Cities 133 (2023) 104129

Contents lists available at ScienceDirect

Cities

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

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

Isabel Georg [a,b,*], Thomas Blaschke [a], Hannes Taubenböck [b]

[a] Department of Geoinformatics, University of Salzburg, 5020 Salzburg, Austria
[b] Earth Observation Center, Remote Sensing Data Center, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany

ARTICLE INFO

Keywords:
Urban corridor
Large urban areas
Urban extent
Spatial delineation

ABSTRACT

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

1. Introduction

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

1.1. Urban corridors: a spatial view

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

According to UN-Habitat (2008, 2013), three main forms of these novel configurations are megaregions (or mega-regions), urban corridors and city regions (or city-regions); in addition, a plethora of terms describing large urban areas of various shapes and sizes exists. Of these, megaregions in particular have been extensively studied (e.g., Banerjee, 2012; Florida et al., 2008; Growe et al., 2015; Hall, 2009), but their spatial delineation remains subjective: There is no common agreement with respect to which indicators (e.g. economic output, population, physical dimensions or functional linkages) or spatial entities (e.g., administrative boundaries or natural regions such as watersheds) allow

* Corresponding author at: Department of Geoinformatics, University of Salzburg, 5020 Salzburg, Austria.
E-mail addresses: isabel@southward-georg.com (I. Georg), thomas.blaschke@sbg.ac.at (T. Blaschke), hannes.taubenboeck@dlr.de (H. Taubenböck).

https://doi.org/10.1016/j.cities.2022.104129
Received 8 August 2021; Received in revised form 23 November 2022; Accepted 27 November 2022
Available online 15 December 2022

a meaningful spatial delimitation. Similarly, other concepts of large urban areas are also not universally defined (see for example Fang & Yu, 2017, who aim to provide a definition of 'urban agglomeration' based on a literature review). This includes the construct of urban corridors which, although the term was introduced over 50 years ago (Whebell, 1969) to describe the urban development in Southern Ontario (Canada), has not gained a large amount of influence or empirical evidence.

From a purely spatial perspective, an urban corridor is a linearly-shaped and massively large polycentric agglomeration consisting of several big cities (often megacities) or city clusters and their hinterlands, connected through high-speed road or rail infrastructure (Li & Cao, 2005; Neuman, 2000; Priemus & Zonneveld, 2003; Trip, 2003; UN-Habitat, 2008 & 2013). Other interpretations imply the transportation network itself – as is the case with road corridor (Forman & Alexander, 1998), transport corridor (Curtis, 2006; Jain et al., 2021; Trip, 2003; Van Houtum & Lagendijk, 2001) or freight corridor (Rodrigue, 2004).

Examples include *BESETO* (Beijing–Seoul–Tokyo) (Choe, 1996), GILA (Greater Ibadan–Lagos–Accra) or Lagos–Abidjan (Hertzog, 2020; UN-Habitat, 2008), the *Blue Banana* from Birmingham to Milan (Hospers, 2002), the Delhi-Mumbai Industrial Corridor (DMIC) (Chandrashekar & Aithal, 2021; Jain et al., 2014) or – one of the most typical examples – the *Boswash* area from Boston to Washington (Florida et al., 2008; Gottmann, 1957).

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

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

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

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

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

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

2. Data, study area and method

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

- Without a priori knowledge about the spatial properties of urban corridors and without predefining them, we base our approach on a 'master' corridor that is generally accepted by scholars to be a prototype: the *Boswash* area on the northeastern seaboard of the US (e.g. Florida et al., 2008; Regional Plan Association, 2007; Lang & Dhavale, 2005; Gottmann, 1957). We identify individual thresholds for several spatial input layers to describe the *Boswash* 'master' corridor as one coalescent patch. This basic threshold is then transferred to the whole study area to delineate coherent large urban corridors. Our initial focus is on 13 regions which we identified through a combination of a literature review and questionnaire (Georg, Blaschke, Taubenböck, 2016b). From these, we derive 'urban corridor candidates', calculate their dimensions and check for conformity with criteria describing urban corridors. The Method section below

further describes the generation of probabilistic extents of urban corridors and grouping into four categories according to their spatial dimensions.

- Spatial analyses are inherently ambiguous and contain uncertainties. These include distortions and errors of data scaling and aggregation described as the Modifiable Aerial Unit Problem (MAUP) (Fotheringham & Rogerson, 1993): Even at the same scale, there are numerous possible combinations of spatial units which can make a difference to analysis results (Openshaw, 1984; Pietrzak, 2014; Marceau, 1999). In this study, we aggregate all input data sets to a 1 km^2 grid which we believe provides sufficient detail for a large spatial extent such as our study area. 1 km^2 grids have been used elsewhere for similar-scale projects, such as the INSPIRE Geographical Grid System (INSPIRE, 2014) or Hansen et al. (2000).
- Another issue in spatial analysis are boundary effects (Fotheringham & Rogerson, 1993). Hence, rather than relying on administrative boundaries, we use a homogeneous base without pre-defined borders and edges. Furthermore, our results show variable, fuzzy extents of urban corridors, which we believe is a good fit for the concept. Fuzzy boundaries are more effective for a functional connectedness and flexible cooperation between stakeholders (Nelson & Rae, 2016; Yang & Hillier, 2007). Fuzziness is frequently expressed in Boolean terms through a class membership (such as 'artificial surfaces', 'vegetation', 'water' etc.) of either 0 or 1. Fuzzy logic, on the other hand, deals with the possibility of partial membership to a specific class, without clear boundaries (Zadeh, 1965). Membership to a class is between 0 and 1, with the degree described as numerical information (Malczewski, 2004). For the delimitation of urban corridors, a distinction between "corridor" and "not corridor" is diffuse and a single criterion is not likely to provide adequate information; rather, a combination of characteristics is necessary. This is addressed in this manuscript through the use of different datasets representing different aspects of urbanity. The actual membership of an area to an urban corridor is described through membership values obtained through different input datasets.
- The input data used are from a variety of sources to ensure a multi-dimensional and therefore better empirical robustness of our approach. By this we avoid data related coincidences and smooth out inaccuracies of the datasets. Furthermore, urban corridors are fundamentally based on several indicators (rather than just one) which means that our data approach approximates the actual concept as closely as possible.
- Our approach describes urban corridors from a spatial perspective using physically measurable parameters. Functional or relational aspects are not taken into account due to data availability.

2.1. Input data

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

- *Night-time light data (NTL)*: We used data from the Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS) cloud-free stable lights composite from 2010 (Elvidge, n.d.), with a resolution of 30 arcsec and gas flares removed. Night-time imagery has been used for urban studies since the 1990s (Elvidge et al., 1997; Imhoff et al., 1997; Small et al., 2005). Apart from urban area extent, these night-time lights are an indicator for energy consumption, population density and GDP (gross domestic product) (e.g. Florida et al., 2008; Loveland et al., 2000).
- The availability of consistent global records for human settlements is quite limited and includes datasets such as the Global Human Settlements Layer (GHSL) (Pesaresi et al., 2015) and *Global Urban Footprint (GUF)* (Esch et al., 2012) which has been used in this study. The GUF is a data product from the German Aerospace Centre (DLR) derived from TerraSAR-X imagery. Based on 2010–2012 data, the 12 m product shows man-made settlement structures with a vertical dimension (particularly buildings) (Esch et al., 2012).
- *Global Land Cover (GLC-share)*: The dataset provided by the United Nation's Food and Agriculture Organization (FAO) consists of eleven different land use classes from different sources such as medium-resolution satellite sensors – e.g., MERIS (Medium Resolution Imaging Spectrometer) or MODIS (Moderate-Resolution Imaging Spectroradiometer) – or national mapping schemes (Latham et al., 2014). For this study, we only use the artificial surfaces layer.
- *UCLouvain's GlobCover*: The land cover dataset from ESA is obtained from MERIS data (Bontemps et al., 2011; European Space Agency (ESA), n.d.). It consists of 22 land cover classes and similar to the GLC data, we only select the artificial surfaces class. This is defined as having an 'urban' percentage of more than 50 % for each pixel.
- *Global Impervious Surface Area (ISA) dataset*: It is based on DMSP-OLS night-time imagery and population count with a resolution of 1000 m. It includes man-made surfaces including buildings and roads (NOAA (National Oceanic and Atmospheric Administration), n.d.).
- *Population density data*: these are available from the Center for International Earth Science Information Network at Columbia University (CIESIN)'s Socioeconomic Data and Applications Center (SEDAC) (CIESIN, 2017). The version 4 (GPWv4) datasets consists of population counts from national censuses and population registers.

Each of the datasets offers a different interpretation of urban extents and their spatial relationships. High *NTL emissions* are a proxy for economic activity (Mellander et al., 2015) and thus used here for the spatial delineation of contiguous areas of increased economic activity. An increased *settlement density* as shown in the GUF implies coherence and

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

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

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

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

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

2.2. Study area

Our study area covers the whole connected United States territory plus several hundred km to the north and south. With an extent ranging north of Edmonton, the most densely populated parts of Canada are included. To the south, the study area extends south of Aguascalientes in Mexico and comprises the northern part of Cuba and most of the Bahamas. The northern and southern boundaries were chosen to include possible urban corridor areas which transgress the administrative boundaries of the US.

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

2.3. Method

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

2.3.1. Identification of urban corridor candidates

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

As an initial step, we define how large urban areas can be best described uniformly. Since we chose not to make a priori any predetermination of threshold values for delineation, we model this on an

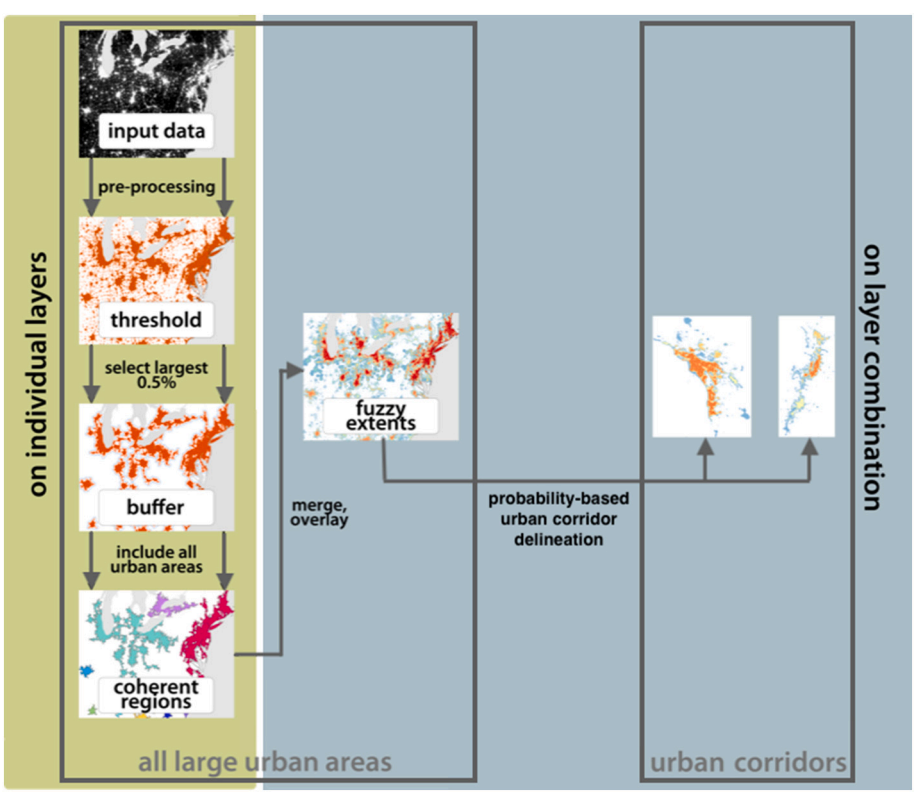

Fig. 1. Workflow of the urban corridor delineation. The green part shows the analysis performed on each individual input layer separately (pre-processing, selection based on a threshold, buffer and identification of coherent regions). The steps in the blue part are completed on all layers combined. The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

area generally accepted in scientific studies as a prime example of an urban corridor: the well-documented *Boswash* area from Boston, MA, to Washington, DC, which we use as a 'master'. Also known as *Megalopolis* (Gottmann, 1957) or *Northeast Megaregion* (Lang & Dhavale, 2005), the area covers around a fifth of both US population and GDP, but only around 2 % of the land area (Lang & Dhavale, 2005). Because it can be separated from surrounding, less urbanized space, the area is suitable as a model from which to derive a spatial delimitation of large urban areas. Georg et al. (2018) developed a method for a fuzzy delimitation of Boswash, using multi-source earth observation, infrastructure and socioeconomic data including night-time lights, settlement classifications, road density and income data.

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

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

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

2.3.2. *From candidates to corridors*

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

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

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

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

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

Finally, we eliminate areas identified through this approach which do not show characteristics typical for urban corridors (a specific

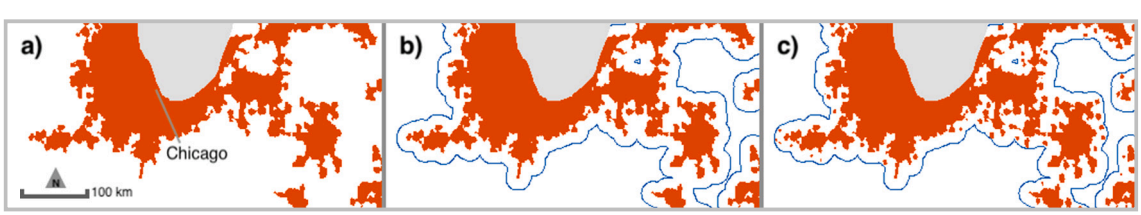

Fig. 2. The Chicago urban area identified by the largest 0.5 % of urban patches (a) in the NTL dataset. A buffer is applied (b) and all omitted smaller patches within this buffer are included again (c). The bottom right patch is an example of the buffer size (b, c): it is not part of the buffer around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

minimum length, width, length-to-width ratio or area) as suggested by Georg, Blaschke, Taubenböck (2016b). The remaining areas represent those regions best classified as urban corridors. To account for the varying extents of these, we group them into four categories based on the area measurements for three agreeing input layers.

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

3. Results

3.1. Mapping of urban corridor candidates

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

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

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

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

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

From these separate large urban area clusters, we select for each layer the areas identified as urban corridors (cf. Georg, Blaschke,

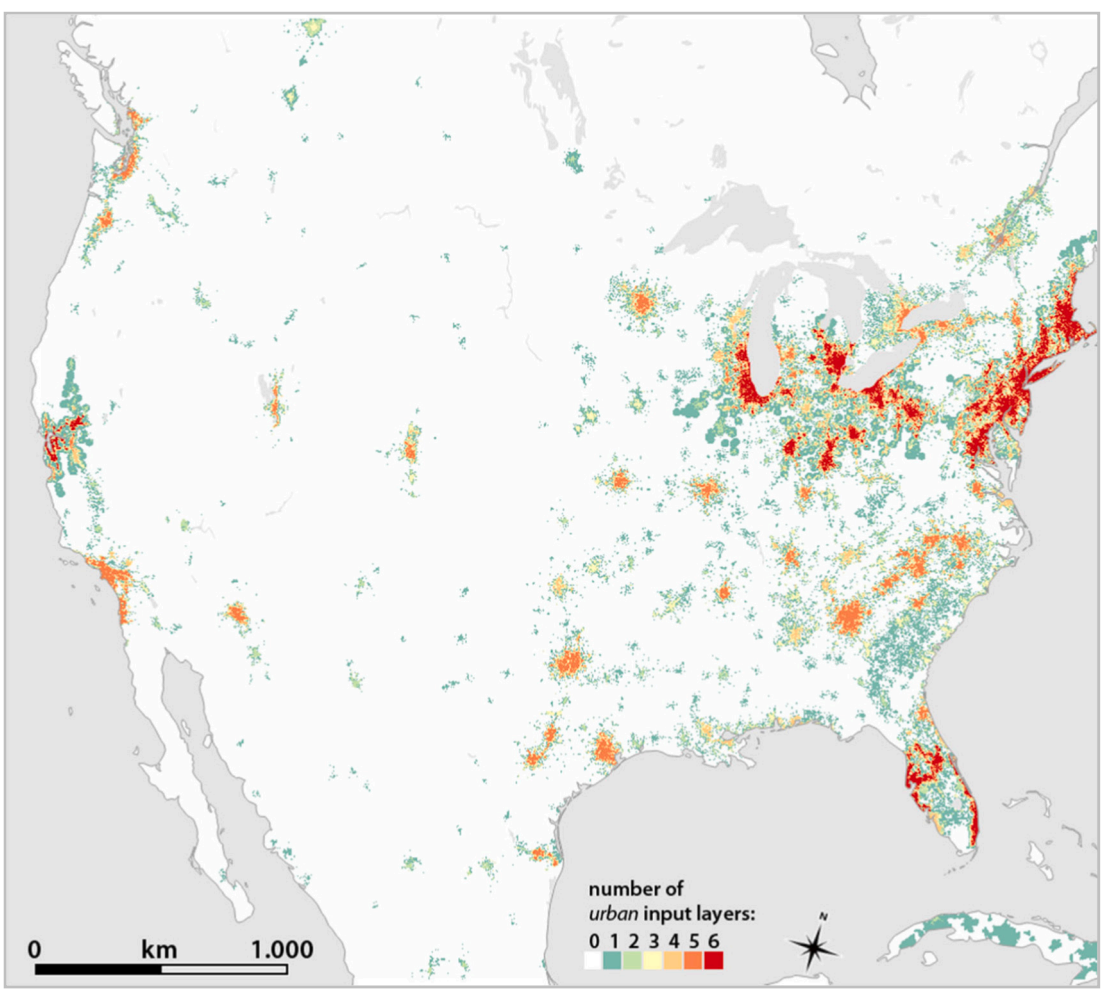

Fig. 3. Overlay of top 0.5 % of large urban areas, color-coded according to the number of input layers defining 'urban'.

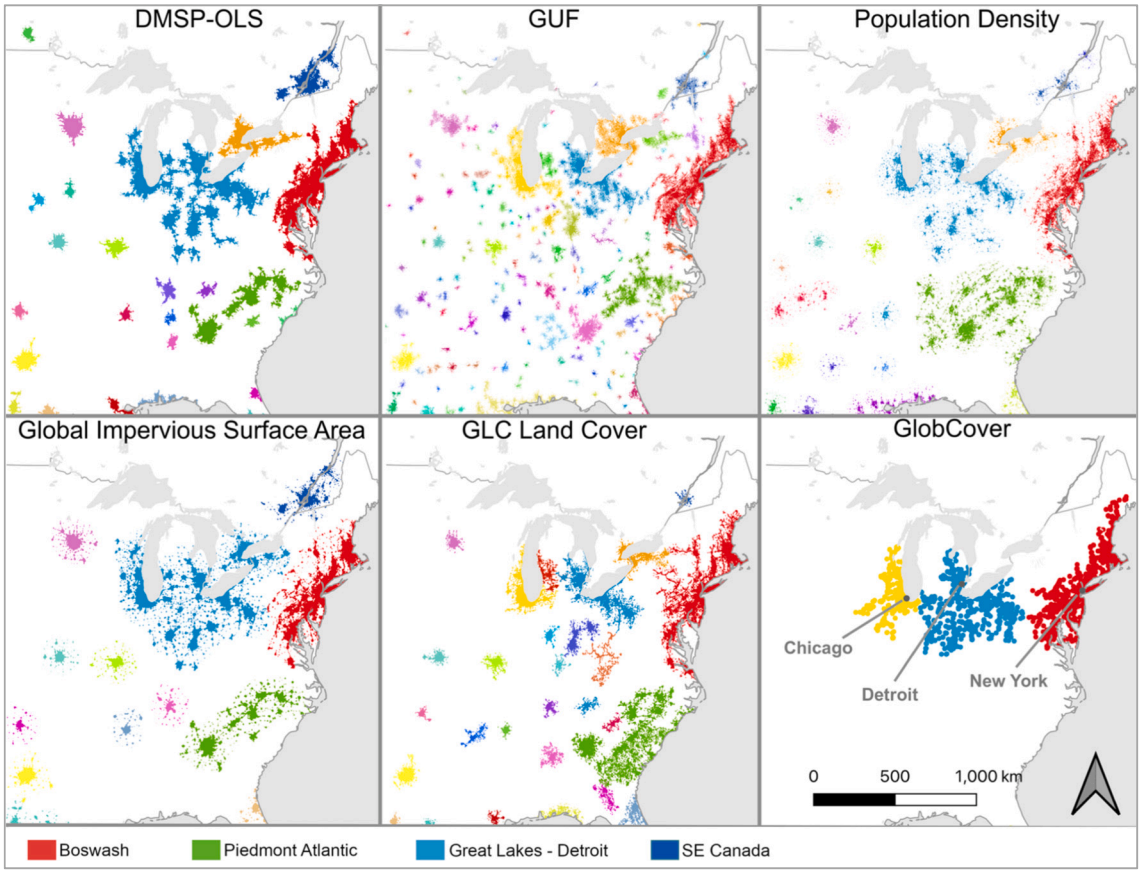

Fig. 4. Grouped urban clusters for each input layer (north-eastern part of the study area). While Boswash is a separate connected patch in each input layer, other patches are less concordant. Main colors: red (Boswash), green (Piedmont Atlantic), blue (Great Lakes - Detroit), dark blue (SE Canada), random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

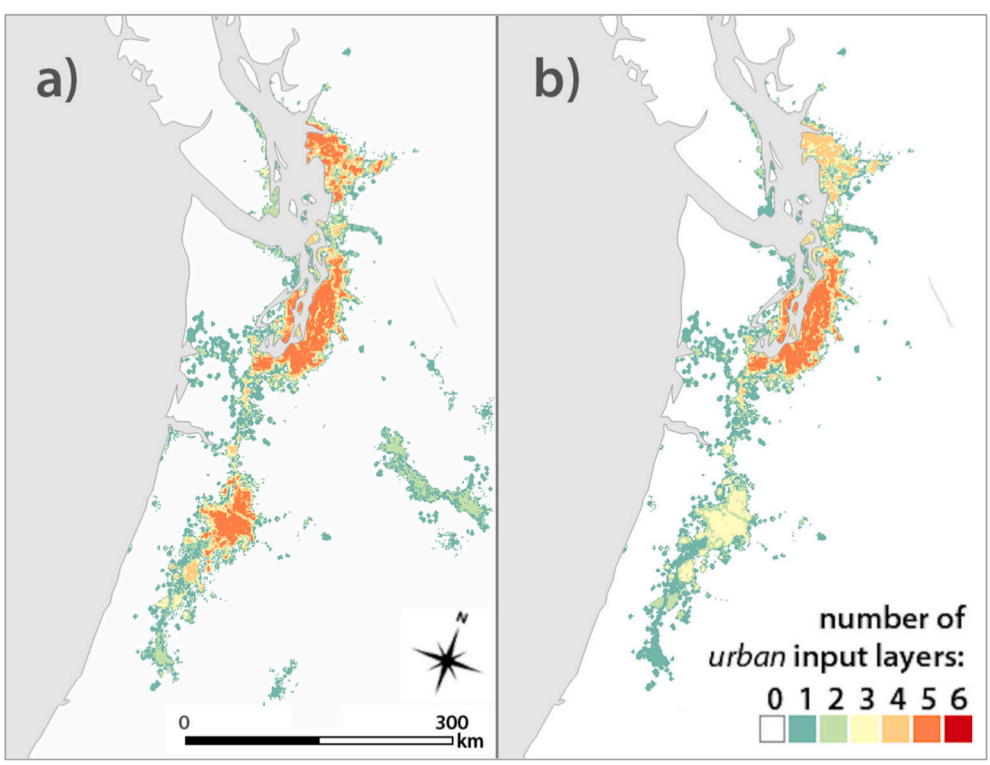

Fig. 5. Cascadia: a) all urban areas classified, b) urban areas belonging to the urban corridor based on our methodology.

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

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

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

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

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

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

Boswash and South California display a relatively consistent length-to-width ratio, Florida and North California a slight decrease. Quebec-Montreal and – despite a peak in between – Cascadia show a significant ratio decrease, indicating a stronger focus on one central urban area (here: Montreal and Seattle) with a higher compactness.

Some areas are only identified through a maximum of three (Gulf Coast) or four (Austin-San Antonio, I-35 Oklahoma) matching layers. With the exception of Oklahoma (which has been excluded above), all show a relatively lengthy shape, indicating a possible urban corridor. In Texas, it is noticeable that a connection between Houston and Austin-San Antonio can be seen for one match, while for several matches, this

Table 2
19 corridor candidates with their respective extents.

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

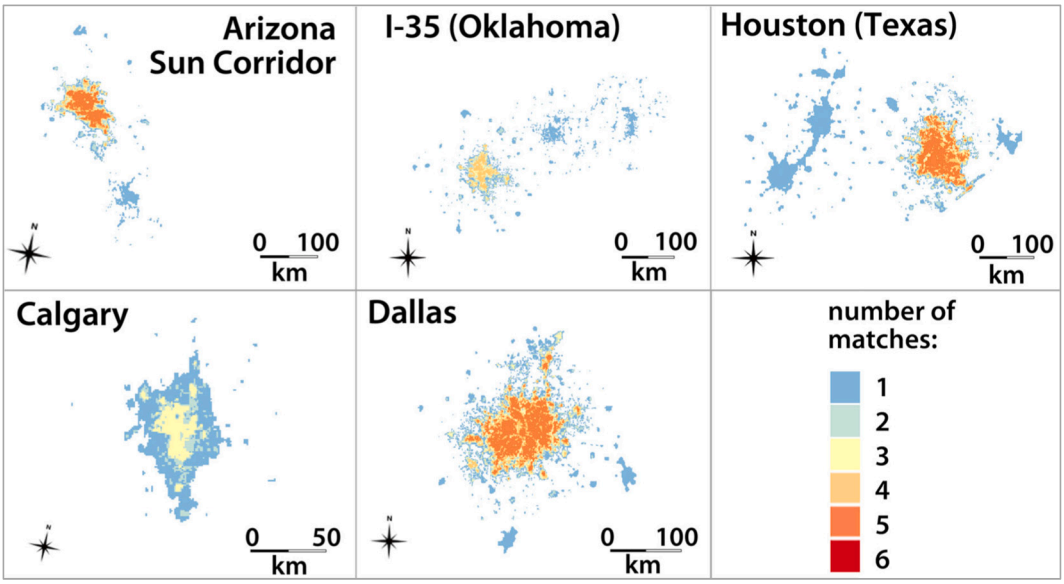

Fig. 6. Omitted corridor candidates. These five areas do not match the criteria to qualify as urban corridors due to their compact shape and/or small size and are therefore excluded from further analysis.

coherence is no longer given and Austin and San Antonio are no longer part of the Houston cluster.

Table 3 displays the corridor candidates and the resulting urban corridors with their respective length, width and area measurements for the examples of one and three matching input layers. It is notable that, for one match, all corridors exceed the originally derived dimensions by far, due to the influence of single individual input layers. An exception is the I-35 (Austin–San Antonio) region because of the subdivision into several smaller areas. In order to reduce the influence of a single input layer, we re-examine the resulting areas for 3 or more matches. That way, we disregard the extents where only one or two (i.e., one third) of the input layers return a positive result. The resulting measurements show three different sets of characteristics:

(1) Urban corridors smaller than the candidate estimate. This mainly includes areas we had to split into smaller individual parts (Montreal/Toronto, the Gulf Coast or the I-35 sections).
(2) Urban corridors moderately larger than the candidate estimate which is the case for a large number of areas. Due to a different approach in measuring urban corridor extent, most areas identified through three overlapping input layers are about two to three times larger than previously measured.
(3) Urban corridors significantly larger than the candidate estimate. This comprises the massive extents of the Great Lakes region with Chicago and Detroit which have been greatly underestimated before. This is mainly due to the night-time lights, impervious surface area and population density input layers which generate a very large coherent patch for the Great Lakes area.

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

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

4. Discussion

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

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

As scattered and small urbanized areas are seen as not decisive in forming an urban corridor, we selected, for each input layer, the largest 0.5 % of all urban patches. While this seems a subjective and arbitrary number, a closer inspection reveals that a percentage less than 0.5 % would have omitted some areas which have been categorized as possible urban corridors (Georg, Blaschke, Taubenböck, 2016b), while a larger share would have included a considerable number of relatively small areas which are not considered relevant for our study. Also, using a percentage is, in contrast to an absolute number like for example 'the 25 biggest patches', or a manual selection of only certain areas, transferable across scales and input data. We have not systematically tested the effects of alternating thresholds, but believe that the strength of this approach is in its consistency.

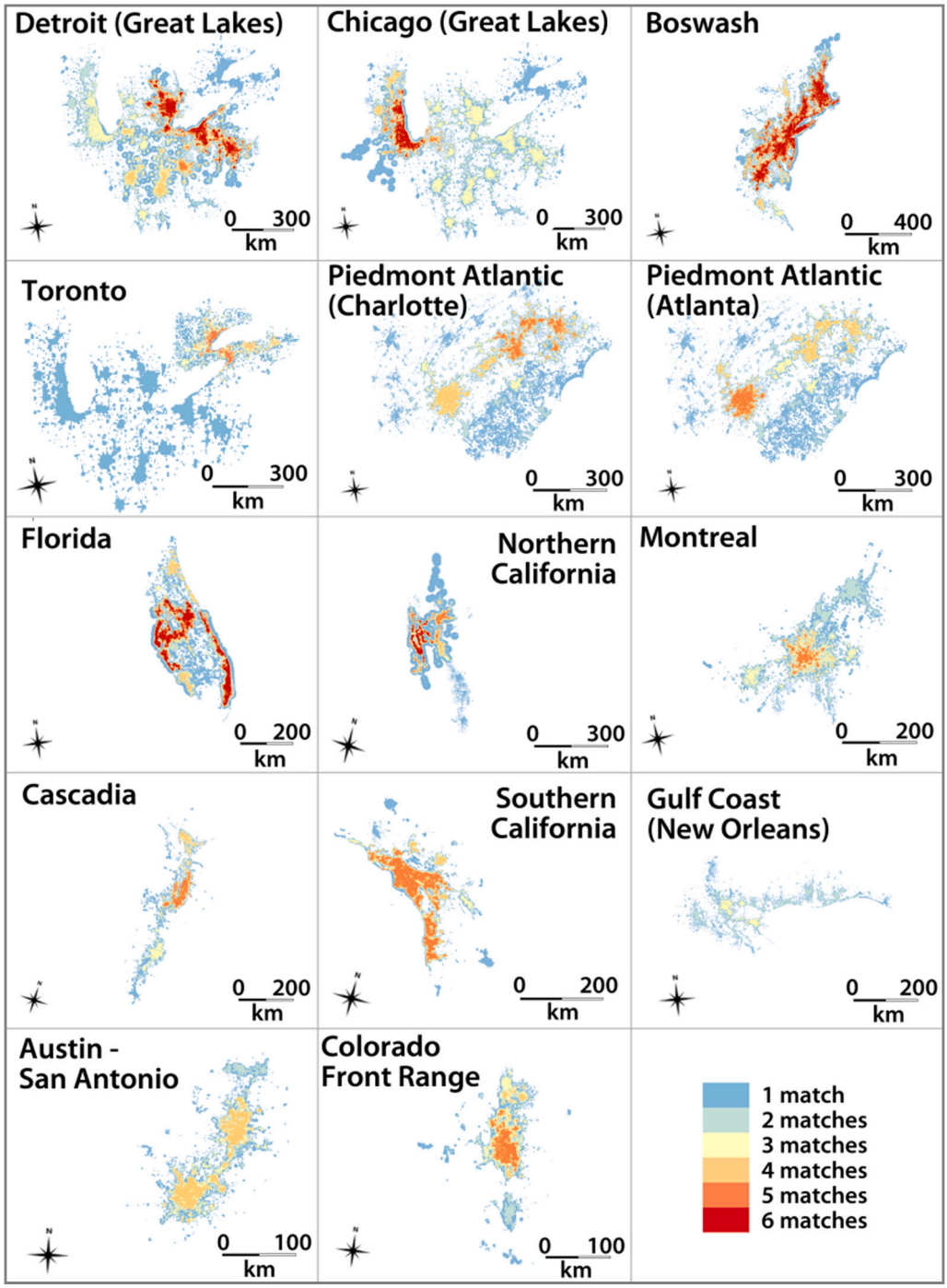

Fig. 7. Urban corridors identified through overlay of input layers.

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

All input data were aggregated to a 1 km^2 grid. This may seem quite coarse, particularly when considering that there is a variety of high-resolution, often free imagery available. However, for large spatial extents such as urban corridors, we believe that a higher resolution is not required for our purposes. A 1 km^2 grid is also used by transnational programs such as the European INSPIRE initiative (INSPIRE, 2014) or Hansen et al. (2000). Furthermore, when using densities (such as road or

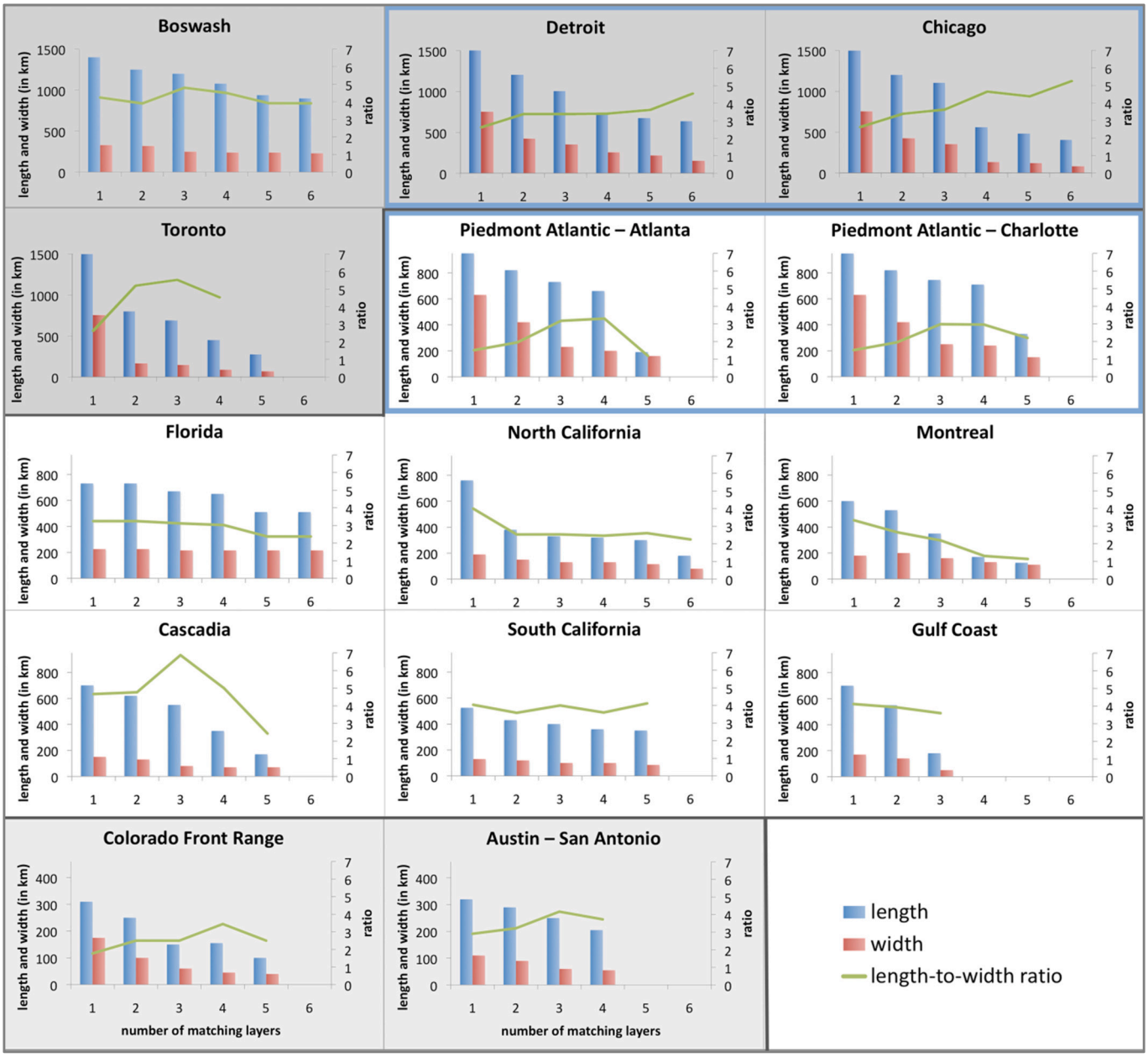

Fig. 8. The 14 corridors identified and their physical characteristics with respect to the number of matching layers. Note the different scales (marked by background color).

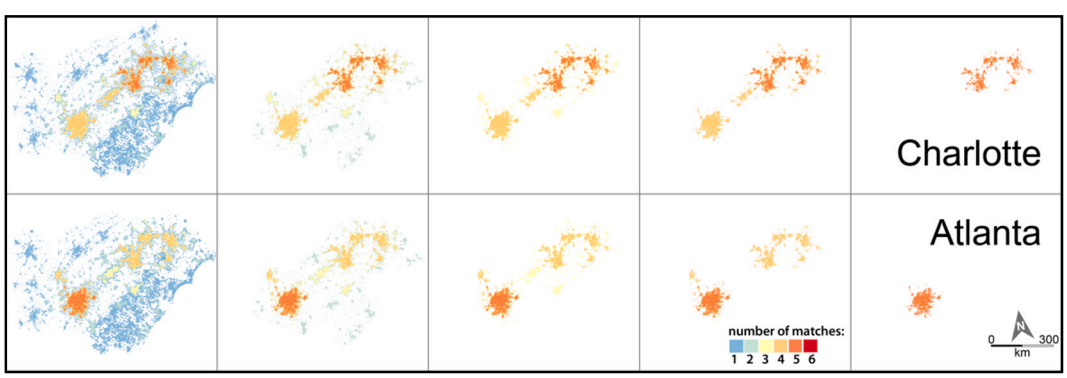

Fig. 9. Piedmont Atlantic: varying extents for Atlanta (top row) and Charlotte (bottom row) as main centers.

Table 3
Urban corridor dimensions (candidates and results) for one and three matches.

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

Corridor candidates are marked with 'cand.' and highlighted in light grey. No highlight and italics: excluded corridor candidates. Dark grey and bold: final corridors.

rail densities as used by Georg et al., 2018), data aggregation is necessary. As Wei et al. (2017a) describe, this aggregation does not necessarily imply less appropriate outcomes. It has to be kept in mind that spatial statistics are hardly ever unambiguous and results depend on input data and parameters used.

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

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

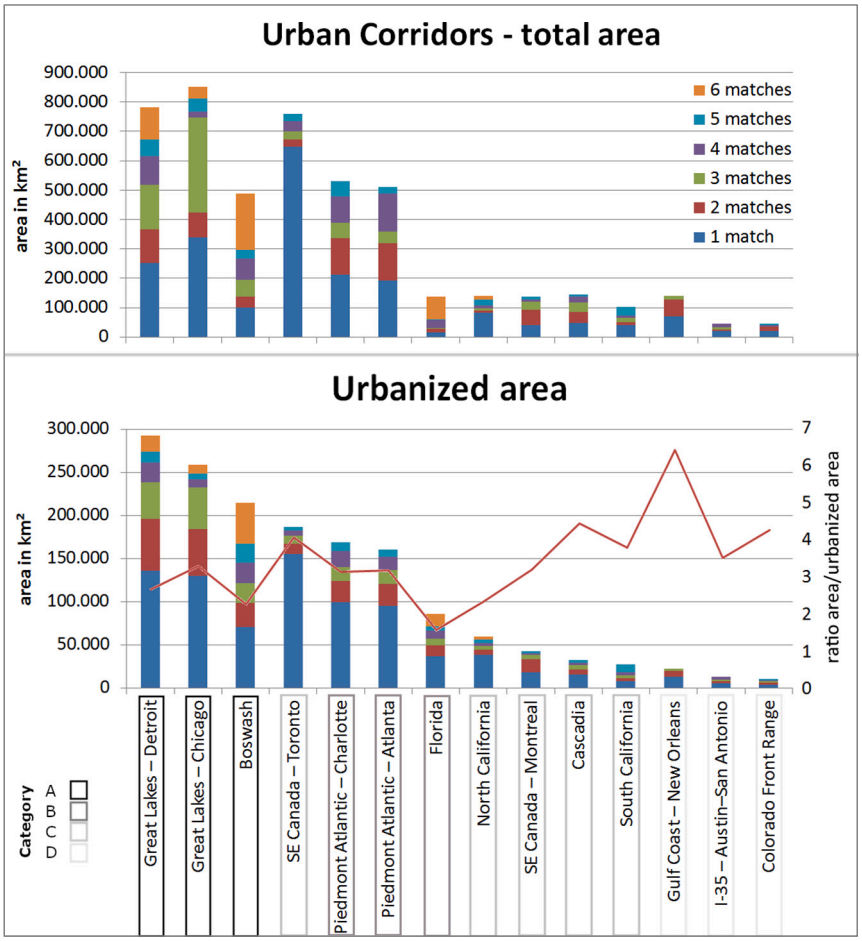

Fig. 10. Urban corridors: total area (top) and urbanized space within that area plus the ratio between total and urbanized space (bottom). Note the different scales. The boxes around the corridor names indicate the respective urban corridor category.

prevalent than for more matches (see Tables 2 and 3).

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

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

We acknowledge that this approach is following a 'space of place' logic (Castells, 2006). This is of relevance as these areas define the places where living and working is established. However, this physicalism does not account for the 'spaces of flows' (Castells, 2006). Thus, some areas integrated in these national or continental to global networks which might not have the related physicalism cannot be identified using these data and approaches. Functional aspects are of importance when describing large urban areas, particularly concepts such as 'Functional Urban Region (FUR)'. However, these functional connections cannot be measured globally and consistently – this can only be achieved for the spatial patterns. We are aware that, while spatial patterns help 'describe' urban areas, they cannot 'explain' them. We acknowledge that for a comprehensive understanding of the dynamics behind urban change, economic and other functional data would have to be taken into account.

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

5. Summary and outlook

In ever-expanding urban landscapes, new physical forms of cities and connected cities evolve. However, what constitutes the urban in comparison to the rural is fuzzy and far from any unambiguity (Taubenböck et al., 2022), and the spatial delineation of conceptual terms such as urban corridors or the like is just as challenging due to different conceptual approaches, varying datasets available and variables used or methods applied. Beyond, the delineation usually relies on administrative boundaries or other, fixed extents. The reality, though, is more complex, as urbanization processes have, over time and in many

locations, outgrown those fixed boundaries. This work aimed to take into account the underlying fuzziness of a concept such as an 'urban corridor', the variability of different input data and a specific thresholding approach resulting in a probability-based, malleable definition for each area investigated here.

This paper examined the extents of large urban areas within the study area covering the connected United States territory plus several hundred km to the north and south, allowing to disregard administrative boundaries. Particular focus was on 'urban corridors', i.e. massively large, lengthy and spatially coherent urbanized areas. The input data for our analysis consist of six globally available datasets (night-time lights, population density and different artificial surfaces layers) – thus allowing for transferability to other parts of the world. For each input layer, we derived an individual extent of each region and combined the individual results for a fuzzy delineation of the area independent of administrative boundaries: the likelihood that an area is part of an urban corridor is a factor of how many input layers define an area as urban. Our results reveal that 14 areas in Northern America show spatial properties of an urban corridor as a specific type of large urban areas.

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

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

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author statement

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the Corresponding Author is the sole contact for the Editorial process. She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

CRediT authorship contribution statement

Isabel Georg: Conceptualization, Methodology, Formal analysis, Writing – original draft, Visualization. **Thomas Blaschke:** Conceptualization, Writing – review & editing. **Hannes Taubenböck:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data will be made available on request.

References

Angel, S., Parent, J., & Civco, D. L. (2012). The fragmentation of urban landscapes: Global evidence of a key attribute of the spatial structure of cities, 1990–2000. *Environment and Urbanization, 24*(1), 249–283.

Angel, S., Sheppard, S., Civco, D. L., Buckley, R., Chabaeva, A., Gitlin, L., Kraley, A., Parent, J., & Perlin, M. (2005). *The dynamics of global urban expansion*. Washington, DC: World Bank, Transport and Urban Development Department.

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., & Batty, M. (2015). Constructing cities, deconstructing scaling laws. *Journal of the Royal Society Interface, 12*, 20140745.

Banerjee, T. (2012). Megaregions or megasprawls? Issues of density, urban design, and quality growth. In C. L. Ross (Ed.), *Megaregions. Planning for global competitiveness* (pp. 83–106). Washington: Island Press.

Bontemps, S., Defourny, P., Bogaert, E. V., Arino, O., Kalogirou, V., & Perez, J. R. (2011). *GLOBCOVER 2009 - Products description and validation report*.

Cai, J., Huang, B., & Song, Y. (2017). Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sensing of Environment, 202*, 210–221.

Castells, M. (2006). *The space of flows*. Wiley Online Library.

Center for International Earth Science Information Network - CIESIN - Columbia University. (2017). Gridded population of the world, version 4 (GPWv4): Population density, revision 10 [WWW Document]. http://sedac.ciesin.columbia.edu/data/collection/gpw-v4/sets/browse accessed 09 Aug 22.

Cervero, R., & Murakami, J. (2010). Effects of built environments on vehicle miles traveled: Evidence from 370 US urbanized areas. *Environment and Planning A, 42*(2), 400–418.

Chandrashekar, C. M., & Aithal, B. H. (2021). Impact assessment of corridor oriented development a case of urban agglomerations of India. *International Review for Spatial Planning and Sustainable Development, 9*(2), 172–194.

Choe, S.-C. (1996). The evolving urban system in North-East Asia. *Emerging world cities in Pacific Asia*, 498–519.

Association, R. P. (2007). In *, 489. Northeast Megaregion 2050: a common future* (p. 28). New York: Regional Plan Association.

Crooks, A., Pfoser, D., Jenkins, A., Croitoru, A., Stefanidis, A., Smith, D., Karagiorgou, S., Efentakis, A., & Lamprianidis, G. (2014). Crowdsourcing urban form and function. *International Journal of Geographical Information Science, 29*, 720–741. https://doi.org/10.1080/13658816.2014.977905

Curtis, C. (2006). Network city: Retrofitting the Perth metropolitan region to facilitate sustainable travel. *Urban Policy and Research, 24*, 159–180.

Davoudi, S. (2007). Polycentricity: Panacea or pipedream. In N. Cattan (Ed.), *Cities and networks in Europe: A critical approach of polycentrism* (pp. 65–73). John Libbey Eurotext: Montrouge, France.

Dewan, A. M., & Yamaguchi, Y. (2009). Land use and land cover change in greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization. *Applied Geography, 29*, 390–401.

Elvidge, n.d.C.D. Elvidge n.d. Version 4 DMSP-OLS Nighttime Lights Time Series [WWW Document]. URL https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (accessed 09 Aug 22).

Elvidge, C. D., Baugh, K. E., Kihn, E. A., Kroehl, H. W., & Davis, E. R. (1997). Mapping city lights with nighttime data from the DMSP operational linescan system. *Photogrammetric Engineering and Remote Sensing, 63*, 727–734.

Elvidge, C. D., Baugh, K. E., Zhizhin, M., & Hsu, F.-C. (2013). Why VIIRS data are superior to DMSP for mapping nighttime lights. In *, 35. Proceedings of the Asia-Pacific advanced network* (pp. 62–69).

Esch, T., Taubenböck, H., Roth, A., Heldens, W., Felbier, A., Thiel, M., Schmidt, M., Müller, A., & Dech, S. (2012). TanDEM-X mission: New perspectives for the inventory and monitoring of global settlement patterns. *Journal of Applied Remote Sensing., 6*(1). pp. 21 pages.

European Space Agency (ESA), n.d.European Space Agency (ESA), n.d. GlobCover 2009 (Global Land Cover Map) [WWW Document]. URL http://due.esrin.esa.int/page_globcover.php (accessed 09 Aug 22).

Fang, C., & Yu, D. (2017). Urban agglomeration: An evolving concept of an emerging phenomenon. *Landscape and Urban Planning, 162*, 126–136.

Florida, R., Gulden, T., & Mellander, C. (2008). The rise of the mega-region. *Cambridge Journal of Regions, Economy and Society, 1*, 459–476.

Forman, R. T., & Alexander, L. E. (1998). Roads and their major ecological effects. *Annual Review of Ecology and Systematics, 29*, 207–231.

Fotheringham, A. S., & Rogerson, P. A. (1993). GIS and spatial analytical problems. *International Journal of Geographical Information Science, 7*(1), 3–19.

Georg, I., Blaschke, T., & Taubenböck, H. (2016a). New spatial dimensions of global cityscapes: From reviewing existing concepts to a conceptual spatial approach. *Journal of Geographical Sciences, 26*, 355–380.

Georg, I., Blaschke, T., & Taubenböck, H. (2016b). A global inventory of urban corridors based on perceptions and night-time light imagery. *ISPRS International Journal of Geo-Information, 5*, 233.

Georg, I., Blaschke, T., & Taubenböck, H. (2018). Are we in boswash Yet? A multi-source geodata approach to spatially delimit urban corridors. *ISPRS International Journal of Geo-Information, 7*, 15.

Glocker, D. (2018). The rise of megaregions: Delineating a new scale of economic geography. In *OECD regional development working papers 2018/04*.

Gottmann, J. (1957). Megalopolis or the urbanization of the northeastern seaboard. *Economic Geography, 33*, 189–200.

Growe, A., Harrison, J., & Hoyler, M. (2015). Megaregions. In *, 73. Globalization's new urban form? Raumforschung und Raumordnung* (pp. 305–307).

Hall, P. (2009). Looking backward, looking forward: The city region of the mid-21st century. *Regional Studies, 43*(6), 803–817.

Hansen, M. C., DeFries, R. S., Townshend, J. R., et al. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing, 21*(6/7), 1331–1364.

Henderson, M., Yeh, E. T., Gong, P., Elvidge, C., & Baugh, K. (2003). Validation of urban boundaries derived from global night-time satellite imagery. *International Journal of Remote Sensing, 24*, 595–609.

Hertzog, A. (2020). *The Lagos Abidjan corridor: Migration driven urbanisation in west Africa*. Zurich: ETH. Doctoral dissertation.

Hospers, G.-J. (2002). *Beyond the blue banana? Structural change in Europe's geo-economy*.

Imhoff, M. L., Lawrence, W. T., Stutzer, D. C., & Elvidge, C. D. (1997). A technique for using composite DMSP/OLS "city lights" satellite data to map urban area. *Remote Sensing of Environment, 61*, 361–370.

Innes, J. E., Booher, D. E., & Di Vittorio, S. (2010). Strategies for megaregion governance: Collaborative dialogue, networks, and self-organization. *Journal of the American Planning Association, 77*, 55–67.

INSPIRE. (2014). INSPIRE - Infrastructure for spatial information in Europe. In *Data specification on geographical grid systems – Technical guidelines*.

Jain, M., Korzhenevych, A., & Basu, A. M. (2021). Integrating spatial development with infrastructure provision along an envisioned transport corridor: A conceptual framework and its application to India. *Land Use Policy, 104*, Article 105364.

Jain, S., Panda, J., & Kant, S. (2014). Possible socio-scientific issues of land-use and land-cover change impact and associated tools of study with a special reference to Delhi-Mumbai industrial corridor region. *International Journal of Earth and Atmospheric Science, 1*(2), 58–70.

Jing, W., Yang, Y., Yue, X., & Zhao, X. (2015). Mapping urban areas with integration of DMSP/OLS nighttime light and MODIS data using machine learning techniques. *Remote Sensing, 7*, 12419–12439.

Kar, A., Varshney, M., & Varshney, M. (2021). The emergence of urban corridors - a brief look in Indian context. *International Research Journal of Engineering and Technology (IRJET), 8*(11).

Khanna, n.d.P. Khanna n.d. A new map for America [WWW document]. URL https://www.paragkhanna.com/2016-4-16-a-new-map-for-america/ (accessed 30 Aug 22).

Klotz, M., Kemper, T., Geiß, C., Esch, T., & Taubenböck, H. (2016). How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sensing of Environment, 178*, 191–212.

Lang, R. E., & Dhavale, D. (2005). *Beyond megalopolis: Exploring America's new "megapolitan" geography*.

Latham, J., Cumani, R., Rosati, I., & Bloise, M. (2014). *Global land cover share (GLC-SHARE) database beta-release version 1.0-2014*. Rome, Italy: FAO.

Li, P., & Cao, X. (2005). Evolution and development of Guangzhou-Hong Kong corridor. *Chinese Geographical Science, 15*, 206.

Lin, J., Huang, B., Chen, M., & Huang, Z. (2014). Modeling urban vertical growth using cellular automata—Guangzhou as a case study. *Applied Geography*, 172–186.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing, 21*, 1303–1330.

Malczewski, J. (2004). GIS-based land-use suitability analysis: A critical overview. *Progress in Planning, 62*, 3–65.

Marceau, D. J. (1999). The scale issue in the social and natural sciences. *Canadian Journal of Remote Sensing, 25*, 347–356.

Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-time light data: A good proxy measure for economic activity? *PloS one, 10*, Article e0139779.

Miller, B.C., 2021. *Cluster-based Delineation of Megaregions in the United States: Identifying administrative boundaries that reflect meta-communities to improve the effectiveness and efficiency of government* (Doctoral dissertation, Georgia Institute of Technology).

Mukhopadhyay, C. (2018). Is the Delhi Mumbai industrial corridor (DMIC) an emerging 'Megaregion' in India? *Planning Theory & Practice, 19*(2), 305–309. https://doi.org/10.1080/14649357.2017.1392136

Nelson, G. D., & Rae, A. (2016). An economic geography of the United States: From commutes to megaregions. *PloS one, 11*, Article e0166083.

Neuman, M. (2000). Regional design: Recovering a great landscape architecture and urban planning tradition. *Landscape and Urban Planning, 47*, 115–128.

NOAA (National Oceanic and Atmospheric Administration), n.d.NOAA (National Oceanic and Atmospheric Administration), n.d. Global distribution and density of constructed impervious surfaces. [WWW document]. URL https://www.ngdc.noaa.gov/eog/dmsp/download_global_isa.html (accessed 09 Aug 22).

Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich, UK: Geo Abstracts, University of East Anglia.

Pesaresi, M., Ehrlich, D., Ferri, S., Florczyk, A., Freire, S., Haag, F., Halkia, M., Julea, A. M., Kemper, T., & Soille, P. (2015). Global human settlement analysis for disaster risk reduction. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40*, 837.

Pesaresi, M., Gerhardinger, A., & Kayitakire, F. (2008). A robust built-up area presence index by anisotropic rotation-invariant textural measure. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1*, 180–192.

Pietrzak, M. B. (2014). The modifiable areal unit problem–Analysis of correlation and regression. Equilibrium. *Quarterly Journal of Economics and Economic Policy, 9*(4), 113–131.

Potere, D., Schneider, A., Angel, S., & Civco, D. L. (2009). Mapping urban areas on a global scale: Which of the eight maps now available is more accurate? *International Journal of Remote Sensing, 30*(24), 6531–6558.

Priemus, H., & Zonneveld, W. (2003). What are corridors and what are the issues? Introduction to special issue: The governance of corridors. *Journal of Transport Geography, 11*, 167–177.

Qin, J., Fang, C., Wang, Y., Li, G., & Wang, S. (2015). Evaluation of three-dimensional urban expansion: A case study of Yangzhou City, Jiangsu Province, China. *Chinese Geographical Science, 25*, 224–236. https://doi.org/10.1007/s11769-014-0728-8

Regional Plan Association (RPA). (2007). In *, 489. Northeast Megaregion 2050: A common future* (p. 28). New York: Regional Plan Association.

Rodrigue, J.-P. (2004). Freight, gateways and mega-urban regions: The logistical integration of the Bostwash corridor. *Tijdschrift voor Economische en Sociale Geografie, 95*, 147–161.

Ross, C. L. (2008). *Megaregions: Literature review of the implications for U.S. infrastructure investment and transportation planning*. Georgia: Georgia Institute of Technology.

Rozenfeld, H. D., Rybski, D., Gabaix, X., & Makse, H. A. (2011). The area and population of cities: New insights from a different perspective on cities. *American Economic Review, 101*, 2205–2225.

Rustiadi, E., Pravitasari, A. E., Setiawan, Y., Mulya, S. P., Pribadi, D. O., & Tsutsumida, N. (2021). Impact of continuous Jakarta megacity urban expansion on the formation of the Jakarta-Bandung conurbation over the rice farm regions. *Cities, 111*, Article 103000.

Schneider, A., & Woodcock, C. E. (2008). Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Studies, 45*, 659–692.

Seto, K. C., & Fragkias, M. (2005). Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. *Landscape Ecology, 20*, 871–888.

Shi, K., Huang, C., Yu, B., Yin, B., Huang, Y., & Wu, J. (2014). Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. *Remote Sensing Letters, 5*, 358–366.

Small, C., Pozzi, F., & Elvidge, C. D. (2005). Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sensing of Environment, 96*, 277–291.

Soja, E., & Kanai, E. (2008). The Urbanization of the World. In *The endless city* (pp. 54–69). London: Phaidon Press.

Taubenböck, H., Bauer, P., Geiß, C., & Wurm, M. (2017). Mega-regions in China—A spatial analysis of settlement patterns using Earth observation data. In *Proceedings of the IEEE-CPS joint urban remote sensing event (JURSE), Dubai, UAE* (pp. 6–8).

Taubenböck, H., Droin, A., Standfuß, I., Dosch, F., Sander, N., Milbert, A., Eichfuss, S., & Wurm, M. (2022). To be, or not to be 'urban'? A multi-modal method for the differentiated measurement of the degree of urbanization. *Computers, Environment & Urban Systems, 95*.

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 International Journal of Geo-Information, 6*(2), 1–27.

Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J., & Dech, S. (2019). A new ranking of the world's largest cities - Do administrative units obscure morphological realities? *Remote Sensing of Environment, 232*, Article 111353.

Taubenböck, H., & Wiesner, M. (2015). The spatial network of megaregions-types of connectivity between cities based on settlement patterns derived from EO-data. *Computers, Environment and Urban Systems, 54*, 165–180.

Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., & Dech, S. (2014). New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. *Applied Geography, 47*, 137–153.

Trip, J. J. (2003). The corridor as a transport network. Megacorridors in North West Europe. Investigating a new transnational planning concept. *Housing and Urban Policy Studies, 27*, 17–23.

UN-DESA (Ed.). (2019). *World urbanization prospects: The 2018 revision*. New York: United Nations. https://desapublications.un.org/publications/2018-revision-world-urbanization-prospects. (Accessed 8 September 2022).

UN-Habitat. (2008). *The State of African cities: A framework for addressing urban challenges in Africa*.

UN-Habitat (Ed.). (2013). *State of the world's cities 2012/2013. Prosperity of cities*. Nairobi: Routledge & Earthscan.

Van Houtum, H., & Lagendijk, A. (2001). Contextualising regional identity and imagination in the construction of polycentric urban regions: The cases of the Ruhr area and the Basque Country. *Urban Studies, 38*, 747–767.

Wei, C., Blaschke, T., Kazakopoulos, P., Taubenböck, H., & Tiede, D. (2017). Is spatial resolution critical in urbanization velocity analysis? Investigations in the Pearl River Delta. *Remote Sensing, 9*(80).

Wei, C., Taubenböck, H., & Blaschke, T. (2017). Measuring urban agglomeration using a city-scale dasymetric population map: A study in the Pearl River Delta, China. *Habitat International, 59*, 32–43.

Wheeler, Stephen (2009). Regions, Megaregions, and Sustainability. *Regional Studies, 43* (6), 863–876. https://doi.org/10.1080/00343400701861344

Whebell, C. F. (1969). Corridors: A theory of urban systems. *Annals of the Association of American Geographers, 59*, 1–26.

Xie, Y., & Weng, Q. (2016). Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sensing of Environment, 187*, 1–13.

Xu, H. (2008). A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing, 29*, 4269–4276.

Yang, T., & Hillier, B. (2007). The fuzzy boundary: the spatial definition of urban areas. In *Proceedings, 6th international space syntax symposium, İstanbul, 2007*. Istanbul Technical University. pp. 091–01.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*, 338–353.

Zhou, Y., Smith, S. J., Elvidge, C. D., Zhao, K., Thomson, A., & Imhoff, M. (2014). A cluster-based method to map urban area from DMSP/OLS nightlights. *Remote Sensing of Environment, 147*, 173–185.