section 3 we introduce the data used and the developed methodology. In Section 4, we present the empirical results for 1692 cities across the globe, which are listed larger than 300,000 inhabitants in 2016. In Section 5, we critically discuss the capabilities and limitations of data and methods and foremost the implications of these findings, and in Section 6, we conclude the study.

2. Background and conceptualization

What defines the size of a city? Is it the administrative unit, the functional urban region, the physical extent of the built environment, the total population, the volume of trade, the economic power, the interlacing space, the functional move-in area, the perception of people, or other related factors?

These different concepts to approach the size of the city show, as Jessop et al. (2008) remark, that a city, metropolitan area or region can be imagined and constructed in varying ways with different methods and indicators for region-building: e.g., from tightly sealed areas with a territorially-embedded thinking of regions, to porous nodes in a networked space of flows (Castells, 1999; Harrison and Growe, 2014). As Castells (2000) discusses, space consists in the conflict between the ‘space of flows’ and the ‘space of places’. The ‘space of flows’ is defined as regular flows of people, goods, or information between separate but networked locations, while the ‘space of place’ refers to the physical boundaries of locations. In consequence, there is neither an “all-purpose definition”, nor is there a universal truth for the size of the city. Definitions and delimitations will differ depending on the specific purpose, the data and the methodologies used.

A fundamental basis for assessing the size of a city is to define what is meant by ‘urban’. In the scientific discourse on this subject no standard and internationally accepted definition of ‘urban’ or ‘urban population’ has yet prevailed. Each country uses its own definition, and collects data accordingly. The statistic that currently 55% of the world's population is urban dwellers (UN, 2018) relies on adding up these figures coming from often incomparable definitions, which are even based on arbitrary and thus inconsistent spatial admin units (Deuskar, 2015). Research studies delimiting urban space rely on various methods and geodata such as street networks (Masucci et al., 2015), remote sensing (e.g. Liu et al., 2016; Esch et al., 2014), demographic (Rozenfeld et al., 2011), among other data. Multi-criteria attempts using indicator combinations of minimum population size and density, travel times to the central place, among others are suggested and in use (e.g. World Development Report, 2009; Abed and Kaysi, 2003; Dijkstra and Poelman, 2014). Georg et al. (2018) take account of this by delimiting urban space using remote sensing, infrastructure, population and economic data constructing several possible spatial forms illuminating the fuzziness of the delimitation of territorial space. The definition and delimitation of the urban and the related urban-rural transition is subject to many years of scientific discussion, without, however, arriving at a uniformly accepted approach (Simon, 2008). Ross (2011) notes that boundaries are malleable as the transition is complex, indistinct, irregular, and dynamic and they should include some capacity for flexibility depending on the purposes. In spite of the large body of literature, Masucci et al. (2015) remark, that the very concept of cities remains obscure, hidden or assumed.

In consequence, any classifications of city sizes are not innocent. The concept, criteria and methods used might change details, and thresholds applied might be manipulated with often undocumented effect. In recommendations by the United Nations it is even formulated that due to different characteristics and understandings of ‘urban’ and ‘rural’ across the globe, a global definition is not possible (Dijkstra and Poelman, 2014). In this paper, however, we argue that, even though local forms of settlement, flows of goods, and life vary widely, a consistent methodological approach generating statistically comparable spatial units based on one consistent data source brings to light novel perspectives and facts about city sizes. Against this background we understand ‘urban’ in this work in reference to Schneider et al. (2009) and Taubenböck et al. (2012) as places dominated by the built environment. The built environment includes human-construct elements, roads, buildings, or the like. Non built land (e.g. vegetation, bare soil) is not considered urban, even though they may function as urban space.

We are building on this definition because we approach the delimitation of ‘urban’ using a globally consistent data source of satellite-based earth observation (EO). This data has developed to a tool providing area-wide spatial information on the location, spatial distribution and characteristics of settlement features at global scale with high geometric resolution (e.g. Esch et al., 2012; Pesaresi et al., 2013) and high accuracies (Esch et al., 2018a, b; Klotz et al., 2016; Taubenböck et al., 2011). We rely on the Global Urban Footprint suite (GUF) (Esch et al., 2018a, b) capturing physical attributes of settlements allowing us to focus on urban form. Building on this, we construct territorially bounded city areas based on similar characteristics of the built city.

Thus, we use the 'space of place' logic to construct physical boundaries of cities. While many EO-based studies approached the global urban extent (e.g. Schneider et al., 2009; Elvidge et al., 2007; Pesaresi et al., 2013; Esch et al., 2012), only few studies, mostly relying on night light emission and/or gridded population data have approached the delimitation of city extents on a global scale. In these studies the power law scaling of rank-size distributions is documented (Decker et al., 2007; Small et al., 2011; Small and Sousa, 2016) and consistency is confirmed over time (Small et al., 2018). Small et al. (2011) found spatial networks of urban development are vastly larger than the administratively-defined cities they contain. At higher resolutions, but only for a comparatively small subset of the world's cities, Fragkias and Seto (2009) revealed oscillations of rank-size distributions over time due to the coalescence of settlements across administrative city boundaries. This scarcity of studies at global scales originates from the conceptual complexity for a consistent and comparable delimitation of city boundaries, any widely accepted definition of what constitutes a coherent urban area and the necessary globally consistent data set.

So, what does this imply regarding the delimitation of cities to determine their sizes? Unlike the above mentioned studies, which used night-time light emissions or population data as proxy for settlement extents, this study is based on another proxy: the empirical relationship between spatial variability of X-band backscatter and assumed structure of built morphology mapped at a comparatively high geometric resolution. We assume that the urban-rural transition can be defined by the slope of the gradient of physical density of built-up structures towards the periphery. Although reality is often more like a fuzzy transition in the urban-rural continuum, we aim at a clear-cut, but compared to conventional administrative units, consistently derived boundary. In this way we try to spatially designate the built city and to re-territorialize the existing measure of the administrative spatial units more sensibly. Based on these new re-territorialized urban entities, we use WorldPop data (Tatem, 2017) for calculating related population figures. The effects of potential changes of re-territorialized city sizes (in extent and population) compared to administrative spatial units (in extent and population) are represented by rank size distributions and related statistics.

3. Data and methodology

3.1. Data

Our analysis bases on data from *remote sensing* (Esch et al., 2012; Esch et al., 2018a), geodata from *OpenStreetMap* (OSM, 2018), the *Global Administrative Areas* dataset (GADM, 2018), data from the United Nations' *populations statistics* (UN, 2015), and the global population dataset from *WorldPop* (Tatem, 2017).

The two remote sensing based mapping products are from the Global Urban Footprint suite: The *Global Urban Footprint* (GUF) and the *GUF Density* (GUF-DenS) classification. The *GUF* is a binary raster layer presenting the worldwide distribution of human settlements in a so far unique spatial resolution of 12 m (Esch et al., 2012; Fig. 1b). Using the operational Urban Footprint Processor framework described by Esch et al. (2013), the vertical built-up structures in urban and rural environments are mapped based on a global analysis of > 180,000 TerraSAR-X and TanDEM-X StripMap radar images collected between 2011 and 2014 (93% of images recorded in 2011–2012, 7% in 2013–2014). Spatial complexity of varying objects within small areas is characteristic for urban areas. This is represented in highly textured image regions of strong directional, non-Gaussian backscatter due to double bounce effects in radar data. This information is used in combination with the intensity information to delineate 'settlements' from 'non-settlements' using an unsupervised image analysis technique. The experimental *GUF-DenS* layer is a semantically enhanced version of the GUF that provides information on the built-up density or – as an inverse, the greenness – in form of the percentage of impervious surfaces within the area assigned as settlement in the GUF (Esch et al., 2018a). The *GUF-DenS* results from a combination of the binary GUF (which is used as a mask) and a TimeScan dataset derived from > 450,000 Landsat images (Esch et al., 2018b). The TimeScan layer is used to actually model the imperviousness/greenness via the temporal characteristics of the Normalized Difference Vegetation Index (NDVI) (e.g., mean, max, min, stdev) over a 3-year period. Since the NDVI is only calculated for the "built-up" regions within the GUF, the potential uncertainties due to mix-ups with bare soil/sand (e.g. in semi-arid regions) are significantly reduced. Assuming a strong inverse relation between vegetated and impervious surfaces, the intensity of vegetation cover defined by the NDVI can be used as a proxy for the percent impervious surface following an approach introduced by Esch et al. (2009). The resulting GUF-DenS is a raster layer in 30 m spatial resolution that shows values between 0 and 100 as an indication of the percentage of impervious surfaces per grid cell (Fig. 1b).

We use *administrative units* for each city provided by the Global Administrative Areas (GADM) database (GADM, 2018). However, there is no internationally agreed definition of metropolitan areas, which are the spatial basis for the World Urbanization Prospects statistics (UN, 2018). The GADM database contains administrative units at different levels: from the national to the municipal or even to the district level. The supposed spatial units of the population figures of the World Urbanization Prospects statistics, however, are not simply mapped in the GADM database. To achieve congruent spatial units, we adjust administrative units to metropolitan areas where the World Urbanization Prospects statistics refer to this spatial entity (cf. example of Tokyo in the Introduction). Beyond, for application of our monocentric sectoral city model (cf. Section 3.2.), we need to define the location of the city center. Here, we rely on the spatial definition of the central points for each city as provided by the United Nations (2014) (Fig. 1b).

For our comparative spatial city analysis, we exclude all areas containing water bodies such as oceans, rivers and lakes as they represent non-buildable areas. To do so, we use the water bodies' dataset from the OpenStreetMap project (OSM, 2018).

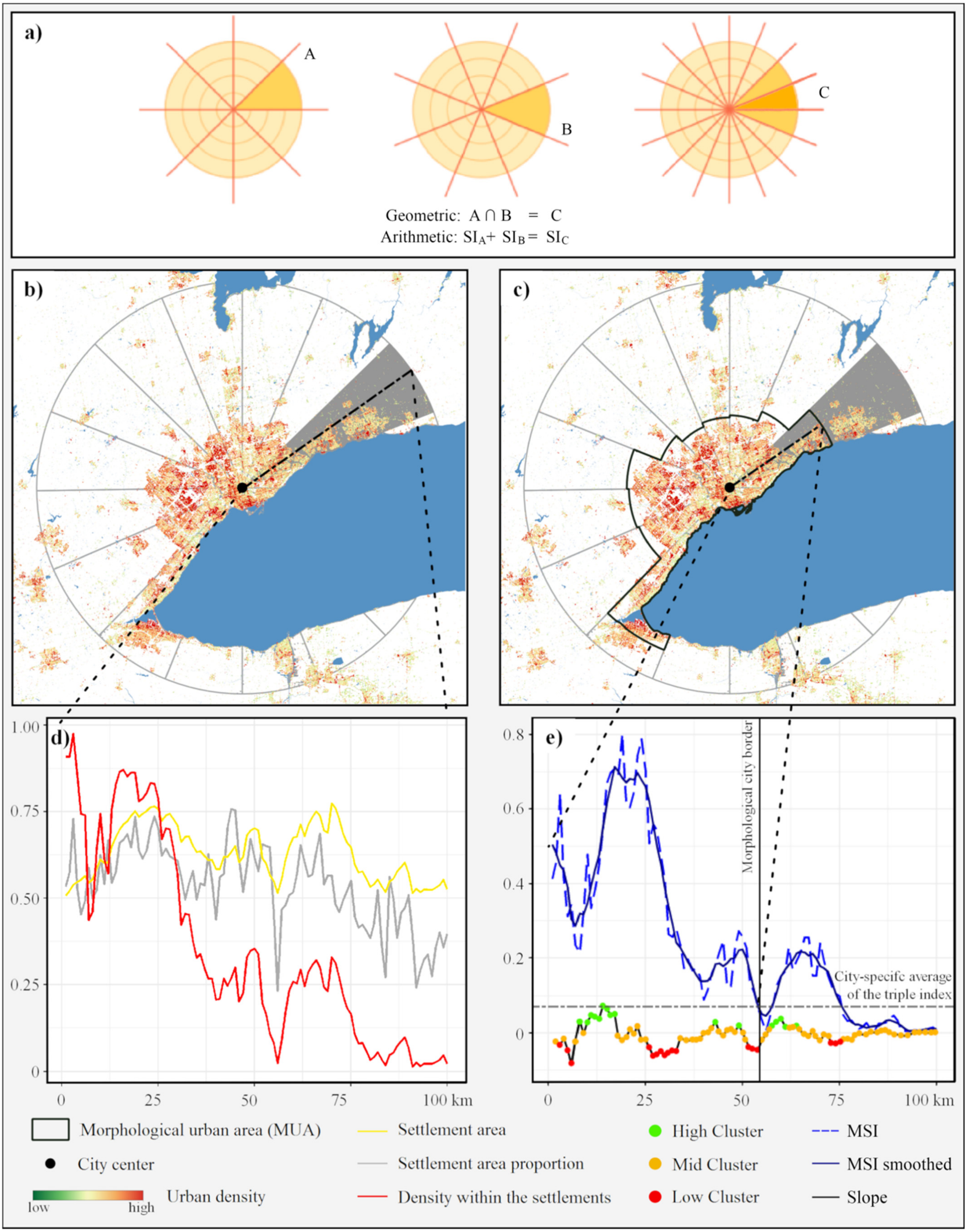
For the selection of the cities under investigation we rely on population statistics from the United Nations (2015). We perform the analysis for all cities on our planet with > 300,000 inhabitants, i.e. a total of 1692 cities are included in this study.

For the calculation of population numbers for the newly re-territorialized cities we rely on *WorldPop* data (Tatem, 2017). WorldPop data are gridded population counts at spatial scales finer than the administrative unit level of census data using a suite of geospatial layers (e.g. Sorichetta et al., 2015). The approach relies on a random forest-based dasymetrically disaggregation of the census counts from administrative units into grid cells (Stevens et al., 2015).

3.2. Spatial delimitation of morphological urban areas (MUAs) from rural environments

Where do patterns of settlements change from urban to rural? There can be no simple answer to this rather philosophical question. To address this question, however, we use a conceptual framework that is tailored to our available global data. We assume that the urban-rural transition is represented somewhere along a decreasing gradient of physical built-up density with rising distance to the urban center. Thus, we measure the built-up urban extent of cities in their physical form, as opposed to metropolitan areas, which are defined in their economic or functional form.

For measurement and mathematical representation of the gradient of built-up density from the center to the periphery we choose a monocentric, sectoral city model as spatial entity. The monocentricity assumed here is, of course, a simplification of the urban spatial patterns. This traditional, monocentric model, known as the "Alonso-Mills-Muth" model, refers to the equilibrium distribution of land use within a city as a function of land rents, which decreases with distance from the



(caption on next page)

Fig. 1. Workflow for delimiting MUAs: **a)** The monocentric, sectoral city model with the two 8 sector models and the 16 corridors derived from it as well as bandwidths of 1 km form the spatial entities for analysis; **b)** The sectoral, monocentric model projected around the defined city center, onto the settlement area from the GUF and the calculated densities of the GUF-DenS classification within the settlements – the example pictures the city of Toronto, Canada; **c)** The delimitation of the MUA based on the morphological settlement index (MSI); **d)** The gradients of all three indicators –settlement area, settlement area proportion and density within settlements– for a sample sector from center to periphery; **e)** The MSI and the slope allowing to classify the three clusters. The first low point of the low cluster below the MSI average of the low cluster is used as cut-off value.

center (Paulson, 2012). Due to its empirical traceability and a proven explanation of a substantial portion (at least 80%) of the variation in urbanized land area across cities (e.g. Spivey, 2008), it continues to serve as a theoretical and empirical core concept of urban land development (e.g. McMillen, 2006). Beyond, it can also be conceptually reconciled with our assumption of an urban-rural transition along a decreasing gradient of built-up density with rising distance to the center. We use ring models with commonly used bandwidths of 1 km to determine the location relative to the defined urban center point. As the urban-rural gradient may vary around the city and urban or rural islands may occur within the cityscape (Simon, 2008), we additionally subdivide the urban area into 16 corridors (sectors) to take account of the local spatial context. Each sector opens radially outwards at 22.5°, starting at the defined city center (cf. Fig. 1a).

For the measurement of the urban-rural transition, we suggest a *morphological settlement index* (MSI) consisting of three indicators for defining the *Morphological Urban Area* (MUA): 1) The *settlement area*, 2) the *settlement area proportion*, and 3) the *density within the settlements* (Fig. 1d).

1) The sectors from the city center to the periphery contain a certain *settlement area*. Along these sectors, the available area per ring, i. e. per 1 km bandwidth, increases with rising distance to the center point and so does the potentially available settlement area. We assume that as long as the absolute settlement area per ring is increasing towards the periphery, the sector still belongs to the MUA of the city.
2) The *settlement area proportion* corresponds to the extents of the settlement areas per respective ring per sector. We assume that a decreasing density hints at a transition area from urban to rural.
3) The settlement area proportion may in some cases be equal at two different locations; however, the *built-up density within the settlement areas* might differ. We assume decreasing densities hint at a transition towards rural environments. Here we calculate the mean built-up density within the respective settlement area per ring and sector.

The Pearson product-moment correlation between the three indicators is low (between settlement area and settlement area proportion $r = 0.14$; between settlement area and built-up density within the settlement areas $r = -0.12$; and between built-up density and settlement area proportion within the settlement areas $r = 0.26$); therefore these indicators are permissible for our model. As it has been shown, we formulate a hypothesis per indicator which points to the urban-rural transition. However, a simple decrease in *settlement area proportion* from one ring to another – to take one example – is obviously given in most cases and, thus, not meaningful for finding an urban-rural boundary. In consequence, we combine all three normalized indicators to a *MSI* (Fig. 1e). We assume if all three indicators in combination suggest a decreasing gradient the probability for an urban-rural transition rises.

We calculate the MSI per sector (✳) and per respective ring area to display the density gradients. For the derivation of the MSI, however, we would also like to add the larger urban context to each sectoral gradient. The following train of thought is the basis for it: if a low dense island occurs within one 22.5° sector embedded in highly dense urbanized areas in the surrounding sectors, the localization of a morphological city boundary should be less likely than if a low dense island in the central sector is flanked by undeveloped open spaces in the surrounding sectors. To account for this consideration, we integrate MSI values of neighboring sectors for the derivations of the final MSI value. To do so, we calculate the MSI values for two separate (+ and ×) sectoral models consisting of 8 sectors each (see Fig. 1a). Since both sectors are rotated by 22.5° (cf. in Fig. 1a the sectors 'A' and 'B' illustrate this), their geometric superposition results in the sector 'C' (cf. Fig. 1a). The final MSI value for sector 'C' is a function of both, the MSI values of sector 'A' and 'B' by addition. Thereby, the central sector 'C' is given a higher weight, since it density values contribute to both sectors ('A' and 'B').

For a reasonable classification of the *MUA* that only defines a city boundary if there is a real morphological change towards low dense, rural structures, we calculate three clusters based on the distribution of MSI values. We establish clusters by initializing MSI values using a *k-means* approach. The three clusters are the *high cluster* indicating a rising morphological density gradient, the *mid cluster* indicating a generally constant density gradient, and the *low cluster* indicating a decreasing density gradient. The high and mid clusters testify to a continuous settlement area. The low cluster, however, suggests decreasing morphological density. In Fig. 1d the three morphologic indicators are illustrated along the rising distance from the city center. In Fig. 1e the *MSI* and the related increase or decrease of the gradient is visualized. Also indicated are the three different clusters, with the low cluster (in red dots) suggesting potential candidates for urban-rural cut-off values. For delimiting the *MUA*, we define the cut-off value as the first dot towards the periphery belonging to the low cluster which features a *MSI* below the particular city average of all low clusters. This in turn, allows that –as we argued above– urban or rural islands can occur within the cityscape. As an example, our approach allows a park area (i.e. very low built-up density) with a continuous built urban landscape before and after the open area to be classified as part of the city as long as the identified low cluster is above the city average of the *MSI* at all cut-off candidates. Compared to other studies, which are working with minimum fixed distances between settlements to be still considered continuously urban (e.g. 200 m in the study of Weber, 2001), our approach is independent from such strict thresholds.

Many urban regions across the globe have experienced a coalescence of multiple, once morphologically separate cities, but remain jurisdictionally separated (e.g. Taubenböck and Wiesner, 2015). Although we chose a monocentric approach based on the city centers of the still jurisdictionally separated cities, we aim to capture morphologically merged cities as one MUA. If MUAs from two (or more) neighboring cities overlap, we combine the MUAs from both (or more) cities into one. We then count them as one urban agglomeration in our statistics and attribute this 'city' as a metropolitan region (M.R.).

3.3. Spatial statistics for analyzing the effects of re-territorializing cities by MUAs

We present the effects of changes in city extents by our morphologic approach compared to the commonly used administrative spatial units at two spatial entities: at *individual city level*, and at *continental level*.

At *individual city level*, we compare all 1692 cities across the globe based on the rank in the rank-size distribution. Rank-size distributions show the relation between rank $\tilde{N} = \tilde{N}(S)$ (Auerbach, 1913). The rank indicates the position of a city if you arrange the cities according to their size. Scholars observed that the rank is proportional to the reciprocal size of the city (e.g. Auerbach, 1913; Nitsch, 2005; Soo, 2005).

At *continental level*, we analyze trends between administrative space units and our re-territorialized MUAs. To do so, we propose the Normalized Difference Area Index (NDAI). It is calculated as follows:

$$NDAI = \frac{MUA - AUA}{MUA + AUA}$$

where *MUA* is the morphological urban area and *AUA* is the administrative urban area. The NDAI ranges from -1 to 1. Negative values indicate a significantly larger AUA in reference to the MUA. Positive NDAI values indicate that the MUA exceeds the AUA. The NDAI allows projecting all cities onto a world map to capture regional differences. In addition we aggregate the results of all cities belonging to a continent using boxplots for comparison.

3.4. Accuracy assessment of input data

The validity of our intended results depends on the accuracy of the input data: For the *GUF* and the *WorldPop data* we do not perform an accuracy assessment here, but rely on measures presented in other studies (Stevens et al., 2015; Taubenböck et al., 2011; Klotz et al., 2016). For the experimental GUF-DenS layer, we perform an accuracy assessment. We relate the density values to aggregates of built-up densities based on independent, highly resolved geodata: We apply OSM data (OSM, 2018) for Paris, France and New York City, Dallas, and Las Vegas, USA. We use building footprints alongside features for rails, roads and industry. For the city of Munich and the entire state of Bavaria we substitute the building footprints from OSM data by Level-of-Detail 1 building models provided by the German Federal Agency for Cartography and Geodesy (BKG) (www.bkg.de). From a geographic point of view, we chose the latter example to encounter for a landscape consisting of highly dense urbanized areas and low dense, rural environments. These samples are picked as for them a complete set of built-up features is available. Using these data bases, we calculate built-up density for four different scenarios: (1) buildings, (2) buildings and roads, (3) buildings, roads and railways, (4) buildings, railways, roads and industrial facilities. For these scenarios, the built-up density is aggregated to square kilometers. The GUF-DenS product is then compared against all four scenarios, to identify the most corresponding thematic relation and to provide a validation of the input data set.

4. Results

4.1. Mapping results: morphological urban areas and administrative urban areas

The proposed approach allows a demarcation of the main morphological urban entity from low dense peripheral and rural areas in a globally consistent manner. In Fig. 2 cartographic results of the derived MUAs are projected onto the input data, i.e. onto the GUF classification and the density values of the GUF-DenS layer within the settlement areas. Administrative boundaries and defined urban center points derived from data from the United Nations (2014) are visualized as well.

The following cases have been identified:

a) The resulting MUAs related to different urban center points spatially overlap and, based on our methodology, are merged to one polycentric spatial entity (in our abbreviation M.R.) consisting of more original cities by UN definition (Fig. 2a).
b) The resulting MUA is smaller than the administrative unit (Fig. 2b).
c) The resulting MUA is larger than the administrative unit (Fig. 2c).
d) The resulting MUA and the administrative unit are comparatively similar and can be considered a 'true-bounded' city (Fig. 2d).
e) The resulting MUA and the administrative unit are comparatively similar in size; however the spatial extents are not congruent (Fig. 2e).

The newly calculated re-territorialized spatial boundaries of MUAs are available for download in vector format in supplementary A-1.

The mapping results rely on multi-sensoral remote sensing data and do not provide accuracies of cadastral data. However, the studies of Klotz et al. (2016), Taubenböck et al. (2011) and Mück et al. (2017) clearly show the improvement of map accuracy of the GUF layer over other global urban mapping products. Especially for cities, areas characterized by high settlement densities, the GUF features high accuracies of about 90%. In turn, we assume this input data set provides a reliable basis. Also the GUF-DenS values show high agreement with very high resolution reference data from OSM or from BKG. At an aggregated grid of $1km^2$, we find the density values correspond best to the built-up density calculated by the combination of the thematic classes 'buildings', 'streets' and 'railways'; in turn, this defines basically the thematic definition of our GUF-DenS input layer. Median absolute errors of 1.67 for Paris, 4.92 for Munich or 3.29 for Dallas prove the validity of the input layer. However, we also observe tendency of slight overestimation of density values in this data set in general, and a tendency of higher overestimation for areas of higher density within cities in particular. This is e.g. true for Las Vegas where the aridity of the location leads to open soil as dominating land cover beyond impervious surfaces. Thus, these two classes seem to get mixed up spectrally causing overestimation with a median absolute error of 29.71 (Fig. 3). However, we also believe that it is highly probable that the measured overestimation is in reality lower because, unlike the BKG data in Germany, we cannot ensure that the OSM are indeed complete. In general, we conclude that our input data set still provide highly even if variable accurate density values for different areas or landscape types across the globe and represents the built-up density of 'buildings', 'streets' and 'railways'.

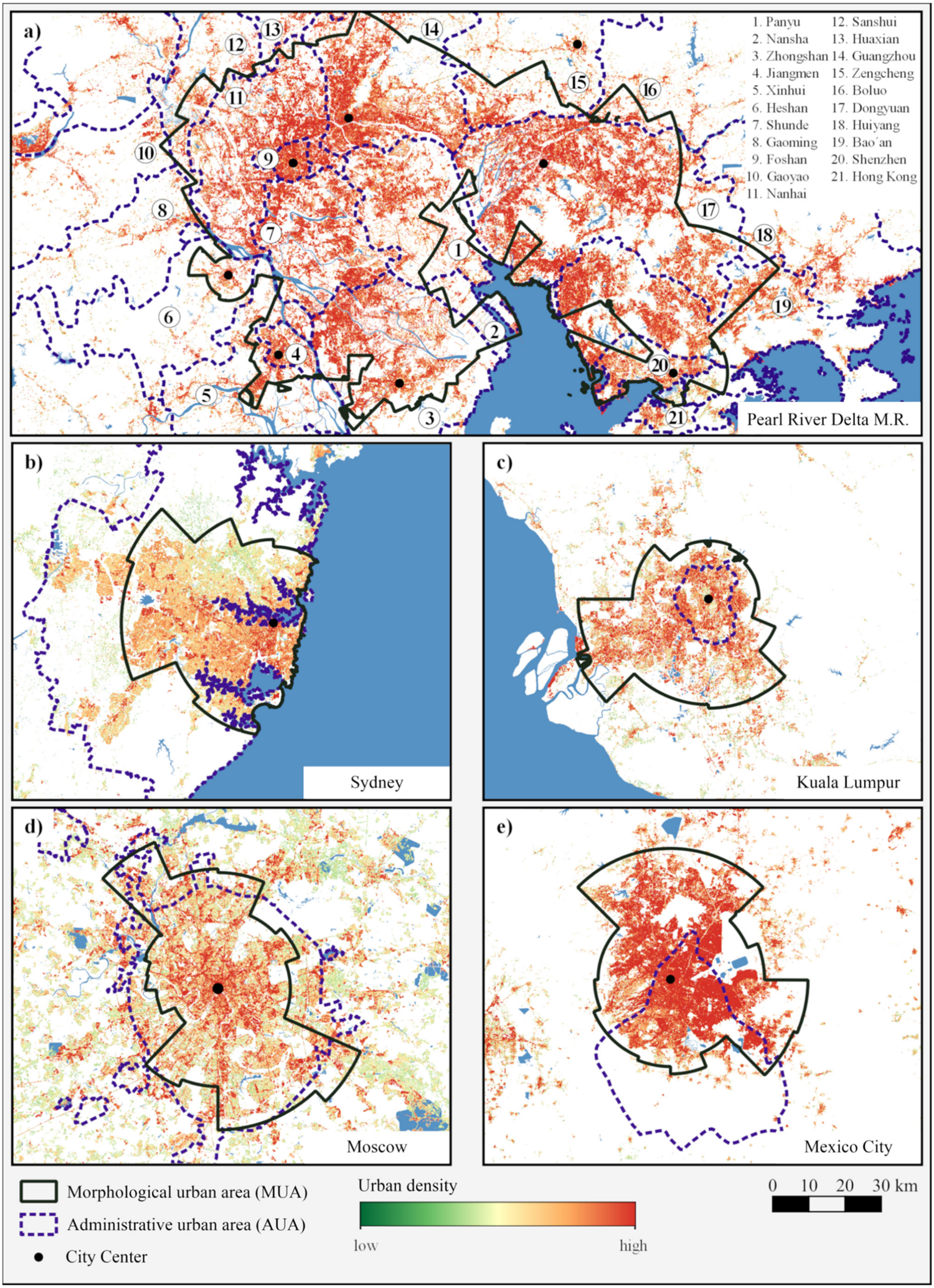
For the WorldPop data high accuracies have been measured for national-scale population distributions presented by Tatem (2017) or Stevens et al. (2015). In turn, we also assume this input data set provides a reliable basis for population assessment.

4.2. Global rank-size distributions at individual city level

With the first sentence of this paper – quoting the 2018 World Urbanization Prospects publication – a clear statement about the ranking of the largest cities in the world is given: Tokyo is the largest city of the world measured by population based on administrative units (Fig. 5a). Our results, however, suggest a different truth. If one determines the MUA of cities and calculates the populations related to them, the currently largest urban agglomeration in the world is the continuous urban landscape in the Pearl River Delta (PRD) metropolitan region (M.R.) in China (Fig. 2a; Fig. 4).

In this PRD M.R. the once physically (and still administratively) separate cities of 21 administrative units (among them are such large cities as Guangzhou, Dongguan, Foshan or Shenzhen which are entirely or partly within the new MUA) morphologically coalesced. The PRD M.R. is home to 42.6 million inhabitants. With 5.6 million more than the city of Tokyo, which has been listed as the largest city to date, this is 15.1% more compared to Tokyo in the United Nations ranking. The dimension of such an urban agglomeration becomes clearer when one compares it with population figures of countries. In a ranking of countries, this single urban agglomeration would rank at 34, larger than populations of e.g. Canada, Poland, or Australia. Considering the large but unknown informal population in the PRD M.R. (assumptions suggest about 20 million (Liang et al., 2014)), the PRD M.R. would be home to about 62.6 million which would rank the city even at 22, in the range of Great Britain, Italy or South Africa. Based on our approach, Tokyo ranks now at number 2; however, its population is with 31.9 million inhabitants far behind the PRD M.R.

If one focusses on the newly calculated MUAs, the PRD M.R. is also measured spatially the largest agglomeration of today. The second largest city by MUAs is the coalescent M.R. of Los Angeles, USA; interestingly, its population with 13.7 million inhabitants ranks L.A. only



(caption on next page)

Fig. 2. a) The largest city in the world by areal size and population: the Pearl River Delta (PRD) Metropolitan Region (M.R.) in China and the many individual administrative units of cities commonly used for statistics; **b)** Sydney, Australia resulting in a smaller MUA than the administrative unit; **c)** Kuala Lumpur, Malaysia resulting in a larger MUA than the administrative unit; **d)** Moscow, Russia with a comparatively small discrepancy between MUA and administrative unit, and **e)** Mexico City, Mexico resulting in a comparable MUA and administrative unit; however, spatially both units are displaced.

at 16. At rank three in MUA is the M.R. of Changzhou (consisting of the cities of Changzhou, Jiangyin, Jingjiang, Suzhou, Wuxi and Zhangjiagang). However, this city is with 14.5 million inhabitants only ranked at 14. Vice versa, the fourth largest city by MUAs is Tokyo, Japan, but ranked second in population. In this context these relationships reveal indirect statements about the particular density of built structures.

We have also specifically included the city of Ad-Damman (Saudia-Arabia) in Fig. 4, as it leads the AUAs size ranking (cf. Fig. 5b). Ranked at 383 by population based on MUAs and ranked 173 using MUAs also reveals clearly how arbitrary spatial units are and how they may differ to structural characteristics.

In general, we find the three largest and 14 out of the largest 30 MUAs are, by our definition, polycentric metropolitan regions where once separated cities coalesced. This testifies to the fact that formerly morphologically separated urban areas have merged into continuous urban landscapes of a new dimension exceeding current spatial control units of AUAs. In our analysis of MUAs, the PRD M.R. is the largest urban agglomeration in the world with the largest population. In previous statistics the PRD M.R. is not listed, but the cities that have grown together in the meantime are still counted individually. As examples, the city of Guangzhou is listed at rank 19, the city of Shenzhen at 26 by population in UN statistics (Fig. 5b). By administrative spatial units Guangzhou ranks only at 699 revealing the skewness of the statistics (Fig. 5a). In Fig. 5 we reveal that the largest extents are found in Saudi Arabia; cities, as indicated above, far from being the largest cities in the world. These areas carry political history in them and we should disregard them in geographical comparisons. Let us take as one example the city of Kuala Lumpur in Malaysia: the city is ranked 40 based on population figures relying on MUAs. The MUA ranks it even on 33. For administrative units Kuala Lumpur ranks at 49 for population, but for administrative space units it is only at 1253. This finding reveals how the common, accepted but somehow artificial and arbitrary AUAs and the related statistics obscure reality and challenge any ranking system due to an unequal and thus generally incomparable denominator.

In general, this alternative approach using MUAs as spatial baseline reveals a striking spatial difference to the extents of the existing AUAs. Considering a city 'true-bounded' if spatial deviation is within 10% between MUAs and AUAs, we find this only in 3.7% of all cases. This seems to be an alarmingly high value that should make us re-think existing administrative units.

Although according to this study a new largest city in the world has been identified, which is with 42.6 million measured larger than the previously assumed 37 million of Tokyo, overall fewer cities achieve mega-city status (mega cities are defined as cities with > 10 million inhabitants (UN, 2018)). Compared to the currently 30 led by the United Nations statistics, only 26 are identified by our approach. Perhaps even more interesting is the fact that these two lists are very different from each other: 10 mega-cities listed by the United Nations – Bangalore, Chennai (Madras), Chongqing, Moskva (Moscow), Tianjin,

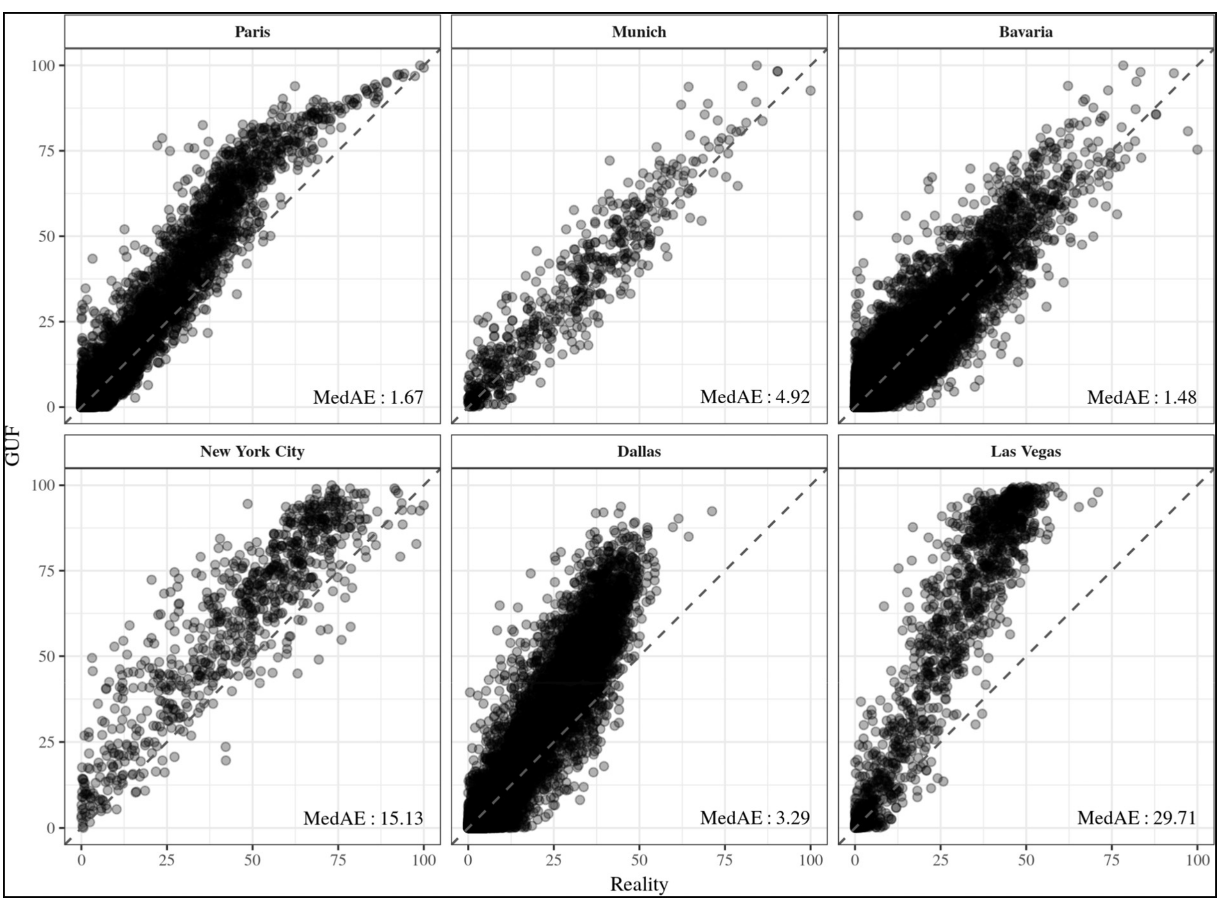


Fig. 3. Validation of the GUF-DenS layer at an aggregated grid of 1 km^2 for different areas: Mean absolute errors (MAE) calculated for Paris, France; Munich, Germany; and the State of Bavaria in Germany consisting of cities and large rural environments and New York City, Dallas und Las Vegas, USA.

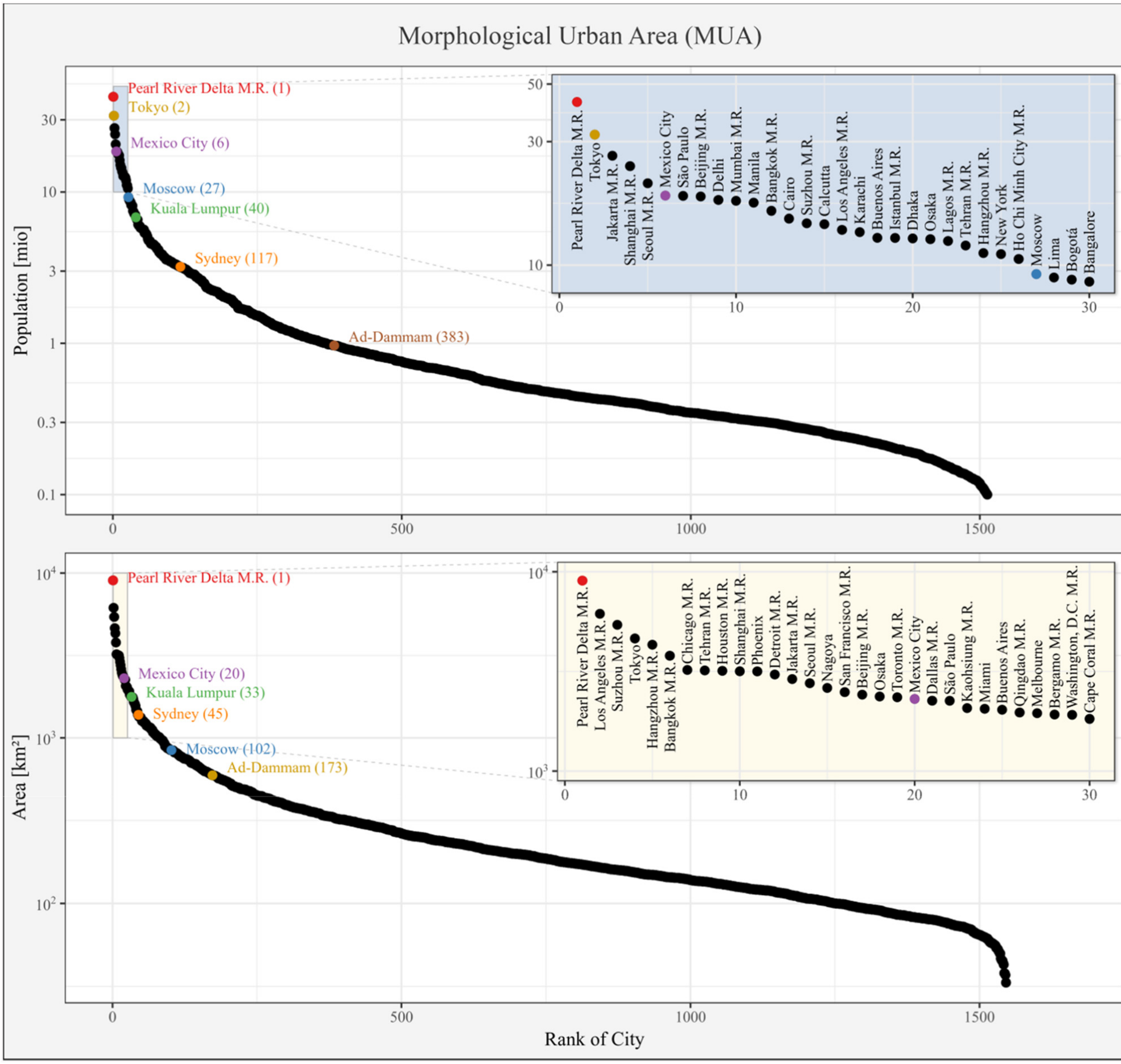


Fig. 4. Rank-size distributions of the largest 1692 cities across the globe: For population derived from re-territorialized MUAs (on top) and for AUAs (below); the details in the boxes present the largest 30 cities.

Kinshasa & Brazzaville, Lima, Rio de Janeiro, London, Paris – are not considered mega-cities based on our newly designed MUAs. In turn, six metropolitan regions do now feature mega-city status based on our approach (Ho Chi Minh City and Bien Hoa M.R. in Vietnam; Seoul, Incheon, Suweon, Seongnam, Goyang Bucheon, Ansan, Anyang, Uijeongbu, Siheung and Gwangmyeong M.R. in South Korea; Suzhou, Wuxi, Changzhou, Jiangyin, Zhangjiagang and Jingjiang M.R. in China; Hangzhou, Shaoxing, Cixi, Yuyao and Shangyu M.R. in China; Tehran, Karaj, Eslamshahr, Malard and Qods M.R. in Iran; Bangkok, Samut Prakan and Nonthaburi M.R. in Thailand). What is striking here is that based on our approach all newly identified mega-cities are located in Asia, while 6 out of 10 mega cities that fall out of this statistic are outside Asia.

4.3. Geographical patterns with regard to the different spatial units (MUAs vs. AUAs and related populations) at continental level

If one looks at the discrepancies between MUAs and AUAs in terms of geographical distribution, some things stand out especially: The previous analysis focused on the largest cities in the world, where some MUAs far exceed their administrative boundaries (as shown e.g. for the PRD M.R.). It might be surprising that in most cases the opposite is measured. The actual tendency reveals that the calculated MUAs are more often much smaller than AUAs. This tendency is particularly pronounced in Asia (Fig. 6). The exception, however, is Europe (Fig. 6). In Europe the spatial units of MUA and AUAs match on median, the distribution, however, indicates that a majority of AUAs is smaller than the calculated MUAs. This reveals a spatial land organization on comparatively small entities. It is also interesting that a gradient from west to east is emerging for Europe. In the west there are many cities that have grown beyond their administrative borders, while in the east this is the other way round. Another example for regional gradients: In China, there are only a few cities that go beyond established administrative boundaries; however, these are found almost exclusively on the east coast. A few other regional peculiarities are the following: in Indonesia, MUAs are measured consistently larger than AUAs, whereas in India, the Arabian Peninsula, as well as in the Middle East, it is predominantly the other way around.

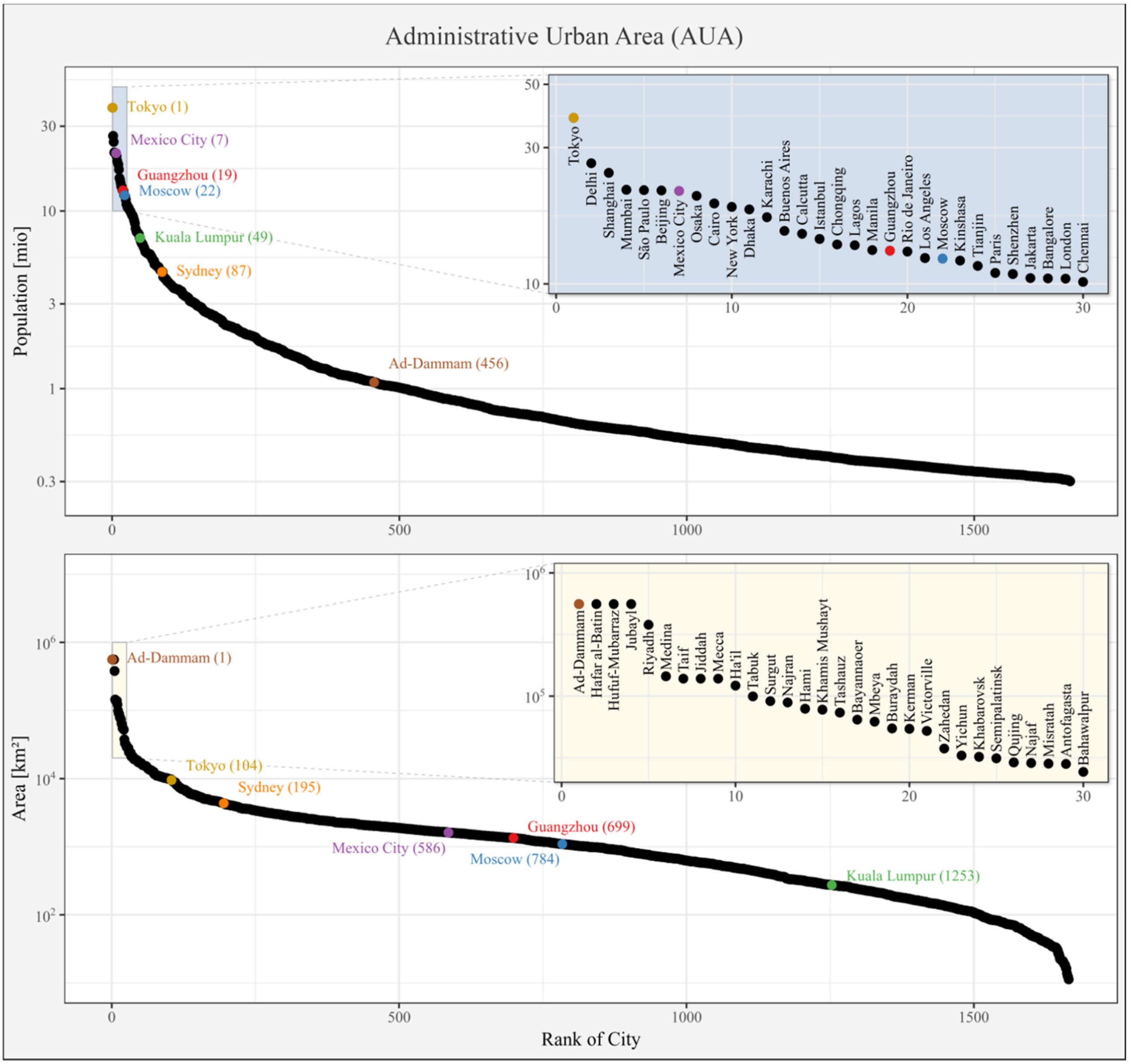


Fig. 5. Rank-size distributions of the conventional administrative space units for the largest 1692 cities across the globe: for population related to AUAs (on top) and for spatial extents (below); the details in the boxes present the largest 30 cities.

If we take the 100 largest cities for comparison, we find that independent from the method of measurement (population or extent based on either MUAs or AUAs) the large share of the largest cities is located in Asia (Table 1). 48 cities based on MUAs and even 63 for population related to MUAs out of 100 are located there. For the AUAs, 73 cities in spatial extent and 58 in population are located in Asia. At the other end, Oceania features only 4 for the MUAs, or even 0 respectively for population; for the AUAs it is 0 or just 2. Remarkable is the difference in city extents in North America: For the MUAs, 32 cities are counted among the largest 100, while for the respective population only 13 belong to this list. This indicates to the extensively large and continuously low dense sprawling cities in the USA with comparatively low dense populations. The detailed list of the 100 largest cities and their attributes are available in supplementary A-2.

5. Discussion and interpretation: Shaky truths and morphologic realities

The morphologically coalesced polycentric metropolitan region in the Pearl River Delta is currently the largest urban agglomeration in the world. Its population is suggested at 42.6 million. Tokyo, however, usually considered the largest city in the world, is ranked second. With 31.9 million it is suggested with fewer inhabitants than in UN statistics. So, in general we see a shake-up in the city sizes and populations and their respective rankings at global scale comparing this alternative, but methodologically and spatially consistent approach using MUAs vs. the common UN statistics relying on administrative units.

So, is in consequence our global urban landscape defined by larger urban entities than we have assumed? There is no simple answer to this question and this must be considered in a differentiated way: On the one hand, conceptual approaches such as ‘mega-regions’ or ‘urban corridors’, which see themselves as an integrative cluster of cities with their surrounding suburban hinterlands forming far larger urban

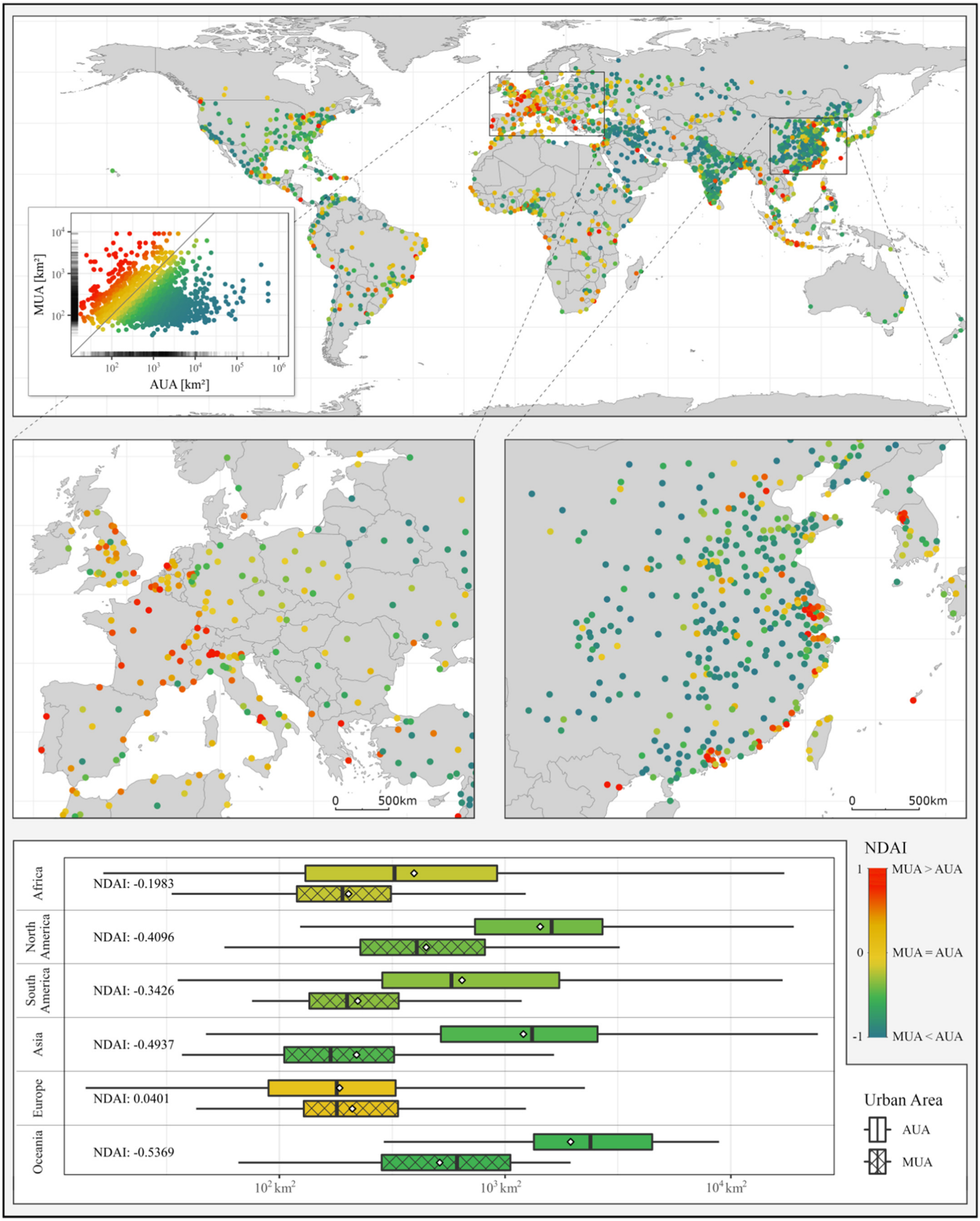


Fig. 6. Global map of cities classified based on the NDAI indicating the relationship of MUAs to AUAs. Boxplots illustrating the MUA distributions versus the administrative units aggregated to continental level. Be aware of a non-linear nature of the NDAI.

Table 1
Geographic trends at continental level: Quantity of cities per continent belonging to the largest 100 cities based on MUAs and AUAs as well as related population figures.

Continent	Population (MUAs)	MUAs	Population (AUAs)	AUAs
Africa	12	5	10	8
Asia	63	48	58	73
Europe	5	6	4	0
North America	13	32	16	16
South America	7	5	10	3
Oceania	0	4	2	0

landscapes, have been identified a long time ago (Gottmann, 1957; Whebell, 1969). In this context, Small et al. (2011) revealed spatial urban extents vastly larger than the administratively-defined cities. These evolving new city patterns have been documented as an interweaving space as well as in their physical shape and are already acknowledged for a new dimension of urban landscapes (e.g. Florida et al., 2008; Taubenböck et al., 2014). On the other hand, within these mega-regions or urban corridors individual cities (e.g. New York City within the Boston to Washington corridor) remain individual lighthouses regardless of being part of a larger urban constellation for the global perception of the city, its economic success, creativity, cultural diversity, and much more. It therefore remains relevant to shed light to these physical dimensions of individual cities within these larger, more loosely bounded urban landscapes. The statistical shift of city sizes measured by MUAs, as done in our study, reveals that we are de facto dealing in parts with larger urban entities than we have assumed in traditional statistics (PRD M.R. vs. Guangzhou within a region as well as PRD M.R. vs. Tokyo in a global ranking). In large parts, however, the statistical shift of city sizes directs us towards smaller entities in MUAs than in administrative units.

From a methodological point of view, this approach aims to translate the agglomeration concept understood as continuous physical built-up landscape into statistical practice. Naturally, the city sizes based on MUAs are indicative rather than definitive. It is clear that there is not a single regional logic, nor a single dimension defining connected spaces, nor is a constructed territorial space "correct" or "incorrect" in absolute measures (Taubenböck et al., 2017). Rather, the approach is sensitive to our assumptions, i.e. the usage of the spatial entity of a monocentric city model, the defined center points, the corresponding indicators and the thresholds set. The monocentric city model is documented empirically robust and analytically tractable (Paulson, 2012), the defined center points are uniformly derived from one database (UN, 2014), the corresponding indicators only show a low correlation and their combination to the MSI makes it more robust and allows for a factual hypothesis, and the thresholds set are also based on a factual approach that produces results that are reasonable in the spatial domain (cf. Figs. 1 and 2). These explanations may be more or less comprehensible; however, we argue here that the decisive added value does not lie in the choice of these specifics, since our results are plausible spatial delimitation of the physical city extents, but in the globally consistent application of these specifics and thus a comparable basis is created. Moreover, it must also be clear to one that the analysis also relates very much to the spatial and thematic scale. The abstract representation of settlements used in our EO-mapping products influences the results; if we consider this mapping product in comparison to lower resolution data such as night-time lights as used by Small et al. (2011) or to geometrically and thematically higher resolved data sets such as three-dimensional city models, which might even be enriched with usage types, different diversities of morphological and functional details allow for different concepts, methods, perceptions and results on delimiting urban from rural. However, the latter higher resolved data sets are not consistently available at global level, if they are available at all and thus do not allow for a global study as performed here. The results of our MUAs also depend on the accuracy of the input data – GUF and GUF Density. We need to be aware that the GUF does not feature a consistent high accuracy across the globe. However, if one considers the studies on this subject, for urban areas, the accuracies are documented consistently high (Taubenböck et al., 2011; Klotz et al., 2016) even in landscape types such as arid regions (Mück et al., 2017). The accuracies of the GUF Density layer are documented also high, however, also with varying precisions for different location. This inconsistency in accuracy for different areas on our planet must remain an unknown, since good and complete reference data are largely unavailable. And even if, as the Las Vegas example shows, the data set tends to be overestimated in certain areas, our approach for delineating cities based on local context information such as the threshold relative to the city average is adaptive and can compensate for misclassification. So in consideration of these aspects, we think the decisive contribution of this study is not 'general truth', but '*a truth through methodological and as far as possible data technical consistency*'.

Geographically, the physical approach of consistently delineating cities using the local context produces MUAs that still contain a large morphological variability in it. It is also clear that measurement of physical contiguity of a city is one-dimensional and does not necessarily imply a high degree of interaction and interdependence within urban landscapes. This means, our analysis focusing on the "space of place" is only one perspective; however, it is in line with the perspective taken by ranking lists on city sizes. For a more comprehensive understanding of the physical and virtual sizes of cities, it needs to be complemented by the "space of flows" within and across city areas. Compared to all these theoretically to be considered data, conceptual and methodical possibilities, the strength of this approach here is to be seen in its global consistency regarding input data, concept, methods, and in consequence geographical results.

So in conclusion, is 42.6 million inhabitants the new correct number for the largest morphologically contiguous urban agglomeration in the world? Unfortunately, this must be questioned, too. *First*, as just discussed, the MUA approach is sensitive to the choice of certain indicators and the input data has its own errors. Manipulating the algorithm would allow us to expand or shrink the MUA resulting in larger or smaller population figures. Let's take the example of the PRD: Our approach separates the MUA of the PRD from the city of Hong Kong, as the topography and the ocean there constitute a natural barrier of low settlement density (or no settlement at all). Our method takes this circumstance into account. Many urban geographers, however, would argue that Hong Kong is functionally part of the PRD metropolitan region. By separating these morphological units, we measure the PRD with 42.6 million and Hong Kong with 4.5 million. With a different conceptual and methodological approach the PRD could therefore also be measured at 47.1 million by adding these numbers. So, we have to recognize, there is no 'absolute truth' to it; we argue if you look closely at the results in Fig. 1c and 2, you see that the derived MUAs capture the morphological urban space well and thus our approach is reasonable, transparent and consistent without claiming to be the only truth. *Second*, while the largest MUA according to our approach comprises 42.6 million people, we find in the extended Shanghai region a very high density of MUAs in close proximity to each other on a comparatively small area of 250 × 300 km. 17 MUAs – among them are the Shanghai, Kunshan, Taicang MUA (with 24.1 million the fourth largest in the world), the Suzhou, Wuxi, Changzhou, Jiangyin, Zhangjiagang, Jingjiang MUA (with 14.5 million the 14th largest) and the Hangzhou, Shaoxing, Cixi, Yuyao, Shanggyu MUA (with 11.1 million the 24th largest) – add up to 61.2 million inhabitants (the entire stretch is even home to 96 million). While we list these 17 individual MUAs in our statistics separately, this clustering of so many large MUAs indicates a larger urban agglomeration developing within the above mentioned conceptual framework of a mega-region. From this conceptual point of view, this mega-region (which is about the size of Austria) could be described with 96 million inhabitants as the largest urban agglomeration in the world. With it the urban agglomeration would be ranked 16 among countries with a population larger than Germany or Turkey. *Third*, as shown in other studies, it is highly probable that today's population figures especially in such dynamically growing large cities are still rather underestimated, mainly due to the difficulty to record informal population groups (e.g. Taubenböck and Wurm, 2015). Ultimately, therefore, the absolute population figures determined must be regarded as uncertain (just take the possibly 20 million informal workers in the PRD M.R. (Liang et al., 2014)). However, the relative population figures appear to be much more consistent than any previous data. This is due to the comparable spatial basis of consistently derived MUAs and the globally consistent input data. In summary this means, 42.6 million might not be the new correct number for the largest city in the world, but if you take a comparable spatial baseline the PRD M.R. is de facto the currently largest individual urban agglomeration in the world.

Consequently, the globally consistent and harmonized approach mapping MUAs provides a comparable basis for geographic research. It thus allows scrutinizing administrative units whether they indicate close to reality statistical information and/or whether they form meaningful areas of political competence. As we find only 3.7% of the 1692 cities 'true-bounded' we can now clearly state that the usual statistics rely on for comparisons actually inadmissible spatial entities and thus obscure morphologic reality. The newly derived MUAs may allow overcoming the sometimes arbitrary effects caused by AUAs, they may reveal misjudgments based on previously accepted statistics or they will make us re-think about whether existing spatial units should or should not be reformed. Ranking Los Angeles M.R. at number 2 (by MUA size) or at 16 (by MUA population) is both correct but global perception is, depending on the particular list, fundamentally different. The results discussed thus make clear that one must critically question every ranking list – knowing that one single truth may not exist.

6. Conclusion and outlook

Yes, administrative units obscure morphologic reality and significantly influence statistics and perceptions of cities. We find our planet already consists of larger individual city entities than generally accepted. The metropolitan region of the Pearl River Delta is currently the largest urban agglomeration in the world instead of Tokyo. We need to re-think current ranking lists on the spatial and demographic dimension of urbanization in a critical manner.

This work is a plea to overcome historical or arbitrary spatial units for generating statistics, and more importantly for managing our living environments from jurisdictional, political and planning perspectives. As we could show, in most cases the administrative spatial units and the true morphologic extents of cities do not match. Our newly generated spatial entities derived in consistent manner may form an admissible spatial baseline for better and more comparable statistics and urban research studies in the future. We propose that this approach needs to be extended systematically by analyzing the influence of different data and methods for city delineations as well as by a multi-temporal analysis of city size development over time for further geographic findings.

Isn't it interesting how old questions that seemed answered can now be re-examined in a more objective way, but remain unanswerable? By this re-examination, however, we believe that based on our analysis the opening quote of this article would be closer to the truth if changed to "*the Pearl River Delta M.R. is the world's largest city with an agglomeration of 42.6 million inhabitants, followed by Tokyo with 31.9 million, Jakarta M.R. with 26.5 million, Shanghai M.R. with 24.1 million and Seoul M.R. with 20.7 million Today, Mexico City, Sao Paulo, Delhi, Mumbai, Beijing M.R., and Manila have all close to 20 million inhabitants*".

Acknowledgements

We would like to thank our supporters for their commitment to this project: Kevser Basdas, Carla Madueno, Nicole Osterkamp, and Thomas Spieß.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2019.111353.

References

Abed, J., Kaysi, I., 2003. Identifying urban boundaries: applications of remote sensing and GIS technologies. Can. J. Civ. Eng. 30 (6), 992–999.
Anas, A., Arnott, R., Small, K., 1998. Urban spatial structure. J. Econ. Lit. 36 (3), 1426–1464.
Angel, S., Parent, J., Civco, D.L., Blei, A.M., Potere, D., 2011. The dimensions of global urban expansion: estimates and projections for all countries, 2000—2050. Prog. Plan. 75 (2), 53–108.
Auerbach, F., 1913. Das Gesetz der Bevölkerungskonzentration. Petermanns Geogr. Mitt. 59, 74–76.
Batty, M., Besussi, E., Maat, K., Harts, J.J., 2004. Representing multifunctional cities: density and diversity in space and time. Built Environ. 30 (4), 324–337. https://doi.org/10.2148/benv.30.4.324.57156.
Berry, J.L., Okulicz-Kozaryn, A., 2012. The city size distribution debate: resolution for US urban regions and megalopolitan areas. Cities 29, 17–23.
Castells, M., 1999. Grassrooting the space of flows. Urban Geogr. 20, 294–302.
Castells, M., 2000. The Rise of the Network Society, 2nd ed. Wiley-Blackwell, Oxford, UK.
Decker, E.H., Kerkhoff, A.J., Moses, M.E., 2007. Global patterns of City size distributions and their fundamental drivers. PLoS One 2 (9), e934. https://doi.org/10.1371/journal.pone.0000934.
Deuskar, C., 2015. What does "urban" mean? http://blogs.worldbank.org/sustainablecities/what-does-urban-mean.
Dijkstra, L., Poelman, H., 2014. A harmonized definition of cities and rural areas: the new degree of urbanization. Working Papers WP01/2014. http://ec.europa.eu/regional_policy/sources/docgener/work/2014_01_new_urban.pdf.
Eeckhout, J. (2004). Gibrat's law for (all) cities. Am. Econ. Rev., 94(5), 1429–1451. https://doi.org/10.1257/0002828043052303 http://pubs.aeaweb.org/doi/10.1257/0002828043052303.
Elvidge, C., Tuttle, B.T., Sutton, P.C., Baugh, K.E., Howard, A.T., Milesi, C., Nemani, R., 2007. Global distribution and density of constructed impervious surfaces. Sensor 7 (9), 1962–1979.
Esch, T., Himmler, V., Schorcht, G., Thiel, M., Conrad, C., Wehrmann, T., Bachofer, F., Schmidt, M., Dech, S., 2009. Large-area assessment of impervious surface based on integrated analysis of single-date Landsat-7 images and geospatial vector data. Remote Sens. Environ. 113 (8), 1678–1690 2009.
Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felbier, A., Thiel, M., ... Dech, S., 2012. TanDEM-X mission—new perspectives for the inventory and monitoring of global settlement patterns. J. Appl. Remote. Sens. 6, 1–21.
Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A., Dech, S., 2013. Urban footprint processor – fully automated processing chain generating settlement masks from global data of the TanDEM-X Mission. IEEE Geosci. Remote Sens. Lett. 10 (6), 1617–1621. November 2013. ISSN 1545-598X. https://doi.org/10.1109/LGRS.2013.2272953.
Esch, T., Marconcini, M., Marmanis, D., Zeidler, J., Elsayed, S., Metz, A., Müller, A., Dech, S., 2014. Dimensioning urbanization - an advanced procedure for characterizing human settlement properties and patterns using spatial network analysis. Appl. Geogr. 55 (2014), 212–228.
Esch, T., Bachofer, F., Heldens, W., Hirner, A., Marconcini, M., Palacios-Lopez, D., Roth, A., Üreyen, S., Zeidler, J., Dech, S., Gorelick, N., 2018a. Where we live – a summary of the achievements and planned evolution of the Global Urban Footprint. Remote Sens. 10 (6), 895 2018. (18 pp.;).
Esch, T., Üreyen, S., Zeidler, J., Hirner, A., Asamer, H., Metz-Marconcini, A., Tum, M., Boettcher, M., Kuchar, S., Svaton, V., Marconcini, M., 2018b. Exploiting big earth data from space – first experiences with the TimeScan processing chain. Big Earth Data. https://doi.org/10.1080/20964471.2018.1433790. (ISSN: 2096-4471 (Print) 2574-5417 (Online).
Florida, R., Gulden, T., Mellander, C., 2008. The rise of the mega-region. Camb. J. Reg. Econ. Soc. 1, 459–476.
Fragkias, M., Seto, K.C., 2009. Evolving rank-size distributions of intra-metropolitan urban area clusters in South China. Comput. Environ. Urban. Syst. 33 (3), 189–199.
GADM, 2018. Database of Global Administrative Areas. GADM, Montreal, QC, Cannada.
Georg, I., Blaschke, T., Taubenböck, H., 2018. Are we in Boswash yet? A multi-source geodata approach to spatially delimit urban corridors. ISPRS Internatl. Journal of Geo-Information 7 (1), 1–15.
Gottmann, J., 1957. Megalopolis or the urbanization of the northeastern seaboard. Econ. Geogr. 189–200.
Harrison, J., Growe, A., 2014. From places to flows? Planning for the new 'regional world' in Germany. Eur. Urban Reg. Stud. 21, 22–41.
Jessop, B., Brenner, N., Jones, M., 2008. Theorizing sociospatial relations. Environ. Plan. D 26, 389–401.
Klotz, M., Kemper, T., Geiß, C., Esch, T., Taubenböck, H., 2016. How good is the map? A multi-scale cross-comparison framework for global settlement layers: evidence from Central Europe. Remote Sens. Environ. 178, 191.212.
Lechner, A.M., Reinke, K.J., Wang, Y., Bastien, L., 2013. Interactions between landcover

pattern and geospatial processing methods: effects on landscape metrics and classification accuracy. Ecol. Complex. 15, 71–82.
Liang, Z., Li, Z., Ma, Z., 2014. Changing patterns of the floating population in China during 2000–2010. Popul. Dev. Rev. 40 (4), 695–716.
Liu, Y., Delahunty, T., Zhao, N., Cao, G., 2016. These lit areas are undeveloped: delimiting China's urban extents from thresholded nighttime light imagery. Int. J. Appl. Earth Obs. Geoinf. 50, 39–50.
Masucci, A.P., Arcaute, E., Hatna, E., Stanilov, K., Batty, M., 2015. On the problem of boundaries and scaling for urban street networks. J. R. Soc. Interface 12 (111).
McMillen, D., 2006. Testing for monocentricity. In: Arnott, R.J., McMillen, D.P. (Eds.), A Companion to Urban Economics. Blackwell Publishing, Malden, MA, pp. 128–140.
Mück, M., Klotz, M., Taubenböck, H., 2017. Validation of the DLR global urban footprint in rural areas: a case study for Burkina Faso. In: IEEE-CPS Joint Urban Remote Sensing Event (JURSE). VAE, Dubai.
Nitsch, V., 2005. Zipf zipped. J. Urban Econ. 57 (1), 86–100.
Openshaw, S., 1983. The Modifiable Areal Unit Problem. Geo Books, Norwick [Norfolk].
OSM, 2018. OpenStreetMap. https://www.openstreetmap.org.
Parr, J, B. (2007): Spatial definition of the city: four perspectives. Urban Stud., vol. 44, no. 2, pp. 381–392.
Paulson, K., 2012. Yet even more evidence on the spatial size of cities: using spatial expansion in the US, 1980–2000. Reg. Sci. Urban Econ. 42, 561–568.
Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Zanchetta, L., 2013. A global human settlement layer from optical HR/VHR RS data: concept and first results. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6 (6), 2102–2131.
Riitters, K.H., O'Neill, R.V., Hunsaker, C.T., Wickham, J.D., Yankee, D.H., Timmins, S.P., Jones, K.B., Jackson, B.L., 1995. A factor analysis of landscape pattern and structure metrics. Landsc. Ecol. 10, 23–40.
Ross, C.L., 2011. Literature Review of Organizational Structures and Finance of Multi-jurisdictional Initiatives and the Implications for Megaregion Transportation Planning in the U.S. 2011. http://www.fhwa.dot.gov/planning/publications/megaregions_report_2012/megaregions2012.pdf>.
Rozenfeld, Hernán D., Rybski, Diego, Gabaix, Xavier, Makse, Hernán A., 2011. The area and population of cities: new insights from a different perspective on cities. Am. Econ. Rev. 101 (5), 2205–2225.
Schneider, A., Friedl, M.A., Potere, D., 2009. A new map of global urban extent from MODIS satellite data. Environ. Res. Lett. 4.
Siedentop, S., Fina, S. (2010): Monitoring urban sprawl in Germany. Towards a GIS-based measurement and assessment approach. In: J. Land Use Sci. 5 (No.2, June 2010), S. 73–104.
Simon, D., 2008. Urban environments: issues on the peri-urban fringe. The Annual Review of Environment and Resources 33, 167–185.
Small, C., Sousa, D., 2016. Humans on earth: global extents of anthropogenic land cover from remote sensing. Anthropocene 14, 1–33.
Small, C., Elvdigde, C.D., Balk, D., Montgomery, M., 2011. Spatial scaling of stable night lights. Remote Sens. Environ. 115, 269–280.
Small, C., Sousa, D., Yetman, G., Elvidge, C., MacManus, K., 2018. Decades of urban growth and development on the Asian megadeltas. Glob. Planet. Chang. 165, 62–89.
Soo, K.T., 2005. Zipf's law for cities: a cross-country investigation. Reg. Sci. Urban Econ. 35 (3), 239–263.
Sorichetta, A., et al., 2015. High-resolution gridded population datasets for Latin America and the Caribbean in 2010, 2015, and 2020. Sci Data 2, 150045.
Spivey, C., 2008. The Mills–Muth model of urban spatial structure: surviving the test of time? Urban Stud. 45 (2), 295–312.
Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J., 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PLoS One 10 (2), e0107042. https://doi.org/10.1371/journal.pone.0107042.
Tatem, A.J., 2017. WorldPop, open data for spatial demography. Scientific Data 4. https://www.nature.com/articles/sdata20174.
Taubenböck, H., Wiesner, M., 2015. The spatial network of megaregions – types of connectivity between cities based on settlement patterns derived from EO-data. Computers, Environment & Urban Systems 54, 165–180.
Taubenböck, H., Wurm, M., 2015. Ich weiß, dass ich nichts weiß – Bevölkerungsschätzung in der Megacity Mumbai. In: Taubenböck, H., Wurm, M., Esch, T., Dech, S. (Eds.), Globale Urbanisierung – Perspektive aus dem All. SpringerSpektrum, Berlin/Heidelberg, pp. 171–178.
Taubenböck, H., Esch, T., Felbier, A., Roth, A., Dech, S., 2011a. Pattern-based accuracy assessment of an urban footprint classification using TerraSAR-X data. IEEE Geosci. Remote Sens. Lett. 8 (2), 278–282.
Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., Dech, S., 2012. Monitoring urbanization in mega cities from space. Remote Sens. Environ. 117, 162–176.
Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., Dech, S., 2014. New dimensions of urban landscapes: the spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. Appl. Geogr. 47, 137–153.
Taubenböck, H., Standfuß, I., Klotz, M., Wurm, M., 2016. The physical density of the city – deconstruction of the delusive density measure with evidence from two European megacities. ISPRS Internatl. Journal of Geo-Information 5 (11), 1–24.
Taubenböck, H., Ferstl, J., Dech, S., 2017. Regions set in stone – classifying and categorizing regions in Europe by settlement patterns derived from EO-data. ISPRS Internatl. Journal of Geo-Information 6 (2), 1–27.
Taubenböck, H., Wurm, M., Geiß, C., Dech, S., Siedentop, S., 2019. Urbanization between compactness and dispersion – designing a spatial model for measuring 2D binary settlement landscape configurations. International Journal of Digital Earth 12 (6), 679–698.
United Nations, Department of Economic and Social Affairs, Population Division, 2014. World Urbanization Prospects: The 2014 Revision, CD-ROM edition. Data publication. https://esa.un.org/unpd/wup/CD-ROM/WUP2014_XLS_CD_FILES/WUP2014-F12-Cities_Over_300K.xls.
United Nations, Department of Economic and Social Affairs, Population Division, 2015. World urbanization prospects: the 2014 revision, (ST/ESA/SER.A/366). Supplementary file 22. https://population.un.org/wup/Publications/Files/WUP2014-Report.pdf.
United Nations, Department of Economic and Social Affairs, Population Division, 2018. 2018 Revision of the World Urbanization Prospects. https://www.un.org/development/desa/publications/2018-revision-of-world-urbanization-prospects.html.
Weber, C., 2001. Urban agglomeration delimitation using remote sensing. In: Donnay, J., Barnsley, M., Longley, P. (Eds.), Remote Sensing and Urban Analysis. U.K. Taylor and francis, London, pp. 145–159.
Whebell, C.F., 1969. Corridors: a theory of urban systems. Ann. Assoc. Am. Geogr. 59 (1), 1–26.
World Development Report, 2009. Reshaping Economic Geography. World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/5991 (License: CC BY 3.0 IGO.).
Zipf, G.K., 1941. National Unity and Disunity - the Nation as a Bio-Social Organism. Bloomington, Indiana. https://babel.hathitrust.org/cgi/pt?id=mdp.39015057175484;view=1up;seq=5.