

Article

# Are We in Boswash Yet? A Multi-Source Geodata Approach to Spatially Delimit Urban Corridors

Isabel Georg <sup>1,\*</sup> , Thomas Blaschke <sup>1</sup>  and Hannes Taubenböck <sup>2</sup> 

<sup>1</sup> Department of Geoinformatics, University of Salzburg, 5020 Salzburg, Austria; thomas.blaschke@sbg.ac.at

<sup>2</sup> Earth Observation Center, Remote Sensing Data Center, German Aerospace Center (DLR), 82234 Oberpfaffenhofen, Germany; hannes.taubenboeck@dlr.de

\* Correspondence: isabel@southward-georg.com; Tel.: +49-178-3064774

Received: 20 November 2017; Accepted: 23 December 2017; Published: 4 January 2018

**Abstract:** The delimitation of urban space is conceptually elusive and fuzzy. Commonly, urban areas are delimited through administrative boundaries. These artificial, fixed boundaries, however, do not necessarily represent the actual built-up extent, the urban catchment, or the economic linkage within and across neighboring metropolitan regions. For an approach to spatially delimit an urban corridor—a generically defined concept of a massive urban area—we use the Boston to Washington (Boswash) region as an example. This area has been consistently conceptualized in literature as bounded urban space. We develop a method to spatially delimit the urban corridor using multi-source geodata (built-up extent, infrastructure and socioeconomic data) which are based on a grid rather than on administrative units. Threshold approaches for the input data serve to construct Boswash as varying connected territorial spaces, allowing us to investigate the variability of possible spatial forms of the area, i.e., to overcome the simple dichotomous classification in favor of a probability-based differentiation. Our transparent multi-layer approach, validated through income data, can easily be modified by using different input datasets while maintaining the underlying idea that the likelihood of an area being part of an urban corridor is flexible, i.e., in our case a factor of how many input layers return positive results.

**Keywords:** urban corridor; spatial analysis; remote sensing; urban areas; pixel-based classification; fuzzy boundaries

## 1. Introduction

Are we in Boswash yet? A simple question—however, also a contentious one in urban geography. A distinct answer is conceptually and spatially challenging and consensus on how to delimit an urban corridor—such as the very densely populated area stretching from Boston to Washington D.C.—is absent. In times where global population growth is—as has been commonly observed (e.g., [1,2])—mainly absorbed by urban areas, there is a need for discussing new regional phenomena. As new large urban constellations emerge [3] and cities constantly change their spatial appearance, both vertically and horizontally (e.g., [1,4–6]), the individual city as self-contained entity has been overcome by these developments. There is a multitude of terms describing and conceptualizing larger multi-nuclei urban areas (e.g., urban corridor, mega-region, conurbation), some of which show considerable overlap in their spatial extents [7]. The boundaries of such larger urban areas are often fuzzy and the terminology is ambiguous. For example, a conurbation describes more or less the same spatial constellation as a functional urban region or a metropolitan area. Generally, administrative borders are used for the delimitation of urban areas and for spatial planning or governance. While this makes sense for many scenarios, frequently, though, administrative borders do not reflect the spatial extent or connectedness of an urban area adequately. Spatially capturing large urban constructs, thus, is a complex task and other data and methods to describe these need to be established.

### 1.1. Large Urban Areas I: The Case of Urban Corridors

Urban corridors are an example of a fuzzy concept of a large urban area without a clear description or conceptualization: The term—introduced by [8] who described different stages of urban development of Southern Ontario (Canada)—has been repeatedly used in scientific literature, but a universal definition does not yet exist [3]. Spatially, an urban corridor is a collection of large cities or clusters which are aligned in a linear structure, generally along high-speed surface transport infrastructure (road or rail) [9–12]. Corridor features include indicators such as high population density, heterogeneous land use, and “connection” (i.e., people, goods or information flow) [13]. The area from Boston to Washington (*Boswash*) along the northeastern seaboard of the US is—based on this definition—a prime example of such a massive area identified by many scholars (e.g., [14–18]). Other areas include BESETO (Beijing–Seoul–Tokyo) [19], GILA (Greater Idaban–Lagos–Accra) [20], or the *Blue Banana* from Birmingham to Milan [21], which can be subdivided into smaller individual urban corridors (e.g., RhineRuhr or Randstad), among others.

In the context of urban corridors it seems obvious that information from crisp administrative boundaries will rarely be appropriate. The construction of space itself (here: an urban corridor) depends on different indicators and logic, which are expressed through intrinsically diverse areal delimitations. In fact, spatial contiguity is the critical criterion of a region, as opposed to type [22]. Ref. [23] reiterate earlier findings that spatial contiguity usually only has a small effect on the loss of detail during aggregations. The geographic principle of spatial autocorrelation often serves as the mechanistic construct for this human-thinking centered view of fuzzy boundaries. Region building assumes the interdependency of spatial structure and spatial variables [24]; a more or less uniform behavior or an impression of a landscape or cultural features would be altered with transitions to a different location along a crisp boundary.

Manuel Castells [25,26] maintains that the concept of space consists in the conflict between the ‘space of flows’ and the ‘space of places’. He defines the ‘space of flows’ 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. Facilitated by telecommunications and rapid computerized transport systems, large agglomerations such as Randstad (Netherlands), the Pearl River Delta (China) or Tokyo–Kobe–Kyoto (Japan) have become highly interrelated functional areas in which spatial concentration and decentralization happen simultaneously [26]. These massive regions “without a name, without a culture, and without institutions” [26] face a number of challenges including political accountability, lack of citizen participation, management, or cooperation between organizations, resulting in a discrepancy between the actual spatial units and the institutions to manage these.

### 1.2. Large Urban Areas II: Earth Observation

For the description as well as monitoring of spatially large urban constellations (thus taking a ‘space of place’ perspective), one data source—earth observation imagery—has provided valuable information over the last decades. In particular, satellite imagery has been used in a number of studies to explicitly describe the very large forms of connected urban territories and related processes of their development. Night-time lights such as the DMSP-OLS (Defense Meteorological Satellite Program’s Operational Linescan System) have proven beneficial for such urban studies (e.g., [27–30]). A newer night-time light sensor, the Visible Infrared Imaging Radiometer Suite sensor (VIIRS) has been used in combination with social media data to identify main centers and subcenters in three Chinese cities [31] or, in combination with population data, to assess the number of people affected during a humanitarian crisis [32].

For worldwide inventories of large urban constellations, globally consistent datasets are ideally suited and with increasing data availability, a number of global urban area maps have been created [33–37]. Also, there are several global land cover classification products from multi-source earth observation imagery which are relevant for urban area delimitation. Earlier, lower-resolution examples include the Global Land Cover (GLC2000) [38], GlobCover [39], IGBP-DISCover [40] or

MODIS land-cover map [41] products. More recently, the high-resolution Global Human Settlement Layer (GHSL) [42] and Global Urban Footprint (GUF) [37] datasets have become available for urban area observation.

Quite a few studies focus on the definition and spatial delineation of regional phenomena: In the US, two of the most influential studies to describe large-scale urban regions are the Regional Plan Association (RPA)'s *Megaregions* [16] and Metropolitan Institute at Virginia Tech's *Megapolitans* [15]. These concepts have been identified independent of each other but result in the same ten regions with similar extents, taking into account criteria such as population, settlement patterns, economic performance, infrastructure, or compactness as well as complexity. Ref. [43] measure the demographic spatial structure of mega-regions in China, basing the boundaries on commonly accepted definitions. Ref. [44] use given boundaries for a multi-temporal gridded estimate of population density in the Pearl River Delta. Using settlement patterns gained from Landsat and TanDEM-X imagery, Ref. [6] delineate megaregions without using administrative boundaries.

At the same time, urban areas have frequently been described without using administrative boundaries. For example, Ref. [45] spatially characterize mega-regions through multi-source, multi-temporal satellite imagery and develop a metric to assess spatial connectivity between urban areas. They demonstrate how over time, separated cities can converge into polycentric, massive regions without clear delimitations and boundaries. In further studies, settlement densities and sizes from Global Urban Footprint data are applied to classify and categorize urban nodes and their connectivity to find regional phenomena across Europe [46] or China [47]. Ref. [48] derive "natural cities" through clustering OpenStreetMap road data, without any census information, aiming to validate Zipf's law for the United States. Ref. [49] use commuting and population data to describe cities, revealing linear correlations between urban indicators and city size.

The number of *global* efforts to actually *delineate* large urban areas through their spatial extent—as opposed to *describing* these—using remote sensing imagery is relatively small. Ref. [17] use night-time imagery to classify urban corridors and mega-regions across the globe. They estimate population, economic activity, and technological and scientific innovation through additional datasets (population data, GDP, patents, and scientific citations). Ref. [50] quantify spatial extent, expansion and urbanization patterns of 25 global cities using remote sensing imagery, pattern metrics and census data. Ref. [7] identify urban corridors on a global scale using night-time imagery and OpenStreetMap data.

Building on the results of the latter study, in this paper we present and discuss a spatial approach to outline—as one example—the Boswash urban corridor in order to establish the territorial connectedness based on different variables: Built-up extent, infrastructure and socioeconomic data. While there are a number of studies on the spatial extent of this area (see Section 2), to date it has not been delimited through a similarly transparent method like the one described in this paper. We strive to overcome one-dimensional, crisp administrative boundaries through the use of homogeneously acquired remote sensing data and spatial disaggregation of the socio-economic information from the administrative units. Using a number of different data sources, we acknowledge the functional interrelations of cities through economic ties as well as physical interrelations of cities through settlements or street networks. This approach tackles the boundary problem, which is one of several issues in spatial analysis identified by [51]. A further issue is the Modifiable Areal Unit Problem (MAUP) [52] which encompasses (1) a scale problem and (2) an aggregation problem. The former occurs through the aggregation of data into fewer and larger units, while the latter deals with the variation of possible combinations of spatial units at the same or similar scales [53]. The MAUP has frequently been investigated through correlation and regression methods (e.g., [51,54,55]). As [56] concludes, often, the effect of the MAUP is not very large, but it does make a difference in some cases.

For the purposes of our study, we set individual thresholds for each input layer based on the connectivity of the Boswash area and check our results for plausibility using income data for the area. This method was chosen to allow us to compare the study area to other areas identified as

Urban Corridors in a previous study [7], which is outside the scope of this paper and will have to be demonstrated in future research. Other methods to derive thresholds include (among others):

- A cluster-based method in which the input data are segmented into clusters and a threshold is calculated for each cluster [57],
- a head/tail breaks classification [58] which groups data in heavy-tailed distributions into two parts around the mean value and iterates this process until the distribution of the head values is not heavy-tailed anymore,
- a support vector machine (SVM) method which is a machine-learning algorithm requiring only a small number of training samples to find the best separating hyperplane between classes [59,60].

The focus of this study is on the following research question: Can the conceptually vague definition of an urban corridor be transferred into a flexible territorial region of constructed interlinked urban spaces? The paper's main scientific contribution lies in the fuzzy delineation of a study region independent of administrative boundaries through multi-source geodata using one consistent, transferable method. The related structure of the paper is presented in Figure 1.

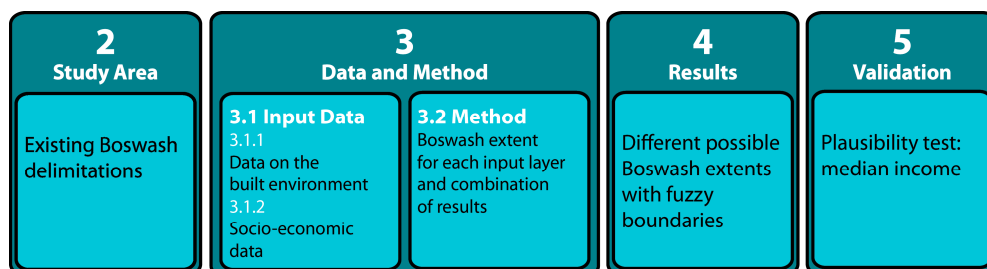


Figure 1. Structure of the paper.

## 2. The Study Area and Its Conceptual Background

Around 60 years ago, French geographer Jean Gottmann [14,61] described the term *Megalopolis* using the example of the coherent urbanized stretch between Boston and Washington. According to Gottmann, a megalopolis is a polynuclear urban system with a minimum population of 25 million people. This system is cohesive from a transport and communication aspect but with each city as its individual system [62,63]. Further, it is separable from other large urban regions by less urbanized space in between. The term itself, although generally attributed to Gottmann, has been used since the 1820s and also appears in essays from Geddes and Mumford [14,64].

Although the indicators for the delimitation of this urban corridor are more or less subjective (cf., Table 1), it is widely accepted in literature that the area forms a connected urban space. Now commonly known under its acronym coined by [65]—*Boswash*—the urban stretch has been studied extensively from a variety of disciplines, particularly transport planning [66], economic performance [17] and urban studies [15,16,67]. Different approaches result in a great variation as to both actual extent (Table 1) and terminology: Bos-Wash, East Coast Metroplex, Northeast Megaregion, I-95 Corridor, BostWash (e.g., [68]). Not even the original description is unambiguous—[69] identified at least six modifications used by Gottmann. In Gottmann's map, the area stretches from southern New Hampshire to Richmond and Norfolk (Virginia) (see Figure 2a).

**Table 1.** Ambiguous descriptions and delimitations: Different spatial and socio-economic indicators of Boswash presented in various studies.

Term	Area (sq. mi.)	Population (2000)	Number of Counties	Number of States	Length (in miles)	Source
Megalopolis, Northeastern Seaboard	-	n/a (30 million in 1950, about 37 million in 1960)	-	-	600	Gottmann (1957, 1961)
Northeast Megaregion (core only—similar to Megapolitan)	62,440	47.6 million	-	-	450	Penn School, 2005
Northeast Megaregion (core plus support zone)	188,380	69.2 million	405	14 incl. D.C.	-	Penn School, 2005
Megapolitan	70,062	49,182,941 (of which 89.3% are “urbanized”)	145	12 plus 1 district (D.C.)	-	Morrison Institute Report by Lang & Dhavale (2005)
Northeast Megapolitan	62,612	51,601,118	-	-	-	Morrison Institute, 2008
Northeast Megaregion	61,942	49,563,296	142	12 plus 1 district (D.C.)	450	Hagler (America 2050), 2009; Regional Plan Association, 2007
Megalopolis	-	-	-	10	450	Miller, 1975 (in Oswald et al., 2009)
Megalopolis	52,310	48,720,108	117	12 plus 1 district (D.C.)	-	Vicino, Hanlon & Short, 2007
Megalopolis	52,310	Almost 50 million	124	12 plus 1 district (D.C.)	600	Short, 2009; Vicino, Hanlon & Short, 2007
Megalopolis	13,490	42.4 million; 45.4 million (2010)	-	-	-	Morrill, 2006 & 2012
New Mega, Mega-Region, Corridor	-	54.3 million	-	-	500	Florida et al., 2008

Boswash almost covers 18% of the US population, around 20% of the GDP, but only around 2% of the land area [70]. According to [71], the GDP for the Northeast region is \$2.2 trillion (similar to Brazil’s), and the average wages of \$76,000 are \$19,000 more than the US average.

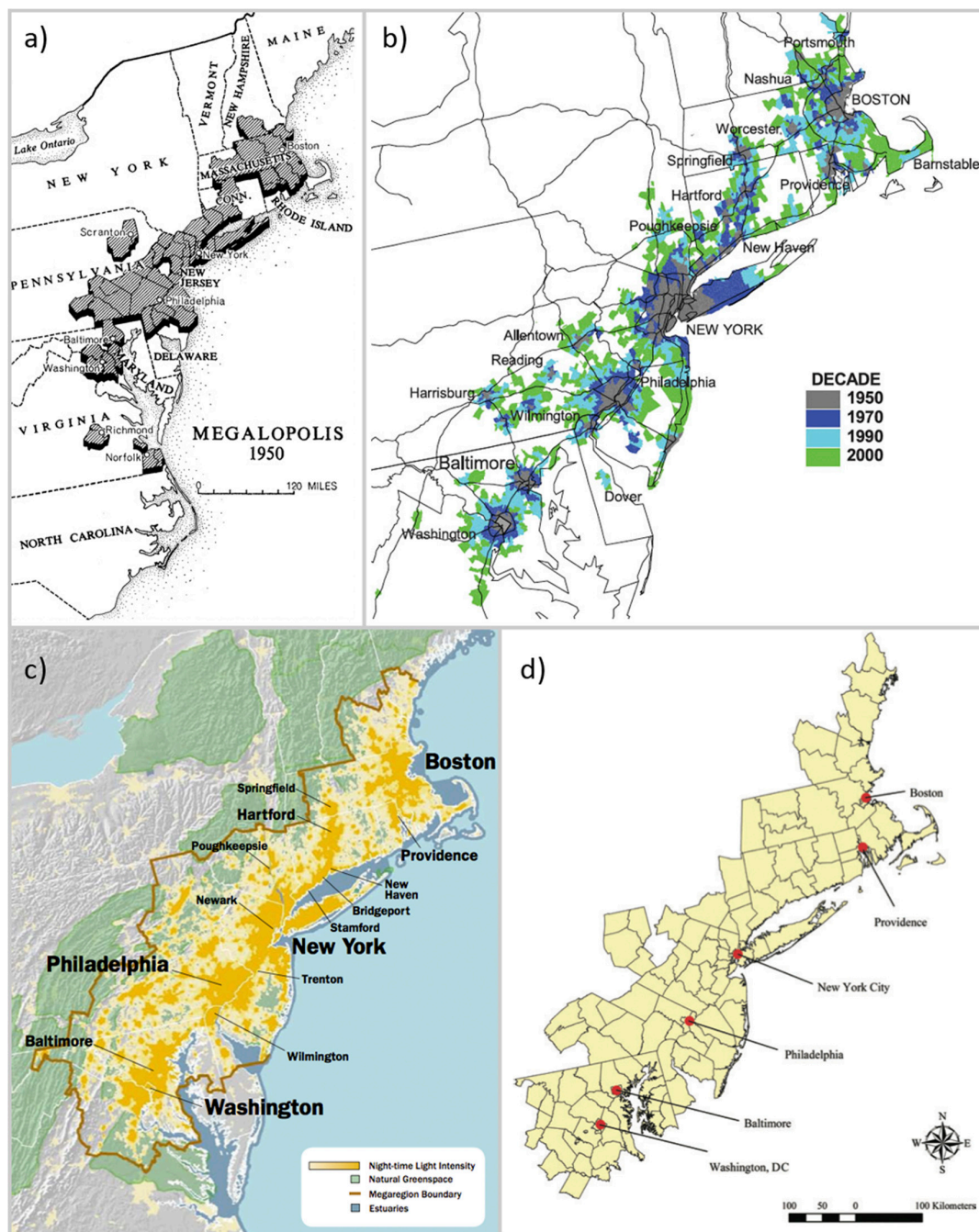
While [69] base their population figures on a consistent area for different periods of time (Table 2), Morrill [67,72] uses a dynamic spatial extent (Table 3): Over the course of 50 years, the Boswash area has quadrupled in size and almost doubled in population. However at the same time, the population density has more than halved, indicating a high level of land consumption. Suburbanization and urban sprawl account for the main share of this decreased [66]; the total figure of the urban cores hardly changed at all [69]. Despite this variability in population and area, Megalopolis is still the densest region in the US, among the largest city regions in the world and a main driver of both national and international economy [67]. The global influence was already recognized by Gottmann who also saw the potential of other parts of the US developing into similar massive constellations [62,73].

**Table 2.** Growth of Megalopolis 1950–2000 (after Vicino, Hanlon and Short, 2007).

Year	Population (Megalopolis)	Population (Metro) *	Population of Metro Centers	Population of Suburban Counties	Area (sq. mi) (Megalopolis)	Density (Megalopolis)
2000	48,720,108	47,681,719	16,453,217	31,228,502	52,310	931.3
1950	31,924,488	22,720,346 **	16,435,953	6,284,393	52,310	610.2

\* Metro includes the four metropolitan areas of New York, Boston, Washington and Philadelphia. \*\* Corrected figure after [69].





**Figure 2.** Maps of Megalopolis. (a) Gottmann, 1957; (b) Morrill, 2006; (c) Regional Plan Association, 2007; (d) Vicino, Hanlon and Short, 2007.

**Table 3.** Growth of Megalopolis 1950–2010 (after Morrill, 2006, 2012).

Year	Population (in 1000)	Area (sq. mi.)	Density	Population Change (%)
2010	45,357	-	-	7
2000	42,374	13,490	3155	15.8
1990	36,580	10,185	3590	6.4
1980	34,365	8390	4100	1.2
1970	34,005	7006	4768	18.5
1960	29,441	5348	5285	20
1950	24,534	3283	7315	n/a

A comprehensive update on the extent of Boswash was carried out by [67] in his revision of a 1974 map by Clyde E. Browning (see Figure 2b). Morrill chose to use the census tract (in its 2000 geography) as basic unit which better represents the functional connectivity of the area, rather than block groups or blocks despite their better resolution. Morrill makes out three distinct periods:

1. Post-war housing boom and rapid suburbanization (1950–1970),
2. inner metropolitan decline and slower population growth but massive increase in area, combined with continuous improvement of the highway system (1970–1990),
3. downtown revitalization and gentrification as well as low-density extra-urban sprawl (1990–2000).

The RPA's *Megaregion* [16] delimitation of Boswash (see Figure 2c) is similar to the *Megapolitan* extent established by Virginia Tech [74] but excludes Richmond and Norfolk (Virginia). In their analysis of the world's largest metropolitan regions, [75] list different parts of Boswash as separate economic regions (New York–Philadelphia–Newark, Boston, and Washington–Baltimore). Even as separate parts, these three regions rank among the top five economic regions in the US.

Although the Boswash urban corridor has been extensively studied, [66] see the need for further research into the area to ensure its manageability and sustainability. While the general spatial delimitations are similar, they still differ due to data, methods and spatial entities applied. Our study, therefore, aims to contribute to the body of this research by delineating the region independent of administrative borders or census definitions, allowing—due to the use of multi-source data—for a flexible delimitation of the area using one consistent method. Our aim is not to create merely another delineation using newer and more detailed datasets down to 3D data or individual building level but to create a basis for a global identification of large urban areas using a consistent database.

### 3. Data and Method

#### 3.1. Multi-Source Geodata Sets

There are a number of spatial entities and thematic variables to base a territorial delimitation of large urban areas on—administrative boundaries, census tracts or blocks, built-up area, population densities etc. In order to identify a suitable representation of the Boswash urban corridor without administrative borders, we choose different earth observation and infrastructure data (which we use to determine the built-up space) as well as socioeconomic data. Since not all datasets are available for a more recent date, we use data from around 2010 for our analysis to ensure a meaningful result.

We are aware that the datasets we apply (Section 3.1.1) differ significantly in their resolution and quality—from 12 m to ~1 km and with known over- or underrepresentation of urban areas. We do, however, accept this unreliability of the input data since it is not our aim to produce an unambiguous delineation of the study area. Rather, our method serves to establish just how fuzzy and variable an area's extent is depending on the input data. We test this method for feasibility on the Boswash area but will, in future steps, apply it on a larger scale. Therefore, we chose particular datasets with the following condition: Their worldwide availability (with the restriction that road and rail data would need to be obtained from a different data source such as OpenStreetMap).

Furthermore, we use income data. As they are not available globally from one data source or sometimes outdated or not existent at all, we apply them only as validation layer of the case study.

##### 3.1.1. Data on the Built Environment (Including Infrastructure)

In this study, we use DMSP-OLS stable lights from 2010. The dataset is available as cloud-free composite from <http://ngdc.noaa.gov/eog/dmsp.html>, with a resolution of 30 arcsec. A downside of night-time imagery, particularly DMSP-OLS data, though, is their tendency to overestimate urban areas because of a blooming effect, of the sensor properties, and/or cumulative georeferencing errors [76]. The newer VIIRS sensor outperforms the DMSP-OLS system through improved spatial, radiometric

and spectral resolution [77] as well as accuracy [78]. However, since VIIRS data have only been available since 2012, we use DMSP-OLS stable lights from 2010 in this study.

Global Urban Footprint (GUF) data have been produced at the German Aerospace Centre (DLR) from TerraSAR-X imagery. The product with a resolution of 12 m is based on 2010–2012 imagery and shows man-made urban structures with a vertical extent—i.e., particularly buildings. Flat areas such as streets or parking areas are not included even though they may function as urban space [79].

The United Nation’s Food and Agriculture Organization (FAO) provide a Global Land Cover dataset (GLC-share) with a spatial resolution of 30 arc seconds [80]. It uses information from different data sources including national mapping projects as well as satellite-based datasets (e.g., medium-resolution sensors such as MERIS or MODIS). The database (available for the year 2013) consists of 11 global land cover layers from which we selected the artificial surfaces layer. This class contains urban or related features including urban parks [81].

ESA and UCLouvain’s GlobCover product consists of 22 land-cover classes [39,82]. The land cover map with a 300 m resolution is obtained from MERIS (Medium Resolution Imaging Spectrometer) data. Relevant for this study is the artificial surfaces class which is defined as an urban area percentage of at least 50% for each pixel.

Another dataset is a global inventory of impervious surface area (ISA) with a 1 km resolution containing manmade surfaces including roads, parking lots and buildings. It is based on DMSP-OLS night-time imagery and population count [83].

Road and railroad data serve as a proxy for urban extent in this study. Data from the year 2010 can be obtained through the US Census Bureau’s TIGER (Topologically Integrated Geographic Encoding and Referencing) geodatabase [84]. While railroads are available for the combined United States, primary and secondary roads can only be accessed on a per-state basis.

Table 4 lists all input layers with technical details used for this study.

**Table 4.** Input layers: Built environment (1–5), infrastructure (6,7), socio-economic data (8,9) and one socio-economic data layer for plausibility (10).

	Layer	Category	Source	Original Resolution	Year	Threshold	Threshold Info
1	Global Urban Footprint (GUF)	Built environment	DLR	12 m	2010–2012	$\geq 5$	max. 250
2	Night time lights DMSP-OLS	Built environment	NOAA	30 arcsec (~1 km)	2010	$\geq 15$	max. 63
3	GLC share—artificial surfaces	Built environment	FAO	30 arcsec (~1 km)	2013	$\leq 2000$	all areas within 2 km of GLC artificial surfaces
4	GlobCover (MERIS)	Built environment	ESA 2010 and UCLouvain	300 m	2009	$\leq 16,000$	all areas within 16 km of GlobCover artificial surfaces
5	Estimated Impervious Surface Area	Built environment	NOAA	1000 m	2010	$\geq 3.3$	max. 100
6	Rail lines	Built environment (Infrastructure)	TIGER Line data (US Census)	n/a (vector data)	2010	$\geq 0.09$	max. 0.83
7	All roads	Built environment (Infrastructure)	TIGER Line data (US Census)	n/a (vector data)	2010	$\geq 295$	max. 897
8	GPWv4 Population Count	Socio-economic	CIESIN (SEDAC)	30 arcsec (~1 km)	2010	$\geq 38$	max. 35,848
9	GPWv4 Population Density	Socio-economic	CIESIN (SEDAC)	30 arcsec (~1 km)	2010	$\geq 65$	max. 55,068
10	Income	Socio-economic	American Fact Finder	Census Tract level (administrative unit)	2010	n/a	n/a



### 3.1.2. Socio-Economic Data

The socioeconomic data used in this study include population and income data. We obtained population count and density from the Socioeconomic Data and Applications Center (SEDAC) [85,86] and per capita income on Census Tract level from the US Census Bureau's American FactFinder database [87]. Although population count and density show a high level of redundancy, we nevertheless use both layers in order to apply separate thresholds.

### 3.1.3. Data Pre-Processing

The spatial resolutions of the input layers (Table 4) vary significantly. For a spatially consistent and transparent implementation of a more comprehensive, homogeneous delineation of the Boswash corridor, we reproject all input data to an Albers Equal Area (North America) projection and aggregate them to a 1 km<sup>2</sup> grid. For example, in the case of GlobCover data, this is done through a simple majority rule, which means that an aggregated 1 km<sup>2</sup> pixel is urban if more than 50% of the original input data are also urban (with an original 300 m resolution,  $3.33 \times 3.33$  pixels are aggregated to our 1 km grid).

The final 1 km<sup>2</sup> grid of the area of interest has a total area of 1,107,882 km<sup>2</sup>. Water bodies and other no data areas are removed from this analysis, leaving a total study area of 520,143 km<sup>2</sup>.

Since the built environment and transportation networks of a city are closely related [88], we calculate the respective line densities for road and railroad data. The road data, at the level of detail used in this study, are only available on a per-state bases; therefore, we first merged the data and removed all duplicates (which occurred frequently at state boundaries). We acknowledge that the choice of cell size (here: 1000) and search radius (here: 1000) will influence the output.

For a spatial representation of the mean income on census tract level, we join the income data with the TIGER geodatabase [84].

Layers 1–9 from Table 4 are used in all steps of the analysis; layer 10—the only input dataset not available on a global scale—serves to validate the results.

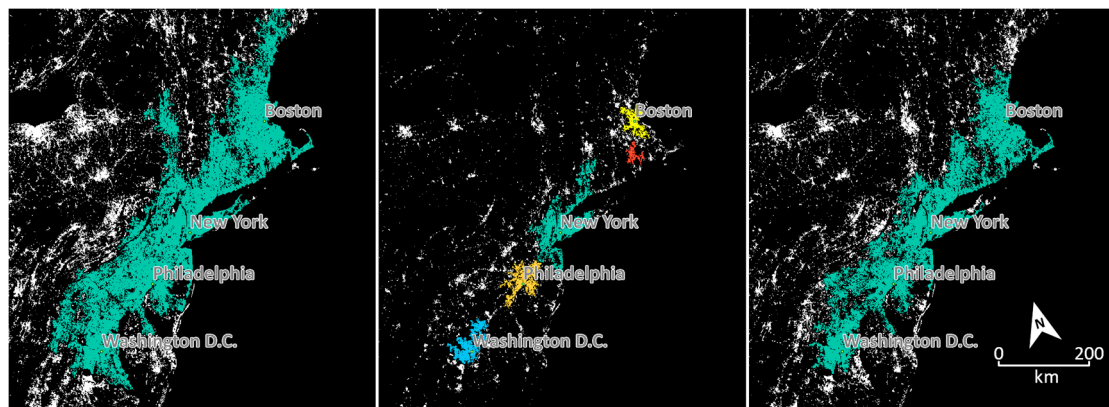
## 3.2. Method for a Flexible Delineation of an Urban Corridor

For each input layer, we establish an individual, data-specific extent of the Boswash corridor using the following method: Our main condition is that we assume the urban corridor from Boston to Washington must be a spatially connected, continuous (thus spatially uninterrupted) area. This spatial contiguity is, according to [17], a common characteristic. Therefore, we set a threshold for each individual input layer where this connection is given: All pixels within this area are above (or below, depending on the input dataset) a certain value, i.e., these pixels fulfilling the threshold condition can be combined to form one single large patch without spatial interruptions between Boston and Washington; the respective threshold equals the lowest (or highest) value where this one single patch exists. This way, we are able to determine how Boswash can be separated from its surroundings using the same method for each layer individually.

To illustrate our method, we use the variable 'population density' as an example. A spatially continuous connection between Boston and Washington applies for a minimum population density of >65 people/km<sup>2</sup> (Figure 3, right); thus, we use this value as threshold for our analysis. For a threshold lower than 65, the stretch from Framingham to Worcester (west of Boston) would not be spatially bounded.

For the DMSP-OLS imagery, the threshold we established is 15. Compared to other studies, this value is relatively low: In their object-based approach to describe urban areas in China using night-time imagery, [89] found that a single threshold underestimates smaller urban areas while overestimating large areas. The optimal threshold values, for their study, were between 5 (small urban objects) and 60 (large objects) for DMSP/OLS images from the year 2000. [76] set the DMSP threshold

for Lhasa, Beijing and San Francisco to 19, 30 and 51 respectively which corresponds best to their Landsat TM-derived urban boundaries for each study area.



**Figure 3.** Threshold options for the population density layer. From left to right: 33, 360 and 65 people/km. The largest patch is in green; other colors denote spatially separate large metropolitan areas (blue: Washington D.C., orange: Philadelphia, green: New York, red: Providence, yellow: Boston).

Table 4 shows the thresholds identified for all input raster layers based on the particular lowest (or highest, respectively) value resulting in one single patch. The weakest link is generally identified along the Boston–Providence–New London–New Haven or the Boston–Worcester–Springfield–Hartford stretches but occasionally in the southern part of the area between Baltimore and Philadelphia or Harrisburg. For the road and railway line density network, the resulting threshold also includes Pittsburgh (Pennsylvania) as well as Richmond and Norfolk (Virginia). Raising the threshold, though, would cut off the greater Boston area.

In consequence of this methodological approach, we construct 9 different spatial forms of the urban corridor, one from each individual input layer. This spatially variable view on the urban corridor shows that regions are not fixed, but fuzzy territorial entities (soft spaces) and boundaries are malleable depending on data and methods applied (e.g., [74]). To account for this fuzziness in the conceptual idea of urban corridors, we stack the resultant entities and calculate the number of “Boswash matches” for each pixel. This shows the Boswash area with fuzzy boundaries where higher matches indicate that more layers show agreement. With this approach we provide a flexible territorial form of the urban corridor with fuzzy boundaries. In a final step, we break down this summary layer to obtain different levels of agreement. This ranges from one (if for at least one layer a pixel lies ‘within Boswash’) to nine (all input layers classify a pixel as being ‘within Boswash’), where ‘within Boswash’ means ‘within the threshold identified as possible physical shape of the area’.

### 3.3. Plausibility of the Method

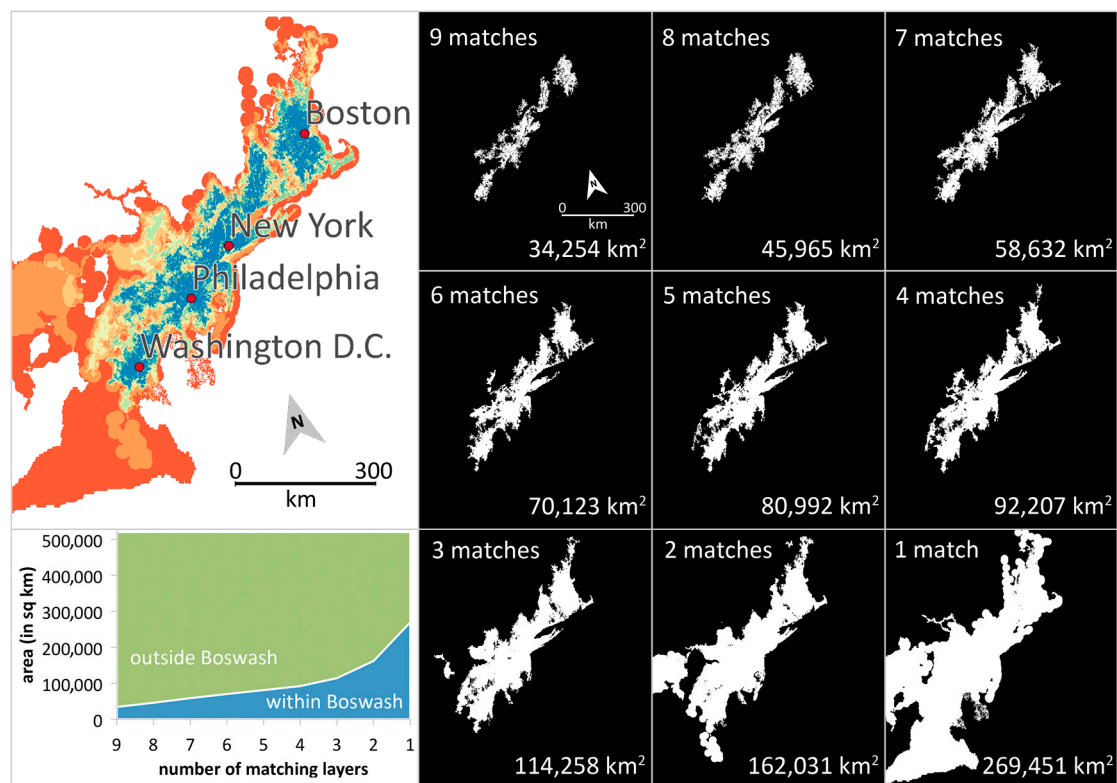
The construction of regional spaces such as an urban corridor is, due to the inexistence of a ‘reference truth’, never ‘correct’ or ‘incorrect’. Thus, an assessment of accuracy in a traditional sense is not meaningful. However, since urban corridors, as an accepted term in urban geography, are proving to be economically relevant and dominant, we check the plausibility of our (fuzzy) territorial space against our hypothesis that the income within those spaces is significantly higher than in the surrounding hinterlands. For this, we join income data from the American Fact Finder (AFF) with Census Tract boundaries from the Census Bureau’s TIGER/Line shapefile database. The result is rasterized and resampled to our existing 1 km grid. We then calculate the median income for the nine different urban corridor extents derived through the overlay of our original input layers. In order to establish whether there are significant differences to the surroundings, we also determine the median income for all areas outside the possible Boswash area for each of the different extents. The results

help to quantify the fuzzy outlines of the area, but first and foremost allow to check the plausibility of our flexible territorial urban corridor delineation.

## 4. Results

### 4.1. Probability-Based Spatial Delimitation of the Boswash Urban Corridor

The main results are spatial delimitations of the Boston to Washington urban corridor with various probabilities for being part of the corridor. Thus, the boundaries to delimit Boswash are fuzzy since the extent varies for each input layer. The boundaries are freed from any administrative unit but based on a consistent 1 km grid. Figure 4 shows the different levels of overlap for different numbers of matching layers. This illustrates the fuzziness of the continuously connected urban space of the area: Where fewer input layers agree, the potential Boswash extent is naturally a lot larger. Where all nine input layers agree, the area is considerably smaller, more fragmented and more distinct. This difference in extent ranges from almost 270,000 km<sup>2</sup> (1 match) to under 35,000 km<sup>2</sup> (9 matches).



**Figure 4.** Top left: Probability-based flexible spatial delimitation of the Boswash urban corridor. From red to blue: Increase in layer agreement. Bottom left: Area outside and within Boswash in dependence of matching layers. Right: Different possible forms of the Boswash urban corridor in relation to the number of matching layers.

Blue areas in Figure 4 (top left) show high levels of agreement between all datasets, representing areas best described as part of Boswash through our analysis. Red areas show decreasing probabilities of spaces belonging to the corridor. Overall these results show the complexity when constructing territorial spaces. In addition, the spatial fuzziness of the Boswash urban corridor is further illustrated in Figure 2, where similar elongated extents from other studies are shown.

#### 4.2. Validation of Results: Median Income within and Outside of Boswash

Urban corridors are conceptualized as focal areas of economic activity (e.g., [90,91]). In consequence, we check the plausibility of our spatial delimitations using socioeconomic data in the form of median income. We assume the delimitation is feasible if the median income within the delimited urban corridor spaces area is significantly higher than in peripheral areas. With this, we prove whether the economic dominance of the area can be verified. Therefore, we calculate the median income for each of our nine resulting agreement layers both within and outside Boswash.

Figure 5 illustrates that the median income is generally higher within the delimited urban corridor spaces and thus proves that our assumption is met. We find a gradual decrease of median income with an increasing extent (i.e., a smaller number of matching layers) of the Boswash area—a correlation that the fuzzy boundaries in this multi-source geodata approach in general confirm the hypothesis. The total median income for the study area is \$55,950, which is well above the US median household income of \$49,445 [92]. In the case of all 9 income layers belonging to Boswash, the income within Boswash is just over \$88,000 compared to \$53,690 outside—a plus of about two thirds. With a decreasing number of matching layers, the median income also drops, but even for just one match, the difference between within and outside Boswash is still almost \$16,000 or around one third (\$63,655 versus \$47,669). This illustrates that the lower the agreement between the input layers is, the lower the median income, both within and outside Boswash. For each decrease in agreement, areas with a relatively high income outside Boswash are now attributed to the “within” area. Since the higher income outside is still lower than within, this results in a lower median income both within and outside the Boswash area. The spatial extent of the potential Boswash area varies significantly between the different result layers (see Figure 4): The area with only one match is about eight times larger than where nine input layers agree.

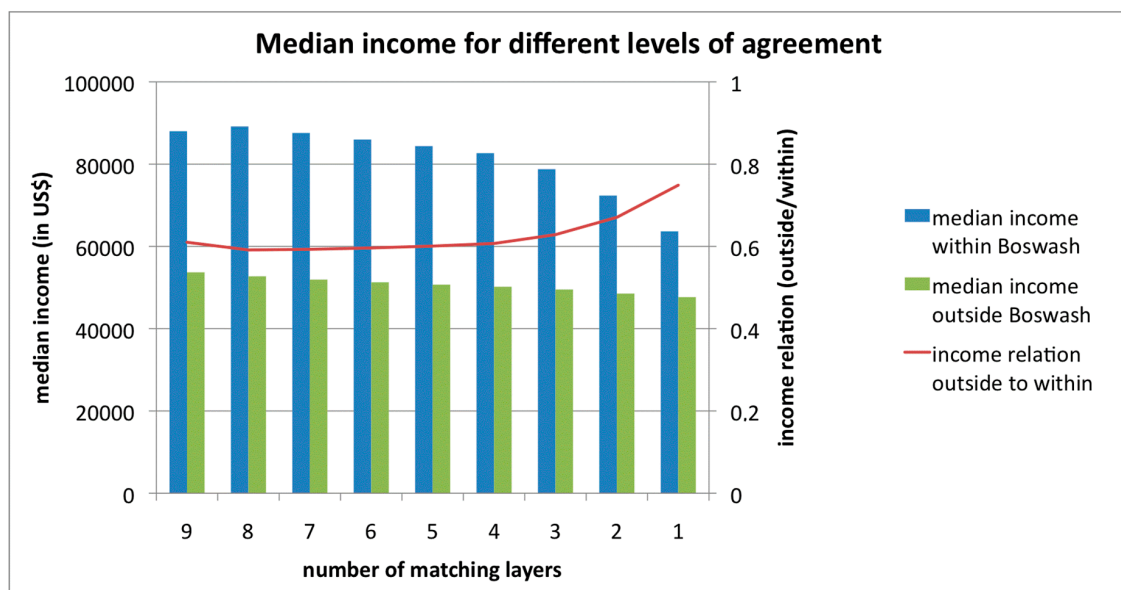


Figure 5. Median income for different Boswash extents.

## 5. Discussion

The spatial delimitation of regions is challenging—clear, crisp borders do not exist in complex real-world landscapes, and assumed or accepted boundaries are malleable due to conceptual logic, datasets or thresholds applied. Conceptual complexity leads to a struggle to construct regions in a consistent territorial layout [93]. This study aimed at delineating the Boswash urban corridor in a transparent and methodologically consistent way, taking the conceptual fuzziness of the term ‘urban

corridor' into account with flexible, probability-based boundaries of such a geographic construct. Therefore, we used a variety of input layers (nine layers), which were treated as equally important. A tenth layer—median income—served as validation layer. This strategy aims for transparency, re-usability over time and consistency within the area of interest due to a grid-based approach. Furthermore, a transfer of our method to other large urban areas worldwide is possible since with the exception of the US Census Bureau data (for income and road/rail data), all datasets are available on a global scale (however, with varying accuracies). While for household income, a globally consistent high-resolution dataset does not exist, road and rail data can alternatively be obtained through the OpenStreetMap (OSM) platform. However, for our purposes, the availability of 2010 data from the US Census was preferable. By adding or removing data layers, the approach presented in this paper can easily be extended or altered while retaining the universal applicability: The likelihood that a particular part of the Earth's surface belongs to an urban corridor (or any other concept for a constructed territorial space) is determined by the amount of layers which yield positive results. Commuter patterns as well as airline, mobile phone or collective sensing data (e.g., [94–98]) would provide further insight into the functional connectedness of cities—maybe even adding a space of flows perspective to the 'space of place' logic applied in this study—however these are not available globally and consistently at this stage.

We construct a flexible and fuzzy spatial entity of an urban corridor using one method on nine different input layers. Our probability-based, spatially flexible way of delimiting an urban corridor allows for a variable rather than fixed perspective onto regional developments—essentially a territorial 'space of place' perspective, as the most visible appearance of success is through local activities. We are aware that regions can also be constructed in other ways, e.g., interpreted as virtual spaces of flows. However, regional phenomena, as [99] argues, are not single, unified phenomena, but a syndrome of processes and activities. In our approach we try to integrate locally relevant indicators for constructing regional entities.

We acknowledge that there are some limitations with the input data used for this study. Night-time lights, for example, represent economic activity rather than urban areas directly. GlobCover imagery underestimates urban extent especially in lower dense peri-urban and rural areas (e.g., [100]), resulting in large gaps between urban areas which impedes our analysis. The thresholds for our line densities calculated for rail and road networks do not show the apparent Boswash outline we were hoping to obtain—large areas towards the south and west are included—which is reflected in the results where only one layer indicates an association with Boswash. Also, there is some data redundancy through the GPWv4 population count and density layers. However, we believe that the input layers can be exchanged or increased through other datasets. For example, the income layer which we used for our validation could have been added as an input layer in the analysis. Overall, each result reiterates a noticeable gap in the economic performance within and outside of Boswash.

Through our method, we obtain a spatial delimitation of the Boswash urban corridor without using administrative boundaries. Our thresholds describe areas with similar characteristics which are then combined to delimit different variations of Boswash depending on the level of agreement between the input layers. We are aware that the use of thresholds may be limiting [101]: Small modifications of a threshold can change the overall result noticeably. Also, they may well work in a sample area but may not be applicable globally because of different socioeconomic status and physical environment [29,59]. Some authors, such as [76], use different thresholds for different urban areas. However, [102] calculated a global threshold for the extraction of "natural cities" on night-time imagery using a statistical head/tail breaks classification developed by [58]. The thresholds used in this paper, though, are derived so that the study area represents a connected patch, which is part of the underlying assumptions of our method.

For the DMSP-OLS data, the threshold value is relatively low compared to other studies [76,89]), resulting in some overestimation of urban areas—a known characteristic of these data [76,89].



For the population density, other possible thresholds include, for example, the average for the United States (33 people/km<sup>2</sup> in 2010; Figure 3, left) [103]. However, Ref. [104] identifies an average population density of about 360 people/km<sup>2</sup> for the Boswash megalopolis. This threshold only returns the area between Hartford–New York–Toms River as largest continuous connected patch but does not include Boston, Philadelphia or Washington (separate patches in Figure 3, center). While these alternative thresholds also produce a possible delimitation, neither of them would, however, allow us to follow our basic conceptual considerations and to use our method consistently for all layers.

At first glance, a raster resolution of 1 km seems 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, a higher resolution is not required, particularly when keeping in mind a possible application of our method to much larger regions. A 1 km grid is also used by transnational programs such as the European INSPIRE initiative [105].

We are aware that each individual binary version of an input layer could be questioned by an expert. The overall construct, however—in our case consisting of nine layers—turned out to be very robust. The spatial resolution of 1 km returned plausible results even though some datasets are available in a higher resolution. As [106] found out, data aggregation does not necessarily imply less appropriate outcomes but can in fact mean better results. Our fuzzy delimitation of Boswash complements existing maps (see Figure 2) through the use of diverse input data and variables. We believe that our method to describe the area as a single connected area is mathematically reasonable and thus objective overall. Interactive GIS applications may even allow planners to overcome the binary view while performing queries like “show all raster cells which yield at least seven positive scores”.

## 6. Conclusions

This study built on the conceptual work of [3,7] and instantiated the generically defined concept of urban corridors for the well-known example of Boswash. Although this is one of the best examples of an urban corridor in the world, this construct has so far not been delineated in such a transparent and flexible way (regarding data sets using one consistent method). Existing map representations differ surprisingly regarding the shape and the extent of the corridor.

We developed consistent way to delineate this urban corridor by using nine different input data layers. Using earth observation, infrastructure, population and economic data, we constructed several possible spatial forms of the Boswash urban corridor. This approach is, to our best knowledge, the first transparent and transferable methodology to delineate an urban corridor while taking the fuzziness of the delimitation of territorial space into consideration through a probability-based approach. In addition to the generic methodology developed in this study, we could reveal some idiosyncrasies of the Boswash corridor. We discussed aspects that influence the robustness of the delineation, as a prerequisite for transferability and potentially for a monitoring of the corridor over time. The proposed approach is a means to a spatially explicit delineation and mapping of complex geographical phenomena which are sometimes hand-drawn by planners. Its usability will be benchmarked in future research regarding the transferability of the approach. Nevertheless, future research needs to demonstrate whether the ‘appropriateness’ of the approach depends on the context of the study or whether the approach is generically applicable. The new geography of spatially connected, massive urban constructs will need new instruments for governance, resource management and planning, for which our method can provide valuable information.

**Acknowledgments:** This research was funded by the Austrian Science Fund (FWF) through the Doctoral College GIScience (DK W 1237-N23) at the University of Salzburg.

**Author Contributions:** I.G. and H.T. designed the method; I.G., T.B. and H.T. reviewed the method; I.G. analyzed the data and wrote the manuscript, H.T. and T.B. supervised and critically reviewed the findings of the paper.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Martine, G. Preparing for sustainable urban growth in developing areas. In *UN-DESA (2011). Population Distribution, Urbanization, Internal Migration and Development: An International Perspective*; United Nations Publication: New York, NY, USA, 2011; pp. 6–30.
2. United Nations Human Settlements Programme (UN-Habitat). *State of the World's Cities 2012/13. Prosperity of Cities*; Routledge: New York, NY, USA, 2013.
3. Georg, I.; Blaschke, T.; Taubenböck, H. New spatial dimensions of global cityscapes: From reviewing existing concepts to a conceptual spatial approach. *J. Geogr. Sci.* **2016**, *26*, 355–380. [[CrossRef](#)]
4. Angel, S.; Sheppard, S.; Civco, D.L.; Buckley, R.; Chabaeva, A.; Gitlin, L.; Krale, A.; Parent, J.; Perlin, M. *The Dynamics of Global Urban Expansion*; Transport and Urban Development Department, World Bank: Washington, DC, USA, 2005.
5. Soja, E.; Kanai, E. The Urbanization of the World. In *The Endless City*; Burdett, R., Ed.; Phaidon Press Limited: London, UK, 2007; pp. 54–69.
6. Taubenböck, H.; Wiesner, M.; Felbier, A.; Marconcini, M.; Esch, T.; Dech, S. New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. *Appl. Geogr.* **2014**, *47*, 137–153. [[CrossRef](#)]
7. Georg, I.; Blaschke, T.; Taubenböck, H. A Global Inventory of Urban Corridors Based on Perceptions and Night-Time Light Imagery. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 233. [[CrossRef](#)]
8. Whebell, C.F.J. Corridors: A theory of urban systems. *Ann. Assoc. Am. Geogr.* **1969**, *59*, 1–26. [[CrossRef](#)]
9. United Nations Human Settlements Programme (UN-Habitat). *State of the World's Cities 2010/2011: Bridging the Urban Divide*; Earthscan: London, UK, 2008.
10. Trip, J.J. The corridor as a transport network. In *Megacorridors in North West Europe. Investigating a New Transnational Planning Concept*; Zonneveld, W., Trip, J.J., Eds.; Housing and Urban Policy Studies 27; Delft University Press: Delft, The Netherlands, 2003; pp. 17–23.
11. Li, P.; Cao, X.S. Evolution and development of Guangzhou-Hong Kong corridor. *Chin. Geogr. Sci.* **2005**, *15*, 206–211. [[CrossRef](#)]
12. Neuman, M. Regional design: Recovering a great landscape architecture and urban planning tradition. *Landsc. Urban Plan.* **2000**, *47*, 115–128. [[CrossRef](#)]
13. Chapman, D.; Pratt, D.; Larkham, P.; Dickins, I. Concepts and definitions of corridors: Evidence from England's Midlands. *J. Trans. Geogr.* **2003**, *11*, 179–191. [[CrossRef](#)]
14. Gottmann, J. Megalopolis or the urbanization of the northeastern seaboard. *Econ. Geogr.* **1957**, *33*, 189–200. [[CrossRef](#)]
15. Lang, R.E.; Dhavale, D. *Beyond Megalopolis: Exploring America's New "Megapolitan" Geography*; Metropolitan Institute, Virginia Technical University: Alexandria, VA, USA, 2005.
16. Regional Plan Association. *Northeast Megaregion 2050: A Common Future*; Regional Plan Association: New York, NY, USA, 2007.
17. Florida, R.; Gulden, T.; Mellander, C. The rise of the mega-region. *Camb. J. Reg. Econ. Soc.* **2008**, *1*, 459–476. [[CrossRef](#)]
18. Laquian, A. The planning and governance of Asia's mega-urban regions. In *UN-DESA (2011). Population Distribution, Urbanization, Internal Migration and Development: An International Perspective*; United Nations Publication: New York, NY, USA, 2011; pp. 302–322.
19. Choe, S.C. The evolving urban system in North-East Asia. In *Emerging World Cities in Pacific Asia*; Lo, F.C., Yeung, Y.M., Eds.; UNU Press: Tokyo, Japan, 1996; pp. 498–519.
20. United Nations Human Settlements Programme (UN-Habitat). *The State of African Cities 2008. A Framework for Addressing Urban Challenges in Africa*; UN-Habitat: Nairobi, Kenya, 2008.
21. Hospers, G.-J. Beyond the Blue Banana? Structural Change in Europe's Geo-Economy. *Intereconomics* **2003**, *38*, 76–85. [[CrossRef](#)]
22. Abler, R.; Adams, J.S.; Gould, P. *Spatial Organization: The Geographer's View of the World*; Prentice-Hall International: London, UK, 1972.
23. Lang, S.; Kienberger, S.; Tiede, D.; Hagenlocher, M.; Pernkopf, L. Geons—Domain-specific regionalization of space. *Cartogr. Geogr. Inf. Sci.* **2014**, *41*, 214–226. [[CrossRef](#)]

24. Berry, B.J.L. Interdependency of Spatial Structure and Spatial Behavior: A General Field Theory Formulation. *Pap. Reg. Sci.* **1968**, *21*, 205–227. [[CrossRef](#)]
25. Castells, M. *The Rise of the Network Society*, 2nd ed.; Wiley-Blackwell: Oxford, UK, 2000.
26. Castells, M. Local and global: Cities in the network society. *Tijdschrift voor Economische en Sociale Geografie* **2002**, *93*, 548–558. [[CrossRef](#)]
27. Elvidge, C.D.; Baugh, K.E.; Kihn, E.A.; Kroehl, H.W.; Davis, E.R. Mapping city lights with nighttime data from the DMSP Operational Linescan System. *Photogramm. Eng. Remote Sens.* **1997**, *63*, 727–734.
28. Imhoff, M.L.; Lawrence, W.T.; Stutzer, D.C.; Elvidge, C.D. A technique for using composite DMSP/OLS “city lights” satellite data to map urban area. *Remote Sens. Environ.* **1997**, *61*, 361–370. [[CrossRef](#)]
29. Small, C.; Pozzi, F.; Elvidge, C.D. Spatial analysis of global urban extent from DMSP-OLS night lights. *Remote Sens. Environ.* **2005**, *96*, 277–291. [[CrossRef](#)]
30. Zhang, Q.; Seto, K.C. Can night-time light data identify typologies of urbanization? A global assessment of successes and failures. *Remote Sens.* **2013**, *5*, 3476–3494. [[CrossRef](#)]
31. Cai, J.; Huang, B.; Song, W. Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sens. Environ.* **2017**, *202*, 210–221. [[CrossRef](#)]
32. Corbane, C.; Kemper, T.; Freire, S.; Louvrier, C.; Pesaresi, M. *Monitoring the Syrian Humanitarian Crisis with the JRC’s Global Human Settlement Layer and Night-Time Satellite Data*; Publications Office of the European Union: Luxembourg, UK, 2016. [[CrossRef](#)]
33. Schneider, A.; Friedl, M.A.; Potere, D. Mapping global urban areas using MODIS 500-m data: New methods and datasets based on ‘urban ecoregions’. *Remote Sens. Environ.* **2010**, *114*, 1733–1746. [[CrossRef](#)]
34. Miyazaki, H.; Shao, X.; Iwao, K.; Shibasaki, R. An automated method for global urban area mapping by integrating ASTER satellite images and GIS data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 1004–1019. [[CrossRef](#)]
35. Gamba, P.; Lisini, G. Fast and efficient urban extent extraction using ASAR wide swath mode data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2184–2195. [[CrossRef](#)]
36. Zhou, Y.; Smith, S.J.; Zhao, K.; Imhoff, M.; Thomson, A.; Bond-Lamberty, B.; Asrar, G.R.; Zhang, X.; He, C.; Elvidge, C.D. A global map of urban extent from nightlights. *Environ. Res. Lett.* **2015**, *10*, 054011. [[CrossRef](#)]
37. Esch, T.; Taubenböck, H.; Roth, A.; Heldens, W.; Felber, A.; Thiel, M.; Schmidt, M.; Müller, A.; Dech, S. TanDEM-X mission: New perspectives for the inventory and monitoring of global settlement patterns. *J. Appl. Remote Sens.* **2012**, *6*, 061702. [[CrossRef](#)]
38. Bartholomé, E.; Belward, A.S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **2005**, *26*, 1959–1977. [[CrossRef](#)]
39. Bontemps, S.; Defourny, P.; Van Bogaert, E.; Arino, O.; Kalogirou, V.; Ramos Perez, J. *GlobCOVER 2009 Products Description and Validation Report*; Université catholique de Louvain: Louvain-la-Neuve, Belgium; ESA: Paris, France, 2011.
40. Loveland, T.R.; Reed, B.C.; Brown, J.F.; Ohlen, D.O.; Zhu, Z.; Yang, L.; Merchant, J.W. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens.* **2000**, *21*, 1303–1330. [[CrossRef](#)]
41. Tateishi, R.; Uriyangqai, B.; Al-Bilbisi, H.; Ghar, M.A.; Tsend-Ayush, J.; Kobayashi, T.; Kasimu, A.; Thanh, H.N.; Shalaby, A.; Alsaadeh, B.; et al. Production of Global Land Cover, GLCNMO. *Int. J. Digit. Earth* **2011**, *4*, 22–49. [[CrossRef](#)]
42. Pesaresi, M.; Huadong, G.; Blaes, X.; Ehrlich, D.; Ferri, S.; Gueguen, L.; Marin-Herrera, M.A. A Global Human Settlement Layer from optical HR/VHR RS data: Concept and first results. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 2102–2131. [[CrossRef](#)]
43. Yang, J.; Song, G.; Lin, J. Measuring spatial structure of China’s megaregions. *J. Urban Plan. Dev.* **2014**, *141*, 04014021. [[CrossRef](#)]
44. Wei, C.; Taubenböck, H.; Blaschke, T. Measuring urban agglomeration using a city-scale dasymetric population map: A study in the Pearl River Delta, China. *Habitat Int.* **2017**, *59*, 32–43. [[CrossRef](#)]
45. Taubenböck, H.; Wiesner, M. The spatial network of megaregions—Types of connectivity between cities based on settlement patterns derived from EO-data. *Comput. Environ. Urban Syst.* **2015**, *54*, 165–180. [[CrossRef](#)]
46. Taubenböck, H.; Ferstl, J.; Dech, S. Regions set in stone—Classifying and categorizing regions in Europe by settlement patterns derived from EO-data. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 55. [[CrossRef](#)]

47. Taubenböck, H.; Bauer, P.; Geiß, C.; Wurm, M. 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, 6–8 March 2017.
48. Jiang, B.; Jia, T. Zipf's law for all the natural cities in the United States: A geospatial perspective. *Int. J. Geogr. Inf. Sci.* **2010**, *25*, 1269–1281. [[CrossRef](#)]
49. Arcaute, E.; Hatna, E.; Ferguson, P.; Youn, H.; Johansson, A.; Batty, M. Constructing cities, deconstructing scaling laws. *J. R. Soc. Interface* **2015**, *12*, 20140745. [[CrossRef](#)] [[PubMed](#)]
50. Schneider, A.; Woodcock, C.E. Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Stud.* **2008**, *45*, 659–692. [[CrossRef](#)]
51. Fotheringham, A.S.; Wong, D.W. The modifiable areal unit problem in multivariate statistical analysis. *Environ. Plan. A* **1991**, *23*, 1025–1044. [[CrossRef](#)]
52. Openshaw, S. *The Modifiable Areal Unit Problem*; Geo Abstracts; University of East Anglia: Norwich, UK, 1984.
53. Marceau, D.J. The scale issue in the social and natural sciences. *Can. J. Remote Sens.* **1999**, *25*, 347–356. [[CrossRef](#)]
54. Pietrzak, M.B. The Modifiable Areal Unit Problem—Analysis of Correlation and Regression. *Equilibrium. Q. J. Econ. Eco. Policy* **2014**, *9*, 113–131.
55. Taubenböck, H.; Standfuß, I.; Klotz, M.; Wurm, M. The physical density of the city—Deconstruction of the delusive density measure with evidence from European megacities. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 206. [[CrossRef](#)]
56. Flowerdew, R. How serious is the Modifiable Areal Unit Problem for analysis of English census data? *Popul. Trends* **2011**, *145*, 106–118. [[CrossRef](#)] [[PubMed](#)]
57. Zhou, Y.; Smith, S.J.; Elvidge, C.D.; Zhao, K.; Thomson, A.; Imhoff, M. A cluster-based method to map urban area from DMSP/OLS nightlights. *Remote Sens. Environ.* **2014**, *147*, 173–185. [[CrossRef](#)]
58. Jiang, B. Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *Prof. Geogr.* **2013**, *65*, 482–494. [[CrossRef](#)]
59. Cao, X.; Chen, J.; Imura, H.; Higashi, O. A SVM-based method to extract urban areas from DMSP-OLS and SPOT VGT data. *Remote Sens. Environ.* **2009**, *113*, 2205–2209. [[CrossRef](#)]
60. Foody, G.M.; Mathur, A. A relative evaluation of multiclass image classification by support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1335–1343. [[CrossRef](#)]
61. Gottmann, J. *Megalopolis: The Urbanized Northeastern Seaboard of the United States*; The Twentieth Century Fund: London, UK, 1961.
62. Gottmann, J. Megalopolitan systems around the world. *Ekistics* **1976**, *41*, 109–113.
63. Pacione, M. *Urban Geography: A Global Perspective*, 3rd ed.; Routledge: London, UK, 2009.
64. Baigent, E. Patrick Geddes, Lewis Mumford and Jean Gottmann: Divisions over 'megalopolis'. *Prog. Hum. Geogr.* **2004**, *28*, 687–700. [[CrossRef](#)]
65. Kahn, H.; Wiener, A. The Next Thirty-Three Years: A Framework for Speculation. *Daedalus* **1967**, *96*, 705–732.
66. Oswald, M.R.; Gayley, R.J.; Ames, D.L.; McNeil, S. *Evaluating the Current State of the BOSWASH Transportation Corridor and Indicators of Resiliency*; Department of Civil Engineering, University of Delaware: Newark, DE, USA, 2009. Available online: [www.ce.udel.edu/UTC/Presentation%2009/MichelleOswald-AISIM5paper.pdf](http://www.ce.udel.edu/UTC/Presentation%2009/MichelleOswald-AISIM5paper.pdf) (accessed on 13 November 2017).
67. Morrill, R. Classic map revisited: The growth of Megalopolis. *Prof. Geogr.* **2006**, *58*, 155–160. [[CrossRef](#)]
68. Rodrigue, J.P. Freight, Gateways and Mega-Urban Regions: The Logistical Integration of the Bostwash Corridor. *Tijdschrift voor Economische en Sociale Geografie* **2004**, *95*, 147–161. [[CrossRef](#)]
69. Vicino, T.J.; Hanlon, B.; Short, J.R. Megalopolis 50 years on: The transformation of a city region. *Int. J. Urban Reg. Res.* **2007**, *31*, 344–367. [[CrossRef](#)]
70. School of Design, University of Pennsylvania. Reinventing megalopolis: The Northeast Megaregion. 2005. Available online: <http://www.america2050.org/2005/11/reinventing-megalopolis-the-no.html> (accessed on 13 November 2017).
71. Kotkin, J.; Schill, M. A Map of America's Future: Where Growth Will Be over the Next Decade. Available online: <https://www.forbes.com/sites/joelkotkin/2013/09/04/a-map-of-americas-future-where-growth-will-be-over-the-next-decade/#67fd1dc3623d> (accessed on 03 January 2018).



72. Morrill, R. Megalopolis and Its Rivals. 2012. Available online: <http://www.newgeography.com/content/002788-megalopolis-and-its-rivals> (accessed on 13 November 2017).
73. Hagler, Y. *Defining US Megaregions*; America 2050 Report; Regional Plan Association: New York, NY, USA, 2009.
74. Ross, C.L.; Yang, J.; Barringer, J. *Megaregions: Literature Review of the Implications for U.S. Infrastructure Investment and Transportation Planning*; Center for Quality Growth and Regional Development (CQGRD), Georgia Institute of Technology: Atlanta, GA, USA, 2008.
75. Florida, R.; Mellander, C.; Gulden, T. Global metropolis: Assessing economic activity in urban centers based on nighttime satellite images. *Prof. Geogr.* **2012**, *64*, 178–187. [[CrossRef](#)]
76. Henderson, M.; Yeh, E.T.; Gong, P.; Elvidge, C.; Baugh, K. Validation of urban boundaries derived from global night-time satellite imagery. *Int. J. Remote Sens.* **2003**, *24*, 595–609. [[CrossRef](#)]
77. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.C. Why VIIRS data are superior to DMSP for mapping nighttime lights. In Proceedings of the Asia-Pacific Advanced Network 2013, Honolulu, HI, USA, 13–16 January 2013; pp. 62–69.
78. Shi, K.; Huang, C.; Yu, B.; Yin, B.; Huang, Y.; Wu, J. Evaluation of NPP-VIIRS night-time light composite data for extracting built-up urban areas. *Remote Sens. Lett.* **2014**, *5*, 358–366. [[CrossRef](#)]
79. Esch, T.; Taubenböck, H.; Felbier, A.; Roth, A.; Müller, A.; Dech, S. The path to mapping the global urban footprint using TanDEM-X data. In Proceeding of the ISRSE 2011, Sydney, Australia, 10–15 April 2011.
80. The Food and Agriculture Organization of the United Nations (FAO). 2016. Available online: <http://www.fao.org/geonetwork/srv/en/main.home> (accessed on 13 November 2017).
81. Latham, J.; Cumani, R.; Rosati, I.; Bloise, M. *Global Land Cover Share (GLC-SHARE) Database Beta-Release Version 1.0-2014*; FAO: Rome, Italy, 2014.
82. The European Space Agency (ESA). GlobCover 2009 (Global Land Cover Map). Available online: [http://due.esrin.esa.int/page\\_globcover.php](http://due.esrin.esa.int/page_globcover.php) (accessed on 13 November 2017).
83. National Oceanic and Atmospheric Administration (NOAA). Global Distribution and Density of Constructed Impervious Surfaces. Available online: [https://www.ngdc.noaa.gov/eog/dmsp/download\\_global\\_isa.html](https://www.ngdc.noaa.gov/eog/dmsp/download_global_isa.html) (accessed on 13 November 2017).
84. United States Census Bureau. TIGER Products. Available online: <https://www.census.gov/geo/maps-data/data/tiger.html> (accessed on 13 November 2017).
85. CIESIN Center for International Earth Science Information Network—CIESIN—Columbia University. *Gridded Population of the World, Version 4 (GPWv4): Population Count*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2016. [[CrossRef](#)]
86. CIESIN Center for International Earth Science Information Network—CIESIN—Columbia University. *Gridded Population of the World, Version 4 (GPWv4): Population Density*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2016. [[CrossRef](#)]
87. United States Census Bureau. American FactFinder AFF. Available online: <http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml> (accessed on 13 November 2017).
88. Borruso, G. Network density and the delimitation of urban areas. *Trans. GIS* **2003**, *7*, 177–191. [[CrossRef](#)]
89. Xie, Y.; Weng, Q. Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sens. Environ.* **2016**, *187*, 1–13. [[CrossRef](#)]
90. Priemus, H. Szenarien und Konzepte für die Entwicklung von Korridoren in den Niederlanden und in der Europäischen Union. *Inf. Raumentwickl.* **2004**, *62*, 35–48.
91. Trip, J.J. The corridor as an axis of economic development. In *Megacorridors in North West Europe. Investigating a New Transnational Planning Concept*; Zonneveld, W., Trip, J.J., Eds.; Housing and Urban Policy Studies 27; Delft University Press: Delft, The Netherlands, 2003; pp. 24–35.
92. DeNavas-Watt, C.; Proctor, B.D.; Smith, J.C. *Income, Poverty, and Health Insurance Coverage in the United States*; U.S. Census Bureau, Current Population Reports 2010; U.S. Government Printing Office: Washington, DC, USA, 2011; pp. 60–239.
93. Harrison, J.; Growe, A. From places to flows? Planning for the new ‘regional world’ in Germany. *Eur. Urban Reg. Stud.* **2014**, *21*, 21–41. [[CrossRef](#)]
94. Nelson, G.D.; Rae, A. An Economic Geography of the United States: From Commutes to Megaregions. *PLoS ONE* **2016**, *11*, e0166083. [[CrossRef](#)] [[PubMed](#)]



95. Derudder, B.; Witlox, F. An appraisal of the use of airline data in assessing the world city network: A research note on data. *Urban Stud.* **2005**, *42*, 2371–2388. [[CrossRef](#)]
96. Ratti, C.; Williams, S.; Frenchman, D.; Pulselli, R.M. Mobile landscapes: Using location data from cell phones for urban analysis. *Environ. Plan. B Urban Anal. City Sci.* **2006**, *33*, 727–748. [[CrossRef](#)]
97. Blaschke, T.; Hay, G.J.; Weng, Q.; Resch, B. Collective sensing: Integrating geospatial technologies to understand urban systems—An overview. *Remote Sens.* **2011**, *3*, 1743–1776. [[CrossRef](#)]
98. Calabrese, F.; Colonna, M.; Lovisolo, P.; Parata, D.; Ratti, C. Real-Time Urban Monitoring Using Cell Phones: A Case Study in Rome. *IEEE Trans. Intell. Trans. Syst.* **2011**, *12*, 141–151. [[CrossRef](#)]
99. Dicken, P. *Global Shift: Mapping and Changing Contours of the World Economy*; Sage Publications Ltd.: London, UK, 2014.
100. Klotz, M.; Kemper, T.; Geiß, C.; Esch, T.; Taubenböck, H. How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sens. Environ.* **2016**, *178*, 191–212. [[CrossRef](#)]
101. Jing, W.; Yang, Y.; Yue, X.; Zhao, X. Mapping urban areas with integration of DMSP/OLS nighttime light and MODIS data using machine learning techniques. *Remote Sens.* **2015**, *7*, 12419–12439. [[CrossRef](#)]
102. Jiang, B.; Yin, J.; Liu, Q. Zipf’s law for all the natural cities around the world. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 498–522. [[CrossRef](#)]
103. World Bank. 2016. Available online: <http://data.worldbank.org/indicator/EN.POP.DNST> (accessed on 13 November 2017).
104. Short, J.R. The Liquid City of Megalopolis. *Documents d’Anàlisi Geogràfica* **2009**, *55*, 77–90.
105. INSPIRE—Infrastructure for Spatial Information in Europe. Available online: <http://inspire.ec.europa.eu> (accessed on 13 November 2017).
106. Wei, C.; Blaschke, T.; Kazakopoulos, P.; Taubenböck, H.; Tiede, D. Is Spatial Resolution Critical in Urbanization Velocity Analysis? Investigations in the Pearl River Delta. *Remote Sens.* **2017**, *9*, 80. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).