Final manuscript published in:
Schmitt, A., P. Uth, I. Standfuß, B. Heider, S. Siedentop, and H. Taubenböck. 2023. Quantitative assessment and comparison of urban patterns in Germany and the United States. Computers, Environment and Urban Systems 100:101920.
DOI: 10.1016/j.compenvurbsys.2022.101920

Quantitative Assessment and Comparison of Urban Patterns in Germany and the United States

Abstract

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

Keywords: urban patterns, remote sensing, Global Urban Footprint, built-up density, homogeneity, Gaussian Mixture Model.

1. Introduction

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

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

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

Such assessments of convergent and/or divergent development patterns are still – primarily – based on rather qualitative, empirically hardly replicable findings. There are only few international comparative research studies with quantitative, spatial approaches (Mahendra & Seto 2019; Batty et al. 2004; Hall & Pain 2006; Taubenböck et al. 2019, Heider & Siedentop 2020). The widespread absence of international comparative research or the predominance of qualitative-descriptive (and mostly single) studies on the development of urban spatial patterns can on the one hand be attributed to the limited availability of appropriate data across countries. Equivalent data in terms of population or jobs often suffer from inconsistent spatial reference units, challenges regarding reporting requirements, or out-dated data. On the other hand, it can be attributed to methodological challenges of scale, units of measurement, thematic variables used or the applied metrics (e.g. Openshaw 1983; Taubenböck et al. 2019; Bartosiewicsz & Marcinczak 2020). Approaches targeting urban spatial structures vary from supra-regional to local spatial levels, with measurements at administrative city to intra-urban block level and thematic variables such as residents (Garcia López 2010), jobs (Giuliano & Small 1991), companies (Arauzo-Carod & Villadecans-Marsal, 2009), commuter flows (Roca Cladera et al. 2009) or the built urban landscape (Taubenböck et al. 2017). Furthermore, there is a great variety of different methods applied for the identification of urban centres such as density thresholds (Giuliano & Small 1991), locally weighted regression (McMillen 2001), or exploratory spatial data analysis (Arribas-Bel & Sanz-Gracia 2014) and a cornucopia of metrics for the determination of urban form such as rank-size distributions (Arribas-Bel et al. 2015), concentration indices (Heider & Siedentop 2020), density functions (Kim et al. 2014), measures of spatial autocorrelation (Tsai 2005), landscape metrics (Fragkias & Seto 2009) or model-based measurements (Taubenböck et al. 2019). There is no consensus on suitable approaches. Rather it has been shown that results strongly depend on the selected units of analysis or – statistically spoken – the reference unit and the method for the measurement of urban form (e.g. Bartosiewicsz & Marcinczak 2020). Thus, it can be stated that there is a methodological ambiguity as to how urban patterns can be quantified and compared optimally. Against this background, our goal is to present a set of methods that can robustly and comparably measure the settlement and centre structure of an urban region given the same input data in a quantitative manner. Spatial structures define to a not inconsiderable extent how we behave, and accordingly they have an influence on many factors such as traffic volume (e.g., Glaeser 2010), land consumption (e.g., Taubenböck 2021), or economic prosperity (e.g., Bettencourt 2007). To know which structural types occur where on our planet and to be able to relate them in a comparable way thus is crucial.

2. Conceptualization and Research Questions

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

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

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

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

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

1. Does the non-parametric multi-scale heterogeneity approach deliver robust and comparable results independent of the reference unit?
2. Does the parametric hierarchical clustering approach deliver robust and comparable results independent of the reference unit?
3. Do these approaches reveal distinct differences between urban patterns in the United States of America and Germany?

The remainder of this paper is organized as follows. *Section 3* introduces the data basis, the study sites and the spatial concept. *Section 4* introduces the applied methods in detail. In *Section 5*, we apply the parametric and non-parametric methods onto urban patterns of study sites in the U.S. as well as Germany and we present the measured similarities and differences across the regions. This is followed by a discussion in *Section 6*, concluding with a perspective on the implications of the results in *Section 7*.

3. Study design

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

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

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

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

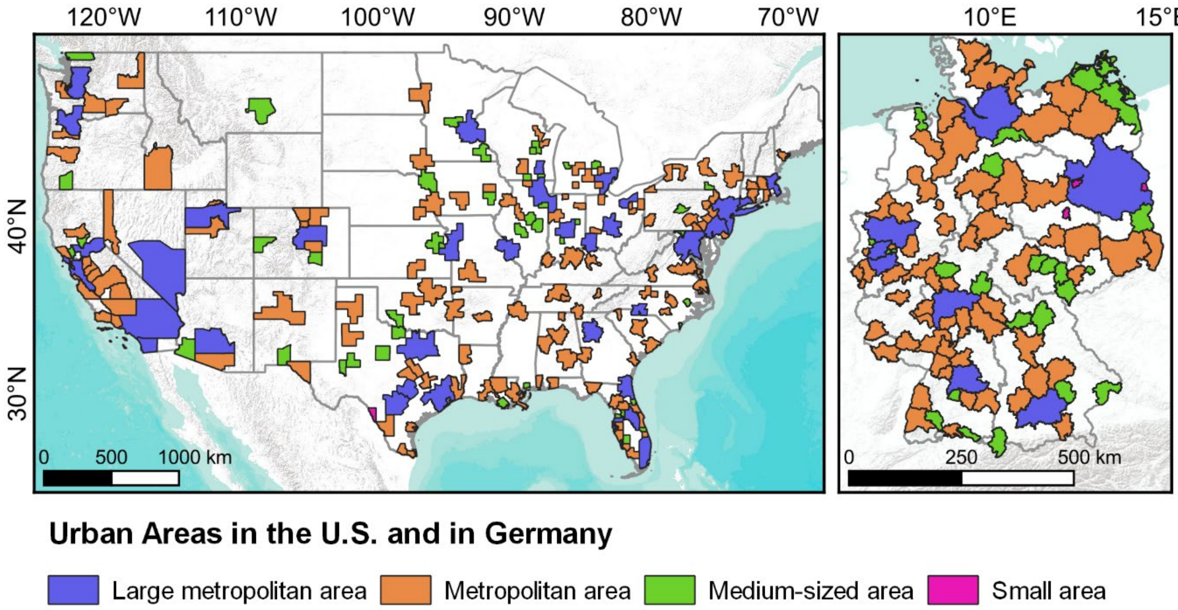


Figure 1 Map of the OECD functional urban areas in the U.S. and Germany

4. Methodology

In this section, we explain the design of the two methodological approaches for measuring urban patterns, i.e., (1) a non-parametric from discrete, model-free statistics and (2) a parametric approach from continuous, model-based statistics. Furthermore, we introduce the geostatistical evaluation of the results. It is well-known that the scale is playing a key role in the spatial analysis (Schmitt et al. 2014, Schmitt et al. 2018). In the following, we consider variations in the scale by the size of patches around one settlement pixel of interest that is included in the calculation of the desired feature. It is typically referred to as level of the Gaussian image pyramid, cf. Figure 2 and Figure 3. The finest scale (smallest patch) of the pyramid starts with one isolated pixel and increases by factor $\sqrt[4]{2}$ (in our implementation) until the coarsest scale (largest patch) is reached, which resembles almost the whole image. The respective patch area in square kilometres can be calculated by the help of the ground sampling distance which slightly varies from image to image according to the latitude of the city in the geographic reference system used.

4.1. Non-parametric approach

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

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

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

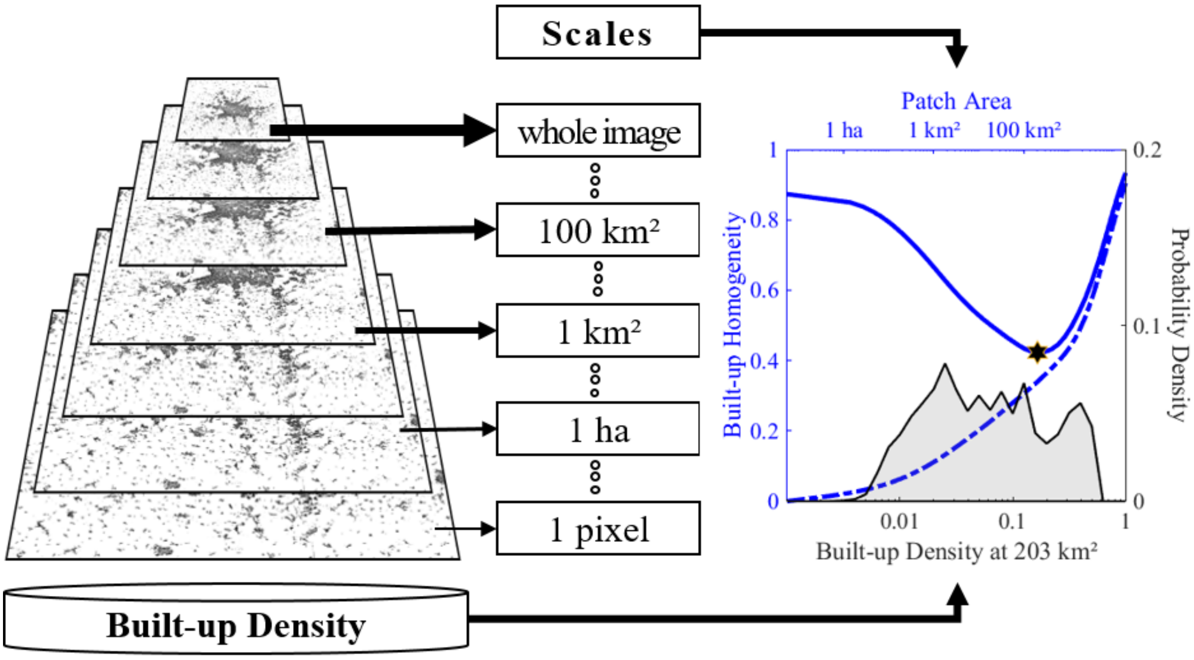


Figure 2 Workflow of the non-parametric approach: multi-scale built-up homogeneity around settled (solid) and non-settled (dash-dotted) areas as well as the density distribution at the scale of minimum homogeneity (star) which is denoted numerically in the bottom label by an area of 203 km² in the visualized case. The blue lines show the homogeneity of built-up areas around built-up pixels (solid line) and around non-built-up pixels (dash-dotted line). The latter indicates if natural areas are clearly separated from settlements (low values) or commonly mixed-up with settlement nearby (high values).

4.2. Parametric approach

For the distinction of individual settlement clusters, we apply a parametric approach. This approach allows us to estimate the centre of clusters by their local modes (coarse, discrete, but robust towards outliers) and gravity centres (fine and continuous via least square estimation) as well as their size by their torques (error ellipses from highly sensitive second order statistics). To do so, we reconstruct the urban landscape as GMM in a hierarchical multi-scale clustering process, which is visualized in Figure 3. The basic idea of the multi-scale hierarchical clustering refers to the Multi-Scale Multi-Looking for the radiometric enhancement of radar intensity images (Schmitt et al. 2014). This approach also runs through the image pyramid from coarser to finer scales and selects the appropriate scale for the local image content. The selection of the appropriate scale is now replaced by the adaptation of a local Gaussian distribution. While the Multi-scale Multi-Looking only knows about circular filter kernels, its multi-directional advancement – called Schmittlets – also takes elongated structures in different directions into account (Schmitt 2016). Instead of searching for Gaussian or TANH-distributions based on a predefined set of parameters as done by the Schmittlets, the approach presented here individually estimates the parameters of the two-dimensional Gaussian Distribution, i.e., mean vector and covariance matrix. Hence, the image is described by a GMM consisting of a varying number of centres with an individual location, extension and orientation. Additionally, each pixel has a certain probability with respect to each Gaussian distribution, i.e., a certain degree of belonging to the respective cluster. In this way, each pixel can be attributed to the cluster showing the highest probability similar to the Maximum Likelihood method. Furthermore, the total probability of belonging to any urban cluster instead of being located in a rural environment, i.e. the probability provided by the resulting GMM, is given by the summed-up probabilities of all clusters found in the image subset.

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

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

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

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

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

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

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

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

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

All operations are performed as simple convolutions in spatial domain using the multi-scale and multi-directional filter bank described above and visualized in Figure 4. The element-wise division by the convolution with the 0th order kernel leads to a weighted mean with the density values as weights. The correction of the variance estimation by the reduced redundancy because the mean value is estimated from the same sample can be regarded as negligible due to the immense number of measurements in one patch. The Gaussian bell curve N of one urban centre then unfolds to

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

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

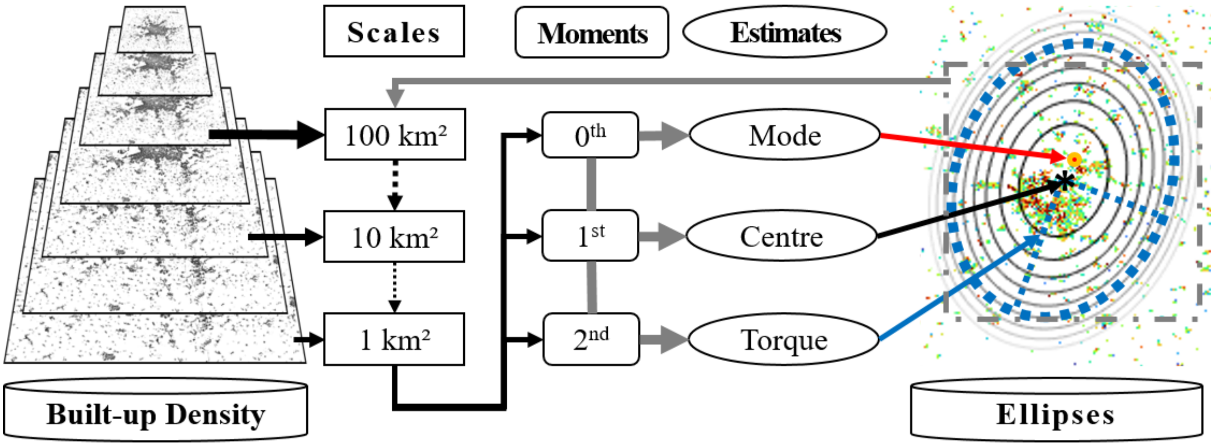


Figure 3 Workflow of the parametric approach: from coarser to finer scales, the local mode, gravity centre and torque are calculated in order to model a two-dimensional Gaussian distribution via least square estimation and to assign the single pixels to a specific cluster. For each image a certain number of ellipses (blue dotted closed line) is derived that are described by their centres, their major and minor half axes (blue dotted straight lines), and their orientation in the coordinate system.

Centre and torque are the common parameters to describe a two-dimensional Gaussian Distribution with its typical error ellipse. All pixels showing more than 1 ppm probability of belonging to the detected cluster are assigned to the this cluster and removed from the further processing. If no more clusters with their mode greater than 10% density are found, the algorithm switches to finer scales until a grid of 5 km by 5 km is reached, i.e., agglomerations covering less than 25 km² are not considered in this study in order to simplify the further evaluation and to not become too fragmented. The choice of the minimum size will be justified by the results of the non-parametric analysis: the least average minimum homogeneity is found at a patch size of 26 km² in FUAs of Germany (Table 1). Doing so, the algorithm gradually approximates the parameters of the Gaussian distribution in a hierarchical (top-down the image pyramid) and robust (from mode via gravity centre to torque) way. For each image, consequently a varying number of centres with varying sizes, extents and orientations is estimated. Furthermore, each pixel can now be assigned to the cluster centre showing the highest probability. The sum of all probabilities per pixel denotes the probability of belonging to an urban agglomeration instead of being located in a rural landscape. The summed-up probability is plotted as contour lines with 0.1, 0.01, 0.001, etc. probability in the corresponding maps, see Figure 5 and Figure 6 for instance. Overlays of several urban clusters are possible, but of low influence because clusters in close vicinity are observed quite rare.

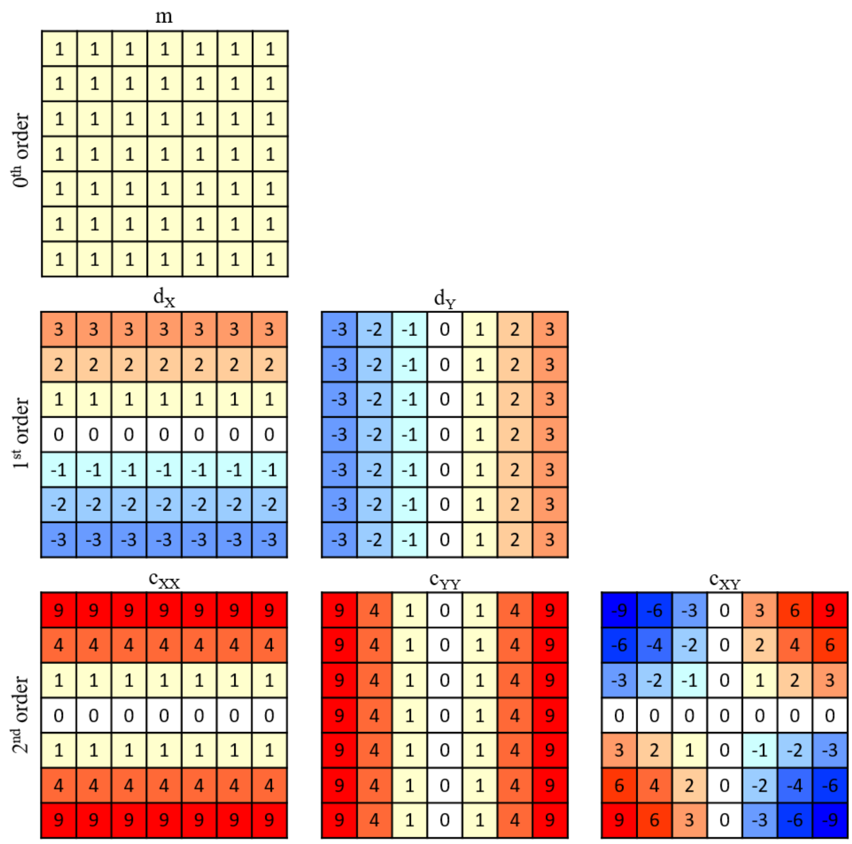


Figure 4 Exemplary filter kernels of the fourth scale covering a 7 x 7 pixel neighbourhood as used in the parametric approach. The 0th order delivers the total sum which is used as normalization factor on the one hand and whose maximum defines the local mode on the other hand. The 1st order calculates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

4.3. Geographic analysis

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

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

Figure 21 and Figure 22).

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

5. Results

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

5.1. Selected results in Germany for reference unit comparison

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

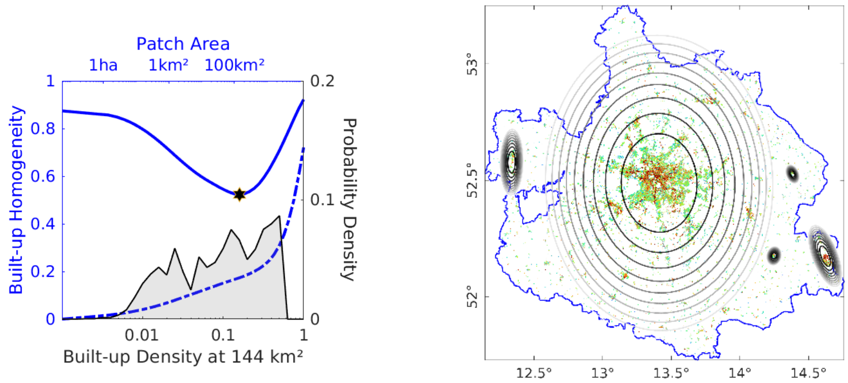


Figure 5 Berlin – functional

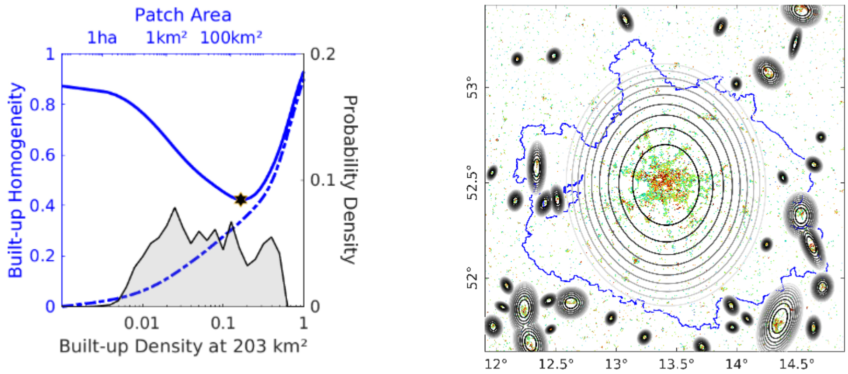
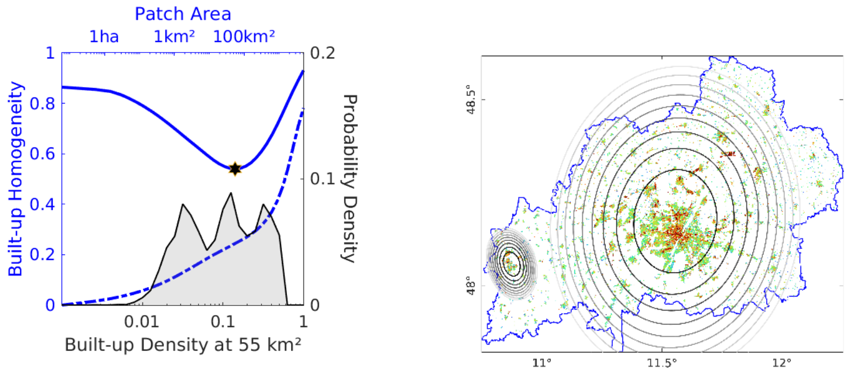
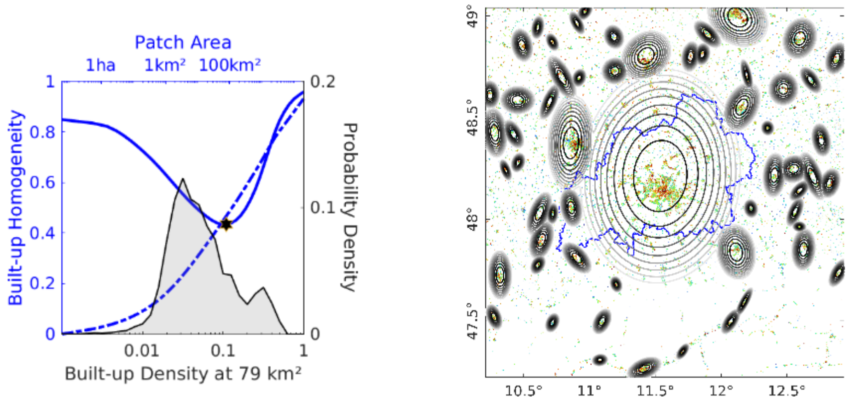


Figure 6 Berlin- standardized





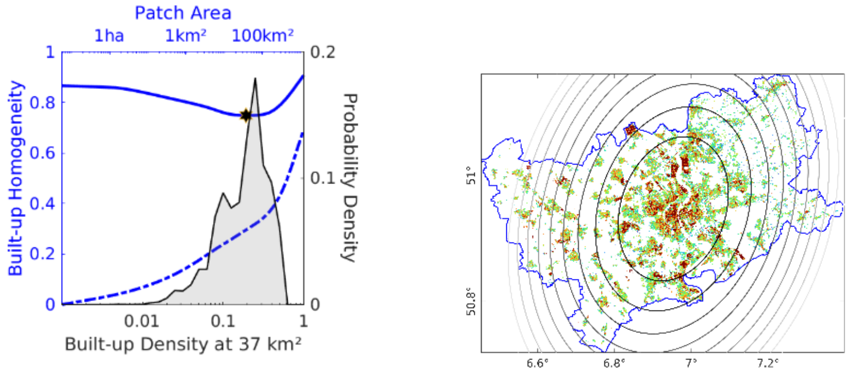


Figure 9 Cologne - functional

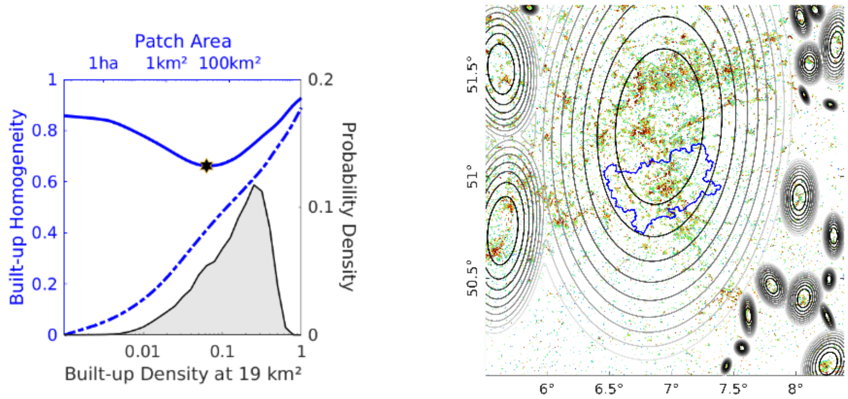
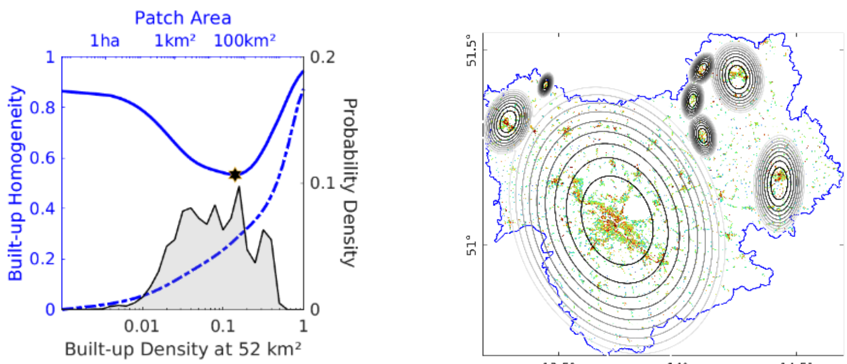
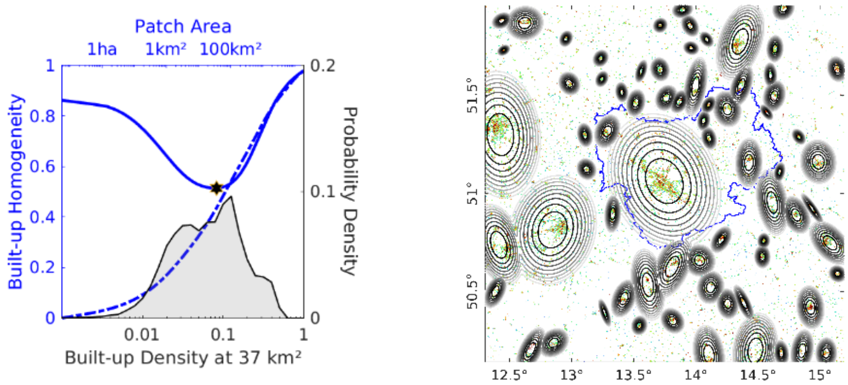


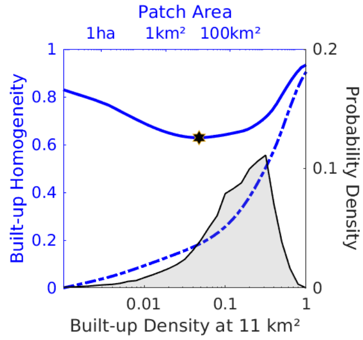
Figure 10 Cologne- standardized

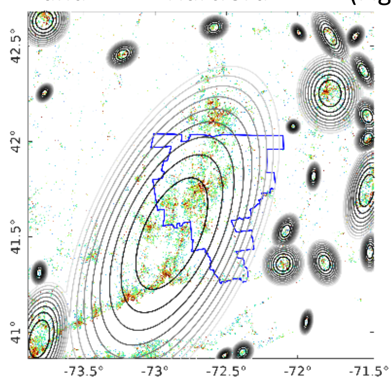


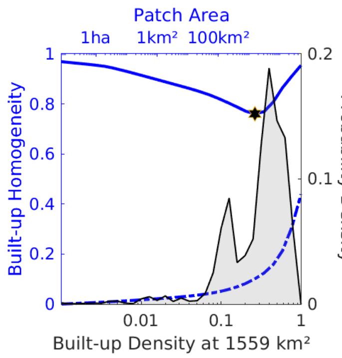


5.2. Selected results in the U.S. for reference unit comparison

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







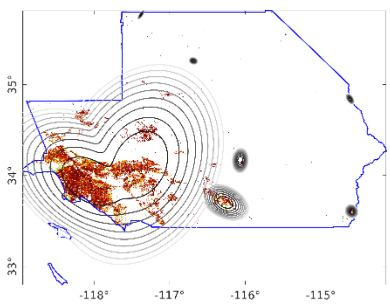


Figure 13 Los Angeles – functional

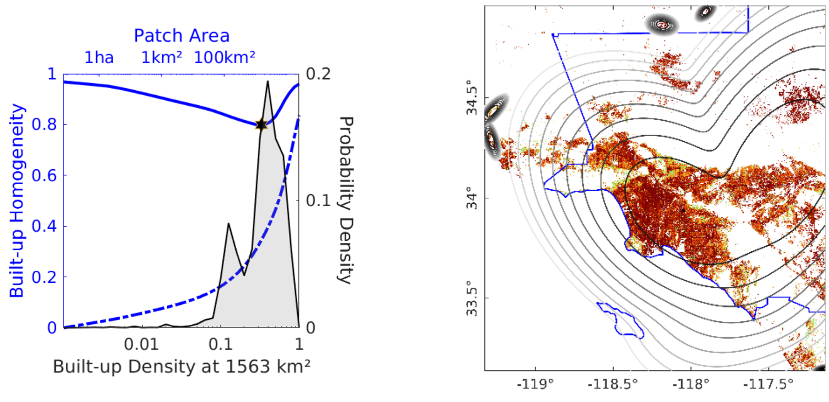


Figure 14 Los Angeles- standardized

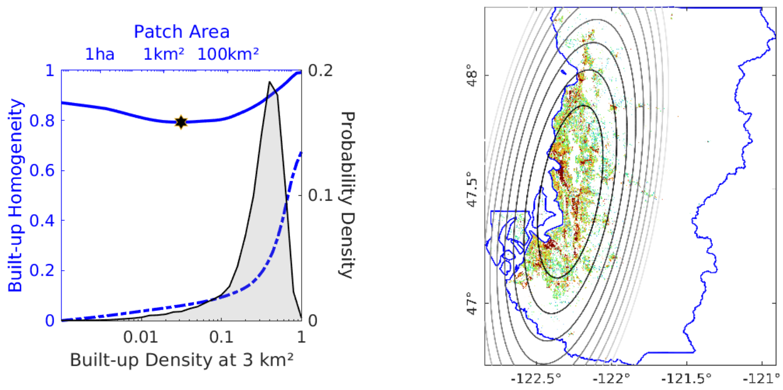
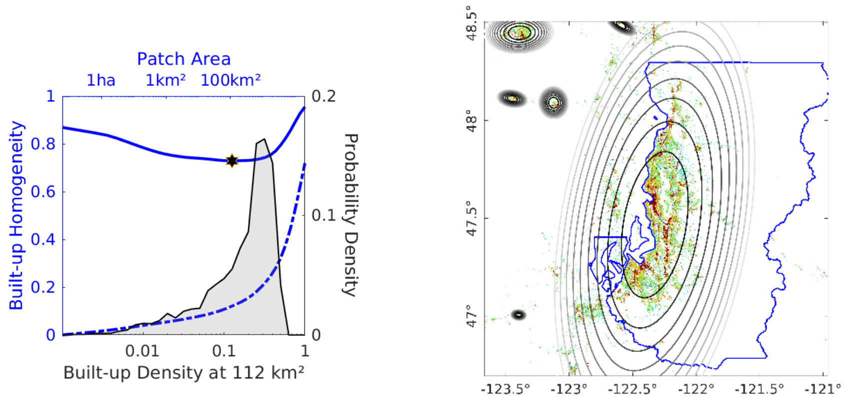


Figure 15 Seattle – functional



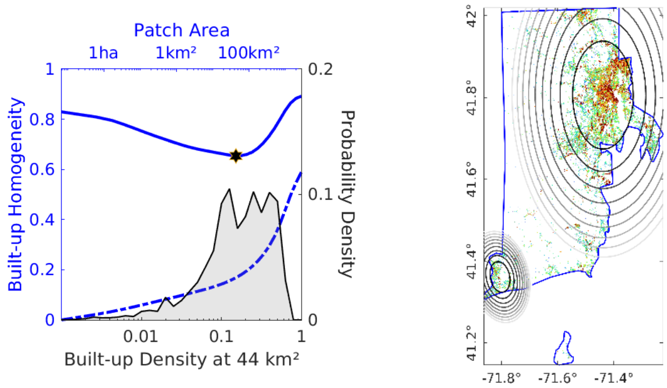


Figure 17 Providence - functional

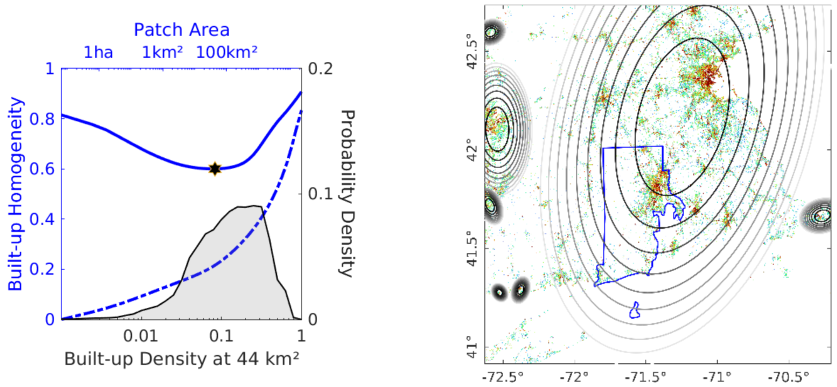
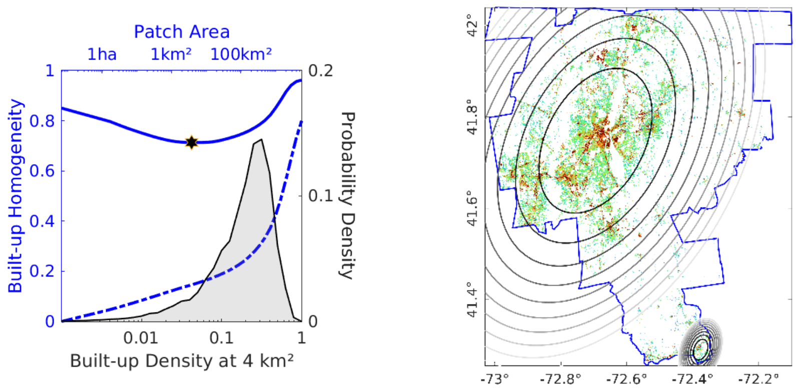
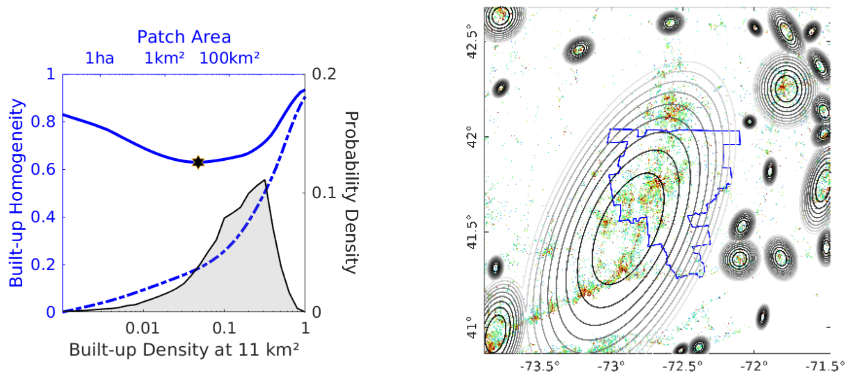


Figure 18 Providence- standardized





5.3. Geographic analysis for intercontinental comparison

To sum up, the figures above indicate that German metropolitan areas are characterized by a steeper minimum of the homogeneity at lower scales than in the U.S. The respective density distribution is almost symmetric while the density distribution of U.S. metropolitan areas shows a high negative skewness. This indicates the polycentric pattern of urbanization in Germany whereas the U.S. urban system is rather characterized by huge conurbations in larger distances to each other. The parametric approach therefore delivers more numerous, but smaller clusters in Germany than in the U.S. In order to proof these assumptions, the results are joined together in one diagram for the functional metropolitan areas in Figure 21 and for the standardized patches in Figure 22.

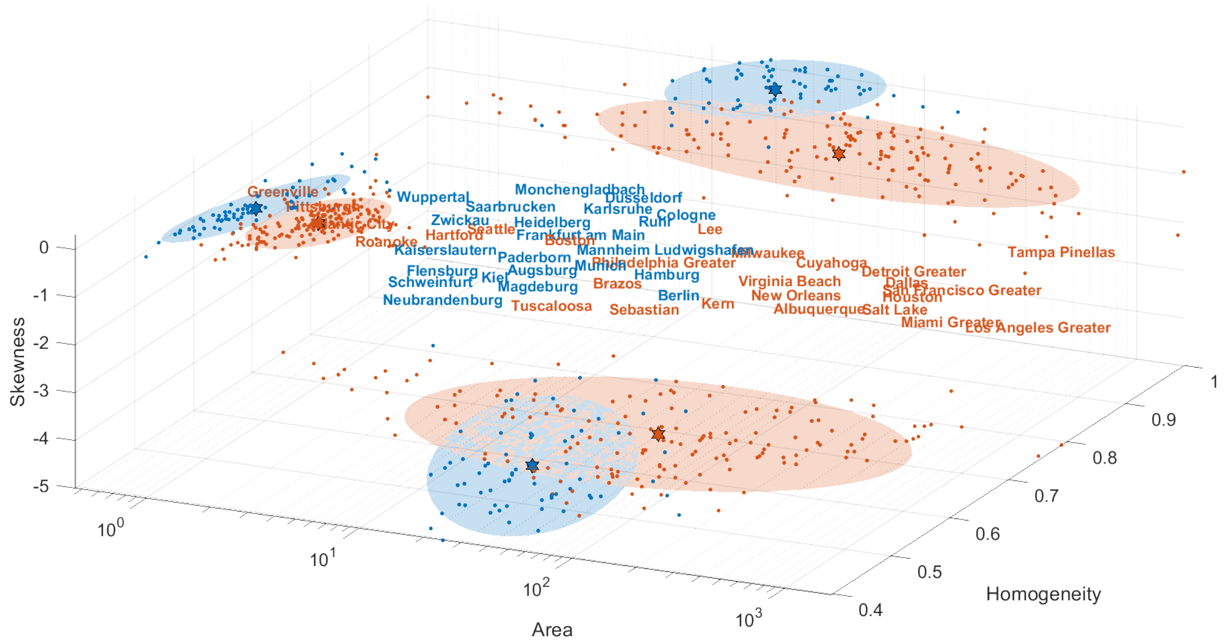


Figure 21 Homogeneity, scale (area), and skewness of the functional units derived by the non-parametric approach: German metropolitan areas are plotted in Blue, U.S. metropolitan areas in Orange. The stars denote the corresponding mean values. The ellipses visualize the theoretical 90%-quantile in the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

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

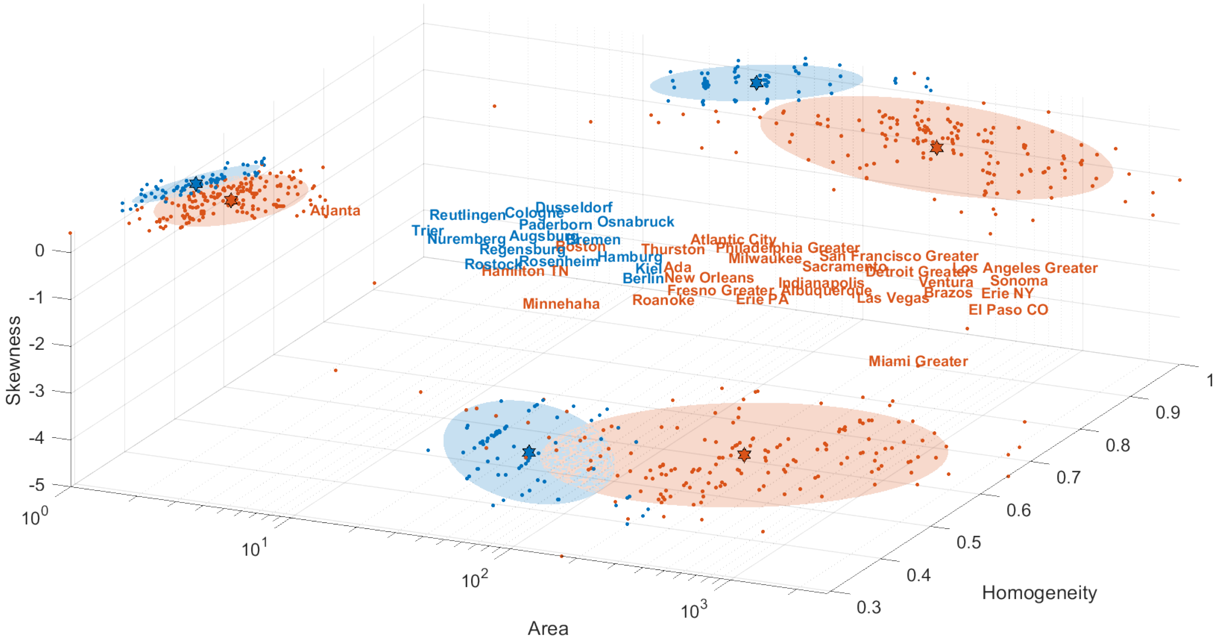


Figure 22 Homogeneity, scale (area), and skewness of the standardized patches derived by the non-parametric approach: German metropolitan areas are plotted in Blue, U.S. metropolitan areas in Orange. The stars denote the corresponding mean values. The ellipses visualize the theoretical 90%-quantile in the respective plane, i.e. the marginal 2D normal distribution. The 3D plot shows only the names of notable cities for better readability, but the points in the marginal 2D distributions represent all cities under study.

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

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

Functional	**Scale [km²]**		**Homogeneity [%]**		**Skewness**	
Germany	26	19	58	56	-0.43	-0.28
U.S.	105	49	69	70	-1.59	-1.53
Standardized	**Scale [km²]**		**Homogeneity [%]**		**Skewness**	
Germany	44	27	54	55	-0.27	-0.31
U.S.	362	243	61	62	-1.11	-1.00
Parameter	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**

The parametric analysis delivers the number of urban centres and the size of the corresponding clusters. For better visualization, the sorted list of clusters by number is presented as area plot in Figure 23 for the functional areas and in Figure 24 for the standardized regions. There is only one dominant cluster in one functional area except for Los Angeles where two clusters of similar size in close vicinity were detected (see Figure 13 and Figure 14). The dominant clusters are much larger in the U.S. with almost 2,000 km² in the case of Los Angeles. The largest dominant clusters in Germany range around 500 km². The number of clusters restricts to twelve at maximum in a few cases whereas the vast majority has less than five clusters both in Germany and in the U.S.

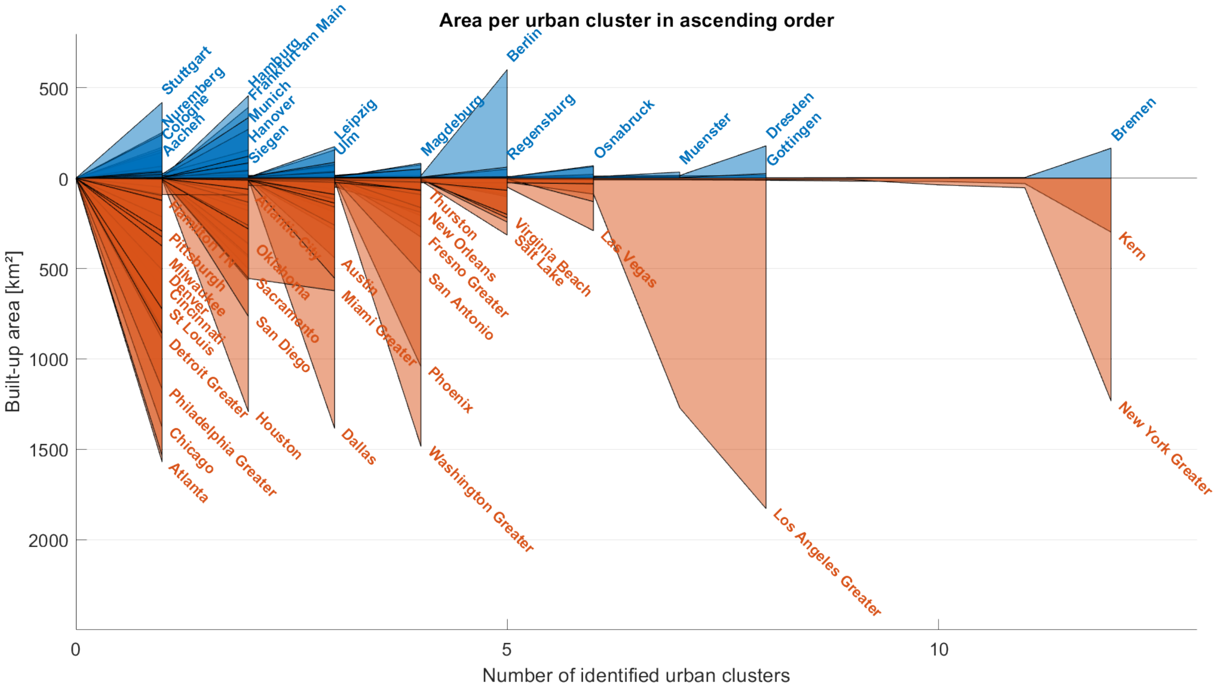


Figure 23 Number of identified clusters and corresponding area of the ellipses in ascending order from the left to the right for the functional areas. German urban areas are marked in Blue upwards, U.S. urban areas in Orange downwards. The plot shows only the names of notable cities for better readability.

The measured values reveal a different result when we switch to standardized areas of interest in Figure 24. The German samples are typically characterized by up to 80 small-sized clusters. Some sample cities in the Ruhr metropolitan area, however, contain only few, but very large clusters that also reach nearly 2,000 km² like the dominant clusters in the U.S. The standardized regions in the U.S. still reveal only one dominant cluster with a large extension in most cases. This underlines again a higher diversity in the German urban environments.

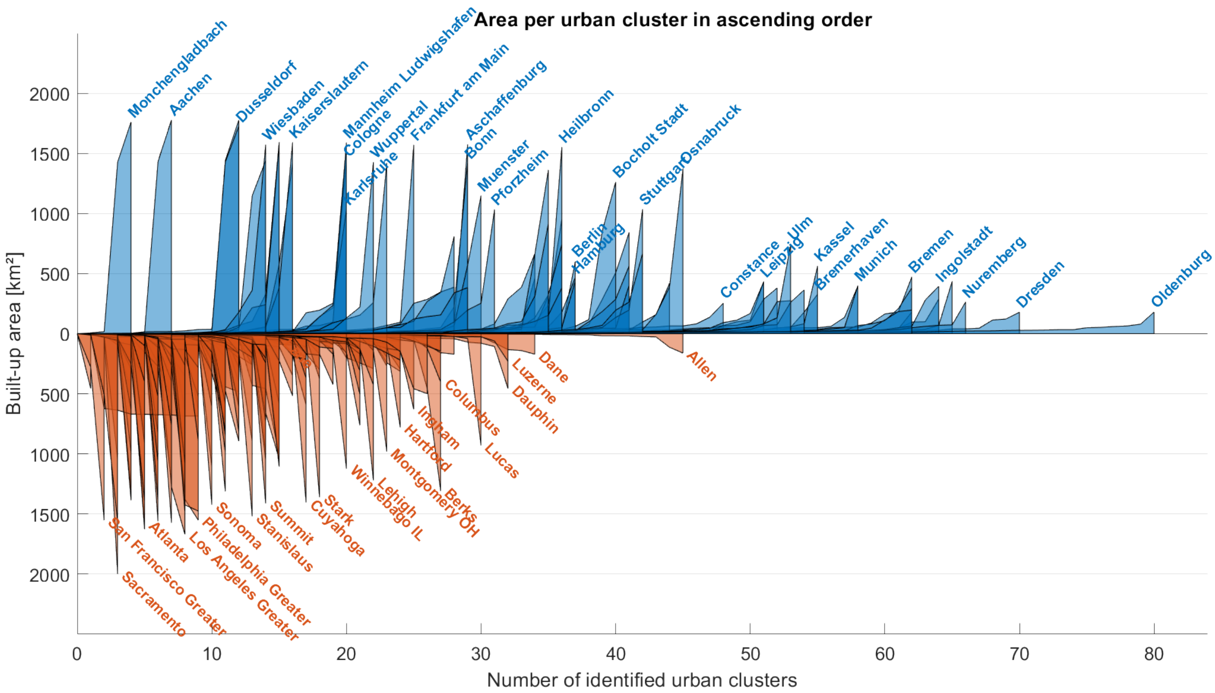


Figure 24 Number of identified clusters and corresponding area of the ellipses in ascending order from the left to the right for the standardized patches. German urban areas are marked in Blue upwards, U.S. urban areas in Orange downwards. The plot shows only the names of notable cities for better readability.

The numerical results of the parametric approach are summed-up in Table 2 and substantiate the above statements. In the functional areas, the expected number of clusters ranges around two in both cases. In the standardized areas, the number rises to around 13 in the U.S. and even to 38 in Germany. The settlement area according to the GMM also rises when switching from functional to standardized areas. Interestingly, the U.S. sites exceed the German sites in the functional areas which indicates that functional areas in Germany are generally smaller. In contrast to that, the settlement areas are almost twice as high in Germany compared to the U.S. which can be referred to the 200 km by 200 km environment which covers mostly rural land in the U.S., but includes many small towns in Germany. The relative standard deviation in the size of the clusters with around 30% is similar for the functional regions both in the U.S. and in Germany. These values reflect the high gradient between the one dominant and the few adjacent clusters. Switching to standardized regions, the relative standard deviation drops down to around 20% for the U.S. and even 10% for Germany. This fact proofs that the standardized areas in Germany are populated with numerous more or less equal small-sized clusters with a mean deviation of size of only 10% whereas the standardized area in the U.S. have a lower number of clusters, but with a higher deviation in size with up to 20%.

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

Functional	**Number**		**Area [km²]**		**Areal Deviation [%]**	
Germany	3	2	113	76	30	32
U.S.	2	2	316	163	30	31
Standardized	**Number**		**Area [km²]**		**Areal Deviation [%]**	
Germany	38	37	1541	1471	10	8
U.S.	14	12	880	714	21	18
Parameter	**Mean**	**Median**	**Mean**	**Median**	**Mean**	**Median**

6. Discussion

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

6.1. Methodological aspects

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

Nevertheless, the choice of two different reference units for the subsequent interpretation points out challenges for geographical comparisons: the functional areas on the one hand and the standardized areas on the other hand, reveal certain, but explainable differences. The fewer number of only two clusters in the functional areas supports the original choice of these regions as catchment areas for workers that commute to the central city more or less daily. Regarding the urban patterns of a whole country, our study shows, that the analysis of individual catchment areas might fall short as each city is embedded in a larger metropolitan territory, at least in Germany. The situation in the U.S. is different because of the long distances in-between neighbouring cities. This becomes obvious in the quite low variation of the expectance values in Table 1 and Table 2 when comparing functional and standardizes regions in contrast to the high variation regarding Germany.

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

6.2. Implementational aspects

The parametric approach requires a certain configuration in order to work efficiently: The minimum patch size is set to 25 km², the minimum density for a cluster is set to 10%, and the minimum probability of belonging to the considered cluster is set to 1 ppm. In order to justify the choice of these three variables, the non-parametric approach is consulted: the mean scale of maximum heterogeneity is always above 25 km², see Table 1. The density threshold of 10% regarding the mode results from the typical site occupancy index in Germany of 20%. Having in mind, that U.S. settlements might be less dense, we reduced this value to the half. The probability threshold is chosen in order to capture almost the whole set of built-up pixels belonging to the respective cluster. We are aware that there is some subjectivity here in the calibration. However, we argue that this is not the crucial point: what is crucial is that it has been carried out consistently for all cities and city regions and thus, becomes comparable. Furthermore, the plausibility of the results confirms that this calibration is appropriate.

The above considerations consequently lead to following kernel sizes in the image pyramid: 1, 3, 5, 7, 9, 11, 13, 15, 19, 21, 25, 31, 37, 45, 53, 63, 75, 89, 107, 127, 151 for the non-parametric approach exclusively, and 181, 215, 255, 303, 361, 429, 511, 607, 723, 861, 1023, 1217, 1447, 1721, 2047, 2435, 2895, 3443 for both approaches. The parametric approach starts at a pyramid level with a patch size of 181 by 181 pixels, which corresponds to an areal coverage of about 20 km² as already explained. Though the number of pixels per dimension is typically doubled with each level of the pyramid, we decided in favour of a much finer sampling of about 1.2 instead of factor 2, which results in four intermediate levels until the number of the pixel per dimension is doubled and still two intermediate levels until the covered area is doubled. Additionally, it is necessary to have an odd integer number of pixels in the patch for implementation reasons. Therefore, several adjacent levels of the image pyramid coincide for small patch sizes below ten pixels. Regarding higher pyramid levels, each level has its own patch size that diverges less than one pixel in each direction from the intended size. The fine scale sampling ensures very stable results along the scale dimension, which is clearly visible in the diagrams (e.g., Figure 18) because the homogeneity curves are very smooth even without any further smoothing algorithm applied for display purposes (e.g.).

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

6.3. Statistical Aspects

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

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

6.4. Geographical aspects

In general, there is a widespread agreement in the urban geographic discourse that American urban regions are characterized by a more polycentric-dispersed structure than European ones (Heider & Siedentop, 2020; Krehl 2016; Guillain et al. 2006; Riguelle et al. 2007). These findings, however, mostly refer to spatial units similar to the functional urban regions considered in our study, i.e., more sub-centres (or "edge cities") are measured within the urban catchment area, and a loss of importance of the core city as an employment hub. Based on our input data, however, referring to centres as density clusters of settlements without a differentiation based on urban functions, jobs or the like, the quantity of centres and subcentres does not differ systematically between Germany and the USA. What is different, however, is that identified centres in the USA are much larger. However, we must acknowledge that the centres identified by the parametric approach can hardly be compared to the centres and subcentres in the above-mentioned literature. Thus, our method seems to be better suited for the analysis of settlement structures at a larger spatial scale such as inter-urban polycentricity (Meijers 2008) than for the measurement of intra-urban polycentricity. When considering the larger, more regional spatial units of 200 km x 200 km around each core city, we measured very different urban patterns between U.S. and German regions. These regional urban patterns in the U.S. have a much stronger dominance of the core cities (including their extended suburban fringe). These have, on the one hand, a larger extent and, on the other hand, fewer sub-centres in the surrounding area than it is measured for German regions. This may be reflective of the much denser settlement system in Germany, which is characterized by a large number of small and medium-sized cities outside of the metropolitan areas. This again shows how vulnerable spatial statistics are due to their specific approaches such as spatial extent, input data, methods, etc. (Thomas et al. 2022). Another key finding of this study is that the homogeneity of American cities is measured as much higher. In order to reach the entire urbanized territory, one has to travel much longer distances in the U.S. cities than in German cities. This finding can be interpreted as the result of different planning cultures and trajectories of urban growth. In the U.S., unregulated sprawl has been the dominant form of urban development, resulting in large and homogenous low-density suburbs and long commuting distances. In stark contrast, urban growth in Germany has always been strongly regulated by land-use planning and a hierarchical system of so-called “central places”, which aims to ensure equal living conditions in all parts of the country (Schmidt & Buehler 2007).

While our study does not treat the multi-temporal aspect, but focuses on the comparison of urban patterns across space at one moment in time, it does not add to the debates about the "Americanization" of European or the "Europeanization" of American urban regions. However, for the specific case of our large spatial units for measurement, at least, it can be stated that the urbanized landscapes around the respective core cities as of today differ significantly between Germany and the U.S.

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

6.5. Future aspects

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

A more methodological refinement is the extension from the two-dimensional image space (with latitude and longitude) to a multi-dimensional feature space. Both, the parametric and non-parametric approach can cope with n-dimensional feature spaces in principle. With an increasing number of features, the number of indistinct features unfortunately increases likewise. So far, the approaches only consider concise descriptors. One solution might be the orthogonal transformation of the feature space to hyper-complex bases in order to select distinct features (Schmitt et al. 2020) or the use of machine learning methods that are known to be resilient towards outliers and indistinct features. Having solved this problem, both approaches could also be utilized for the data-driven clustering in high-dimensional feature spaces, e.g., for image classification and object recognition.

7. Conclusion

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

1. Yes, the multi-scale homogeneity reveals as robust measure to describe urban environments. The influence of the varying reference units in the interpretation is visible, but explainable, and does not impede the inter-continental comparison of the results. The scale of minimum homogeneity and therewith, maximum heterogeneity seems to be an essential variable in the description of spatial proximity. It denotes the spatial distance or the size of the patch that contains the most heterogeneous built-up density.
2. Yes, the decomposition into single Gaussian bell curves is reproducible, even if parts of the respective settlement agglomeration are missing in the actual sample. The method is able to clearly detect density clusters and join neighbouring, but adjacent clusters to one thanks to its hierarchical multi-scale approach that applies step-wise coordinate refinement without any further time-consuming iteration step. The procedure is purely deterministic and thus, easily transferable to other geographical locations and to other, potentially higher-dimensional feature spaces as well.
3. Yes, there are distinct and replicable differences in the urban patterns of the U.S. and Germany independent of the chosen reference unit. Some expectations were confirmed like the large extension of U.S. metropolitan areas or the small-scale urban structures in Germany. Other findings like the scale of the maximum heterogeneity – that also stands for a high quality of living – are completely new. We found a significantly higher heterogeneity in a much smaller

environment in Germany. Furthermore, German metropolitan areas are characterized by numerous small, more or less equal-sized clusters whereas U.S. metropolitan areas contain one large and dominant centre and only few small adjacent urban clusters.

This study provides the methodological basis for a nation-wide, full-coverage comparison between Germany and the U.S. at first and a continental-wide comparison between Europe and North America at second. As the GUF-DenS – the only database needed for this study – assures a global coverage, these two approaches enable even a worldwide comparative study. Including additional data sources like infrastructure or census data will refine the results for dedicated regions where this kind of geo data is available. The fused evaluation of diverse data sets will probably lead to a new perception of spatial proximity in urban environments at local, national, continental, and global scale.

References

Anas, A., Arnott, R. J., & Small, K. A. (1998): Urban Spatial Structure. Journal of Economic Literature, 36(3), 1426–1464.

Angel, S., & Blei, A. M. (2016): The spatial structure of American cities: The great majority of workplaces are no longer in CBDs, employment sub-centers, or live-work communities. Cities, 51, 21–35.

Angel, S., Parent, J., Civco, D., & Blei, A. M. (2011): Making room for a planet of cities. Policy Focus Report. Cambridge, MA: Lincoln Institute of Land Policy.

Arauzo-Carod, J.M. & Villadecans-Marsal, E. (2009): Industrial Location at the Intra-Metropolitan Level: The Role of Agglomeration Economies. Regional Studies, 43(4), 545-558.

Arribas-Bel, D., Ramos, A., Sanz-Gracia, F. (2015): The Size Distribution of Employment Centers within the US Metropolitan Areas. Environment and Planning B: Urban Analytics and City Science, 42(1), 23-39.

Arribas-Bel, D., & Sanz-Gracia, F. (2014): The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. Urban Geography, 35(7), 980-997.

Bartosiewicz, B., & Marcińczak, S. (2020): Investigating polycentric urban regions: Different measures – Different results. Cities, 105, 102855.

Batty, M., Besussi, E., Maat, K., & Harts, J. J. (2004): Representing Multifunctional Cities: Density and Diversity in Space and Time. Built Environment, 30(4), 324–337.

Bertaud, A., & Richardson, H. W. (2004): Transit and Density: Atlanta, the United States and Western Europe. In H. W. Richardson & C.-H. C. Bae (Hrsg.), Sprawl in Western Europe and the United States (S. 293–310). Aldershot, UK: Ashgate.

Bettencourt, L. M. A. (2007). The origins of scaling in cities. Science, 340, 1438–1441.

Bontje, M., & Burdack, J. (2005): Edge Cities, European-style: Examples from Paris and the Randstad. Cities, 22(4), 317–330.

United States Census Bureau (2021): https://www.census.gov/en.html

DESTATIS (2020): ZENSUS2011 - Bevölkerungs- und Wohnungszahen 2011. https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html

Dijkstra L, Poelman H & Veneri P (2019): The EU-OECD definition of a functional urban area. OECD Regional Development Working Papers, N2019/11, Paris, OECD Publishing. https://doi.org/10.1787/d58cb34d-en

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. vol 6., issue 1, 21 pages.

Esch T, Marconcini M, Felbier A, Roth A, Heldens W, Huber M, Schwinger M, Taubenböck H, Müller A & Dech S (2013): Urban footprint processor – Fully automated processing chain generating settlement masks from global data of the TanDEM-X mission. IEEE Geoscience and Remote Sensing Letters. Vol 10, no. 6, November 2013, pp. 1617-1621.

Esch, T., Üreyen, S., Zeidler, J., Hirner, A., Asamer, H., Metz-Marconcini, A., Tum, M., Boettcher, M., Kuchar, S., Svaton, V., Marconcini, M., 2018. Exploiting big earth data from space – first experiences with the TimeScan processing chain. Big Earth Data. ttps://doi.org/10.1080/20964471.2018.1433790. (ISSN: 2096-4471 (Print) 2574-5417 (Online).

Feicht T. & Schmitt, A. (2022): Anamorphic Visualization of the Space-Time-Paradigm by Geodetic Network Adjustment. KN – Journal of Cartography and Geographic Information. Accepted.

Fragkias, M., and K. Seto. (2009): “Evolving Rank-Size Distributions of Intra-Metropolitan Urban Clusters in South China.” Computers, Environment and Urban Systems 33 (3): 189–199.

GADM (2018): Database of Global Administrative Areas. GADM, Montreal, QC, Cannada.

Garcia-López, M.-À. (2010): Population suburbanization in Barcelona, 1991–2005: Is its spatial structure changing? Journal of Housing Economics, 19(2), 119–132.

Garcia-López, M.-À., & Muñiz, I. (2010): Employment Decentralisation: Polycentricity or Scatteration? The Case of Barcelona. Urban Studies, 47(14), 3035–3056.

Garreau, J. (1992): Edge city: Life on the new frontier (1st ed.). New York: Anchor Books.

Giuliano, G., & Small, K. (1991): Subcenters in the Los Angeles region. Regional Science and Urban Economics, 21 (2), 163-182.

Glaeser, E. (2010): Triumph of the city. London: Penguin Press HC 352 pages.

Gordon, P., & Cox, W. (2012): Cities in Western Europe and the United States: Do Policy Differences Matter? The Annals of Regional Science, 48(2), 565-594.

Guillain, R., Le Gallo, J., & Boiteux-Orain, C. (2006): Changes in Spatial and Sectoral Patterns of Employment in Ile-de-France, 1978-97. Urban Studies, 43(11), 2075–2098.

Hall, P., & Pain, K. (Hrsg.) (2006): The polycentric metropolis. Learning from mega-city regions in Europe. London: Earthscan.

Häußermann, H. (1997): Amerikanisierung der deutschen Städte – Divergenz und Konvergenz. In CENTRUM. Jahrbuch für Architektur und Stadt 1997-1998 (S. 92–96). Wiesbaden: Vieweg.

Heider B., & Siedentop S. (2020): Employment suburbanization in the 21st century: A comparison of German and US city regions. Cities. Volume 104, September 2020, 102802

Kim, S. (2007): Changes in the Nature of Urban Spatial Structure in the United States, 1890-2000. Journal of Regional Science, 47(2), 273–287.

Kim JI, Yeo CH, Kwon J (2014): Spatial change in urban employment distribution in Seoul metropolitan city: Clustering, dispersion and general dispersion. International Journal of Urban Sciences 18: 355–372

Krehl, A., & Siedentop, S. (2018): Towards a Typology of Urban Centers and Subcenters—Evidence from German City Regions. Urban Geography.

Krehl, A. (2016): Urban subcentres in German city regions: identification, understanding, comparison. Papers in Regional Science, online first. doi:10.1111/pirs.12235

Le Galès, P., & Zagrodzki, M. (2006): Cities are back in town: the US/Europe comparison. Cahier Européen numéro 05/06 du Pôle Ville/métropolis/cosmopolis. Paris: Centre d'Etudes Européennes de Sciences Po.

Mahendra, A., & Seto, K. C. (2019): Upward and outward growth: Managing urban expansion for more equitable cities in the Global South. Washington, DC, World Resources Institute.

McMillen, D. (2001): Nonparametric Employment Subcenter Identification. Journal of Urban Economics, 50(3), 448-473.

Meijers, E. J. (2008): Measuring Polycentricity and its Promises. European Planning Studies, 16(9): 1313-1323.

Meijers, E. J. (2013): Metropolitan Labor Productivity and Urban Spatial Structure: A Comparison of U.S. Monocentric and Polycentric Metropolitan Areas. In J. Klaesson, B. Johansson, & C. Karlsson (Hrsg.), Advances in spatial sci-ence. Metropolitan regions. Knowledge infrastructures of the global economy (S. 141–166). Berlin, New York: Springer.

Riguelle, F., Thomas, I., & Verhetsel, A. (2007): Measuring urban polycentrism: A European case study and its implications. Journal of Economic Geography, 7(2), 193–215.

Roca Cladera J, Marmolejo Duarte CR, Moix M (2009): Urban structure and polycentrism: Towards a redefinition of the sub-centre concept. Urban Studies 46: 2841–2868

Romero, V., Solís, E., & Ureña, J. M. de. (2014): Beyond the metropolis: new employment centers and historic administrative cities in the Madrid global city region. Urban Geography, 35(6), 889–915.

Schmidt, S.; Buehler, R. (2007): The Planning Process in the US and Germany: A Comparative Analysis. International Planning Studies 12(1): 55-75. DOI: 10.1080/13563470701346592

Schmitt, A.; Wendleder, A.; Kleynmans, R.; Hell, M.; Roth, A.; Hinz, S. (2020): Multi-Source and Multi-Temporal Image Fusion on Hypercomplex Bases. Remote Sens. 2020, 12, 943.

Schmitt, A. , Sieg, T., Wurm, M. & Taubenböck, H. (2018): Investigation on the separability of slums by multi-aspect TerraSAR-X dual-co-polarized high resolution spotlight images based on the multi-scale evaluation of local distributions, International Journal of Applied Earth Observation and Geoinformation, Volume 64, February 2018, Pages 181-198, ISSN 0303-2434.

Schmitt,A. (2016): Multiscale and Multidirectional Multilooking for SAR Image Enhancement, IEEE Transactions on Geoscience and Remote Sensing 54(9): 5117-5134, DOI: 10.1109/TGRS.2016.2555624.

Schmitt, A., Wendleder, A. & Hinz, S. (2014): The Kennaugh element framework for multi-scale, multi-polarized, multi-temporal and multi-frequency SAR image preparation, ISPRS Journal of Photogrammetry and Remote Sensing 102: 122-139, DOI: 10.1016/j.isprsjprs.2015.01.007.

Schollerer, L., Schmitt A., Wendleder, A. & Rogginger, S. (2022): Regular Investigation of Construction Activities from Space based on Open Satellite Imagery [German title: Schritthaltende Baufallerkundung aus dem All mit frei verfügbaren Satellitenaufnahmen], zfv, DVW, 3/2022, 168-180. Wißner-Verlag, DOI 10.12902/zfv-0391-2022.

Shearmur, R., & Alvergne, C. (2003): Regional Planning Policy and the Location of Employment in the Ile-De-France: Does Policy Matter? Urban Affairs Review 39(1), 3-31.

Shearmur, R., & Coffey, W.J. (2002): A Tale of Four Cities: Intrametropolitan Employment Distribution in Toronto, Montreal, Vancouver, and Ottawa – Hull, 1981 – 1996. Environment and Planning A: Economy and Space, 34(4), 575-598.

Siedentop, S., Krehl, A., Guth, D., & Holz-Rau, C. (2016): Morphologische Polyzentralität der Beschäftigung in deutschen Metropolregionen – Aktuelle Befunde und Veränderungen seit 1970. In R. Danielzyk, A. Münter, & T. Wiech-mann (Hrsg.), Planungswissenschaftliche Studien zu Raumordnung und Regionalentwicklung: Vol. 5. Polyzentrale Metropolregionen (S. 45–75). Lemgo: Verlag Dorothea Rohn.

Siedentop, S., & Fina, S. (2012): Who sprawls most? Exploring the patterns of urban growth across 26 European countries. Environment and Planning A, 44(11), 2765-2784.

Taubenböck H, Klotz M, Wurm M, Schmieder J, Wagner B, Wooster M, Esch T & Dech S (2013): Delineation of Central Business Districts in mega city regions using remotely sensed data. Remote Sensing of Environment, 136, 386-401.

Taubenböck H, Standfuß I, Wurm M, Krehl A & Siedentop S (2017): Measuring morphologic polycentricity – A comparative analysis of urban mass concentrations using remote sensing data. Computers, Environment and Urban Systems, vol. 64, pp 42-56.

Taubenböck H, Gerten C, Rusche K, Siedentop S & Wurm M (2019): Patterns of Eastern European Urbanisation in the Mirror of Western Trends – Convergent, Unique or Hybrid? Environment and Planning B: Urban Analytics and City Science. vol. 46(7) pp. 1206-1225.

Taubenböck H (2021): How we live and what that means – A character study with data from space. In: International Geoscience and Remote Sensing Symposium (IGARSS) 12.-16. July 2021, Brussels, Belgium.

Thomas R, Schmidt S, Siedentop S (2021): Toward comparative polycentricity scores: Assessing variations in regional delineation and subcenter identification. Environment and Planning B: Urban Analytics and City Science, online first, December 30, 2021

Tsai, Yu-Hsin (2005): Quantifying Urban Form: Compactness versus 'Sprawl'. Urban Studies, 42(1), 141-161.

Uth, P. (2019): Development of a globally suitable estimation approach for the fully automated determination of city centres and borders applied to Germany and the USA. Master's thesis. Language German. Geoinformatics department, Munich University of Applied Sciences, Munich, Germany.

Appendix

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

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