JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG (JMU)

Institute for Geography and Geology

Department for Remote Sensing, EAGLE Graduate Program

DEUTSCHES ZENTRUM FÜR LUFT UND RAUMFAHRT (DLR)

Deutsches Fernerkundungsdatenzentrum (DFD)

Team Stadt und Gesellschaft

Master Thesis

Analyzing the relationship between urban morphology and economic subcenters with a focus on urban polycentricity using remote sensing and socioeconomic data

Submitted by:

Johannes Mast

Date:

25.05.2021

supervised and

evaluated by

Dr. Hannes Taubenböck (DLR)

Dr. Martin Wegmann (JMU)

Abstract

Polycentricity refers to the existence of more than one center in urban conurbations. While it is a fuzzy concept with no clear definition, it is often approached from a socio-economic perspective using information on the spatial distribution of jobs derived from employment data (EMP). The heterogeneity and shortage of such EMP data is a challenge to the study of polycentricity. Remote sensing of urban morphology could provide a globally available surrogate in the form of TanDEM-X-derived urban mass concentrations (UMC). However, the extent to which UMC are suited to substitute EMP has not yet been systematically analyzed.

To fill this gap, I detect and compare UMC-centers and EMP-centers in four city regions in the United States. I consider the EMP-based centers as a baseline and analyse the degree to which the UMC-centers are congruent with them. In a threefold analysis I quantify the general agreement between UMC and EMP-centers (1), identify morphological patterns and causes of disagreement (2), and assess the feasibility of calculating various measures of polycentricity using UMC data (3).

I find that the UMC approach is able to detect major EMP-centers, although with disagreements that often take the form of spatial overestimations. (1) Agreement is much better in economic (number of jobs) than in spatial (area) terms. (2) The mismatches between UMC- and EMP-centers can in many cases be plausibly explained by a nonlinear relationship between employment density and building volume. (3) Employment-based measures of polycentricity show fair agreement, but most spatial measures suffer from estimations errors.

Altogether, the vast majority of job concentrations is detected by UMC, but precise analyses of distributions are hindered by spatial disagreement. Hence, the results support the careful use of UMC as a substitute for EMP in certain analyses of polycentricity.

Table of Contents

List of Figures

List of Tables

List of Equations

List of Abbreviations

CB	Combined threshold approach
CBD	Central business district
DFG	Deutsche Forschungsgemeinschaft
DFW	Dallas Fort Worth (International Airport)
DSM	Digital surface model
EMP	Employment data
FN	False negative
FP	False positive
GUF	Global Urban Footprint
INSPIRE	Infrastructure for Spatial Information in the European Community
LEHD	Longitudinal Employer-Household Dynamics
LISA	Local Indicators of Spatial Association
LoD-1	Level of detail 1
LODES	LEHD Origin-Destination Employment Statistics LPI
LPI	Largest patch index
LR	Locally weighted regression approach
MAUP	Modifiable areal unit problem
MI	Local Moran's I approach
MNND	Mean nearest neighbor distance
nDSM	Normalized digital surface model
TN	True negative
TP	True positive
UMC	Urban mass concentration data

1. Introduction

1.1 Polycentricity

Polycentricity has become a pivotal concept in the study of cities. It refers to the existence of multiple centers within a conurbation (Riguelle, Thomas, and Verhetsel 2007). The term *monocentricity*, in contrast, describes an urban region with a clear hierarchy between a single urban center and its clearly subordinate hinterland (Krehl 2016a). For most of the 20th century, variations of such monocentric models were used to describe and understand the spatial nature of cities, particularly in the North American region (Davoudi 2003). But over the past decades it has become evident that these standard monocentric models no longer reflect the urban spatial structure of today's metro regions. Still, there is no consensus on what is taking the place of the monocentric structure (Taubenböck et al. 2017; Lee 2007). Some authors consider the possibility of a trend toward total dispersion of economic activity and population (Gordon and Richardson 1996; Lang and LeFurgy 2003). Others believe the most likely development to be the emergence of polycentric structures (Krehl, Siedentop, and Münter 2016). In such polycentric regions, the dominance of the traditional core city decreases as its functions are distributed among several surrounding subcenters (Anas, Arnott, and Small 1998). The debate continues about the future development of our cities. Will they become monocentric, polycentric, or dispersed (Lee 2007)? And is one of these courses politically or economically preferable (Davoudi 2003)? A large number of empirical studies attempt to resolve these pivotal questions by analyzing centers of urban regions (Krehl 2015b).

But although extensive research has been carried out on urban centers, what exactly constitutes such a center is not at all clear. Reviewing the state of research, (Krehl 2016a) finds that the term *subcenter* might be an umbrella term used to refer to several different types of urban spatial densifications which may vary in functional profile, location, or history. The absence of a standard, accepted operational definition of a center diminishes the generalizability and comparability of much published research on the issue of polycentricity (Agarwal, Giuliano, and Redfearn 2012). Consequently, despite its widespread currency (Davoudi 2003), and increasing rigour in research (Roca Cladera, Marmolejo Duarte, and Moix 2009), urban polycentricity is still considered a concept that is at best versatile, at worst vague, and fuzzy (Taubenböck et al. 2017; Meijers 2008; Burger and Meijers 2012). Neither a clear definition, nor a robust theoretical framework have so far emerged (Davoudi 2003). Instead, analyses of polycentricity vary on a number of points (Figure 1):

1) geographic scale
2) concept of polycentricity
3) method for detection and analysis
4) data basis

1) Scholars have observed and analyzed polycentricity at various scales, from intra-city scales to trans-regional and even inter-national scales (Xingjian Liu and Wang 2016; Veneri and Burgalassi 2012). The scale is not simply a matter of the size and ambition of the study (Arribas-Bel, Ramos, and Sanz-Gracia 2015) as polycentricity is not scale-invariant. What appears as polycentricity at a regional scale might be consistent with monocentricity at a local scale (Agarwal, Giuliano, and Redfearn 2012). And, as (Anas, Arnott, and Small 1998) observe, an organized system of subcenters might look like apparently unorganized urban sprawl at a different scale. Concretely, centers may take the form of metro regions at a national scale. At a local scale, centers may take the form of towns. Metro regions and town are affected in their centrality by factors which are unlikely to be scale-invariant. Hence, the choice of the scale likely has a direct impact on the outcome of a study, and not merely its interpretation.

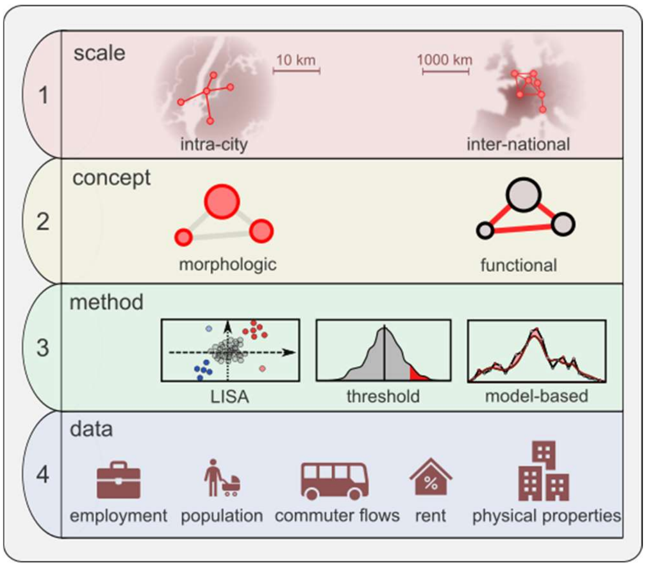


Figure 1: Variations across polycentricity analyses in the literature.

2) So far, there has been no agreement on the conceptualization of polycentricity (Taubenböck et al. 2017). *Functional*, *morphological*, or integrated concepts have been discussed (Veneri and Burgalassi 2012; Burger and Meijers 2012). From a *morphological* perspective, centrality is observed on the **place**-based notions of spatial distribution of economic activity (Krehl 2016a). The morphological centers are commonly explored defined from a socio-economic perspective. In this place-based notion, a region's centers are those places where its economic

activity culminates (Krehl 2016a). This culmination can take the form of local densification of employment (Giuliano and Small 1991). In contrast, the *functional* perspective observes centrality on **flows** of goods, people and information (Veneri and Burgalassi 2012). The functional and morphological concepts of centers do not exclude each other and could be, to some extent, positively correlated (Veneri and Burgalassi 2012) such that functional centers correspond to morphological centers.

3) Most analyses of polycentricity include the algorithmic detection of centers. A considerable amount of literature on the topic has not converged in an agreement on methodologies, but rather produced a large variety of different approaches (Arribas-Bel and Sanz-Gracia 2014). While multiple authors (Krehl 2016a; Grubesic, Wei, and Murray 2014; Agarwal, Giuliano, and Redfearn 2012; Roca Cladera, Marmolejo Duarte, and Moix 2009; Xuejun Liu et al. 2019) have suggested categories of subcenter detection algorithms, so far, no taxonomy has emerged as dominant. For the purpose of this thesis, I synthesize the following three-group categorization:

- *Local Indicators of Spatial Association (LISA).* LISA are a group of exploratory approaches which are developed for the general detection of clusters or hotspots in spatial data. Variations of LISA have been used for urban center detection, for example by Arribas-Bel, Ramos, and Sanz-Gracia (2015).
- *Threshold or cut-off methods*, which identify centers based on fixed thresholds of minimum size or density. Such thresholds can be defined by experts based on prior knowledge of the study area, or dynamically set based on values in the study area. These are among the oldest and simplest methods of center detection but remain widely used (Giuliano and Small 1991; Arribas-Bel and Sanz-Gracia 2014; Taubenböck et al. 2017; Lv et al. 2020).
- *Model-based approaches*, such as the approaches by McMillen (D. McMillen 2001; D. P. McMillen 2004) are based on fitting a model to the distribution of a variable and identifying those objects as centers which offer the greatest statistical support for the fitted model. Often, flexible variants of regressions are used as the basis of these models. Derivatives of such approaches have been applied to the detection of urban centers, e.g. by (Lee 2007; Sun 2020; Krehl et al. 2016; Garcia-López 2010).

The choice of algorithm is not trivial, as it can heavily affect a study's outcome (Anas, Arnott, and Small 1998; Krehl 2016b). Indeed, a meta-analysis by Agarwal, Giuliano, and Redfearn

(2012) shows that even within the same region the number and extent of detected centers varies drastically across applied methods. In this way, the absence of an accepted standard creates a problem for generalizing and comparing results (Taubenböck et al. 2017). Despite this issue, clear criteria by which to select the appropriate algorithm for a certain application are still lacking.

4) Finally, there is no unambiguous data basis for the study of polycentricity. Previous studies have used a variety of proxies for economic activity, such as data on employment or population density, productivity, rents, or wages. Out of these, employment density has been identified as the most suitable proxy for economic center detection by McDonald (1987) and Giuliano and Small (1991). However, traditional sources of employment data are rarely available in a comparable and uniform manner. Consequently, detailed studies have so far been limited to a few case study regions (Heider and Siedentop 2020). To date, there are few studies that have systematically investigated inter-regional and international differences in polycentric development (Heider and Siedentop 2020; Standfuß et al. 2020). Thereby, issues of data heterogeneity, limited data availability and methodological differences still limit our understanding of urban polycentricity.

1.2 Remote sensing of economic activity

A promising solution to the lack of a consistent data basis is the estimation of economic parameters using surrogates derived from remote sensing of physical properties. Over the past decades, advances in remote sensing technology and processing algorithms have been made and allow for the gathering of large-scale information on Earth's land surface, including the expanding urbanized land.

Multiple studies show that it is possible to estimate economic indicators such as GDP using remotely sensed land cover (Faisal and Shaker 2014; Ma and Xu 2010) or night-time light imagery (Doll, Muller, and Elvidge 2000; Doll, Muller, and Morley 2006; Noor et al. 2008; Sutton, Elvidge, and Ghosh 2007; Yue, Gao, and Yang 2014). Keola, Andersson, and Hall (2015) provide an overview of measuring economic development from space. The studies differ in accuracy and operate only at the very coarse scale of administrative regions, which is much too coarse to study intra-urban polycentricity.

Another way to objectively estimate economic parameters is via urban morphology which can be understood as the combined characteristics of the urban area that are immediately derived from their physical properties. Urban morphology can be measured using built-up density, height, or volume (Geiß et al. 2019).

That economy and urban morphology are directly linked is intuitive (Krehl 2015b): Much economic activity is happening at dedicated workplaces, which are often located mainly inside buildings, which in turn shape a city's morphology. While two-dimensional measures of morphology have been used in urban studies, three dimensional measures, notably the built-up volume, provide increased value for the study of polycentricity (Krehl et al. 2016). Several studies have explored the linkage between built-up volume and job concentrations (Krehl 2015b; Taubenböck et al. 2017; Fina et al. 2014; Krehl et al. 2016). These studies used volume information from 3D city models derived from airborne laser scanning (Fina et al. 2014) or urban mass concentrations (UMC) derived from high resolution stereoscopic data combined with building footprints derived from digital topographic maps (Wurm et al. 2014). These products are highly resolved and well suited for small scale studies with high accuracy. Descriptive analyses by Krehl (2015b) and Krehl et al. (2016) explored the interaction between these datasets and socioeconomic indicators and found them to display limited similarities. However, the input data is expensive and not available for larger regions. This creates a challenge for the transfer of these approaches to larger study areas.

New methods promise the derivation of UMC from globally available datasets: Geiß et al. (2015) developed a method to derive UMC at the resolution of 12 m from TanDEM-X radar data. While the spatial accuracy of these UMCs can not compete with highly resolved building models, their potentially global availability makes them highly attractive for large-scale studies of urban regions. The suitability of this data to detect morphological polycentricity has already been investigated in a test case (Standfuß et al. 2020).

Yet, no study has, so far, systematically explored whether this TanDEM-X derived UMC data reflects the socioeconomic reality and, by extension, whether the centers detected based on such data correspond to real economic centers. Hence, it remains unclear in which ways TanDEM-X-derived UMC data can be truly used as a surrogate for the socioeconomic data that is traditionally used in analyses of polycentricity.

1.3 Outline and research objectives

In this study, therefore, I aim to investigate the agreement of employment center detection based on direct employment measurements to detection based on indirect UMC data surrogates.

To accomplish this, I detect employment centers using a set of acknowledged center-detection algorithms on both TanDEM-X derived UMC data and survey-derived employment data for four city regions in the United States.

I select the United States as the geographical region for this study, as it is where a vast majority of research on urban subcenters stems from (Heider and Siedentop 2020; Davoudi 2003). Previous studies using morphological concepts (Giuliano and Small 1991; Giuliano et al. 2007; Lee 2007; McMillen 2004; McMillen 2001; McMillen 2003; Arribas-Bel, Ramos, and Sanz-Gracia 2015; Liu et al. 2019; McDonald 1987) have gathered extensive evidence of polycentricity for metropolitan areas throughout the United States (Agarwal, Giuliano, and Redfearn 2012) where availability of economic data is for such analyses is good (Krehl 2016b; Arribas-Bel and Sanz-Gracia 2014).

As no center detection method has gained full acceptance in literature, **I use a set of three acknowledged center detection methods** instead of a single method, one from each of the identified categories: LISA, threshold-based, and model-based. Those algorithms are separately applied to the EMP and UMC data to identify centers.

After the centers are detected**, I consider the employment-based centers as a baseline** and examine the extent to which the morphologic centers are congruent with them. This comparison of morphological- and employment-based results is also threefold (Table 1). Each part of the analysis is tailored to accomplish a specific research objective:

A) How well can employment centers be detected using UMC data? In an **evaluation of agreement,** I visually assess and quantify the general agreement between detected employment and morphologic centers.

B) Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers? In a systematic **analysis of error causes**, I examine the morphological and geographic properties of correctly and incorrectly detected centers for characteristic patterns.

C) Which analyses of polycentricity are feasible using purely UMC data? In a **feasibility assessment** I compare the agreement between employment and morphologic data specifically for various acknowledged measures of urban center structure.

In conjunction, these analyses target the overarching research question:

In which ways can urban mass concentrations (UMC), derived from globally available TanDEM-X data, be used as a surrogate for employment data (EMP) in the analysis of urban polycentricity?

Throughout the thesis, I aim for transferability and scalability, wherever possible employing such methods which can potentially be applied globally. I also aim for breadth, employing a variety of methods that have been used in other studies of polycentricity, and applying them to multiple test sites. Thereby, I aim to ensure that the findings will be and remain relevant beyond the specific scope of this study. In this way, I attempt to mitigate the issue of comparability that I outlined in the review (Taubenböck et al. 2017; Agarwal, Giuliano, and Redfearn 2012).

Table 1: Components of the three-fold analysis

Part	**Focus**	**Research Question**
A) Evaluation of agreement between detected EMP- and UMC-centers	General agreement	*How well can employment centers be detected using morphologic data?*
B) Systematic analysis of morphological properties of correctly and incorrectly detected employment centers	Error causes	*Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers?*
C) Feasibility assessment of analyzing job distributions in city regions using UMC data	Feasibility	*Which analyses of polycentricity are feasible using purely morphologic data?*

The remainder of this thesis is structured as follows: Chapter 2 introduces the study areas as well as the input data. Chapter 3 presents the three methods of center detection and the three analysis components while Chapter 4 presents their results. Chapter 5 discusses implications of the results and their interpretation and evaluates the influence of the choice of algorithm. Lastly, Chapter 6 provides a concluding summary of the study and gives an outlook to future work.

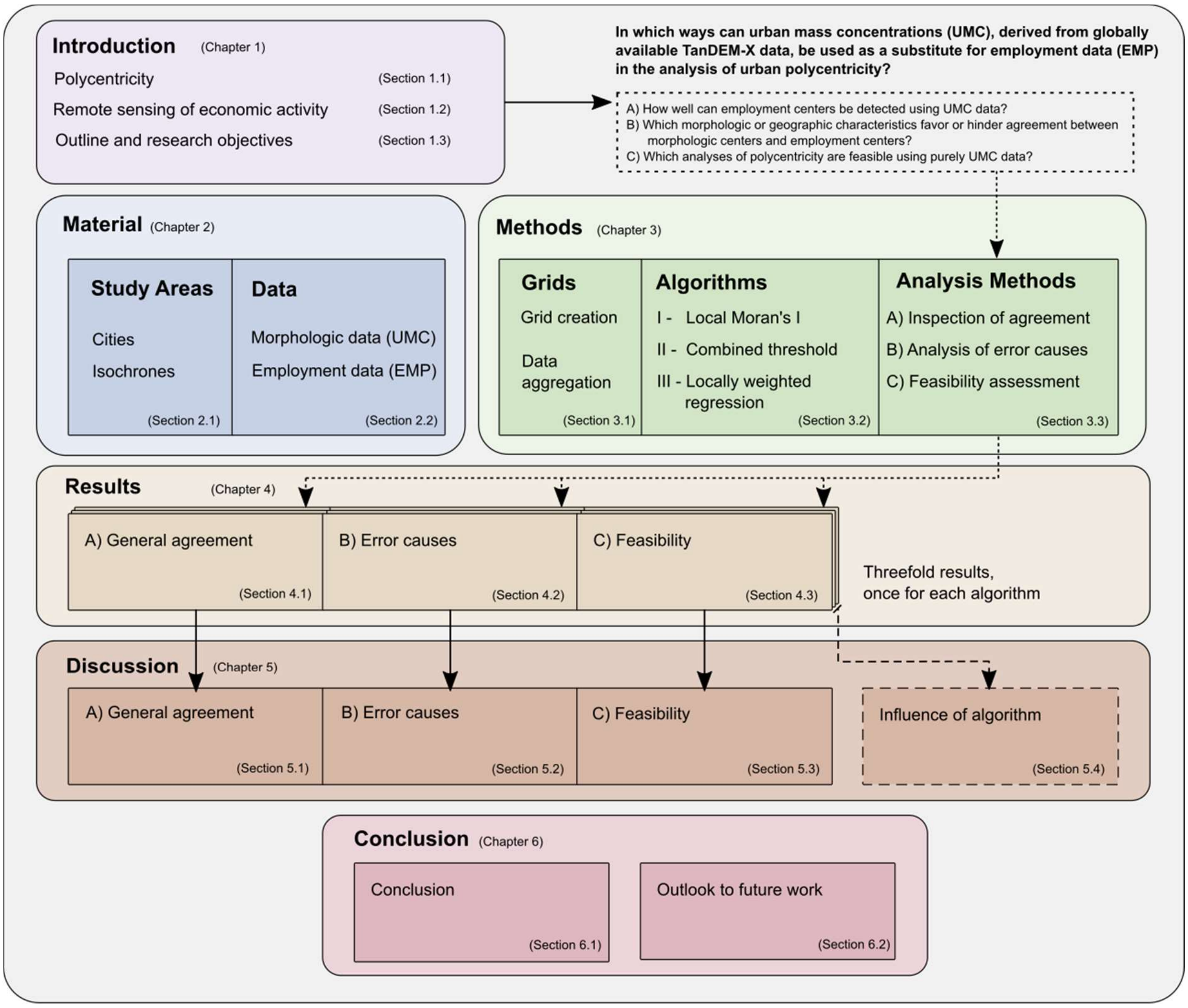


Figure 2: Outline of the thesis.

2. Material

Four cities in the United States were chosen as sites for the study. For each city, I delineated a study area and prepared morphologic and economic data for use in the analysis.

2.1 Study Areas

This thesis is a contribution to the project "*Where are the jobs? Stadtregionale Zentrenstrukturen im internationalen Vergleich*" (engl. '*Where are the jobs? International comparison of urban-regional center structures*') of the Deutsche Forschungsgemeinschaft (DFG), which has been funded since 2018 (for more information see DFG 2021). The project has a dedicated focus on the study of urban regions in the United States and Germany. Up to this point, most literature on polycentricity focused on the United States (Heider and Siedentop 2020). It was, thus, decided to only consider cities in the United States for this thesis. Criteria for the selection of test sites were the presence of monocentrism, moderate to strong growth, the availability of fine-grained information on total employment, and geographic diversity across sites. Following these criteria, the urban areas of Atlanta, Dallas, Pittsburgh, and Seattle were selected by the project committee as test sites for the thesis (Figure 3). In this section, I will briefly describe each of these test sites from a geographic and economic perspective.

Atlanta is the largest city of Georgia and also its capital. It is a major financial and cultural force in the American Southeast, where it has assumed an important position in national and international commerce (Advameg, Inc 2021). The city has emerged as a banking center and boasts the third largest concentration of Fortune 500 companies in the USA (City of Atlanta 2021). In the past two decades, the population of Atlanta metropolitan area has experienced unprecedented growth from 2,9 million to 4,1 million people (City of Atlanta 2021). Hartsfield-Jackson Atlanta International Airport is the world's busiest in daily passenger flights, and beyond its significance as a transport hub, it is also a major employer with more than 63 000 jobs on-site (City of Atlanta 2020). Direct flights to Europe, South America, and Asia have made Atlanta easily accessible to the more than 1 000 international businesses (City of Atlanta 2021).

Located in the rolling prairies of north-central Texas, **Dallas** is separated from its western neighbor Fort Worth by less than 50 kilometers. While each retains a distinct identity, the two cities and their surrounding suburbs are often considered a linked *metroplex* (Advameg, Inc

2021). The city features a diversified economy and is the world headquarters of the U.S. Army and Air Force Exchange Service. Beside this, wholesale and retail trade combine with services to form the backbone of Dallas’ economy. Dallas is also a major transportation hub (Encyclopaedia
Britannica 2021a) and the Dallas/Fort Worth international airport is one of the largest in the world (DFW 2021). Multiple industrial and commercial zones exist in the metro area, such as the business hub around *Addison* (DestinationDFW 2013) or the *Telecom Corridor*, a strip about three miles long on Highway 75 that is home to more than 600 technology companies (Advameg, Inc 2021; REDP 2021). Dallas shows a clear radial expansion, particularly in the northern and western direction, where it connects to its smaller neighbor Fort Worth.

Pittsburgh is nestled among the forested hills of southwestern Pennsylvania at the point where the Allegheny and Monongahela Rivers meet to form the Ohio (Advameg, Inc 2021). The city, which was formerly known as a major industrial center, was forced to reinvent itself after the decline of the steel industry in the second half of the nineteenth century. In the process, the city changed its economic base as its industries and businesses retooled and diversified (VisitPittsburgh 2015). Research, development, and the service sector became increasingly important. By the mid-1980s and again throughout the 2000s, Pittsburgh had gained the reputation as the United States’ *most livable city* (VisitPittsburgh 2015). Moreover, Pittsburgh remains a leading transportation center not least due to the significance of its large inland port (Encyclopaedia Britannica 2021b). The region shows numerous examples of recent development of retail and business parks along its transport and water routes. The city itself has over 70 miles of urban riverfront—more than any other inland port city in America (Advameg, Inc 2021). Geographically, the city is shaped by the relatively strong terrain with steep hills flanking the three major valleys around the core city.

At the southeastern shore of Puget Sound, a deep inland arm of the northern Pacific Ocean, lies the densely populated metropolitan area of **Seattle** (Encyclopaedia Britannica 2021c). While Seattle has in the past been largely dependent on the aerospace industry (Advameg, Inc 2021), an economic shift occurred when Boeing headquarters relocated to Chicago in 2001. From that point on, tech companies began to have a bigger impact on the city's economy, driving a sharp increase in population (City of Seattle 2021). Among the region’s most important economic assets are the ports at Elliot Bay and Tacoma which together form the fourth largest container gateway in North America (Port of Tacoma 2014). The city is built on hills and around water, much shaped in its form by the surrounding geography of Puget Sound in the west, and the Olympic and Cascade mountain ranges to the east and west (City of Seattle 2021).

Beyond the choice of the test sites itself, the choice of the extent of the study area can have great influence on the numerical results of the analysis and their interpretation (Taubenböck et al. 2019). As many center algorithms define a center by its relative density compared to its geographical environment, the definition of what to include in the environment will have an indirect effect on the detection of centers. For this reason, the choice of the study areas should not be arbitrary. From the transferability which I aspired to in this study resulted certain requirements for a study area. Firstly, the area should reflect the area of influence of an urbanized area rather than a strict geospatial buffer zone. Further, the study area should not be defined by historical or political delineations and rather should be dynamically created for any city which may be analysed. These requirements were fulfilled by isochrones of 60-minute travel times, which were created on the basis of transport infrastructure and central points representing the central business district (CBD). The 60 minutes isochrone threshold was judged by the project committee to best reflect the functional extent of the selected study areas. The location of the CBD was uniformly determined using the coordinates representing the geographic midpoints in Open Street Map.

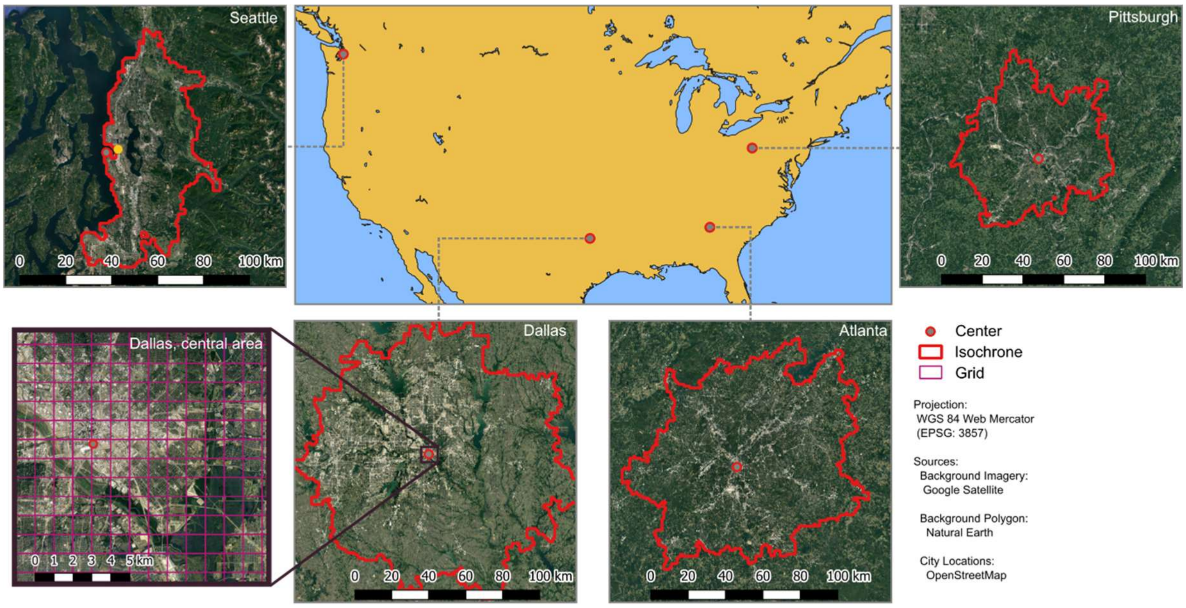


Figure 3: Overview of the test sites.

2.2 Data

At its heart, this analysis is a comparison of two data sources, one of which is morphologic, the other economic. Therefore, I used two distinct datasets in this analysis. Firstly, I used information on built-up volume, which was derived from a combination of remote sensing data. Secondly, as a baseline to compare the morphologic information to, I used information on total employment, which was derived from the LEHD Origin-Destination Employment Statistics (LODES, see Graham et al. 2014). Both datasets were provided to me within the project "*Where are the jobs? Stadtregionale Zentrenstrukturen im internationalen Vergleich*" and are briefly described in this section.

Morphologic Data

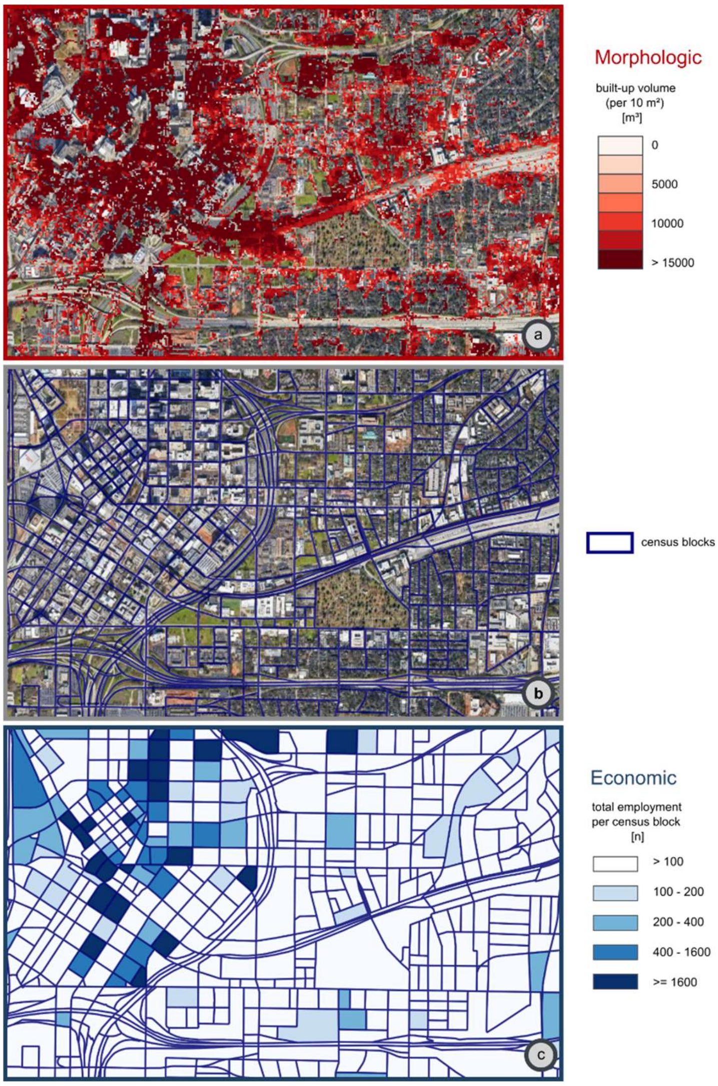


Figure 4: Input data over downtown Atlanta.:
a) TanDEM-X-derived built-up volumes;
b) Census blocks;
c) LODES-derived total employment.

To approximate the urban morphology, I used a processed normalized digital elevation model (nDSM) which comprises elevation information of objects above ground. This data was derived from the globally available TanDEM-X digital elevation model (Zink et al. 2014) using a method developed and validated by Geiß et al. (2015, 2019). In this method, progressive morphological filtering is firstly applied to generate from the DEM a nDSM. Subsequently, pixels containing non-urban objects are masked out using the Global Urban Footprint (Esch et al. 2017), a TanDEM-X-derived binary building mask, as well as vegetation masks derived from Sentinel-2 imagery. The result is a normalized model of urban built-up heights in a raster format with a spatial resolution of 0.4 arcseconds (i.e.,~12 m). Multiplying the height of a pixel with its footprint then results in a continous topographic model of the urban built-up volumes (Figure 4a). Geiß et al. (2019) validated the data by comparing it with LoD-1 building models of major European cities and found it to have overestimations in build-up density in areas with high buildings and underestimations in areas with low buildings, such as suburban areas. They also reported that the method consistently underestimated built-up heights. Yet, altogether, the relative distributions of built-up volumes represented the morphologic structure of the cities well. Further, Geiß et al.

(2019) demonstrate the data's ability distinguish several morphologic classes. This underscores its viability for the purposes of my study.

Economic data

As a counterpart and benchmark to compare this morphologic data to, I required a similarly fine-grained proxy for economic activity. In the United States, annual figures on employment numbers are available even in great detail and at very fine spatial units of aggregation via the Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES, see Graham et al. 2014). LODES data is generated from a variety of administrative and survey data and provides information on the number of employees at the level of census blocks. In this context, jobs are assigned to census blocks based on the location of the place of work, which is in turn defined by the physical or mailing address reported by employers. The LODES dataset excludes certain types of employment, such as the military and other security-related federal agencies, postal workers, some employees at nonprofit and religious institutions, informal workers, and the self-employed (Graham et al. 2014). The census blocks, which serve as the dataset's spatial reference units, are delineated using an automated process based on physical features rather than population or employment. Therefore, there are many census blocks without any population (Rossiter 2011) and the size of the census blocks is varied and often irregular (compare Figure 4b). Nonetheless, the spatial granularity and level of detail make the LODES dataset an appropriate counterpart to the remotely sensed morphologic data.

The **timespans** of the acquisitions of morphologic and economic data overlap. The TanDEM-X elevation model was generated using data acquired between December 12, 2010 and January 16, 2015 while the Sentinel 2 imagery was acquired between 2014 and 2016. The selected LODES data refers to the year 2015. While all jobs in the LODES snapshot are presumed to be held on April 1, some job characteristics may have a different timeframe (Graham et al. 2014). Altogether, the acquisition times of morphologic and economic data are in rough, if not perfect, agreement, that I deem sufficient to support a study for the year 2015.

To make the two datasets more comparable for the purpose of center detection, I applied additional preprocessing. Both this preprocessing and the center detection will be described in the next chapter.

3. Methods

To provide a regular spatial unit for this study, I aggregated both datasets to a spatial grid. To detect centers, I subsequently applied three center detection algorithms on each test site, on both total employment (EMP) and building volume (UMC) per grid cell. On the detected centers, I performed a threefold analysis, each part intended to answer one of my three research questions.

In the first section of this chapter, I present the creation of the spatial grids (section 3.1). In the second section I present and compare the algorithms (section 3.2). In the third and final section, I present the methods of analysis (section 3.3).

3.1 Grids

To begin with, I aggregated both data sources, UMC and EMP, to a uniform spatial basis in the form of a regular square grid. Due to their regularity, grids mitigate effects the modifiable areal unit problem (MAUP, see Openshaw 1983) to an extent. Nevertheless, they are still subject to zoning and scale effects. Using the spatial level of grid cells has several other advantages: It makes it easier to recognize patterns in the data rather than local peculiarities in the spatial unit (Madelin et al. 2009). Furthermore, Krehl (2016) points out additonal advantages for the study of polycentricity: *Firstly*, political considerations and administrative trajectories which may result in varying sizes of spatial units are mitigated. *Secondly*, the notion of absolute numbers (employees per grid cell) and densities (employees per km²) is identical, thus allowing comparisons of the results among all test sites. *Finally*, as Geiß et al. (2019) note, the effect of aggregating morphologic data into larger grid cells can have an averaging effect which mitigates error levels, albeit at a loss of spatial granularity.

I created a regular grid of square cells with an area of 1 km² each for each study area and limited the extent of the grids to the extent of the isochrones around the respective test sites. If the centroid of a grid cell intersected with isochrone, the grid cell was included. The remainder of the study operated on the level of these urban grids within which the 1 km² grid cells were the basic spatial unit. I aggregated both the economic and morphologic data into these grid cells.

These grids match the INSPIRE guidelines which specify a grid of 1 km² cell size as common data basis for spatial analysis in the European Union (INSPIRE 2014). The INSPIRE grid has been used by previous studies in the European region (Krehl 2016a). By sticking close to its spatial structure, I ensured comparability with previous and future studies which use data at the level of INSPIRE grid cells (Taubenböck et al. 2017; Krehl 2015b; Wurm et al. 2014). The detail of a 1 km² grid, Krehl (2016) points out, while not fine-grained enough to detect planned locations at an urban scale, is still well suited to detect regional spatial patterns such as polycentricity.

I aggregated both datasets into the grids, weighting cells or census blocks at the edges of grid cells proportionally to the intersectional area. I subsequently summed up the built-up volume and employment numbers within each grid cell, generating aggregated morphology and employment values for each grid cell.

Table 2: Descriptive statistics of the test sites.

Study Area	Area [km²]	Employment density [n/km²]	Building volume density [km³/km²]	Total employment [n]	Total building volume [km³]
Atlanta	9.745	215,4694	0,000286	2.099.750	2,7902
Dallas	14.558	189,6577	0,000150	2.761.037	2,1848
Pittsburgh	3.939	213,1167	0,000366	839.467	1,4406
Seattle	3.107	474,1575	0,000769	1.473.207	2,3896

Exploratory statistics of the input data reflect the test sites' diversity. The isochrones of the radially expanding Atlanta and Dallas cover much larger areas than Seattle which is much constrained by its geography and exceeds the other sites' densities by far (Figure 5). The total employment within the grids also varies, from between 839.467 employees in the Pittsburgh area to 2.761.037 in the Dallas area (Table 2).

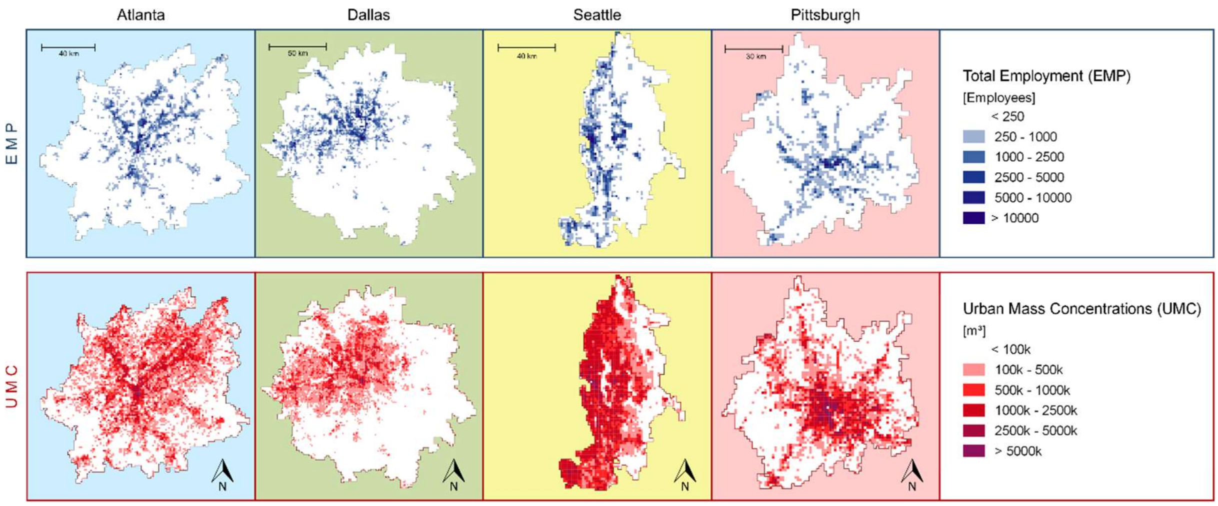


Figure 5: Overview of the test sites' employment (EMP) and built-up volumes (UMC) per grid cell. Note that to enhance detail, the sites are not displayed at matching scales.

To test if the correlation between volume and employment (Krehl 2015b) also holds for American cities, I calculated the correlations on a per-grid cell basis. Across all test sites (excluding cells which contain zeros) the correlation between total built-up volume and employment

exists and is highest compared to any other tested morphological variable (Figure 6). This supports the prior assumption that there is a link between built-up volumes and employment also for our test sites. Henceforth, I only use total built-up volume as the morphologic variable and, following the example of Wurm et al. (2014), I refer to this type of information on building volume as urban mass concentration, UMC. Likewise, I refer to the aggregated total employment per cell as EMP.

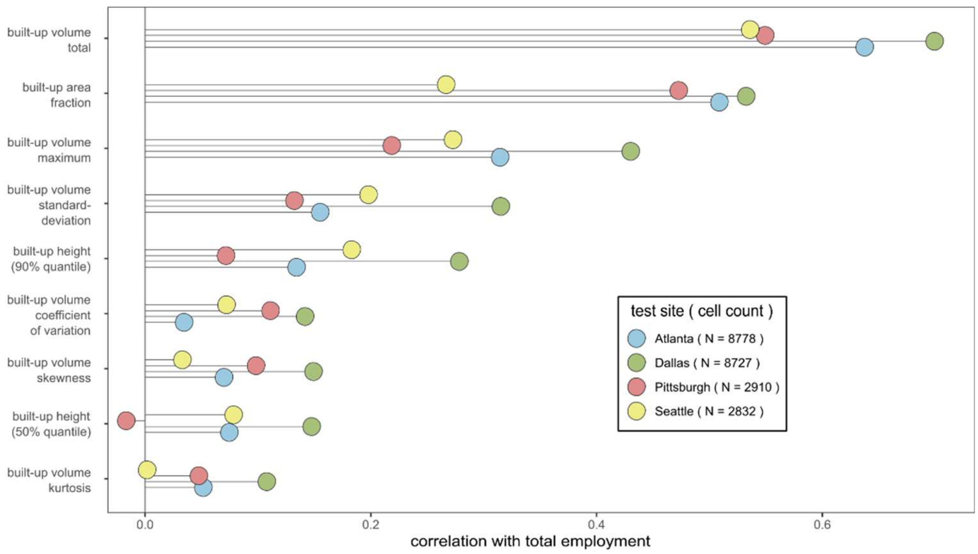


Figure 6: Correlations between total employment per grid cell and statistics of morphological properties per grid cell.

3.2 Algorithms for center detection

Within this harmonized data, I then algorithmically detected those cells which represented UMC-centers or EMP-centers. To accommodate the diversity of approaches found in literature, I separately applied three distinct center detection algorithms. I chose each algorithm to represent one of the previously identified categories:

- Local indicators of spatial association (LISA)
- Threshold approaches
- Model-based approaches

Compared to the use of a single algorithm (Krehl 2015b), my use of multiple algorithms better reflected the diversity of approaches found in literature. Wherever possible, I implemented the algorithms as they were described in previous publications and chose their parameter settings as they are reported by the seminal article, or in absence of such reporting, in the spirit of the article. I only made changes when they were necessary to maintain comparability between the algorithms. My goal was neither to select, create, nor tailor algorithms that were optimized for the task at hand. Rather, I aimed to truthfully represent the approaches as they are described in the literature.

I – Local Moran's I

Local Moran's I (MI) is a commonly used method from the family of Local Indicators of Spatial Association (LISA). It is the local component of the Moran's I statistic of spatial autocorrelation (Anselin 1996).

Local Moran's I was developed as a general exploratory method to find patterns in spatial distributions. It has been used for a wide range of applications in different fields, including image segmentation (Johnson and Xie 2011), Forestry (Fu et al. 2014), Health (Jacquez and Greiling 2003), and Urban Studies (Krehl 2015b). Building on the works of the latter, I used the algorithm as one method to detect urban centers. It is a flexible approach that is capable of identifying two different center concepts which are based on high clusters (HH) and high outliers (HL) respectively. I presumed that in the context of urban structure, high clusters reflect large spatial concentrations of employment that are noteworthy due to their ability to attract and stimulate concentration also in adjacent areas. High outliers, on the other hand, reflect local concentrations of workplaces which are noteworthy due to their local significance within areas of comparative economic weakness.

Figure 7 displays how I used Moran's I for the detection of centers. In the first step (Figure 7a) the local Moran's I statistic relates every cell to its neighbourhood via a weighted cross-product between the value of the cell (V) and the values of the cell's neighborhood (lagged variable, V_{Lag}). High values of this statistic indicate co-occurrence of similar values (high or low) at this location while low values of this statistic indicate dissimilar values (Anselin 1995). The significance of this local autocorrelation is analyzed through comparison with a reference distribution that is derived through permutation of the inputs (Figure 7b). Based on the relation of value of the cell to its global mean, and the relation of the lagged variable to its global mean, the algorithm is capable of detecting four significant types (Figure 7c): Cells with a

significantly high local autocorrelation are clusters of either high (HH) or low values (LL). Cells with a significantly low local autocorrelation are outliers of either high (HL) or low values (LH). Finally, all cells which are of type HH or HL that are also significant are output as centers (Figure 7d).

Implementations of local Moran's I are available in a range of softwares, such as *ArcGIS*, *CrimeStat*, *GeoDA*, *Python*, and *R* (Bivand and Wong 2018). While Bivand and Wong (2018) find that differences between software implementations of local Moran's I exist, to the best of my understanding these are negligible for the purposes of my analysis. In this study, I used the implementation provided by the R package *spdep* (Bivand, Pebesma, and Gomez-Rubio 2013; R Core Team 2020).

As an exploratory method, local Moran's I does not require any prior information about the study area or the distribution of the value (Arribas-Bel, Ramos, and Sanz-Gracia 2015). As with any use of a LISA, the spatial weights that are chosen a priori are very influential (Bivand and Wong 2018) as they determine the considered neighbourhood. I opted for a first order queen-contiguity neighbourhood, meaning that the neighbourhood of every grid-cell consisted of its immediate neighbours with which it shared a corner or edge. An advantage of working on a regular grid is that this neighbourhood is equally large (eight neighbors) for every cell, except for the comparatively few edge cells. While the effects of the neighbourhood choice were not eliminated, in this way, they were spatially homogenous.

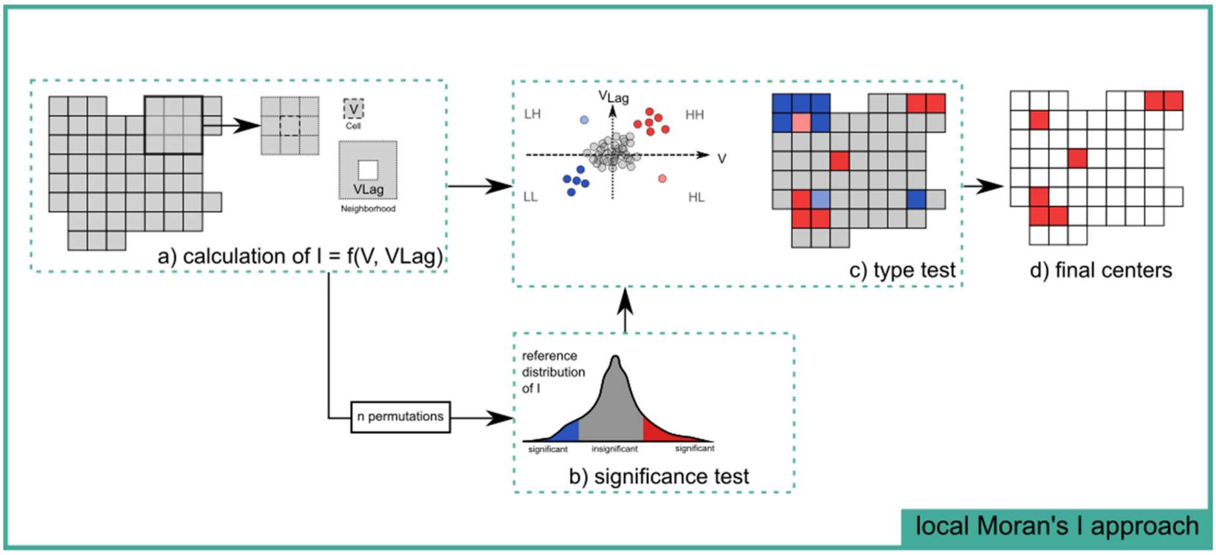


Figure 7: Workflow of the local Moran's I approach to center detection in grids.

The only numeric parameters that need to be defined a priori are the significance threshold and the number of permutations to calculate the reference distribution. I use a significance threshold of 95% computed over 500 permutations as this is the default setting of the *R* package *spdep*. Due to the large and varying number of grid cells, I considered a Bonferroni-correction of the p-value to be impractical (Anselin 1995). I also considered a correction based on the number of neighbors, as offered by *spdep*, to be ineffective in my approach, because of the aforementioned homogeneous neighborhood size.

II – Combined threshold approach

The **combined threshold approach** (CB) was developed by Taubenböck et al. (2017) specifically for the purpose of identifying urban centers based on building volume.

Threshold (or *cut-off*) approaches have a long history in urban studies and are widely used (Liu et al. 2019). They function by identifying all those spatial objects as centers for which either the value of interest or the density of this value exceeds a certain threshold. This threshold can be set in various ways. Fixed thresholds, which are chosen by an analyst based on prior knowledge or intuition (Giuliano and Small 1991), are often favoured for their simplicity, but come at the cost of lacking transferability between cities. An alternative are dynamic thresholds (see, for example, Garcia-López and Muñiz 2010; Lv et al. 2020), which are based on a statistic such as the mean, median, or standard deviation of the value of interest within the region (Lv et al. 2020). These approaches, however, risk that the threshold could be dominated by the influence of few, relatively strong outliers such as a dominant core city. To overcome the shortcomings of individual threshold-based methods, Taubenböck et al. (2017) developed the combined threshold approach (CB).

Unlike the generic Moran's I statistic, the combined threshold approach was specifically designed for the purpose of subcenter identification via their morphologic characteristics. Its two components correspond to two concepts of such centers. The first one, represented by the regional component, is spatially agnostic and defines centers as concentrations that are substantially high compared to the rest of the region. It makes no prior assumptions about the spatial structure and thus functions well for cities that are more polycentric or develop asymmetrically. However, a dominant core city, should it exist, may have dominating influence on the threshold, and as a result prevent the identification of smaller centers. This is the justification for the distance-based threshold. This distance-based threshold is based on the prior assumption that while monocentrism is prevalent, smaller subcenters in the periphery are still relevant. Reflecting this, it evaluates local densifications not in a regional context but within the context of rings with a similar distance from the core city. Thus, it is capable of detecting centers whose values are lower than those of the main center, provided that they exceed typical values of other cells in a similar context defined by the cell's ring. However, the creation of the distance-rings requires that the location of the center of the core city is already defined ex-ante. Within those constrains, the combination of two thresholds results in a robust approach that is both simply and transferable to different regions.

In practice, this combined approach utilizes two dynamic thresholds in disjunction. The first (regional) threshold is derived from the values of all cells in the region (Figure 8a). For the second (distance) threshold, cells are grouped into rings based on their distance to the city center, and the threshold is derived for each ring separately (Figure 8b). In both cases, the thresholds are determined as standard derivations above the mean. A cell can qualify as a center by meeting either one of these thresholds. A factor applied to the standard deviation can serve to adjust the algorithm. From another perspective, this factor itself can be considered the threshold which is applied to the z-scores of the cell, which in turn are calculated twice, once within the regional and once within the distance-ring-based reference population.

In a further processing step, neighboring centers are joined into clusters using a queen contiguity neighborhood. All clusters which fail to each the minimum size of 2 km² are then eliminated (Figure 8c). All remaining clusters are considered centers (Figure 8d).

I implemented the algorithm in *R*. I determined the center location using the coordinates representing the geographic midpoints in *OpenStreetMap* (OpenStreetMap contributors 2020). It is possible that distance-rings at the very center or at the edges of the region may only contain few values, reducing the robustness of the distance-based threshold in those areas.

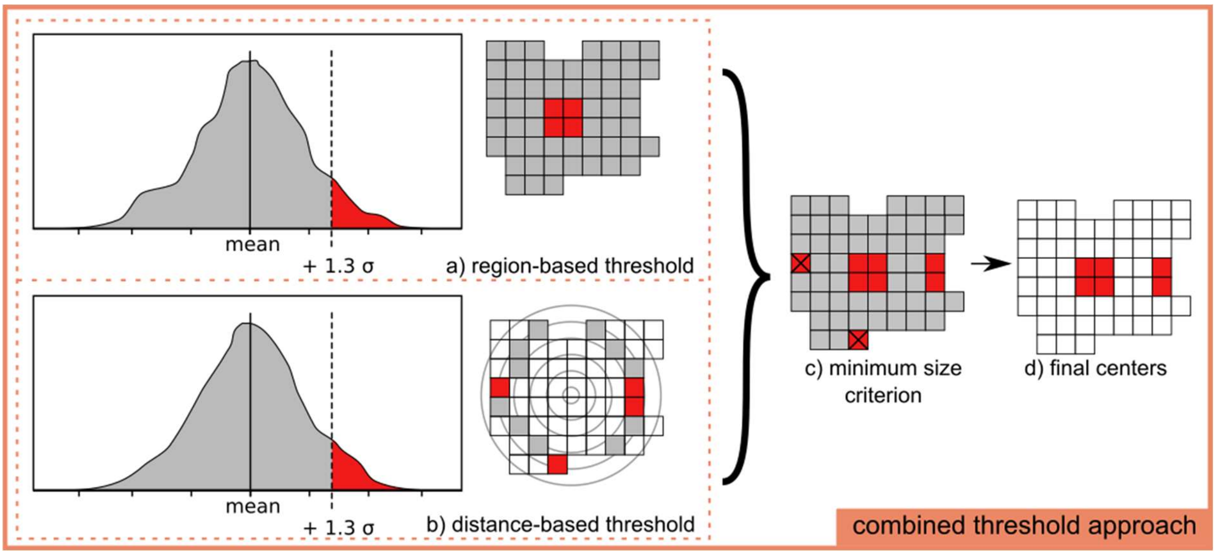


Figure 8: Workflow of the combined threshold approach to center detection in grids.

As the study of Taubenböck et al. (2017), I opted for z-score threshold factors of 1,3 in both the regional and distance-based groups, and use distance-rings of 1 km width. However, I did not adopt the additional, fixed threshold, which in the original implementation was intended to exclude sparsely built residential areas. As the threshold was originally derived for test sites in Germany, and no plausible fixed threshold likewise applies to UMC and EMP data, I decided that its adoption was not feasible.

III - Locally weighted regression

The **locally weighted regression** approach (LR) by McMillen (2001) is an example of theoretical (Krehl 2016b) and model-based approaches. It has been used in multiple studies of urban structure (Garcia-López, Hémet, and Viladecans-Marsal 2017; Muniz, Galindo, and

Garcia 2003) and notably by Krehl (2016) who adapted it, with some modifications, for the detection of economic centers in German city regions.

This approach is grounded in economic theory (Krehl 2016b). Unlike the generic Moran's I and the morphology-targeted combined approach, the LR approach is specifically designed to identify economic subcenters. McMillen (2001) defines an employment subcenter as a concentration of firms large enough to have significant effects on the overall spatial distribution of population, employment, and land prices. To detect such centers algorithmically, it is necessary to not just look at the values of the centers themselves, but also at values in a window surrounding them, and infer the cell's influence on this window.

Concretely, the approach aims to detect subcenters which fulfil two criteria: Firstly, they display significantly higher values than their surroundings. Secondly, they are significantly influential within the city. To identify such centers within a spatial dataset, McMillen (2001) developed a two-step procedure in which the first step identifies likely candidate subcenters and the second one iteratively eliminates candidates with insufficient explanatory power.

In the first step, a locally weighted regression is used with a large window size to fit a relatively smooth surface to the distribution of the value over the city. In McMillen's original implementation (McMillen 2001) this value is the density of employment. When applied to a grid, due to the regularity of the cells, the notions of density and count are identical. Cells with significant positive residuals on the smoothed surface are then considered subcenter candidates. To reduce patches of neighbouring candidates to their center, a maximum filter excludes all candidates whose predicted log-employment densities are not the highest in a three-mile radius (Figure 9a).

In the second step, a semi-parametric regression is again used to fit a surface. This time, a much smaller window results in a much more variable surface. Iteratively, the candidate subcenters with the lowest coefficients are removed and the regression is repeated, until only significantly influential subcenters remain (Figure 9b).

To allow for comparability with the other two algorithms, I empirically made two modifications. Firstly, in step one, I did not filter candidates using the predicted employment density but rather using their initial values. I found that this created a more robust result that was less reliant on the fit. Secondly, after step two, I reactivated candidates removed by this filter as long as they were spatially connected with a remaining candidate (Figure 9c). They were then considered part of the spatial extent of the identified subcenter, provided that it was identified as significant in step two. This allowed me to identify the geographical area covered by the subcenter, rather than just the central grid cell, as the final center extent (Figure 9d).

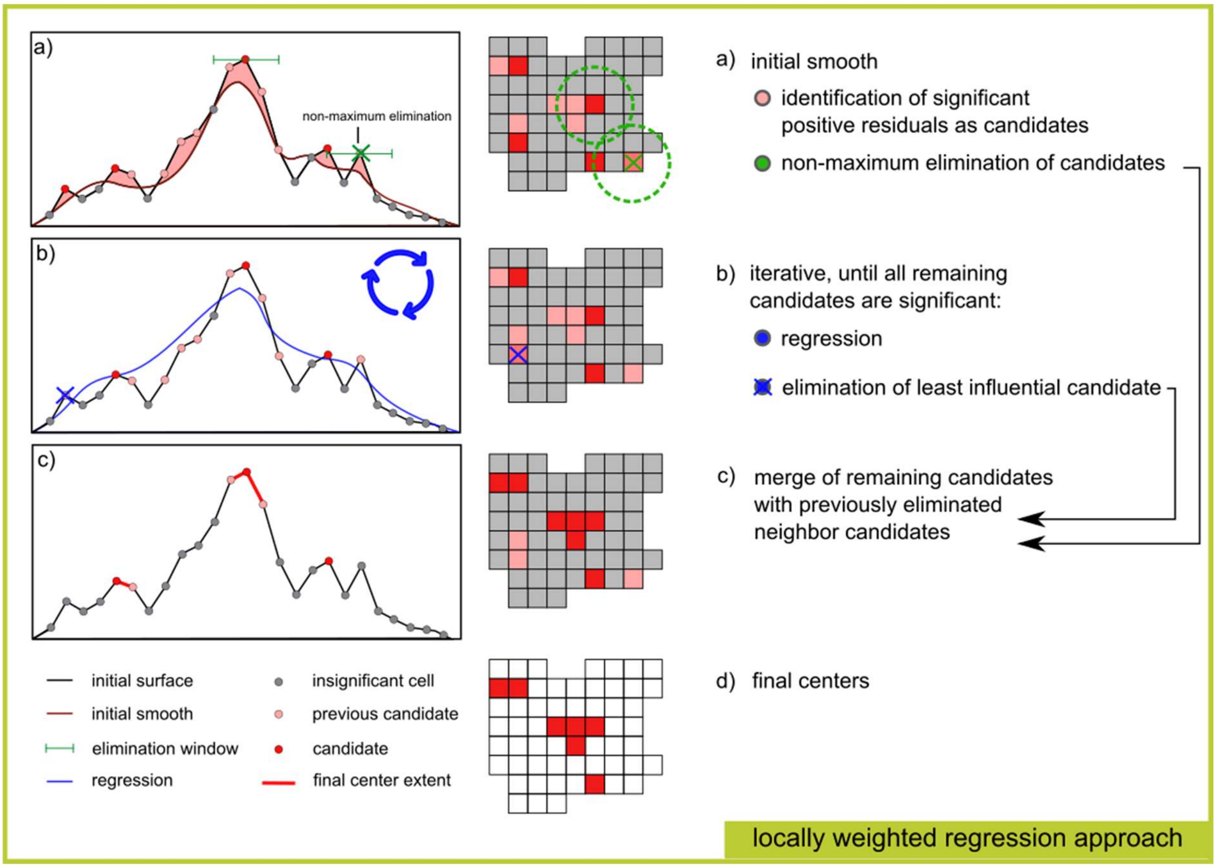


Figure 9: Workflow of the locally weighted regression approach to center detection in grids.

I used the *R* package *McSpatial* (McMillen 2013) to calculate the regression within the framework of my own modifications.

While the approach assumes the presence of a CBD, it does not enforce monocentricity and it allows for locally significant subcenters to be detected, especially if smaller window sizes are used. While the regression model that this approach uses is flexible, the algorithm as a whole requires the definition of a number of parameters, such as the window sizes in the first and in the second step, as well as the p-value to declare significance. Despite that, as Krehl (2016b) points out, it does not require local knowledge of the region that is analysed. Hence, the parameters, once chosen, require no tuning to the test sites. I chose my parameters as in McMillen's original model but made two modifications to the window sizes in accordance with Krehl (2016) using window sizes of 25% and 5% instead of 50% and 25% in the first and second step respectively. This adaptation allowed for the identification of more locally relevant subcenters that my data supports and, thus, a more finely grained view of the city.

Comparison of the Algorithms

All three algorithms are similar in that they can be used to detect urban (sub-)centers. However, differences in their workings reflect different concepts of centers.

The **local Moran's I approach** evaluates the significance of a cells relationship with their neighborhood in a regional context. From this perspective, it allows for the identification of subcenters in the form of *neighbourhoods which are significantly different from their regional context.*

The **combined threshold approach** compares values within their regional and distance-based contexts. In either context, it identifies cells that exceed the normal variance substantially as centers. From this perspective, it allows for the identification subcenters in the form of cells which are *substantially higher than either their regional context or other cells in a similar distance to the main center*.

The **locally weighted regression** approach identifies cells with both significantly high values and significant explanatory power for the overall distribution of values within the region (McMillen 2001). From this perspective, it allows for the identification of subcenters in the form of *cells which have a significant effect on the region.*

Because of these differences it is to be expected that the three algorithms produce different results. As each algorithm is able to identify a particular concept of center in our data, considering all of them may enable a richer and more well-rounded view of the city that is still neither complete, nor absolute. A particular benefit of my method is that it allows me to attribute patterns more clearly to differences between the UMC and EMP datasets, rather than effects of the chosen algorithm or study site.

3.3 Analysis methods

Each of the algorithms yielded a binary classification on a grid cell basis, meaning that each cell was assigned to be either a center or not a center. The application of three algorithms on two data sources and four test sites resulted in a total of 24 binary classifications. Twelve of these contain morphologic centers (UMC-centers) to be evaluated by the twelve matching economic centers (EMP-centers). After the centers were identified in this way, I performed a threefold analysis, each intended to answer one of my three research questions:

A) How well can employment centers be detected using UMC data? To answer this question, I performed a visual inspection and quantified the agreement between detected EMP- and UMC-centers.

B) Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers? I systematically assessed morphological properties of correctly and incorrectly detected employment centers.

C) Which analyses of polycentricity are feasible using purely UMC data? To answer this question, I performed a feasibility assessment of the analysis of economic polycentricity in the test sites using UMC data.

Across these analyses, I considered variations and consistencies across algorithms and test sites. My primary goal was to identify patterns which hold true across all algorithms and test sites and are thus likely to reflect some underlying general relationship between UMC and EMP.

A secondary aim was to identify characteristics, strengths, and weaknesses particular to each algorithm.

A) Evaluation of agreement between detected employment and morphologic centers

Question: How well can employment centers be detected using UMC data?

In the first part of the analysis, I quantified the overlap between the UMC-based and EMP-based centers. I treated the center detection analogous to a classification problem in which the UMC-based centers are the prediction and the EMP-based centers are the ground truth. I find it worth noting that certain flaws of the approach (both in terms of data processing and algorithm) can apply not just to the UMC but also to the EMP. Therefore, the term ground "truth" should not be taken in its most literal sense. Rather, the EMP-based centers are a benchmark to which the UMC-based centers are compared. If UMC are a suitable surrogate for EMP, then the spatial distributions of both datasets' values (UMC and EMP) should be similar, and the overlap between UMC-based centers and EMP-based centers should also be maximal.

Just as for the detection of clusters itself, various ways have been used for quantifying its accuracy (Grubesic, Wei, and Murray 2014; Huang, Pickle, and Das 2008; Cai, Huang, and Song 2017). I captured cluster accuracy on basis of individual grid cells by relying on metrics of *precision*, *recall*, *intersection over union (IoU),* and *Cohen's Kappa*. *Recall,* also called completeness or producer's accuracy (Wurm et al. 2014), relates to the ability of a cluster technique to identify an object of a class as such. It is inversely related to the error of omission. In my application, a large number of undetected EMP-centers would result in a low recall. However, recall alone does not capture the complexity of the agreement's quality. Notably, the creation of a center map with a recall of 100% would be possible by simply classifying all grid cells as centers (Figure 10a). Thus, the precision and IoU scores are a valuable and necessary addition: *Precision*, also called correctness or user's accuracy (Wurm et al. 2014), relates to the ability of a technique to separate a class object from non-class objects. It is inversely related to the error of commission. A large number of falsely committed center cells would result in a low precision (Figure 10b). The *IoU* is a commonly used metric in computer vision that considers both false commissions and omissions (Figure 10c) and is, thus, well suited to quantify how well a classification result overlaps with the ground truth (Csurka et al. 2004).

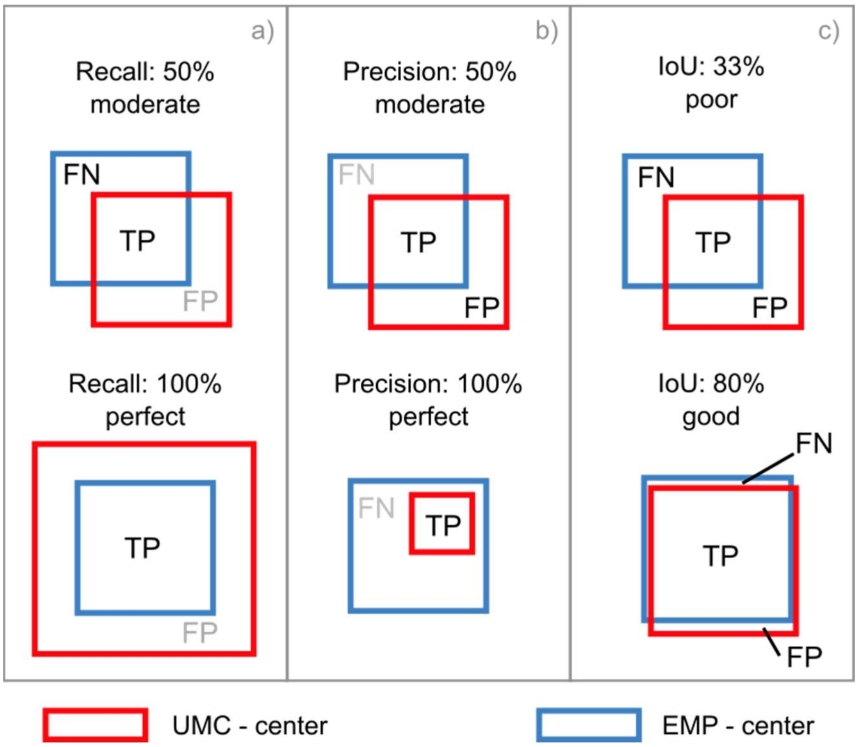


Figure 10: Illustration of recall (a), precision (b), and IoU (c) for different scenarios. The intersection between UMC-centers and EMP-centers generates the cases of true positives (TP), false negatives (FN) and false positives (FP). Cases which do not factor into the calculation of the measure are greyed out. Note how precision and recall can report perfect accuracies even if the agreement is not absolute.
Top row: Secenarios in which the measure returns poor values.
Bottom row: Scenarios in which the measure returns good values.

Based on the agreement between the UMC- and EMP-based center detection, I assigned every cell to be either a true negative (TN), true positive (TP), false negative (FN), or false positive (FP). This was done once for each of the three applied algorithms. From the relation between the counts of those cases (N_{TN}, N_{TP}, N_{FN}, and N_{FP} respectively), I calculated the metrics of recall, precision, and IoU. As all cells are equally large, these metrics can also be considered to reflect the agreement in terms of area:

Area-based precision:

$$Precision_{Area} = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad \text{(Eq. 1)}$$

Area-based recall:

$$Recall_{Area} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad \text{(Eq. 2)}$$

Area-based intersection over union:

$$IoU_{Area} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}} \quad \text{(Eq. 3)}$$

I supplemented the purely cell-based metrics by employment-based measures of precision and recall. These alternative metrics $Precision_{EMP}$ and $Recall_{EMP}$ and IoU_{EMP} consider the number of employees captured by the cases TP, FN, and FP, rather than the cell count. Thereby I was able to effectively mitigate the effect of cells that contained little economic activity.

Employment-based precision:

$$Precision_{EMP} = \frac{EMP_{TP}}{EMP_{TP} + EMP_{FP}} \quad \text{(Eq. 4)}$$

Employment-based recall:

$$Recall_{EMP} = \frac{EMP_{TP}}{EMP_{TP} + EMP_{FN}} \quad \text{(Eq. 5)}$$

Employment-based intersection over union:

$$IoU_{EMP} = \frac{EMP_{TP}}{EMP_{TP} + EMP_{FP} + EMP_{FN}} \quad \text{(Eq. 6)}$$

Where EMP_X is the sum of all the EMP-values of a certain case X (either FP, TP, FN, or FN) at the site.

I further calculated the *Cohen's kappa*, a common metric of interrater reliability (McHugh 2012) that has already been used to evaluate the accuracy of urban center detection (Cai, Huang, and Song 2017). I do not calculate an employment based variant, as the kappa is only intended for categorical variables (Revelle 2020). The kappa can thus be seen as an additional metric of area-based accuracy, and a complement to the IoU_{Area}.

Cohen's kappa:

$$\kappa = \frac{Pr_a - Pr_e}{1 - Pr_e} \quad \text{(Eq. 7)}$$

Where Pr(a) represents the actual observed agreement, and Pr(e) represents chance agreement (McHugh 2012).

By using multiple quantitative measures, I reinforced the results of the analysis and increased their comparability with the outcomes of future studies.

As described in the introduction, the identification of cells as belonging to a center is not the final goal. Rather, it is a necessary processing step for the analyses of the centers. For this reason, it was further necessary to evaluate patterns of error causes (Question B) and the feasibility of the actual analysis methods (Question C).

B) Systematic analysis of morphological properties of correctly and incorrectly detected employment centers

Question: Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers?

In the second part of the analysis, my aim was to identify which morphological properties facilitate the correct (TP, TN) or incorrect (FP, FN) detection of centers.

For this purpose, I firstly investigated the statistical distributions of morphologic and economic properties for each case to identify whether there are clear patterns that are particular to certain types of error. In addition to the UMC and EMP properties, I also used measures of height and built-up density from the TanDEM-X nDSM. I computed height values as the 90% quantile of nDSM pixel values per grid cell, and built-up density as the share of nonzero nDSM pixels.

These measures were then used in explorative methods. For a first investigation, I compared the median values of the height, built-up density, UMC, and EMP measures across the different cases. I further compared the distribution of these metrics in boxplots. My goal in this step was to identify characteristic differences in these measures' distributions between the cases TP, FN, and FP. My presumption was that if the correct detection of a cell was influenced by its own local properties, this may have resulted in a characteristic signature for each of the three cases'. This signature would be reflected in the statistic distributions of the analyzed measures for these cases. However, a purely grid-cell based measure does not adequately capture spatial relations of centers and their environment.

Therefore, I additionally performed a visual screening of the urban regions as a whole but also individual centers in detail, using optical imagery and ancillary data sources to characterize them. In this way, I qualitatively examine the direct influence of the centers' morphologic, economic, and geographic characteristics on their delineation. For such analyses, some normativity and subjectivity cannot be ruled out (Taubenböck et al. 2017) but complements the statistical approach well.

C) Feasibility assessment of analyzing job distributions in city regions using UMC data

Question: Which analyses of polycentricity are feasible using purely UMC data?

In the third part of the analysis, I selected various established measures of polycentricity analysis. I calculated each measure on a EMP and UMC basis. Then I inspected to what extent there was agreement between the two calculated measures. If spatial UMC concentrations reflected spatial EMP concentrations, then I expected that similar statements about the center structures of a city region would be preserved whether I use UMC or EMP-centers. If I detected differences, I additionally investigated if these are systematic and if it is feasible to derive correction factors for these discrepancies. If there was a systematic and consistent difference over test sites, the use of UMC as a surrogate is likely possible with the additional application of a correction factor to the measure.

As pointed out in the introduction, there is a plethora of measures to quantify various concepts of polycentricity. Based on the literature, I identified categories of measures and selected for each category an expressive and simple representative measure. My goal was not to simulate a complete analysis of polycentricity, but rather to test the feasibility of substituting EMP by UMC for these measures.

- **Importance of centers:** I calculated the share of a test site's total employment (*EMP-share*) and the share of a test site's total area (*area-share*) which are located within the test site's UMC- and EMP-centers respectively (Krehl 2016b). A high value of these measures indicates that the identified centers as a whole are important within a city region. Firstly, I compare the EMP-shares of UMC-centers to those of EMP-centers to determine whether they agree in economic importance. Secondly, I compare the area-shares of UMC-centers to those of EMP-centers to determine whether they agree in their spatial extent. Thirdly, I compare the EMP-shares to the area-shares to determine differences between the two measures.
- **Centrality:** I calculated a largest patch index (LPI) of area as a simple method of quantifying the dominance of the city region's largest center. In analogy to Taubenböck

et al. (2017) the LPI was calculated as a percentage of the area covered by the largest connected center divided by the area of all centers.

$$LPI = 100 * \frac{Area_{largest_{center}}}{Area_{all_{centers}}} \qquad \text{(Eq. 8)}$$

A high value of the LPI suggests dominance of the spatially largest center.

- **Hierarchy:** An established method to analyze urban hierarchies is the application of Zipf's law and the corresponding rank-size rule (Krehl 2015a; Taubenböck et al. 2017). I created logarithmic rank size plots and visually evaluated them with regard to Zipf's law (Standfuß et al. 2020). The relation of the plotted centers to the idealised Zipf distribution line allows identification of centers and regions which do conform to Zipf's law.
- **Spatial distribution of centers**: In analogy to Taubenböck et al. (2017) I chose the mean nearest neighbor distance (MNND) as a measure for location-based site-specific pattern analysis. Distances are calculated as the shortest distance between two patches, and they are understood as measures to evaluate whether centers are spatially clustered or dispersed. I calculated the nearest-neighbor distance for each center. The mean of this value across all centers is the mean nearest neighbor distance which serves as a simple measure of the spatial distribution of centers. Specifically, this value can be defined as

$$MNND = \frac{\sum_{i=1}^{n} d_{min_{ij}}}{n} \qquad \text{(Eq. 9)}$$

where $d_{min_{ij}}$ stands for the nearest neighbor center to-center distance and n for the number of connected centers. A high value of this measure indicates that centers tend to be seperated by larger distances, suggesting dispersion. I calculate the MNND twice, once on distances between patch centroids and once on distances between patch edges. The former better considers the location of patches while the latter further considers their extent.

- **Clusteredness:** Lastly, I used the global Moran's I measure as a general measure of clusteredness, which in the urban context can distinguish compactness from sprawl (Tsai 2005). It is defined as

$$I = \frac{\sum_i \sum_j \omega_{ij} z_i \cdot \frac{z_j}{S_0}}{\frac{\sum_i z_i^2}{n}} \qquad \text{(Eq. 10)}$$

with ω_{ij} as the elements of the spatial weights matrix, $S_0 = \sum_i \sum_j \omega_{ij}$ as the sum of all the weights, and n as the number of observations (Anselin 2020). A high I measure suggests clustering of similar values in the study area.

Notably, this measure does not consider the detected centers and is computed directly on the input UMC or EMP values. As (Krehl 2015b) finds that local and global measures are feasible for the objective to detect urban spatial structure, the inclusion of global Moran's I is a valuable addition.

I calculated each of the five measures for each of the four test sites and, except for the global Moran's I, for each of the three algorithms. For the calculation of LPI, MNND and the creation of Zipf plots, I grouped adjacent grid cells into centers based on first order queen contiguity, meaning that cells were considered part of a common center if they share a line or corner.

This chapter has described the study's methods of data processing and analysis of which Figure 11 provides a graphical overview. The next chapter will present the results of this threefold analysis. Both the methods and the results will be discussed in chapter 5.

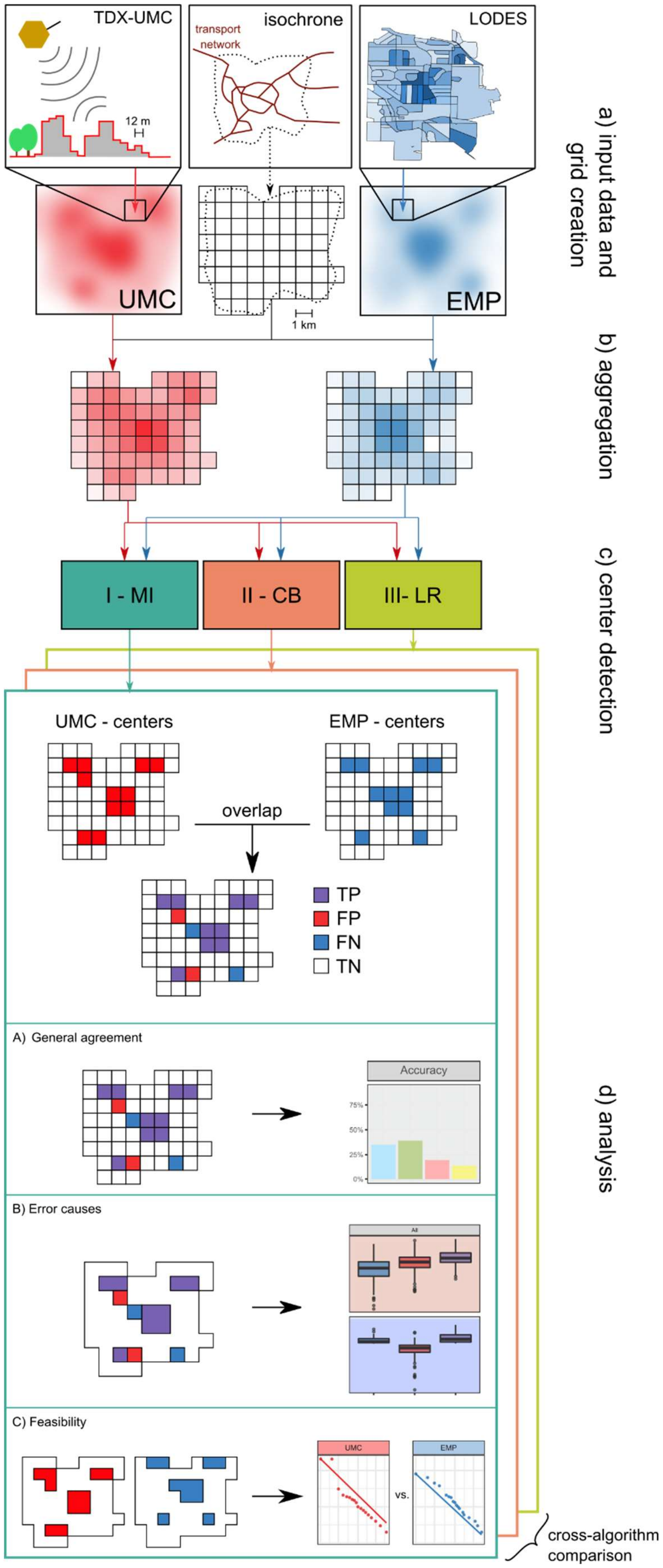


Figure 11: Workflow of the study.

4. Results

As described in the previous section, my analysis is threefold:

- General agreement between detected EMP and UMC-centers.
- Systematic analysis of morphological properties of correctly and incorrectly detected employment centers.
- Feasibility assessment of analyzing job distributions/structure in city regions using morphologic data.

Accordingly, this results section is organised in three consecutive parts, each corresponding to one part of the analysis. The section concludes with a summary of all findings. Throughout, I refer to patches of cells detected as centers via the UMC data as UMC-centers. Likewise, I refer to patches of cells detected as centers via the EMP data as EMP-centers. When not qualified, the generic term "centers" refers to the economic concept outside the scope of the algorithmic results.

4.1 Evaluation of agreement between detected EMP- and UMC-centers

Question: How well can employment centers be detected using UMC data?

Initially, I map the detected centers for all test sites and for all algorithms (Figures 12 & 13): the local Moran's I approach (MI), the combined threshold approach (CB), and the locally weighted regression approach (LR). The visual analysis reveals several spatial patterns.

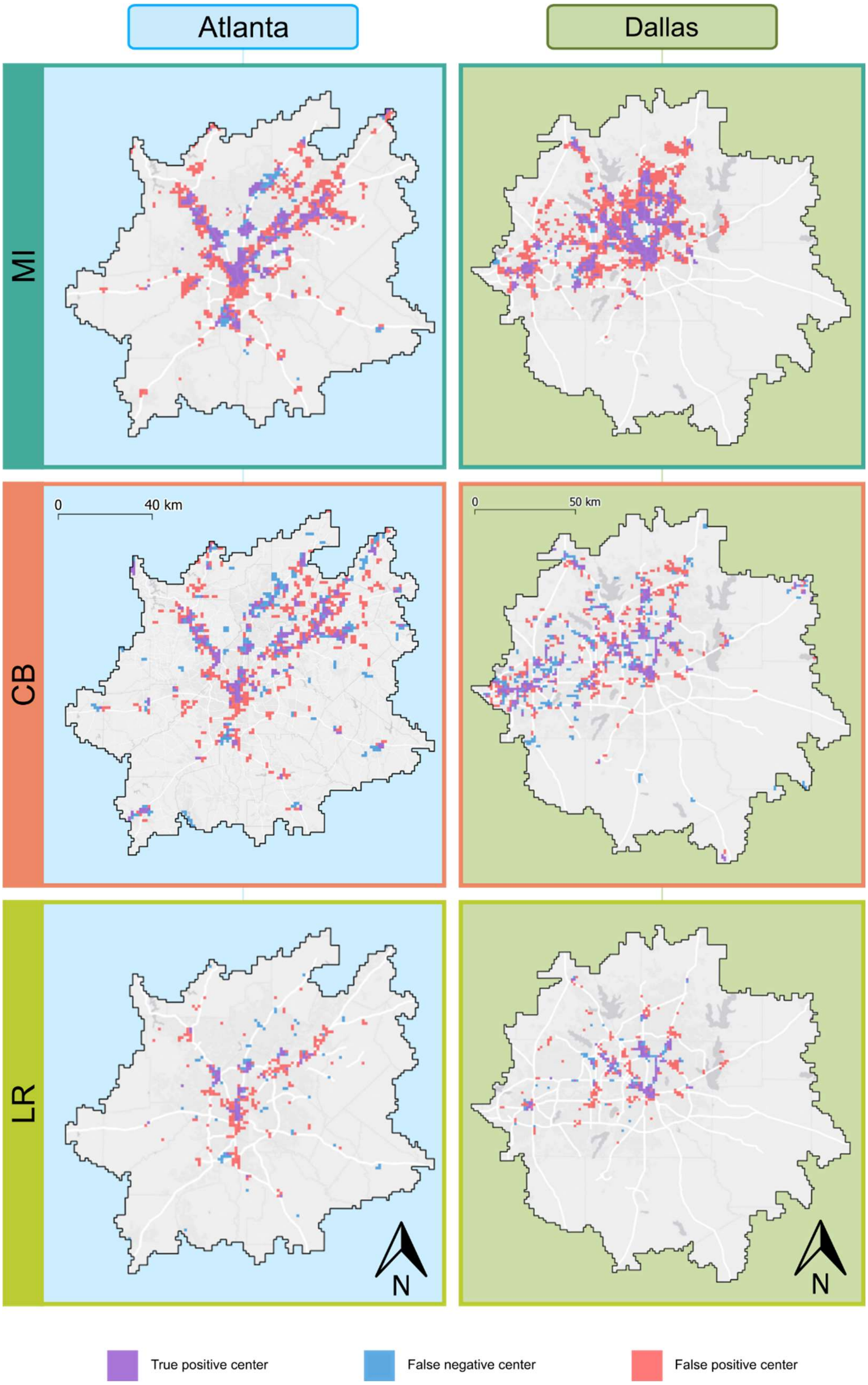


Figure 12: Detected centers by algorithm and test site. Atlanta, Dallas. MI: Local Moran's I approach; CB: combined threshold approach; LR: locally weighted regression.

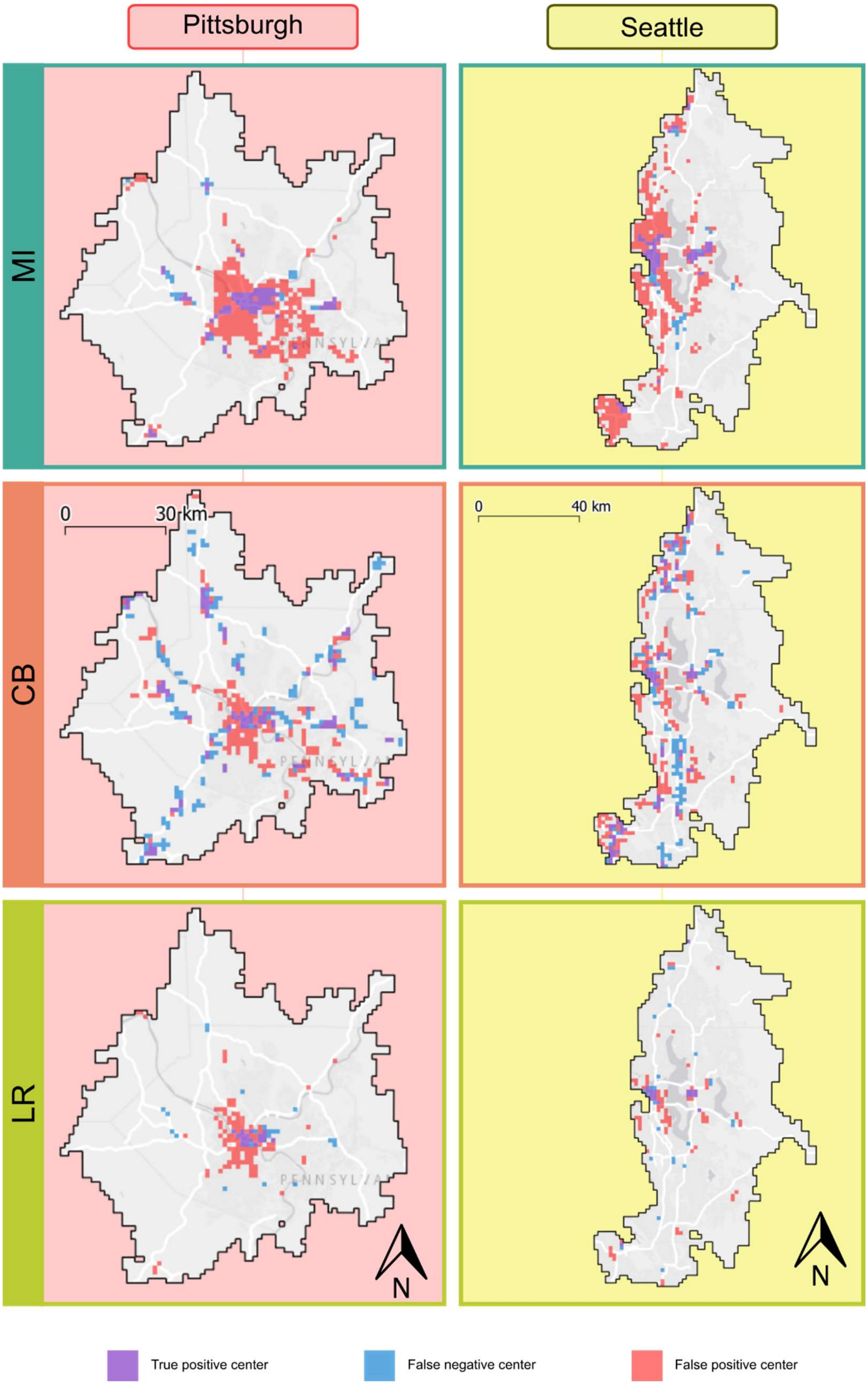


Figure 13: Detected centers by algorithm and test site. Pittsburgh and Seattle. MI: Local Moran's I approach; CB: combined threshold approach; LR: locally weighted regression.

I find that across all algorithms and all sites, there is agreement between EMP-centers and UMC-centers at the location of the core city and the major subcenters. Often the UMC-centers overestimate the area of the EMP-centers, particularly the largest center that contains the central business district (CBD) of the city. Agreement appears to be worse for the smaller and outlying centers.

In addition to these similarities, there are also notable differences between the patterns produced by the different algorithms. In the case of **MI**, the false positives (FP) form large, connected patches that often match the general location of an EMP-center, but overestimate its extent. Compared to the other algorithms, the **CB** identifies larger numbers of both UMC- and EMP-centers in peripheral areas, while the more central EMP- and UMC-centers appear fragmented. Most EMP- and UMC-centers identified by the **LR** are smaller, and beyond the primary centers they consist of individual patches of one to three cells.

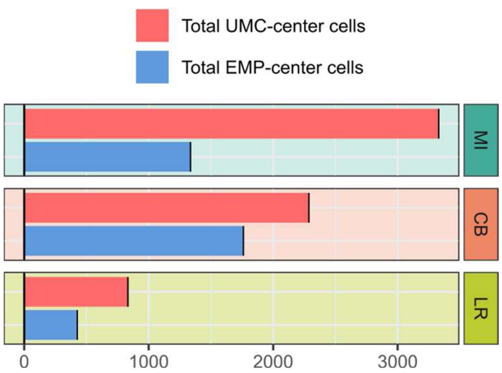


Figure 14: Total number of the study-wide sum of center cells as detected by each algorithm. For comparison, the total number of cells within all grids is 23247.

I supplement the visual impression by counts of the EMP-center cells and UMC-center cells. For complete counts per site, I refer to appendix A.

In line with my visual observation, all algorithms identify a remarkable gap between UMC-center and EMP-center counts (Figure 14). Consistently, the UMC-center counts are noticeably higher. It is interesting that this gap in counts is smallest for the CB.

Of course, counts of the two groups alone, without reference to their spatial location, only paint an incomplete picture. Hence, in addition, it is important to investigate whether the two groups agreed spatially. To quantify this agreement, I use the recall, precision, kappa, and IoU metrics of accuracy (Figure 15).

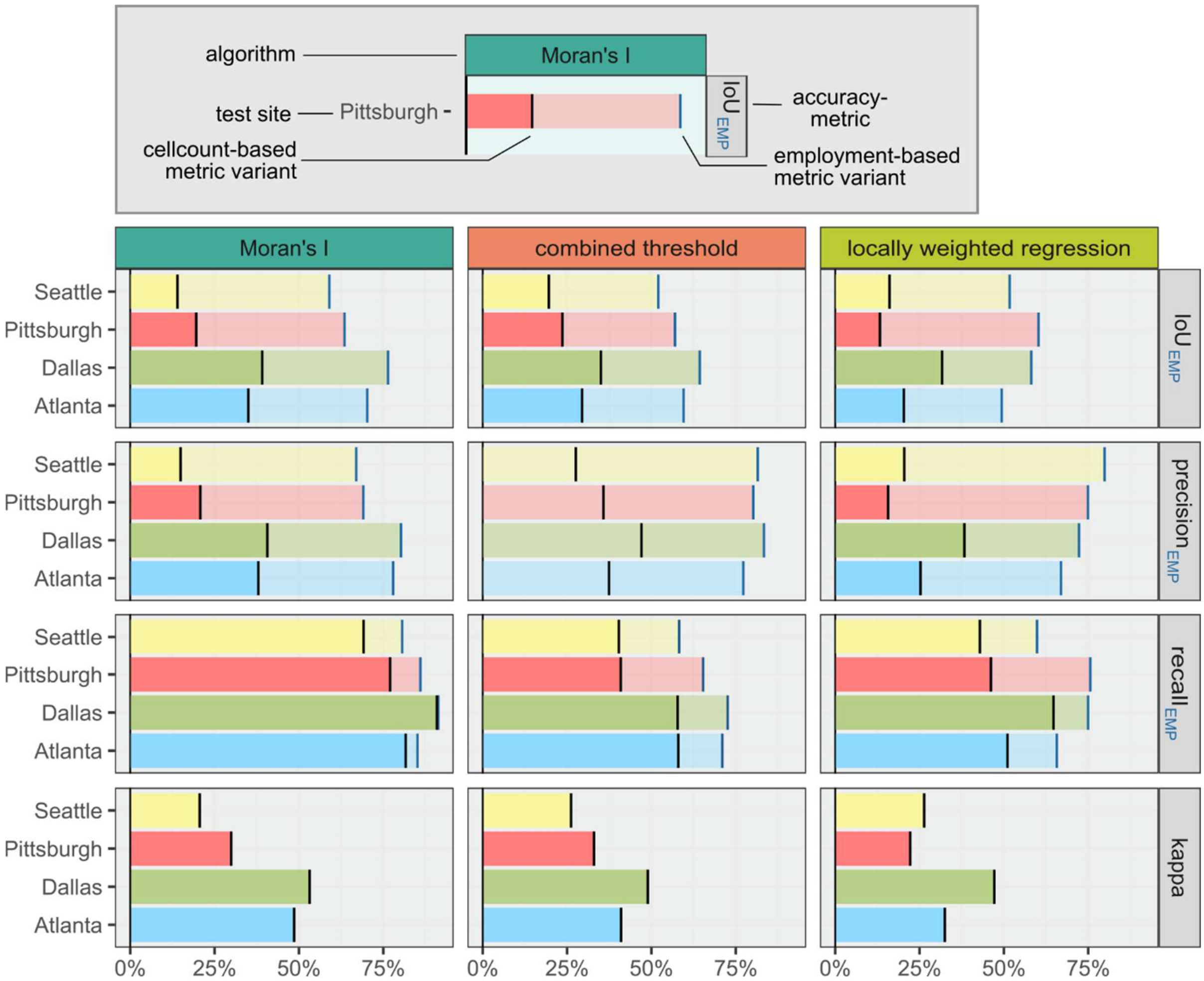


Figure 15: Cell-based accuracies as Intersection over Union (IoU), precision, recall, and kappa. For IoU, precision, and recall, the employment-based accuracies are indicated by pale bars and blue lines.

Again, several patterns emerge independently of the algorithm:

Even a cursory glance reveals a striking difference between the cellcount-based metrics and the employment-based metrics. The latter are consistently much higher. This indicates that the errors are much less significant in terms of employment than in the area they cover.

Comparing the cell-based metrics IoU_{Area}, $\text{precision}_{\text{Area}}$, and $\text{recall}_{\text{Area}}$ reveals that the errors vary considerably by metric. The $\text{recall}_{\text{Area}}$ is high, thus suggesting that the omission errors are few in number. It follows that of the EMP-center cells, most were captured by the UMC-centers. On the flipside, the comparatively low $\text{precision}_{\text{Area}}$ does indicate that commission errors are considerably more common. There are, apparently, many UMC-centers which falsely include cells that are not detected as EMP-centers.

Both of these findings align with my visual impression that while the cores of the centers are rarely not detected by the UMC-centers, the extent of the EMP-centers is often overestimated.

Turning now to the employment-based $precision_{EMP}$ and $recall_{EMP}$, I find that they do not display a similar disparity as their cell-based counterparts $precision_{Area}$ and $recall_{Area}$. The difference between the former is not substantial. Likewise, they express much lower differences between test sites.

By contrast, the cellcount-based accuracies are subject to considerable differences in agreement between test sites, as measured by IoU_{Area} and Kappa. The highest agreement is reached for Dallas, with values of 39.1%, 35.0%, and 31.7% for MI, CB, and LR respectively. The lowest agreement was found for Seattle, with IoU_{Area} values of 19.6%, 16.1%, and 14.0%. The kappa values confirm this pattern. An important observation is that no single algorithm outperforms both others in terms of accuracy. As measured by IoU_{Area}, the MI performed best for Atlanta and Dallas while the CB performed best for Pittsburgh and Seattle. The CB is also the most robust of the algorithms, with the lowest variation across test sites, and the smallest differences between precision and recall.

So far, I used the cases of true positives (TP), false positives (FP), and false negatives (FN) merely to calculate measures of accuracy. In the next part of the analysis, I consider them and their characteristics in more detail.

4.2 Systematic analysis of morphological properties of correctly and incorrectly detected employment centers

Question: Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers?

Each case of a TP, FP, or FN is a square kilometer of urban fabric with its own morphologic characteristics and geographic situation. There can be little doubt that each of these cells contains elements and patterns that facilitate or hinder algorithmic detection of EMP-centers. By examining the characteristics of these cells in detail, I seek to reveal these elements.

Table 3: Economic (blue) and morphologic (red) properties of the processed cells

	case	number of cells [n]	90% height quantile [m]	built-up density [%]	UMC [m³]	EMP [n]
MI	FN	195	(G) 7,66	(F) 0,11	453.040	(B) 1.624
	FP	2.189	9,45	0,18	943.768	326
	TN	27.824	4,68	0,01	25.808	9
	TP	1141	(G) 9,16	(F) 0,25	1.156.905	(B) 1.819
CB	FN	856	7,05	0,12	483.438	978
	FP	1.373	(D) 11,08	(E) 0,20	1.135.929	264
	TN	28.206	4,73	0,01	28.321	9
	TP	914	(D) 10,18	(E) 0,26	(A) 1.313.805	(A) 1.795
LR	FN	186	8,16	0,20	836.831	2.691
	FP	593	11,99	0,28	(C) 1.614.996	845
	TN	30.330	5,02	0,01	42.364	12
	TP	240	11,88	0,34	(C) 1.939.354	4.421

Firstly, without consideration for the spatial situation, I calculate the different cases' median values of morphologic and economic characteristics on a cell basis, again, by algorithm (Table 3). Comparing morphological and economic characteristics of TP, FP and FN reveals a pattern which can be clearly seen across all algorithms: Correctly detected center cells (TP) exhibit both a pronounced morphology as well as high employment (e.g. Table 3, A). The cases of omission errors, the FN, differ in that they also exhibited high (though slightly lower) employment numbers, (e.g. Table 3, B) but with much weaker morphology than TP or FP. By contrast, the FP, cases of commission errors, exhibit a strong morphology that is almost as high as for TP (e.g. Table 3, C), but contained lower numbers of jobs than TP or FN.

What is interesting is that the FP actually surpass the TP in height, but not in density (e.g. Table 3, E). Also, the comparative morphologic deficiency of FN is expressed in density and volume (e.g. Table 3, F) much more than in height (e.g. Table 3, G).

As aggregating morphologic and employment properties of a large number of cells to their median is a simplification that may conceal crucial insights, I further examine the full distributions in boxplots. For brevity, I only present the aggregated distributions across all sites in Figure 16. The full distributions for each of the test site can be found in appendix B.

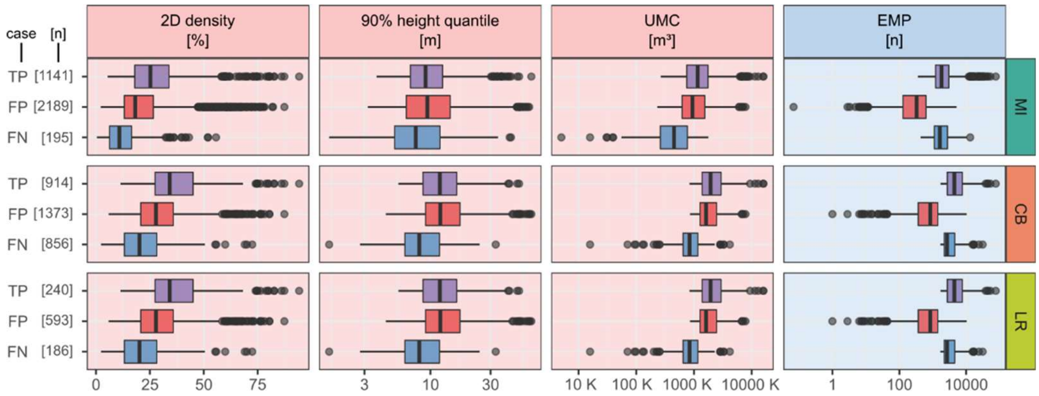


Figure 16: Distribution of morphological and employment properties of true positive (TP), false positive (FP), and false negative cells (FN).

This inspection of the distributions further confirms the previous findings. The TP display strong morphologic and economic properties, while, comparatively, the FP exhibit lower employment and the FN manifest lower morphologic measures. This pattern holds for all test sites and across all algorithms. An exception is found only for the Pittsburgh site, where, in the case of MI and LR, the built-up density distributions of the FP and FN centers are not as clearly separable.

In addition, the boxplots reveal noteworthy outliers: Particularly, there are multiple cases of FP centers which contained very low numbers (0-10) of employees. Some of the FN contain built-up densities of close to zero, meaning that they do not contain any significant measured built structures.

The explorative analysis of statistical properties, morphologic and economic, already reveals several patterns, such as the FP's tendency toward built-up areas with little employment. But it only provides first hints of understanding. What do these patterns look like in reality? And which geographic environment produced them? Answering these questions requires a more detailed examination that goes beyond local statistics. To illustrate my findings in sufficient detail, I present for each of the cases TP, FN, and FP a set of three sites (Figure 17), which exemplify morphological properties that may contribute to agreement or disagreement between UMC and EMP-centers.

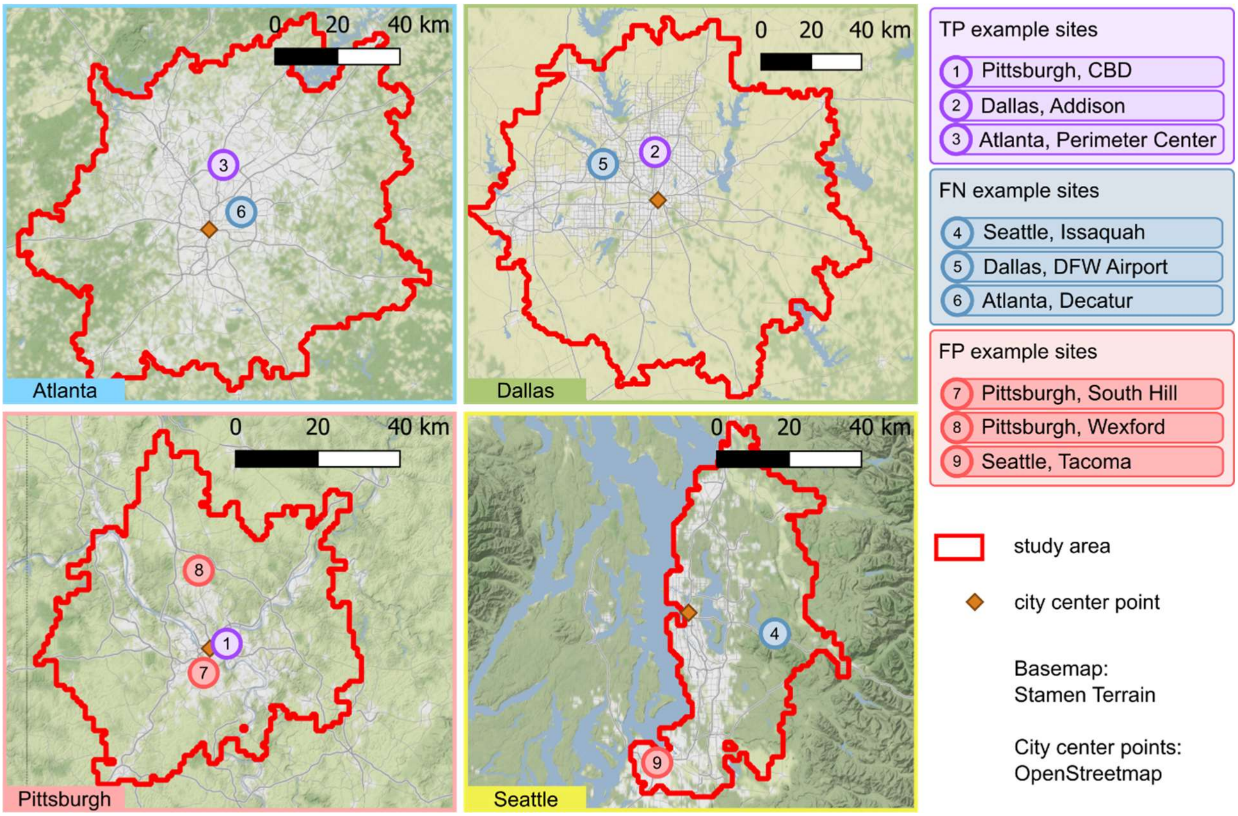


Figure 17: Locations of the example sites.

I begin with three sites which show representative examples of agreement (TP).

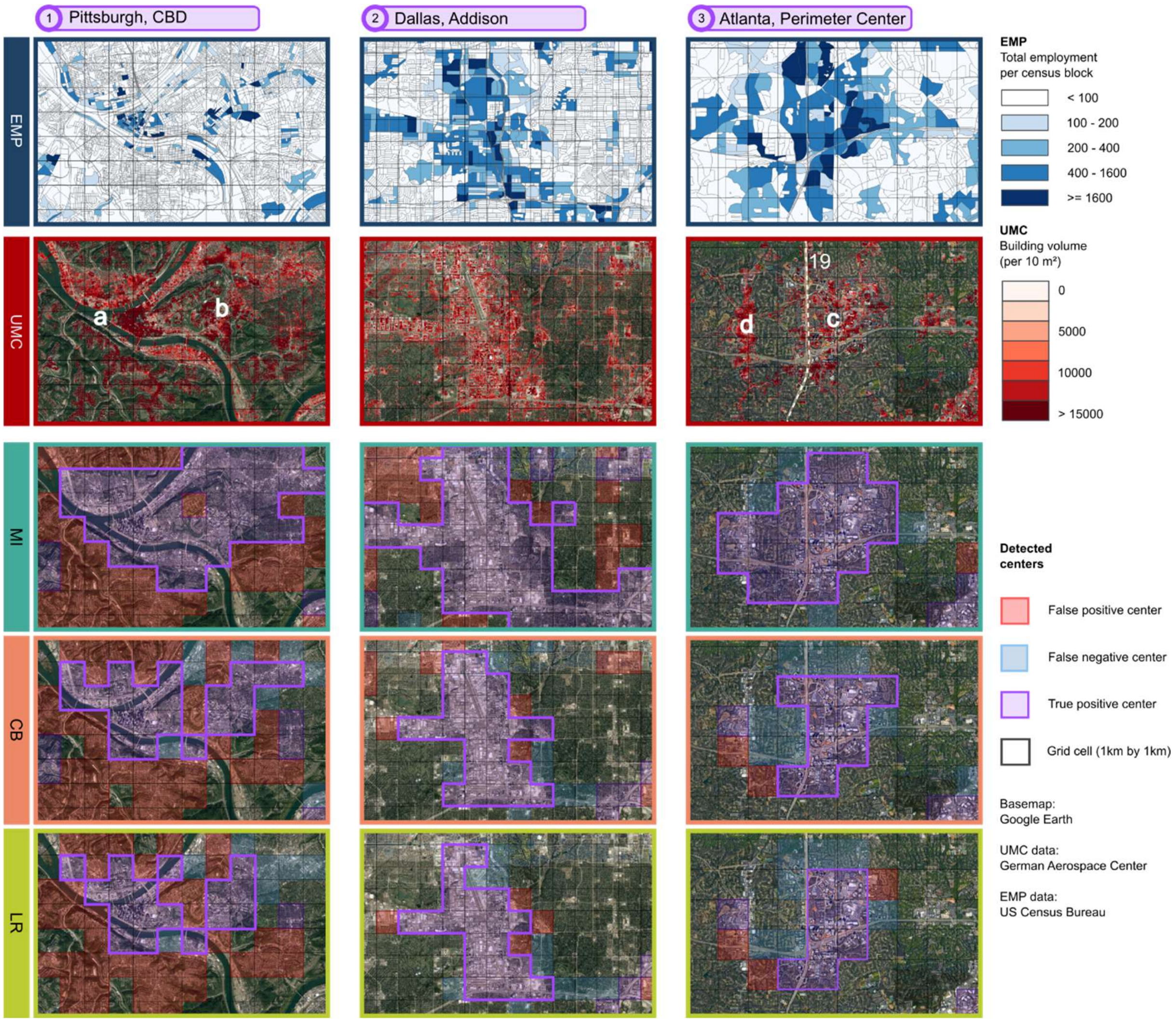


Figure 18: Example sites of true positive centers.

The *central business district* (CBD) of Pittsburgh (a in Figure 18-1) lies at the confluence of two major rivers and features high rise buildings and a dense transport network. East of the CBD lies a secondary center, the neighborhood of Oakland (b in Figure 18-1). EMP numbers are high for census blocks in these two areas, which also contain large, connected patches of high UMC.

Addison, Dallas (Figure 18-2) is a major business hub where, due to commuting, the daytime population noticeably exceeds the residential population (Destination DFW 2013; Cleargov.org 2021). Addison's built-up areas are not particularly high, but they are expansive. Census blocks within these built-up areas also show moderate to high employment numbers.

Perimeter Center (c in Figure 18-3) is a large commercial center north of Atlanta. Lying at the interchange of two important highways, it displays all the signs of an edge city, down to its name (Garreau 1992). A mix of well connected, expansive low rise retail buildings and high rise, spaced out office buildings provides workplaces for several thousand employees. All algorithms show agreement between UMC-centers and EMC-centers for the Perimeter Center east of the highway (19), the MI further shows agreement for the central axis of the town of Sandy Springs (d).

These three cases exemplify a typical pattern that is consistent across all algorithms: Most cases of correct detection are significant commercial hubs, concentrations of economic activity that are mirrored morphologically in concentrations of mid-rise to high-rise buildings. Particularly the core areas of these centers are usually well identified, while differences between algorithms exist in the quality at the fringes of these centers, where all algorithms output several FP and FN.

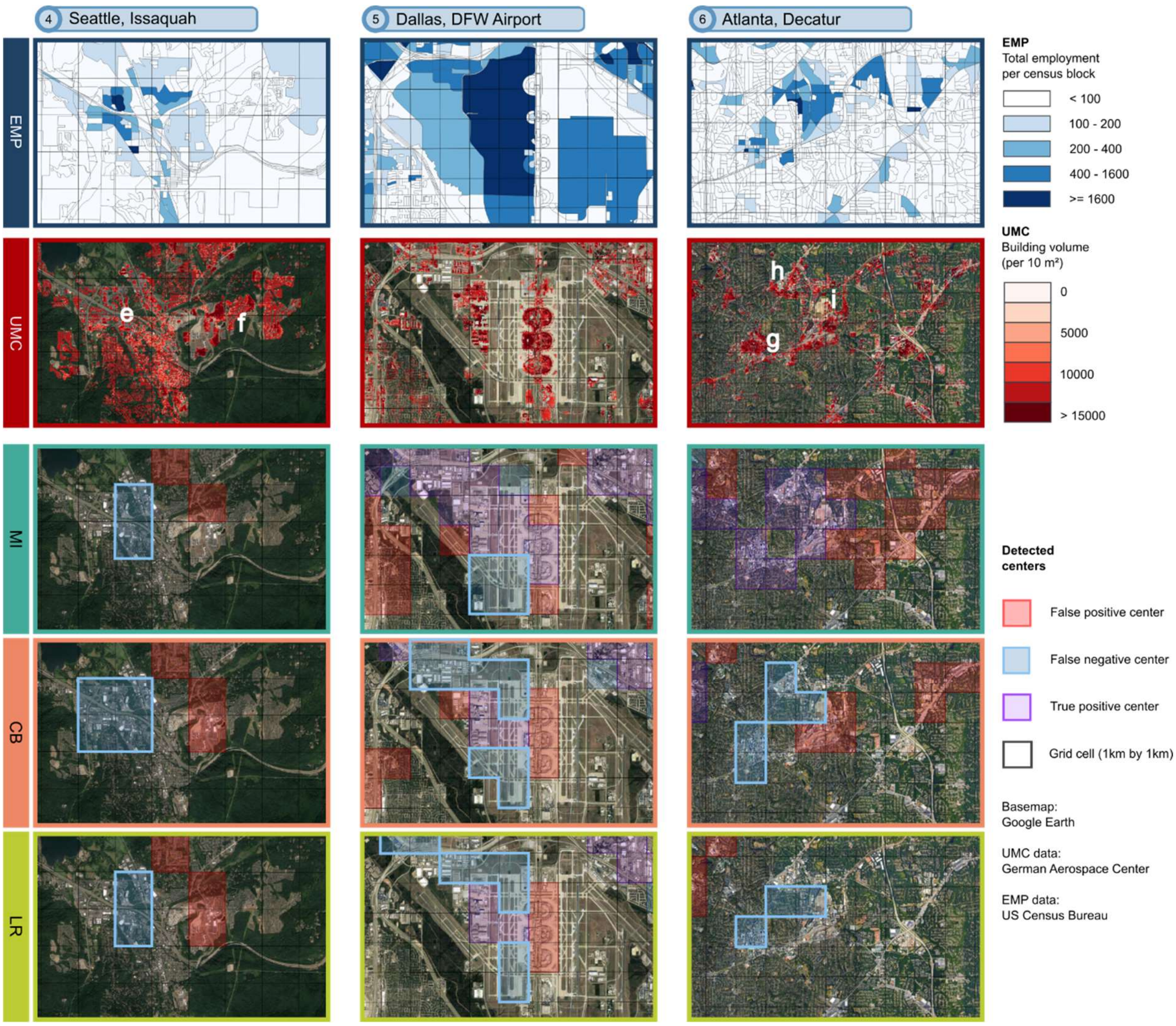


Figure 19: Example sites of false negative centers.

The second set of sides illustrates errors of omission, the FN. *Issaquah* (Figure 19-4) is a town located in a mountainous area southeast of Seattle. Along the highway at the western valley floor (e) lie the town's commercial areas surrounded by urban green– with high employment numbers compared to the other parts of the town. Densely built residential neighborhoods with very little employment lie on the surrounding hills (f). All algorithms fail to correctly detect the commercial areas as centers based on UMC data. It is remarkable that instead, they consistently falsely identify the build-up residential areas in the proximity as centers.

Dallas/Fort Worth International Airport (DFW, Figure 19-5, one of the busiest airports in the world (DFW 2021) with its own zip code, police, and fire departments, covers an expansive (~70 km²) area. Built-up volume is concentrated on the terminals. EMP values are high for the airport. While the eastern terminals report few to no jobs, a single census block, which includes multiple km² of the unbuilt western runways in addition to one western terminal, contains most of the jobs of the airport. It appears that, while built-up volumes are concentrated at the terminal's location, employment numbers are registered for the whole area covered by this large census block. In all algorithms, many cells covering this block are FN, and only some cells which include parts of additional built-up structures are identified as FP.

Decatur (g in Figure 19-6) is a city approximately 9 km east of Atlanta and the seat of DeKalb County. The downtown of Decatur presents itself as a smaller concentration of built-up volume in the form of low-rise and mid-rise buildings, intersected by parking spaces. In addition to retail and services, downtown Decatur notably registers the high number of 6 000 employees in DeKalb County's administration building (DeKalb County 2019). Other commercial areas are found in North Decatur (h) and its eastern neighbor Scottdale (i). All algorithms consider Decatur and North Decatur to be EMP-centers, but only MI also recognizes it as an UMC-center, while CB and LR produce falsely negative detections.

I find that FNs tend to be situated in spaces where a significant number of jobs was registered for relatively few, rather scattered buildings which are intersected by non-built-up areas, such as parking lots and green spaces. These configurations appear to be common for smaller, peripheral centers that are of administrative, logistic, or cultural significance. With remarkable frequency, FNs are also located in proximity to other morphological densifications.

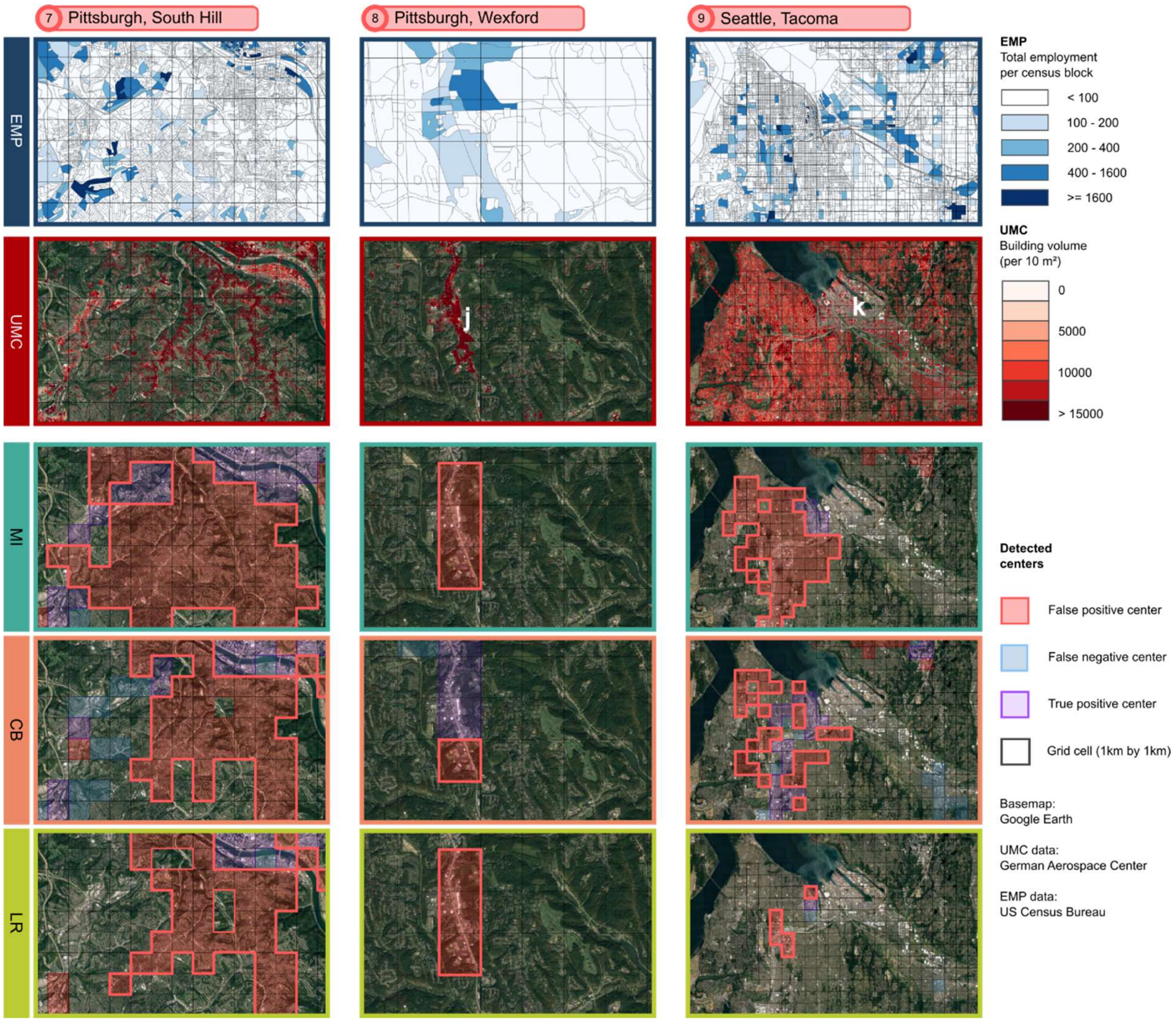


Figure 20: Example sites of false positive centers.

The last set of cases exemplifies the errors of comission, the FP. The *South Hill* neighborhoods of Pittsburgh (Figure 20-7) lie south of the CBD (compare a in Figure 20-1) and are covered with expansive residential areas. While the latter are not particularly dense, very high building heights result in comparatively high UMC values per cell. Most census blocks report no employment. Across all algorithms, I find these residential neighborhoods to be extensive patches of FP.

On the other side of Pittsburgh, 20 km north of the CBD, lies the community of *Wexford* (j) along a highway (Figure 20-8). While the residential areas are sparsely built and interspersed with greenery, the commercial and retail areas which lie on a strip along the highway reach very high UMC densities as well as heights comparable to Pittsburgh's CBD, but at most moderate employment. This strip is falsely committed as a center by all algorithms. Only CB recognizes parts of it also as an EMP-center.

Tacoma (Figure 20-9) is a port city southwest of Seattle, at the southwestern edge of the Seattle study area. Tacoma covers many km^2 of purely residential neighborhoods that are made up of buildings of moderate height arranged along regular street patterns. Most census blocks in the residential areas report no employment, although employers such as schools and local retail exist along the main roads. A major employer in the region is the Port of Tacoma (k in Figure 20-9, Port of Tacoma 2014) and much of Tacoma's economy is situated along the transport axes connected to the port. As in the South Hill neighborhoods of Pittsburgh (Figure 20-7) parts of the residential land are falsely positive. The picture varies depending on the algorithm, with the MI producing a large, connected patch of FP, the CB producing scattered patches of FP around correctly identified centers, and the LR only detecting comparatively few cells.

My review of the FP reveals that they often are areas with high to moderately high buildings at a moderate built-up density. These are often surprisingly homogenous in their structure and occur in a variety of different settings. FP can occur at the fringes of true economic centers, where they appear like an overestimation of its area. However, they also occur, even in large patches, in settings where no significant economic centers are present. Among algorithms, the MI approach appears the most prone to producing the latter. Residential suburbs are the most common representative of this type of setting. In some cases, individual cells appear as single, isolated commission errors. These are most commonly produced by the LR.

In conclusion, the analysis of the morphological and geographical properties reveals that true positives combine strong morphology with high employment and are common in CBDs or edge cities. There, economically relevant buildings are densely packed, often over large areas. False positives exhibit only high morphology with few jobs and are often produced by residential suburbs. These suburbs can contain surprising built-up heights and often register no employees. False negatives exhibit high employment with comparatively weak or disperse morphology, common for peripheral subcenters of logistic or administrative significance. In these subcenters, buildings are still of considerable height, but they are often interspaced with non-built-up areas such as parks and parking lots.

So far, the analysis of the results focused mainly on the qualities of the center detection. In the next section I finally consider the application for which the centers are detected in the first place: The analysis of polycentricity.

4.3 Feasibility assessment of analyzing job distributions in city regions using UMC data

Question: Which analyses of polycentricity are feasible using purely UMC data?

As explained in the introduction, there are many different ways to approach polycentricity and, accordingly, a vast number of measures to analyze the center structure of cities has been developed over the years. In this study, I calculate one representative measure for each of the following center structure characteristics:

- **Importance of centers**, as measured by the share of a site's employment captured by its centers.
- **Centrality**, as measured by the largest patch index (LPI).
- **Center structure and hierarchy**, evaluated by the center's rank-size distribution compared to Zipf's law.
- **Spatial distribution of centers**, as measured by the mean nearest neighbor distance.
- **Clusteredness**, as measured by global Moran's I.

All these measures can be calculated on the basis of both UMC and EMP data. The primary focus of this assessment is the agreement between the two results as the key indicator of feasibility. Only if the two results are similar across test sites and across algorithms, UMC may be an adequate substitute to EMP for a certain measure. Differences which are systematic across test sites could be compensated for by a correction factor. Thus, such systematicity is a secondary result. Context for the scale of the differences and consistencies between EMP-measures and UMC-measures is provided by the measure's variability across test-sites, which ideally is much higher. The leading question could thus also be phrased as: *For which measures is the variability between EMP and UMC clearly lower than the variability across test sites?*

Importance of centers

Firstly, I compare the *EMP-shares* of UMC-centers to those of EMP-centers (Figure 21). A large share of employment numbers is captured within both UMC- and EMP-centers in all test sites. Agreement between the EMP-share of UMC-centers and EMP-centers is, in most cases, moderate to good. Although some differences between the two exist, these differences are much more consistent than the measure's variations across cities and algorithms. Amongst algorithms, the MI overestimates the importance of centers, as indicated by the EMP-share of UMC-centers being higher than the EMP-share of EMP-centers, while the CB underestimates it. The LR shows remarkably good agreement, with the exception of Seattle, which shows among the largest disagreements also for MI and CB.

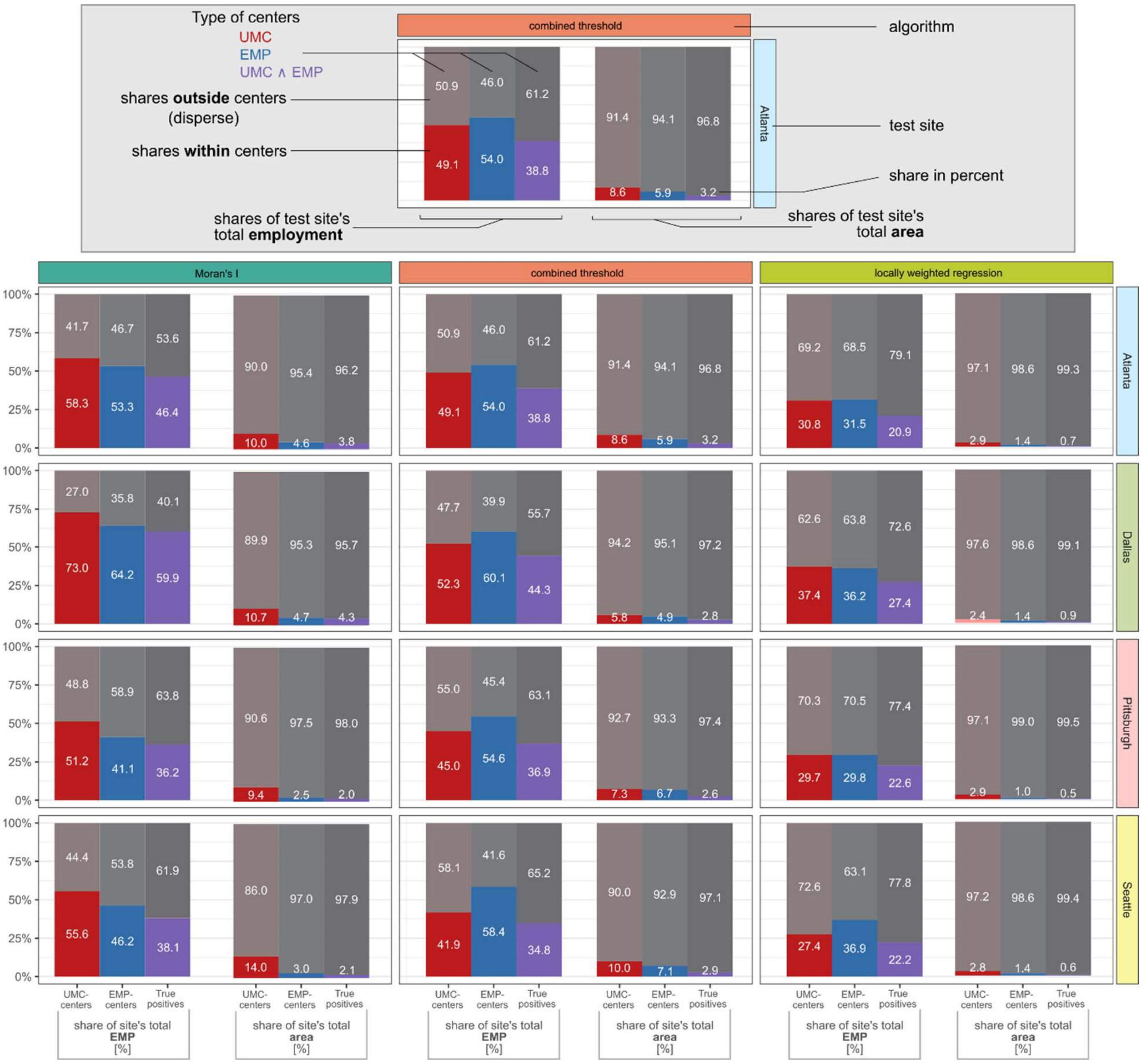


Figure 21: Shares of employment and area covered by centers.

Secondly, I compare the *area-shares* of UMC-centers and EMP-centers. Unlike the good agreement in EMP-share, the area-share shows large differences between the two sets of centers: The area-share of UMC-centers is commonly more than two times higher than the area-share of EMP-centers.

Thirdly, I compare the employment-share measure to the area-share measure. The agreement between UMC and EMP is much better for the former (Figure 21). That the area-based measure displays higher disagreement between EMP and UMC inputs is in line with the previous finding (compare section 4.1) that employment-based accuracy metrics report much higher accuracy than cellcount-based metrics. Further, I consistently observe that the center's area share is much lower than their employment-share. For instance, the EMP-share of UMC-centers detected by MI in Dallas is 73,0% while the corresponding area-share is only 10,7%. Both UMC-centers and EMP-centers capture remarkable shares of employment in comparatively small areas, suggesting that the identified centers have high employment densities.

Altogether, this measure shows that, across algorithms, the UMC-centers and EMP-centers are comparable in economic importance. This supports, though not explicitly, the feasibility of direct estimation of employment-share via UMC.

Centrality

As a complement to the general importance of centers, the LPI measures the importance of the largest center. Again, similar values of the UMC-based LPI and the EMP-based LPI would support the idea that the dominance of the largest center can be measured by UMC as well as EMP.

But no such similarity can be consistently observed. As is apparent from the Figure 22, for all algorithms, there are obvious gaps between the UMC-based LPI (red) and the EMP-based LPI (blue). Furthermore, there is obvious variation also between the algorithms as well as between test sites. In some cases, the UMC based LPI is clearly larger (CB Pittsburgh, MI Atlanta), while in other cases (Dallas CB, Dallas LR) the EMP based LPI is larger. Overall, these results display great variation and little consistent agreement between UMC and EMP.

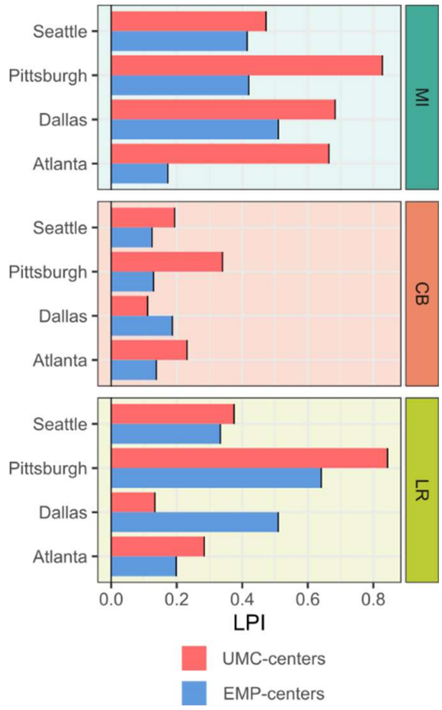


Figure 22: Largest patch index (LPI) for each algorithm and test site. The LPI is defined as the share of the area largest center among the area of all centers.

Center structure and hierarchy

In the plots of rank-size distributions, the degree of alignment between the centers and the idealized Zipf line indicates to what degree the detected centers follow Zipf's law of polycentricity. If centers lie above the line, that is an indicator for polycentricity; vice-versa, centers below the line indicate monocentricity with a dominance of the core city. Figure 23 presents the Zipf plots by the EMP value and EMP rank per center.

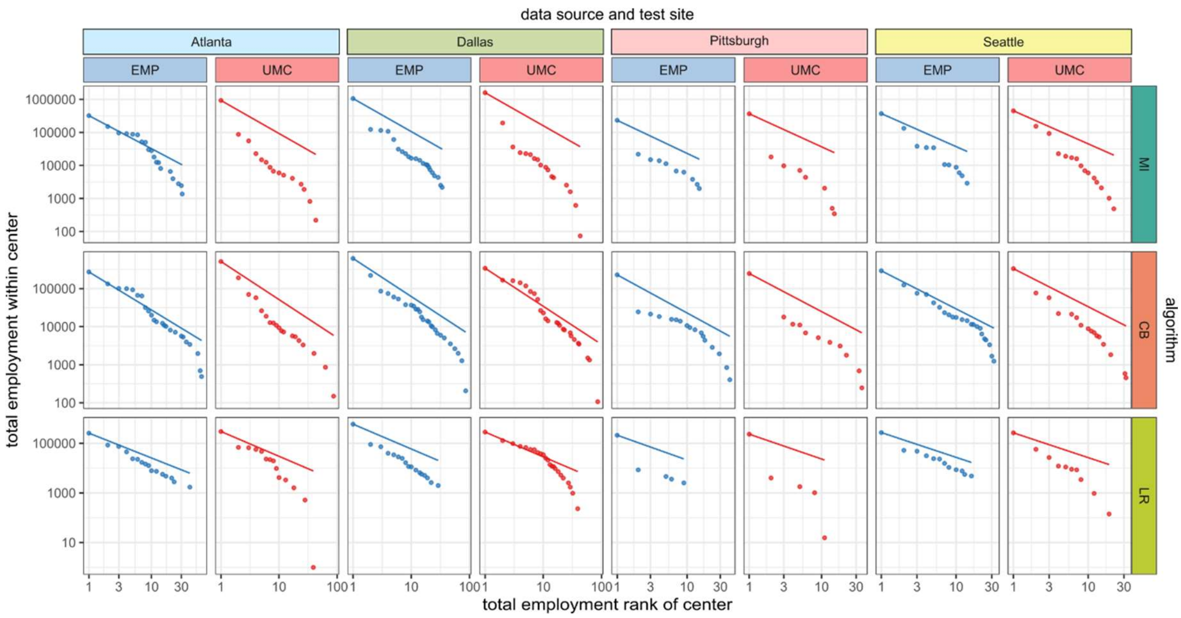


Figure 23: Rank size distributions on a per-center basis, for each algorithm and test site. The line indicates a distribution according to Zipf's law.

My findings generally indicate that distributions based on EMP-centers and those based on UMC-centers do not display similar relationships with regard to Zipf's law. In almost all cases, the majority of points (centers) lie below the line, thus indicating monocentricity. Yet, despite this apparent agreement, the distributions are still subject to high variations. In the case of LR-Dallas, the EMP-centers lie clearly below the line, suggesting monocentricity, while the distribution of the larger UMC-centers fits the Zipf line very well, suggesting polycentricity. In the case of LR-Pittsburgh, both the EMP-base and UMC-based center distributions are clearly monocentric. Due to such differences in results across algorithms and sites, the dissimilarity can also not be considered systematic. Appendix C contains alternative rank-size distributions by area and UMC. In no case, however, do consistently similar patterns for EMP-centers and UMC-centers appear.

Spatial Distribution of centers

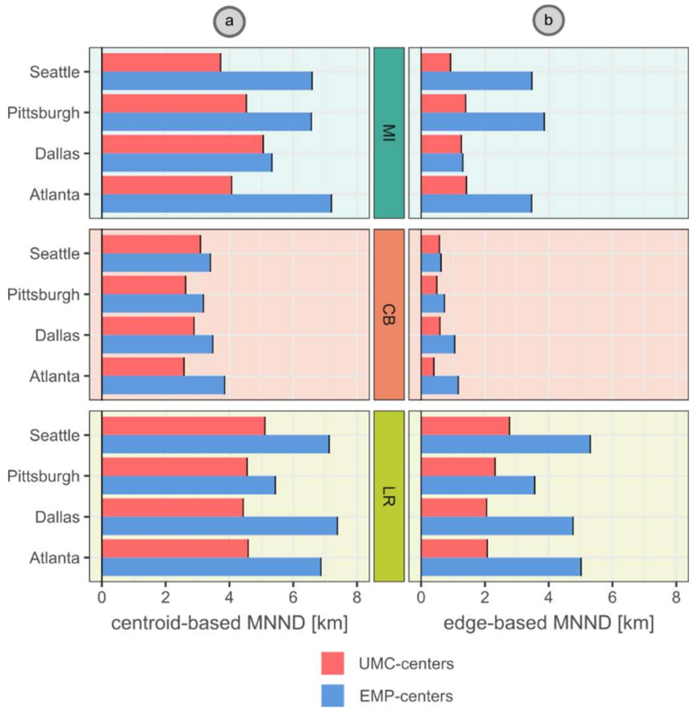


Figure 24: Mean nearest neighbor distances for each algorithm and test site.

The mean nearest neighbor distance (MNND) was calculated once between the center's centroids (Figure 24a) and once between their edges (Figure 24b).

Results show that, consistently, EMP-centers are spaced further apart than between UMC-centers.

Interestingly, agreement is much better for the centroid-based MNND than for the edge-based MNND.

Differences between algorithms are also much lower for the distance-based MNND than for the edge-based MNND. which shows great differences between the CB and the LR results, with CB centers being, on average, much closer than the LR centers.

Altogether, the agreement between EMP and UMC is not consistent for the edge-based MNND and better and more consistent for the centroid-based MNND.

Clusteredness

Independently from the detected centers, I tested whether the spatial distribution of EMP or UMC is rather dispersed or clustered using the global Moran's I. The results reveal that UMC (I of 0.3 to 0.5) are more clustered than EMP (I of 0.5 to 0.6). The magnitude of this relationship is not consistent across sites, however. Hence, it does not provide sufficient evidence of agreement between EMP and UMC. While the overestimation displays a certain consistency, it can not be assumed to be systematic.

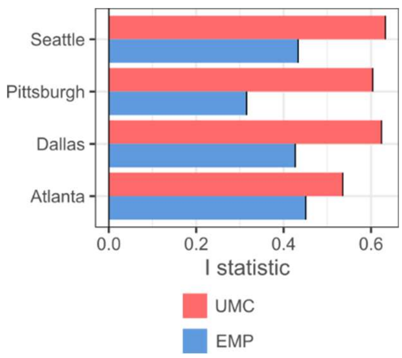


Figure 25: Global Moran's I statistic for each test site.

The results in this chapter indicate that some measures show fair, though imperfect, agreement between EMP-based and UMC-based measures. For other measures, the variabilities far exceed the consistencies.

The most salient findings of the results section can be summarized as follows:

Question: How well can employment centers be detected using UMC data?

1. More UMC-centers than EMP-centers are detected.
2. The commission error greatly exceeds the omission error, suggesting overestimation of centers.
3. Errors are much less significant in terms of employment than in area.
4. No algorithm exceeded all others in accuracy consistently across all test sites.

Question: Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers?

5. True positives combine strong morphology with high employment. False positives manifested only high morphology; false negatives exhibited only high employment.
6. True positives are common in downtowns or edge cities. False positives are produced by dense suburban neighborhoods, while most false negatives correspond to peripheral subcenters of logistic or administrative significance.

Question: Which analyses of polycentricity are feasible using purely UMC data?

7. Most measures show only limited consistency between UMC-centers and EMP-centers.
8. Employment-based measures, and, to a lesser extent, location-based measures show fair and consistent agreement.

Altogether, the results in this chapter indicate that agreement between EMP-based and UMC-based clustering is dependent on both methodology and geography, as well as the chosen measure. In the next chapter, therefore, I discuss the interpretation of the results under consideration of these aspects.

5. Discussion

The goal of this study was to explore whether remotely sensed UMC are a suitable surrogate for employment data in the analysis of urban polycentricity. Previous studies (Krehl 2015b; Taubenböck et al. 2017; Standfuß et al. 2020) explored the feasibility of polycentricity analyses based on morphologic data. They showed the potential of remote sensing to detect centers on a morphological basis. However, the degree to which the distribution of remotely sensed morphology indeed matches that of employment data, was, so far, unknown. To fill this knowledge gap, the present study detected centers based on UMC and EMP data and compared the resulting spatial distribution in a three-fold analysis.

Overall, I find that a large proportion of centers are detected by the morphological approach. There is, however, a significant spatial overestimation which negatively affects measures of area and polycentricity. Location-based and employment-based metrics, on the other hand, are in much better agreement, and indicate that major economic centers are well identified.

In the following sections, I will address each part of the analysis in turn. Unless otherwise specified, I only discuss patterns which are consistently found across all algorithms. Differences between the algorithms will be specifically discussed in section 5.4.

5.1 Evaluation of agreement between detected EMP- and UMC-centers

The first part of the analysis explored whether UMC-centers generally agree with, or overlap, EMP-centers. To this purpose, I used quantitative metrics of agreement supplemented by a visual analysis.

Table 4: Key findings of the evaluation of agreement

Question: How well can employment centers be detected using UMC data?
1. More UMC-centers than EMP-centers are detected.
2. The commission error greatly exceeds the omission error, suggesting overestimation of centers.
3. Errors are much less significant in terms of employment than in area.
4. No algorithm exceeded all others in accuracy consistently across all test sites.

This visual analysis reveals that the main center and the major subcenters at every site are always well identified: Their locations plausibly align with the background information I gathered on the test sites (compare Section 2.1). Nonetheless, it is apparent that the extent of centers is usually overestimated by the UMC approach.

The quantitative results concur with this visual impression of overestimation. *Firstly*, the simple count of UMC-center cells is much higher than its EMP equivalent. *Secondly*, the comparison of the cellcount-based accuracy metrics shows a recall$_{Area}$ which is consistently much higher than the precision$_{Area}$, particularly for the results of the MI. This means that while the UMC approach omits only few economic centers, it also falsely identifies many other cells as centers. Does this trade-off constitute a good agreement? As to my knowledge, no other study has attempted a quantitative validation of morphologic data with economic reference data, I cannot make direct comparisons. However, results of center detection have been validated in various cases. One possible benchmark is the study by Cai, Huang, and Song (2017) who use a combination of nighttime imagery and social media check-in maps. They report a recall of 83.3% - 88.2 % by comparing the centers detected by their method with governmental master plans. Further, via an additional validation on POI points they report best kappa values of 0.49. This is exceeded by the best kappa values in my study which reach up to 0.53, which indicate moderate agreement according to Landis and Koch (1977) and weak agreement according to McHugh (2012). Another possible benchmark is provided by Grubesic, Wei, and Murray (2014) who evaluate the performance of cluster algorithms on a synthetic dataset. For most algorithms they test in their idealized environment, the precision and recall reach over 90%. Such high levels of recall$_{Area}$ are reached by my study in some cases, such the MI's detection on the Dallas site. Yet, in contrast to the benchmark study, this high recall$_{Area}$ coincides with a comparatively low precision$_{Area}$ of under 50% that indicates comparatively high numbers of commission errors.

The results of the cellcount-based metrics stand in contrast with the strikingly good employment-based accuracies, which regularly exceed them more than twofold (Figure 15). Of particular note is the good precision$_{EMP}$. What to make of this discrepancy between spatial and economic accuracy? One interpretation is that the spatial disagreements between EMP-centers and UMC-centers are much less significant when seen through an economic lens. In particular, the spatial overestimation does not seem to have great economic relevance in terms of employment numbers. By extension, this also suggests that in reality, many cases of the aforementioned commission errors are cells with little economic activity. And indeed, my findings of the morphological analysis (section 4.2) confirm this by finding FP predominantly in exclusively residential areas of surprising density. But further, even the erroneously omitted EMP-centers (FN) are, at large, less economically strong than the correctly detected EMP-centers (TP), as is indicated by the $\text{recall}_{\text{EMP}}$ being higher than the $\text{recall}_{\text{Area}}$. Altogether, the contrast between the cellcount-based and employment-based metrics strongly indicates that despite considerable spatial disagreement, the detection is actually good for the major economic centers, and, vice versa, that most of the errors are of lower economic than of spatial relevance.

Whether these error levels are acceptable in practice depends not on general benchmarks but on the specific application the detected UMC-centers are used for. Certain applications may have different requirements and tolerances for commission and omission errors (Czakon 2019). The results indicate that UMC-centers may adequately replace EMP-centers in some applications. One such application in particular, the analysis of polycentricity, will be discussed in Section 5.3. Based on my investigations so far, I suggest that the accuracy may generally support three types of use-cases: *Firstly*, applications which approach centers via their economic importance rather than their spatial properties. *Secondly*, applications that allow a certain degree of fuzziness and do not require precise overlap. *Thirdly*, applications which have a high tolerance for commission errors.

However, while the use of UMC data is plausible for these applications, the results of the study do not encourage universal substitution of EMP by UMC data. This is because despite the conditionally good agreement, the substantial spatial disagreements are not sufficiently systematic to be easily accounted for.

One advantage of the cell-based analysis which produced these quantitative results is that it is an efficient way to assess agreement between two datasets. Yet, as became apparent, a purely cell-based analysis cannot adequately capture the spatial context within which detected centers appear. Nor does it consider connectivity and distances between cells. For cell-based metrics, there is no difference between a UMC-center that was located directly next to an EMP-center -and thus might well be a result of minor scale or aggregation errors- and one that was located further away. Both cases are equally false in the cell-based metrics of accuracy (a and b in Figure 26). In practice, it is unlikely that centers conform to the binary structure imposed by a grid-based classification. Rather, like many spatial objects, they have indeterminate and fuzzy boundaries (Burrough and Frank 2020).

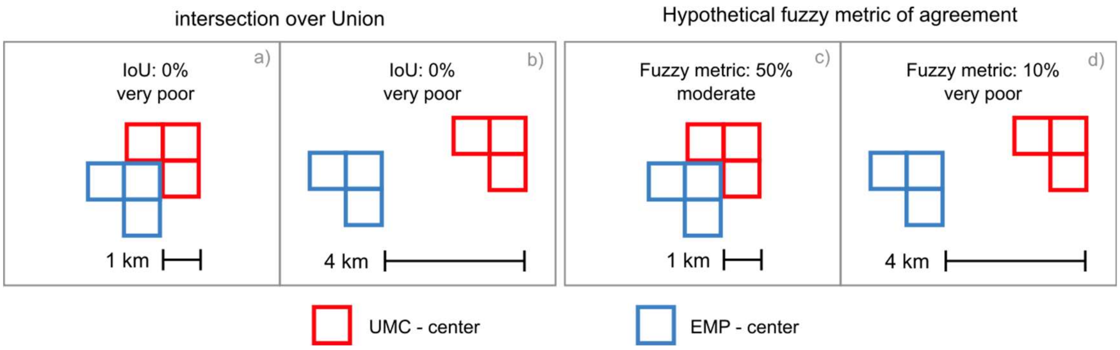


Figure 26: IoU compared to a hypothetical fuzzy metric of agreement. While the IoU is zero for anything but direct overlap, a fuzzy metric could distinguish between varying degrees of disagreement.

As an alternative which could capture this spatial complexity, I considered the development of a fuzzy metric (c and d in Figure 26) which could recognize varying degrees of agreement (Schneider 2000; Dilo 2006). However, while such a metric may be less sensitive, I find no simple way to implement it without imposing additional parameters and normativity. I instead find the simple visual analysis of center maps provides a good, non-parametric complement to the quantitative metrics - at the cost of introducing a degree of subjectivity (Taubenböck et al. 2017) and reduced reproducibility. Still, it may be that the investigation of a fuzzy concept that is polycentricity could benefit from fuzzy metrics. Future studies could develop such metrics and test them on UMC data.

5.2 Systematic analysis of morphological properties of correctly and incorrectly detected employment centers

As mentioned in the previous section, identifying the properties of errors can have practical implications. This is particularly the case if, beyond mere co-occurrence, underlying causalities can be found which explain how the errors result from local morphology or geography.

Table 5: Key findings of the analysis of error causes

Question: Which morphologic or geographic characteristics favor or hinder agreement between morphologic-centers and employment-centers?
5. True positives combine strong morphology with high employment. False positives manifested only high morphology; false negatives exhibited only high employment.
6. True positives are common in downtowns or edge cities. False positives are produced by dense suburban neighborhoods, while most false negatives correspond to peripheral subcenters of logistic or administrative significance.

The cell-based statistics of TP, FP, and FN already give some indication. They reveal that those centers which were not recognized via the morphology-based approach (FN) simply do not manifest very strong morphologic properties. Likewise, the false commissions (FP) do in fact manifest morphologies similar to most economic centers (TP and FN). This indicates that certain disparities are already present in the input data and propagate to the center detection. By extension, it appears that for certain urban areas the correlation between UMC and EMP is disturbed and diminished. In combination with the previous accuracy assessment, the morphological analysis suggests the that the proposition by Krehl (2015b) that employees work in buildings may be extended thusly: Concentrations of employees are often located in concentrations of buildings, as indicated by the high recall.

But the inverse does not hold: There are cases where concentrations of buildings are used for other things than work. This is not particularly surprising and is in line with the findings of Krehl (2016a) and Taubenböck et al. (2017) who observe that the relationship between employees and floor space is one of correlation, but not a universal perfect match. What I did not expect is the observation that these cases, represented as FP, are more common than the true economic centers (TP).

The exploratory analysis reveals that in many cases, these FP are extensive, planned residential suburbs of surprising built-up volume that are found in all four test sites. Frequently, these areas contain few to no blocks which register any employment. This agrees with the previous finding that the economic relevance of commission errors is much lower than their spatial extent (section 5.1). In terms of their location, no single pattern emerges. In some cases, they envelop true economic centers (TP), and it appears that the extent of these is merely overestimated. But in other cases, if they are sufficiently dense, they also may qualify as morphologic centers of their own. The presence of such dominant residential areas may even lead to omission errors on economic centers in their vicinity, especially if these do not themselves display strong morphology. All algorithms, but in particular the LR allows for competition between centers. In this, smaller EMP-centers with weaker morphology appear to be eclipsed by morphologically dense residential suburbs with surprising frequency. But how can this comparative weakness of some economic centers, that results in FN, be explained? It seems possible that it is in part due to certain jobs allowing for higher concentration (Krehl 2015b). This proposition warrants further explanation: Beyond the CBD and the major edge cities, many economic centers are peripheral business parks and administrative, retail, or cultural centers of limited size (<5 km^2). There, many service jobs may be concentrated in relatively small buildings dedicated to and optimized for a single use. Examples of such buildings may be malls, schools, or high-rise office buildings. The space between these buildings is left open for transport, parking spaces, and parks – which increase the areas attractivity to employees and customers, but do not contribute to the UMC. That this can be observed particularly in smaller, peripheral centers could be linked to the influence of land prices, which are lower in the periphery. In contrast, higher land prices in larger centers may discourage extensive open spaces there, provided that the traditional land value gradient holds, which is not a given in polycentric environments (Heikkila et al. 1989; Dubin and Sung 1987). This is supported firstly by the visual observation that economic centers closer to the CBD tend to be well detected, and secondly by the previous finding that the omitted EMP-centers (FN) contain fewer employees than detected EMP-centers (as indicated by the high $Recall_{EMP}$). Not all FN can be plausibly explained in this way. In some cases, certain census blocks register employment numbers far above what appears plausible considering their physical volume, as in the case of the Dallas airport terminals (Figure 19-5). It could be suggested that this is at least partly due to the LODES' nature of nominal workplace designation (Graham et al. 2014), which allows for jobs to be far more densified than buildings. Certainly, there are other explanations for such disagreements between morphology and employment, such as individual preferences for certain types of housing or varying production needs of different sectors (Krehl 2016a; Krehl et al. 2016). Future studies could investigate the underlying mechanisms, which are likely manifold. My study contributes to this effort by suggesting a blueprint for the systematic error assessment.

The causes of uncertainty that I identified in the previous section may be used as idealized patterns by an algorithm or human expert to enhance the results of center detection on UMC data substantially. Nevertheless, such *postprocessing* would reduce objectivity, and perhaps limit the comparability of the study. As a form of *preprocessing*, Krehl (2015b) experimented with ancillary land-use data to identify and eliminate residential areas from the input data. However, she found that mixed land use and limits in available data limit the power of this approach. For future research, it will be interesting to see how this study compares across different regions, particularly those with different levels of government-sponsored regulations, such as zoning policies and building restrictions (Krehl 2015b). In their absence, the relationship between residential space and commercial space may be less heterogeneous, as a split between economy and residence will be the result of various economic factors rather than enforced by policy (Anas, Arnott, and Small 1998). On the one hand, this reduces the feasibility of filtering out residential areas via land use maps, as done by Krehl (2015b). On the other hand, mixed land use might show a stronger spatial link between employment and building volume, thus increasing the suitability of UMC as a proxy for employment (Figure 27). This is, at this point, speculation. Future research is required to examine the relationship of morphology and economy across different land use types. TanDEM-X derived UMC could be used alongside detailed land use information to realize these studies.

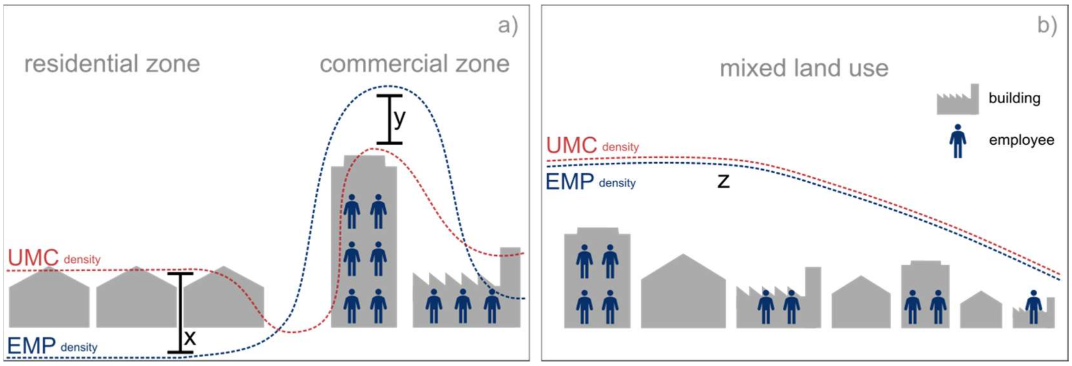


Figure 27: Idealized scenarios illustrating potential effects of zoning on the relationship between UMC and EMP, represented by smoothed densities.
Left scenario (a): Strong zoning reduces correlation between UMC and EMP. High concentration of employees in the commercial zone leads to areas of relative UMC dominance (x) and relative EMP dominance (y).
Right scenario (b): Mixed land use distributes employment and results in high correlation between UMC and EMP distributions (z).

Of course, the UMC data is also subject to errors. In their validation of the TanDEM-X derived input data, Geiß et al. (2019) report systematic underestimation of built-up heights. Particularly, they admit that, due to layover of building over ground, height information cannot be gathered in places where high buildings are at low density. Thereby, they may offer an alternative explanation of the non-detection of loosely built-up economic centers.

Notably, they also report that the estimation error of built-up density varies depending on morphology, underestimating built-up density where built-up heights are low. At first glance, this may seem counter to my observation that residential areas – if presumed to comprise buildings of low height - exhibit strong morphology. However, in my evaluation of the data set I find that some of these residential areas are not low-rise at all; rather, they exhibit built-up heights which match those of CBDs. It is possible that the aforementioned underestimation of built-up density which Geiß et al. (2019) found does not apply to these areas. That may explain why I observe high UMC values for these areas – and why they are prone to produce FP.

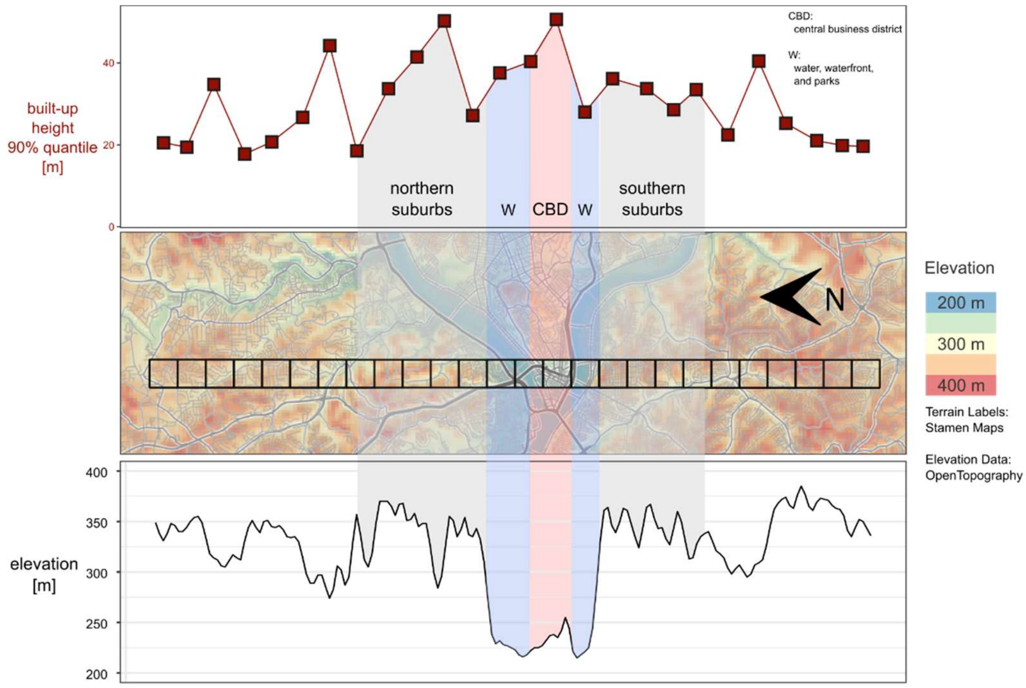


Figure 28: Built-up height and ground elevations in a transect over Pittsburgh. High measurements for built-up height occur in areas with strong elevation changes.

Why these residential areas do occasionally produce such strong height values may at least partially be explained by the application of the morphological filtering to hilly terrain. If steep hills cause local variations in elevation, non-systematic errors in nDSM estimations can arise. This is particularly the case if the bare earth pixels, between which the nDSM is interpolated, are irregularly spaced and not sufficiently dense. As a result, the nDSM may be over-or underestimated and elevation changes may be attributed to built-up height rather than nDSM elevation. This possibly occurred in Pittsburgh (Figure 28). North and south of the CBD in the valley, residential suburbs cover steep hills (which have already been discussed, see Figure 20-7). There, comparatively large elevation changes occur with local rises in built-up heights that match the heights of the CBD. The effect of such errors is limited if they are spatially homogeneous, as no algorithms which I use in this study applies a fixed threshold.

A further limitation of the dataset is that while it can be potentially derived globally, the underlying TanDEM-X data is only available for a timespan of 2012-2016, and not with the temporal regularity of survey data. This may make it less suitable for the analysis of highly dynamic regions such as China.

Finally, the economic reference must be discussed. This study relies on the assumption that centers derived from employment data represent a valid ground-truth against which UMC-derived centers can be compared. This assumption must be questioned as well: *Firstly*, the socioeconomic data may itself be flawed. Issues in data gathering and processing can occur and introduce errors into the data. *Secondly*, the plausibility of reducing the economic activity to a single location of employment is not self-evident. Unlike building volumes, economic activity is not a physical reality but a spatial phenomenon with an ambiguous spatial location. Ambiguity of workplace is already present in many sectors. It is likely to increase as skill-intensive services make up increasing parts of the economy (Buera and Kaboski 2012) and the Covid-19 pandemic is leading to an increase in remote work that can be expected to last in at least some countries and sectors (Brynjolfsson et al. 2020; EU Science Hub 2020). In some ways, this is not a new phenomenon. The separation of living and working activities developed jointly with the industrial economic model (Doling and Arundel 2020) and in many professions was not the norm for long periods of history. Yet, the otherwise very detailed LODES data does not accommodate even regular telework. All teleworking employees are assigned to the company's main address, just as if they were commuting every day. And while reports of multiple worksites are possible, they are infrequent (Graham et al. 2014). That is not to say that the LODES data is an inadequate mirror of economic activity. But it is important to be aware that the researcher's choices on how the economic activity is defined and spatially located will strongly influence the outcome of a study. In my study, there are some cases in which single locations registered curiously high or low numbers of employees. One such example is the case of DFW airport, in which one terminal was assigned 29 000 employees while others registered zero. Likely, such cases contributed to disagreement between the detected EMP- and UMC-centers. *Finally*, the algorithms, if flawed, would impose their flaws on the EMP-based centers just as well as on the UMC-based centers. Hence, EMP-centers might not constitute a perfect reference even if the EMP data itself is technically flawless.

Beyond the contribution of the input data and the algorithms, this study's data structure certainly had an impact on the results and must be discussed. The aggregation of morphologic and economic data into grid cells is likely subject to the MAUP, which can distort results and impede understanding if it is not recognized and dealt with explicitly (Jelinski and Wu 1996; Weigand et al. 2019). The use of a grid is an established method to try and reduce the impact of the MAUP and provides clear benefits for the purposes of this study (see section 3.1). However, it does not constitute a general solution and the choice of grid-cell size, while not arbitrary, is normative. Similarly, the definition of the study area requires an isochrone whose time-distance is manually chosen by the analyst. The grid cells and the study area together define a singular scale of the study. But polycentricity can appear at many different scales and in different forms (Liu and Wang 2016), thereby the choice of scale invariably affects the study's outcomes (Bartosiewicz and Marcińczak 2020). To give an example, reducing the study area of Dallas not include Fort Worth would have undoubtably affected results. And while alternative methods of defining study areas exist, they come with their own caveats (Taubenböck et al. 2019). Altogether, the influence of normativity that affects many spatial data analyses of polycentricity is also evident in this study and not easily resolved (Lee 2007). I acknowledge this limitation and discuss its impact on this study.

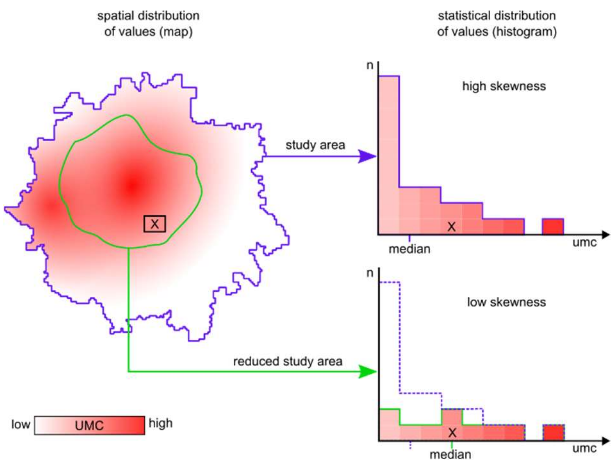


Figure 29: Responsiveness of the distribution to the chosen study area extent. The distance between the values at site X and the median changes if the study area is reduced.

The most significant impact is likely on the statistic distributions of the input data. To stay with the previous example, reducing the extent of the Dallas test site would have meant the exclusion of large numbers of unbuilt or sparsely built grid cells, thereby changing the UMC distribution substantially (Figure 29). The same is, in principle, true for the EMP distribution, although it cannot be expected to be affected in precisely the same way.

The distributions of the input data in turn affect the result of center detection algorithms (Zhang et al. 2008). Responding to these effects, researchers have proposed stabilizing the distribution by excluding outliers and applying transformations to the data (Fu et al. 2011; Zhang et al. 2008; Gimond 2013;Fu et al. 2014). I performed preliminary tests which indicated that this indeed has the potential to increase agreement between UMC and EMP. However, I suspect that such adjustments would have affected certain areas more than others, and thereby created geographic distortions. Further, effects would not have been uniform across analysis methods (such as those presented in section 5.3). Tuning the distribution to maximize agreement on a target measure would reduce the generalizability of the results in a way that could be considered *overfitting*. As the aim of my study was to compare the performance two datasets in already established methods, I did not investigate such tuning. However, I acknowledge that there is abundant room for future studies to explore preprocessing or tuning methods.

5.3 Feasibility assessment of analyzing job distributions in city regions using UMC data

The results of the feasibility assessment paint a diverse picture (Table 6).

Table 6: Key findings of the feasibility assessment

Question: Which analyses of polycentricity are feasible using purely UMC data?
7. Most measures show only limited consistency between UMC-centers and EMP-centers.
8. Employment-based measures, and, to a lesser extent, location-based measures show fair and consistent agreement.

Certain measures, such as employment-share and centroid-based MNND show good agreement that suggests they may feasibly be calculated using UMC as a proxy. Other measures, such as LPI and edge-based MNND show low agreement and high variations. Hence, UMC can not universally substitute EMP. Rather, I suggest that there are certain types of measures for which the use of UMC is feasible. The findings of the quantitative analysis, as discussed in section 5.1, also support this notion. Consequently, it is advisable that before any analysis that uses UMC, the chosen measure should be considered with care. Some aspects which should be part of such considerations are suggested by the results of this study:

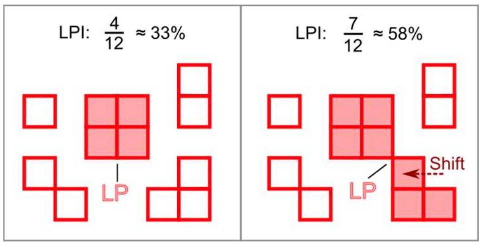


Figure 30: Responsiveness of the LPI to shifts in center detection. The size of the largest patch (LP) and hence the largest patch index (LPI) can be strongly affected by minor shifts.

Area-based measures, such as the area share of centers, the LPI, and the edge-based MNND, show poor agreement with high variability. It is likely that the high variability of the LPI is partially explained by the rather arbitrary merging of centers, which can strongly affect center size (Figure 30). Thus, the LPI has a requirement for spatial precision that is not supported by the UMC data, as can be seen by the low precision$_{Area}$ (Section 4.1). The edge-based MNND is also affected strongly by the center extent, and, thus, the previously identified overestimation. The centroid-based MNND is less affected by overestimation (as illustrated by Figure 31) and shows much better agreement. This suggests that the core location of the centers can be feasibly approximated. However, the MNND considers differences in centroid location only indirectly and in an aggregated manner, and

hence, only allows for limited conclusions about the location accuracy. Alternative measures which
quantify differences in centroid location directly would give further insight, but I did not consider them for this analysis as they do not themselves constitute a measure of polycentricity.

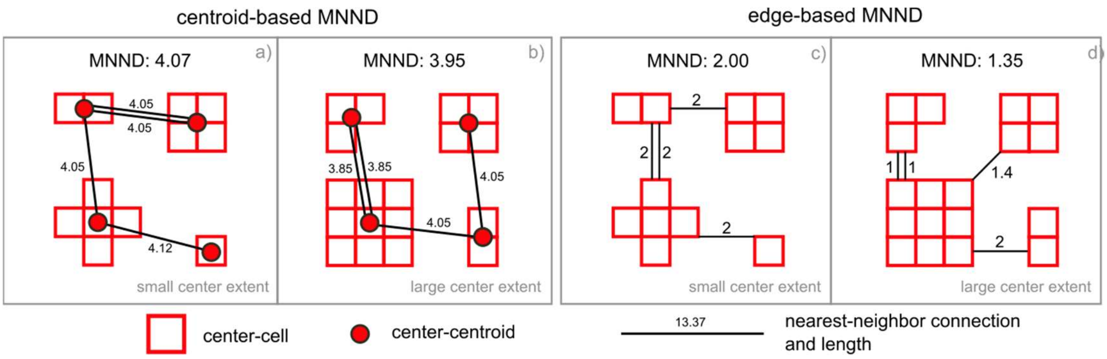


Figure 31: Responsiveness of the two MNND variants to variations in extent of detected centers. The centroid-based MNND (a&b) only shows minor responsiveness to center extent. The edge-based MNND (c&d) responds much more strongly.

A similar indifference to overestimation as observed for centroid-based MNND is shown by the global Moran's I, as it does not depend on center detection at all, but rather the distribution of the input data. When evaluating the importance of subcenters via the employment shares, the UMC-centers and the EMP-centers agree well in terms of employment share covered. Indeed, they agree much better than in area covered. This also accords with my previous observation that employment-based accuracies are much higher than cellcount-based accuracies. The good agreement can be interpreted as an indication that UMC-centers, despite being larger, match EMP-centers in importance. A note of caution is due here, as this measure cannot truly be computed on UMC alone. Thus, these findings do not indicate that the UMC-based center importance quantitively matches the EMP-based center importance. However, they do implicitly support the feasibility of subcenter importance estimation based on UMC. Altogether, these findings imply that for measures which do not depend on center area, the use of UMC as a surrogate may be possible using a correction factor. To determine this factor and its reliability, however, tests on larger numbers of cities are absolutely required and are worth targeting in future research.

In conclusion, despite substantial variations across cities and algorithms, certain measures display good agreement between EMP-based and UMC-based results, especially for measures

which relate to the employment or location of centers. This is consistent with the results of the first part of the analysis, in which the visual analysis showed good agreement for the general pattern of center distribution and the employment-based accuracies.
However, for measures which relate to area cover, the variations exceed the agreement substantially. As such, I cannot clearly attribute the occasional good agreement to causality when it might be caused by chance. While certain patterns appear to be consistent across test sites, the statistical power of four sites is insufficient to assess significance statistically. To some extent this is mitigated by the inclusion of different algorithms. Nevertheless, I refrain from statistical significance tests in favor of an exploratory perspective. I highly recommend that further studies with larger numbers of test sites perform such statistical significance tests. Further, regressions could identify which variations are caused by differences between the UMC and the EMP data, rather than test sites or center detection algorithms. Additionally, regressions could relate measures to additional explanatory factors (for an example, see Lv et al. 2020). If it can be confirmed that differences between UMC- and EMP-based measures show consistency, it is conceivable that a correction factor may be identified to the UMC. However, such a correction factor is still likely to be subject to regional variations.

5.4 Influence of algorithm

Within my test setting, characteristic patterns emerge for the centers detected by each algorithm: the **MI** identifies large patches of adjacent cells. With a particularly high recall, it identifies almost all economic centers at the cost of committing many commission errors in the process. The **CB** reaches the highest precision and appears to be particularly robust. Its results show lower variation across test sites, which could indicate that they are less influenced by site-specific variables and the dominance of local centers or the core city – as indeed was the intention behind the distance component of the CB (Taubenböck et al. 2017). The **LR** succeeds at identifying the core sites of few very significant centers, becoming less robust for smaller and more outlying centers.

I believe that the study provides sufficient evidence to suggest a link between the algorithms underlying concepts and the outcome they produce. For example, the LR enforces a certain competition between proximate centers by explicitly eliminating center candidates based on their distance or their explanatory power. No such constraint is enforced by the MI. I believe that this shows in the outcomes in which LR centers tend to be more regularly spaced while

MI centers occur in groups (as exemplified by the large differences in the MNND, Figure 24). This suggests that the detected centers follow the structure imposed by the method and reflect the algorithms' concepts of center and neighborhood. This is certainly not a new finding, as already in 1998, Anas, Arnott, and Small (1998) observed that the number and extent of subcenters are sensitive to their definition. This definition of centers, in turn, varies by algorithm: Krehl (2016b) finds that LISA and LR are based on different notions of subcenters. From these observations, I conclude that it is crucial for any study to consider the possibility that its chosen measures of urban structure are affected by the algorithm. To give an example: When using the CB algorithm to detect whether an urban region contains dispersed centers at its outer edges, one should be aware that the algorithm implicitly encourages detection of smaller, outlying centers through its distance-based component. If the analysis then reports large numbers of dispersed centers, it must be discussed to what extent the outcomes are due to true dispersion or the underlying model specification (Taubenböck et al. 2017).

To investigate such algorithm-specific details is beyond the scope of my work. But my findings are encouraging further tests using larger numbers of sites. These, in turn, would allow a comparison with sufficient statistical power to support general conclusions about the algorithms. Based on my findings within this study, I am confident that the TanDEM-X-derived UMC data would support such analyses, albeit in the scope of its own data inherent uncertainties.

In terms of accuracy, no algorithm proved to be clearly superior to all others. This is in good agreement with the general finding of Grubesic, Wei, and Murray (2014) that no spatial clustering algorithm is objectively superior. It also reinforces the idea that the choice of algorithms should be application specific. An additional consideration is comparability. Comparisons between studies should only be drawn if the algorithm is identical, as strong variations across algorithms strongly discourage comparison of results across algorithms. This is in line with the findings of Agarwal, Giuliano, and Redfearn (2012) who note significant differences across different studies with different algorithms in the Los Angeles region, with the number of detected centers ranging from 13 to 120. Even in my study, which uses precisely the same data basis for all algorithms, visible differences exist. That is not to say that there is no agreement between algorithms. In all test sites, all algorithms agree on the locations of the main center. That is not surprising as two of the algorithms (CB and LR) even presuppose the existence and the location of such a center (see section 3.2). Yet in addition, the algorithms also agree on most major subcenters.

Due to the limited number of test sites used in my study, the statistical power of these findings is limited, and they may not be generalized to other regions, particularly if those have different sizes and geographic characteristics.

In conclusion, the choice of algorithm should depend on the application. However, considerably more work is required to determine objective criteria for the algorithm selection. Furthermore, any study on urban form should also consider the algorithm's influence on the target measure. Finally, my study shows that using multiple algorithms is a plausible alternative. Beyond reinforcing the generated results, it produces a richer, more nuanced view of the city.

As conclusion to this discussion chapter, I suggest several recommendations concerning the further use of the TanDEM-X derived UMC (Table 7).

Table 7: Recommendations concerning the further use of TanDEM-X derived UMC data.

Quantitative agreement Centers detected based on UMC data can be expected to be sufficiently accurate for certain applications, such as: • Applications which focus on economic properties of centers, rather than spatial properties. • Applications with a tolerance for commission errors.
Patterns of disagreement Despite best efforts to mitigate issues of scale, aggregation, and normativity, they still affect the results, and should be expected to do so in future studies of polycentricity. Disagreements between UMC-centers and EMP-centers should not be considered as simply erroneous but reflect real and interpretable differences in UMC and EMP data. The exploration of these differences, through contrast of UMC and EMP data, may reveal valuable insights into urban spatial structure and tackle questions about our notion of workplace. Postprocessing of centers, preprocessing of UMC data through transformation and filtering based on land use, and tuning of algorithms have the potential to increase agreement further and may be fields of future research.
Analyses of spatial structure UMC can substitute EMP in certain analyses of spatial structure. The measures should, however, be chosen with care. Certain guidelines are: • Measures are preferable if they are not particularly sensitive to variability in the centers' spatial extents. • Measures of employment and location are preferable over measures of area.
Algorithms No algorithm is objectively preferable over the others. Analyses should consider the ways in which the algorithm's specification may impose its underlying center concept to the target measures. UMC data could be used to investigate algorithm-specific effects on polycentricity analyses at a large scale and with sufficient statistical power to draw generalizable conclusions.

6. Conclusion and Outlook

6.1 Conclusion

To provide the missing link between morphologic remote sensing data and socioeconomic survey data, I have devised a systematic procedure to compare the results of employment-based and morphology-based center detections. The morphology-based approach reveals plausible centers which, despite capturing the economic structure of the cities well, show substantial spatial disagreement with employment centers. I conclude that TanDEM-X-derived morphologic data can serve as a surrogate for traditional employment data in certain applications, but not universally.

Further, I find that many cases of disagreement follow systematic patterns and I suspect that there is a fair probability that many can indeed be plausibly explained by a characteristic mismatch between the spatial distributions of morphology and employment. For this reason, my work has led me to conclude that the morphologic perspective, as revealed by remote sensing, also provides a valid alternative view of the city. Its value, I believe, lies not simply in its agreement with economic data, but also in its contrast. Studying this contrast could itself provide important insights into the changing economic structure of cities in a time where there is disagreement about the future of work. I propose to see cases of a disagreement between morphologic and employment centers as not simply an error, but as a challenge to our assumptions. In that light, my results encourage us to reconsider our established concepts of workplace and of the distinction between commercial and residential spaces.

6.2 Outlook to future work

There is potential to significantly improve on this study by including more test sites. A greater number of sites could provide statistical evidence for this study's quantitative results and test their sensitivity to choices of grid size and algorithmic parameters.

As my findings also reveal differences between algorithms, it the performance of different algorithms might be a fruitful target for such a statistical analysis. This could identify which algorithms are best suited for a particular application, and to what extent the centers detected by an algorithm reflect the algorithm's underlying model architecture and specification and are biased to replicate prior notions and assumptions.

Ideally, additional test sites would be located in different geographical regions where comparable socioeconomic data is also available. This would be supported by the global availability and consistency of the remote-sensing data. As I find non-negligible variation across test sites within the same country, the performance in other geographic environments should be carefully evaluated. My study could provide a blueprint for such evaluations.

Research on cross-region comparisons is ongoing. If the global availability of TanDEM-X data can be leveraged in this manner, it opens the door for a new type of global studies of urban development patterns.

In conclusion, I recommend that research on urban economy should further explore the use of globally available remote-sensing morphologic data while being aware of – and, indeed, utilizing - its contrast with socioeconomic data.

References

Print

Agarwal, Ajay, Genevieve Giuliano, and Christian L. Redfearn. 2012. "Strangers in Our Midst: The Usefulness of Exploring Polycentricity." *The Annals of Regional Science* 48 (2): 433–50.

Anas, Alex, Richard Arnott, and Kenneth A. Small. 1998. "Urban Spatial Structure." *Journal of Economic Literature* 36 (3): 1426–64.

Anselin, Luc. 1995. "Local Indicators of Spatial Association-LISA." *Geographical Analysis* 27 (2): 93–115.

———. 1996. "The Moran Scatter Plot as an ESDA Tool to Assess Local Instability in Spatial Association." In *Spatial Analytical Perspectives on GIS*. Routledge.

———. 2020. "GeoDA Documentation - Global Spatial Autocorrelation." 2020. https://geodacenter.github.io/workbook/5a_global_auto/lab5a.html#.

Arribas-Bel, Daniel, Arturo Ramos, and Fernando Sanz-Gracia. 2015. "The Size Distribution of Employment Centers within the US Metropolitan Areas." *Environment and Planning B: Planning and Design* 42 (1): 23–39.

Arribas-Bel, Daniel, and Fernando Sanz-Gracia. 2014. "The Validity of the Monocentric City Model in a Polycentric Age: US Metropolitan Areas in 1990, 2000 and 2010." *Urban Geography* 35 (7): 980–97.

Bartosiewicz, Bartosz, and Szymon Marcińczak. 2020. "Investigating Polycentric Urban Regions: Different Measures – Different Results." *Cities* 105 (October): 102855.

Bivand, Roger S., Edzer Pebesma, and Virgilio Gomez-Rubio. 2013. *Applied Spatial Data Analysis with R, Second Edition*. Springer, NY. http://www.asdar-book.org/.

Brynjolfsson, Erik, John Horton, Adam Ozimek, Daniel Rock, Garima Sharma, and Hong-Yi TuYe. 2020. "COVID-19 and Remote Work: An Early Look at US Data." w27344. Cambridge, MA: National Bureau of Economic Research.

Buera, Francisco J, and Joseph P Kaboski. 2012. "The Rise of the Service Economy." *American Economic Review* 102 (6): 2540–69.

Burger, Martijn, and Evert Meijers. 2012. "Form Follows Function? Linking Morphological and Functional Polycentricity." *Urban Studies* 49 (5): 1127–49.

Burrough, Peter A, and Andrew Frank. 2020. *Geographic Objects with Indeterminate Boundaries*. CRC Press.

Cai, Jixuan, Bo Huang, and Yimeng Song. 2017. "Using Multi-Source Geospatial Big Data to Identify the Structure of Polycentric Cities." *Remote Sensing of Environment* 202 (December): 210–21.

Csurka, Gabriela, Diane Larlus, Florent Perronnin, and F. Meylan. 2004. "What Is a Good Evaluation Measure for Semantic Segmentation?" *IEEE PAMI* 26 (1).

Davoudi, Simin. 2003. "European Briefing: Polycentricity in European Spatial Planning: From an Analytical Tool to a Normative Agenda." *European Planning Studies* 11 (8): 979–99.

Dilo, Areti. 2006. *Representation of and Reasoning with Vagueness in Spatial Information: A System for Handling Vague Objects*. ITC Dissertation 135. Enschede: ITC.

Doling, John, and Rowan Arundel. 2020. "The Home as Workplace: A Challenge for Housing Research." *Housing, Theory and Society*, 1–20.

Doll, Christopher N H, Jan-Peter Muller, and Jeremy G Morley. 2006. "Mapping Regional Economic Activity from Night-Time Light Satellite Imagery." *Ecological Economics*, 18.

Doll, Christopher N. H., Jan-Peter Muller, and Christopher D. Elvidge. 2000. “Night-Time Imagery as a Tool for Global Mapping of Socioeconomic Parameters and Greenhouse Gas Emissions.” *AMBIO: A Journal of the Human Environment* 29 (3): 157–62.

Dubin, Robin A, and Chein-Hsing Sung. 1987. “Spatial Variation in the Price of Housing: Rent Gradients in Non-Monocentric Cities.” *Urban Studies* 24 (3): 193–204.

Esch, Thomas, Wieke Heldens, Andreas Hirner, Manfred Keil, Mattia Marconcini, Achim Roth, Julian Zeidler, Stefan Dech, and Emanuele Strano. 2017. “Breaking New Ground in Mapping Human Settlements from Space–The Global Urban Footprint.” *ISPRS Journal of Photogrammetry and Remote Sensing* 134: 30–42.

Faisal, K., and A. Shaker. 2014. “The Use of Remote Sensing Technique to Predict Gross Domestic Product (GDP): An Analysis of Built-Up Index and GDP in Nine Major Cities in Canada.” *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL–7 (September): 85–92.

Fina, Stefan, Angelika Krehl, Stefan Siedentop, Hannes Taubenböck, and Michael Wurm. 2014. “Dichter dran! Neue Möglichkeiten der Vernetzung von Geobasis-, Statistik- und Erdbeobachtungsdaten zur räumlichen Analyse,” *Raumforschung und Raumordnung*. 72: 179-194.

Fu, Weijun., Peikun K. Jiang, Guomo Zhou, and Keli Zhao. 2014. “Using Moran’s I and GIS to Study the Spatial Pattern of Forest Litter Carbon Density in a Subtropical Region of Southeastern China.” *Biogeosciences* 11 (8): 2401–9.

Fu, Weijun, Keli Zhao, Chaosheng Zhang, and Hubert Tunney. 2011. “Using Moran’s I and Geostatistics to Identify Spatial Patterns of Soil Nutrients in Two Different Long-Term Phosphorus-Application Plots.” *Journal of Plant Nutrition and Soil Science* 174 (5): 785–98.

Garcia-López, Miquel-Àngel. 2010. “Population Suburbanization in Barcelona, 1991–2005: Is Its Spatial Structure Changing?” *Journal of Housing Economics* 19 (2): 119–32.

Garcia-López, Miquel-Àngel, Camille Hémet, and Elisabet Viladecans-Marsal. 2016. “Next Train to the Polycentric City: The Effect of Railroads on Subcenter Formation,” *Regional Science and Urban Economics* 67: 50-63.

Garcia-López, Miquel-Àngel, and Ivan Muñiz. 2010. “Employment Decentralisation: Polycentricity or Scatteration? The Case of Barcelona.” *Urban Studies* 47 (14): 3035–56.

Garreau, Joel. 1992. *Edge City: Life on the New Frontier*. Anchor.

Geiß, Christian, Tobias Leichtle, Michael Wurm, Patrick Aravena Pelizari, Ines Standfuß, Xiao Xiang Zhu, Emily So, Stefan Siedentop, Thomas Esch, and Hannes Taubenbock. 2019. “Large-Area Characterization of Urban Morphology—Mapping of Built-Up Height and Density Using TanDEM-X and Sentinel-2 Data.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12 (8): 2912–27.

Geiß, Christian, Michael Wurm, Markus Breunig, Andreas Felbier, and Hannes Taubenböck. 2015. “Normalization of TanDEM-X DSM Data in Urban Environments with Morphological Filters.” *IEEE Transactions on Geoscience and Remote Sensing* 53 (8): 4348–62.

Giuliano, Genevieve, Christian Redfearn, Ajay Agarwal, Chen Li, and Duan Zhuang. 2007. “Employment Concentrations in Los Angeles, 1980–2000.” *Environment and Planning A: Economy and Space* 39 (12): 2935–57.

Giuliano, Genevieve, and Kenneth A. Small. 1991. “Subcenters in the Los Angeles Region.” *Regional Science and Urban Economics* 21 (2): 163–82.

Gordon, Peter, and Harry W. Richardson. 1996. "Beyond Polycentricity: The Dispersed Metropolis, Los Angeles, 1970-1990." *Journal of the American Planning Association* 62 (3): 289–95.

Grubesic, Tony H., Ran Wei, and Alan T. Murray. 2014. "Spatial Clustering Overview and Comparison: Accuracy, Sensitivity, and Computational Expense." *Annals of the Association of American Geographers* 104 (6): 1134–56.

Heider, Bastian, and Stefan Siedentop. 2020. "Employment Suburbanization in the 21st Century: A Comparison of German and US City Regions." *Cities* 104 (September): 102802.

Heikkila, E, P Gordon, J I Kim, R B Peiser, H W Richardson, and D Dale-Johnson. 1989. "What Happened to the CBD-Distance Gradient?: Land Values in a Policentric City." *Environment and Planning A: Economy and Space* 21 (2): 221–32.

Huang, Lan, Linda W. Pickle, and Barnali Das. 2008. "Evaluating Spatial Methods for Investigating Global Clustering and Cluster Detection of Cancer Cases." *Statistics in Medicine* 27 (25): 5111–42.

Jacquez, Geoffrey M., and Dunrie A. Greiling. 2003. "Local Clustering in Breast, Lung and Colorectal Cancer in Long Island, New York." *International Journal of Health Geographics* 2 (1): 1–12.

Jelinski, Dennis E., and Jianguo Wu. 1996. "The Modifiable Areal Unit Problem and Implications for Landscape Ecology." *Landscape Ecology* 11 (3): 129–40.

Johnson, Brian, and Zhixiao Xie. 2011. "Unsupervised Image Segmentation Evaluation and Refinement Using a Multi-Scale Approach." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (4): 473–83.

Keola, Souknilanh, Magnus Andersson, and Ola Hall. 2015. "Monitoring Economic Development from Space: Using Nighttime Light and Land Cover Data to Measure Economic Growth." *World Development* 66 (February): 322–34.

Krehl, Angelika. 2015a. "Polyzentralität in Deutschen Stadtregionen–Eine Integrierte Bestandsaufnahme." In *Globale Urbanisierung*, 159–70. Springer.

———. 2015b. "Urban Spatial Structure: An Interaction between Employment and Built-up Volumes." *Regional Studies, Regional Science* 2 (1): 290–308.

———. 2016a. "Analyzing Polycentricity." Dissertation, Technische Universität Dortmund.

———. 2016b. "Urban Subcentres in German City Regions: Identification, Understanding, Comparison: Urban Subcentres in German City Regions." *Papers in Regional Science* 97: S79–104.

Krehl, Angelika, Stefan Siedentop, and Angelika Münter. 2016. "Polyzentralität – Ein Facettenreiches Raumwissenschaftliches Konzept." *ILS-Trends*, 2016.

Krehl, Angelika, Stefan Siedentop, Hannes Taubenböck, and Michael Wurm. 2016. "A Comprehensive View on Urban Spatial Structure: Urban Density Patterns of German City Regions." *ISPRS International Journal of Geo-Information* 5 (6): 76.

Landis, J Richard, and Gary G Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics*, 159–74.

Lang, Robert E., and Jennifer LeFurgy. 2003. "Edgeless Cities: Examining the Noncentered Metropolis." *Housing Policy Debate* 14 (3): 427–60.

Lee, Bumsoo. 2007. "'Edge' or 'edgeless' cities? Urban spatial structure in U.S. metropolitan areas, 1980 to 2000." *Journal of Regional Science* 47 (3): 479–515.

Liu, Xingjian, and Mingshu Wang. 2016. "How Polycentric Is Urban China and Why? A Case Study of 318 Cities." *Landscape and Urban Planning* 151 (July): 10–20.

Liu, Xuejun, Qisheng Pan, Lester King, and Zhonghua Jin. 2019. "Analysing the Changes of Employment Subcentres: A Comparison Study of Houston and Dallas." *Urban Studies* 56 (12): 2532–48.

Lv, Yongqiang, Zongmin Lan, Changcheng Kan, and Xinqi Zheng. 2020. "Polycentric Urban Development and Its Determinants in China: A Geospatial Big Data Perspective" *Geographical Analysis*, March 2020.

Ma, Yueliang, and Ruisong Xu. 2010. "Remote Sensing Monitoring and Driving Force Analysis of Urban Expansion in Guangzhou City, China." *Habitat International* 34 (2): 228–35.

Madelin, Malika, Claude Grasland, Hélène Mathian, Léna Sanders, and Jean-Marc Vincent. 2009. "Das „MAUP": Modifiable Areal Unit – Problem oder Fortschritt?" *Informationen zur Raumentwicklung* 10 (11): 16.

McDonald, John F. 1987. "The Identification of Urban Employment Subcenters." *Journal of Urban Economics* 21 (2): 242–58.

McHugh, Mary L. 2012. "Interrater Reliability: The Kappa Statistic." *Biochemia Medica* 22 (3): 276–82.

McMillen, Daniel. 2001. "Nonparametric Employment Subcenter Identification." *Journal of Urban Economics* 50 (3): 448–73.

———2003. "Identifying Sub-Centres Using Contiguity Matrices." *Urban Studies* 40 (1): 57–69.

———. 2004. "Employment Densities, Spatial Autocorrelation, and Subcenters in Large Metropolitan Areas." *Journal of Regional Science* 44 (2): 225–44.

Meijers, Evert. 2008. "Measuring Polycentricity and Its Promises." *European Planning Studies* 16 (9): 1313–23.

Muniz, Ivan, Anna Galindo, and Miguel Angel Garcia. 2003. "Cubic Spline Population Density Functions and Satellite City Delimitation: The Case of Barcelona." *Urban Studies* 40 (7): 1303–21.

Noor, Abdisalan M, Victor A Alegana, Peter W Gething, Andrew J Tatem, and Robert W Snow. 2008. "Using Remotely Sensed Night-Time Light as a Proxy for Poverty in Africa." *Population Health Metrics*, 13.

Openshaw, Stan. 1983. *The Modifiable Areal Unit Problem*. Norwick. Norfolk: Geo Books.

Riguelle, F., I. Thomas, and A. Verhetsel. 2007. "Measuring Urban Polycentrism: A European Case Study and Its Implications." *Journal of Economic Geography* 7 (2): 193–215.

Roca Cladera, Josep, Carlos R. Marmolejo Duarte, and Montserrat Moix. 2009. "Urban Structure and Polycentrism: Towards a Redefinition of the Sub-Centre Concept." *Urban Studies* 46 (13): 2841–68.

Schneider, Markus. 2000. "Metric Operations on Fuzzy Spatial Objects in Databases." In *Proceedings of the Eighth ACM International Symposium on Advances in Geographic Information Systems - GIS '00*, 21–26. Washington, D.C., United States: ACM Press.

Standfuß, I., C. Geiß, M. Kühnl, M. Wurm, H. Taubenböck, S. Siedentop, and B. Heider. 2020. "Deriving Urban Mass Concentrations Using TanDEM-X and Sentinel-2 Data for the Assessment of Morphological Polycentricity." In *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, 4219–22.

Sun, Tieshan. 2020. "A Longitudinal Study of Changes in Intra-Metropolitan Employment Concentration in Beijing: Decentralisation, Reconcentration and Polycentrification." *Urban Studies* 57 (4): 748–65.

Sutton, Paul C, Christopher D Elvidge, and Tilottama Ghosh. 2007. "Estimation of Gross Domestic Product at Sub-National Scales Using Nighttime Satellite Imagery" *International Journal of Ecological Economics & Statistics*, 8(S07), 5-21.

Taubenböck, H., I. Standfuß, M. Wurm, A. Krehl, and S. Siedentop. 2017. "Measuring Morphological Polycentricity - A Comparative Analysis of Urban Mass Concentrations Using Remote Sensing Data." *Computers, Environment and Urban Systems* 64 (July): 42–56.

Taubenböck, Matthias Weigand, Thomas Esch, Jeroen Staab, Michael Wurm, Johannes Mast, and Stefan Dech. 2019. "A New Ranking of the World's Largest Cities—Do Administrative Units Obscure Morphological Realities?" *Remote Sensing of Environment* 232: 111353.

Tsai, Yu-Hsin. 2005. "Quantifying Urban Form: Compactness versus 'Sprawl.'" *Urban Studies* 42 (1): 141–61.

Veneri, Paolo, and David Burgalassi. 2012. "Questioning Polycentric Development and Its Effects. Issues of Definition and Measurement for the Italian NUTS-2 Regions." *European Planning Studies* 20 (6): 1017–37.

Weigand, Matthias, Michael Wurm, Stefan Dech, and Hannes Taubenböck. 2019. "Remote Sensing in Environmental Justice Research—A Review." *ISPRS International Journal of Geo-Information* 8 (1): 20.

Wurm, Michael, Pablo dAngelo, Peter Reinartz, and Hannes Taubenbock. 2014. "Investigating the Applicability of Cartosat-1 DEMs and Topographic Maps to Localize Large-Area Urban Mass Concentrations." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (10): 4138–52.

Yue, Wenze, Jiabin Gao, and Xuchao Yang. 2014. "Estimation of Gross Domestic Product Using Multi-Sensor Remote Sensing Data: A Case Study in Zhejiang Province, East China." *Remote Sensing* 6 (8): 7260–75.

Zhang, Chaosheng, Lin Luo, Weilin Xu, and Valerie Ledwith. 2008. "Use of Local Moran's I and GIS to Identify Pollution Hotspots of Pb in Urban Soils of Galway, Ireland." *Science of The Total Environment* 398 (1–3): 212–21.

Zink, Manfred, Markus Bachmann, Benjamin Brautigam, Thomas Fritz, Irena Hajnsek, Alberto Moreira, Birgit Wessel, and Gerhard Krieger. 2014. "TanDEM-X: The New Global DEM Takes Shape." *IEEE Geoscience and Remote Sensing Magazine* 2 (2): 8–23.

Software

Bivand, Roger S., and David W. S. Wong. 2018. "Comparing Implementations of Global and Local Indicators of Spatial Association." *TEST* 27 (3): 716–48.

McMillen, Daniel. 2013. *McSpatial: Nonparametric Spatial Data Analysis*. https://CRAN.R-project.org/package=McSpatial.

R Core Team. 2020. "R: A Language and Environment for Statistical Computing". Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Revelle, William. 2020. *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University. https://CRAN.R-project.org/package=psych.

Online

Advameg, Inc. 2021. "Cities of the United States." 2021. http://www.city-data.com/us-cities/index.html.

City of Atlanta. 2020. "About ATL | ATL | Hartsfield-Jackson Atlanta International Airport." 2020. https://www.atl.com/about-atl/.

———. 2021. "History | Atlanta, GA." 2021. https://www.atlantaga.gov/visitors/history.

City of Seattle. 2021. "Quick Information - CityArchives | Seattle.Gov." 2021. https://www.seattle.gov/cityarchives/seattle-facts/quick-city-info.
Cleargov.org. 2021. "Cleargov." Addison, TX | Demographic Details | ClearGov. 2021. https://www.cleargov.com/texas/dallas/town/addison/2018/demographics.
Czakon, Jakub. 2019. "F1 Score vs ROC AUC vs Accuracy vs PR AUC: Which Evaluation Metric Should You Choose?" Neptune.Ai. November 4, 2019. https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc.
DeKalb County. 2019. "Home | DeKalb County, GA." 2019. https://www.dekalbcountyga.gov/.
DestinationDFW. 2013. "Addison Texas." Destination DFW - Dallas Fort Worth Relocation Guide. August 27, 2013. https://destinationdfw.com/addison-texas/.
Deutsche Forschungsgemeinschaft - DFG. 2021. "Where Are the Jobs? Stadtregionale Zentrenstrukturen Im Internationalen Vergleich." 2021. https://gepris.dfg.de/gepris/projekt/398024191?context=projekt&task=showDetail&id=398024191&.
DFW. 2021. "About DFW & Fast Facts." 2021. https://www.dfwairport.com/business/about/facts/.
Encyclopaedia Britannica, Editors of. 2021a. "Dallas." Encyclopedia Britannica. 2021. https://www.britannica.com/place/Dallas.
———. 2021b. "Pittsburgh | Pennsylvania, United States." Encyclopedia Britannica. 2021. https://www.britannica.com/place/Pittsburgh.
———. 2021c. "Seattle | Geography, History, & Points of Interest." Encyclopedia Britannica. 2021. https://www.britannica.com/place/Seattle-Washington.
EU Science Hub. 2020. "Telework in the EU before and after the COVID-19: Where We Were, Where We Head To."
Gimond, Manuel. 2013. *RobustI* (version 0.9). http://personal.colby.edu/~mgimond/R/RobustI.R.
Graham, Matthew R, Mark J Kutzbach, Brian McKenzie, and others. 2014. "Design Comparison of LODES and ACS Commuting Data Products." Working Paper CES-WP-14-38. Center for Economic Studies, U.S. Census Bureau.
INSPIRE. 2014. "Data Specification on Geographical Grid Systems – Technical Guidelines." INSPIRE Thematic Working Group Coordinate Reference Systems & Geographical Grid Systems. https://inspire.ec.europa.eu/file/1725/download?token=2-WQia90.
OpenStreetMap contributors. 2020. "Https://Planet.Osm.Org."
Port of Tacoma. 2014. "History." Port of Tacoma. January 14, 2014. https://www.portoftacoma.com/100.
REDP, Richardson Economic Development Partnership. 2021. "About Telecom Corridor." Richardson Economic Development Partnership. 2021. https://www.telecomcorridor.com/.
Rossiter, Katy. 2011. "What Are Census Blocks?" The United States Census Bureau. 2011. https://www.census.gov/newsroom/blogs/random-samplings/2011/07/what-are-census-blocks.html.
VisitPittsburgh. 2015. "Pittsburgh History." Visit Pittsburgh. March 11, 2015. https://www.visitpittsburgh.com/media/news-releases/pittsburgh-history/.

Appendix A: Site-specific cell counts

N	n_{TP}	n_{FP}	n_{TN}	n_{FN}	$n_{UMC\text{-}center}$	$n_{EMP\text{-}center}$	algorithm	test site
31349	1141	2189	27824	195	3330	1336	MI	
31349	914	1373	28206	856	2287	1770	CB	
31349	240	593	30330	186	833	426	LR	All
9745	369	603	8690	83	972	452	MI	
9745	314	523	8648	260	837	574	CB	
9745	72	213	9391	69	285	141	LR	Atlanta
14558	630	921	12947	60	1551	690	MI	
14558	409	442	13402	305	851	714	CB	
14558	132	213	14141	72	345	204	LR	Dallas
3939	77	294	3545	23	371	100	MI	
3939	102	185	3492	160	287	262	CB	
3939	18	97	3803	21	115	39	LR	Pittsburgh
3107	65	371	2642	29	436	94	MI	
3107	89	223	2664	131	312	220	CB	
3107	18	70	2995	24	88	42	LR	Seattle

Detailed cell counts by algorithmic center detection and test sites.

N: Total number of cells in test site

$n_{UMC\text{-}center}$: Number of detected UMC-center cells

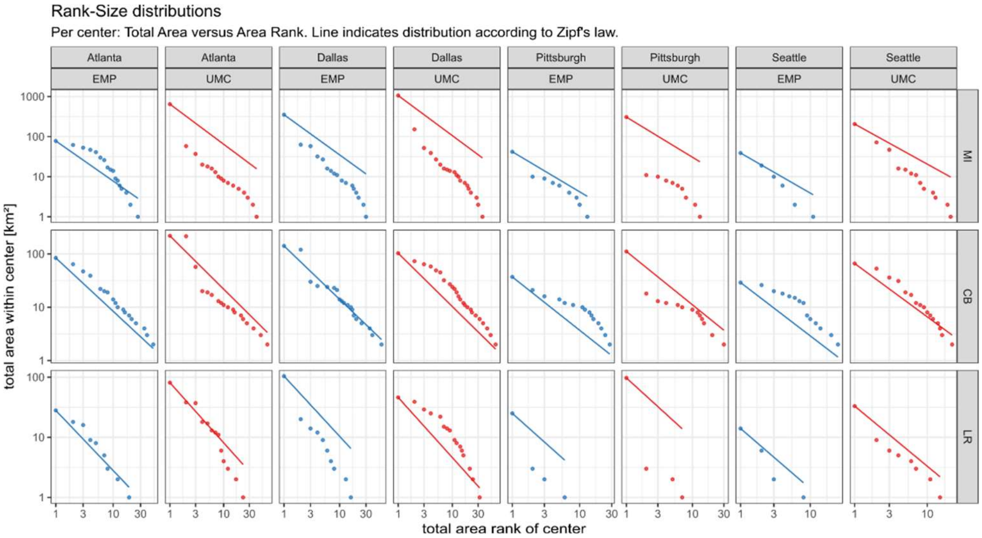
$n_{EMP\text{-}center}$: Number of detected EMP-center cells

n_{TP}, n_{FP}, n_{TN},n_{FN} : Number of cells of the respective case resulting from the comparison of EMP-centers and UMC-centers.

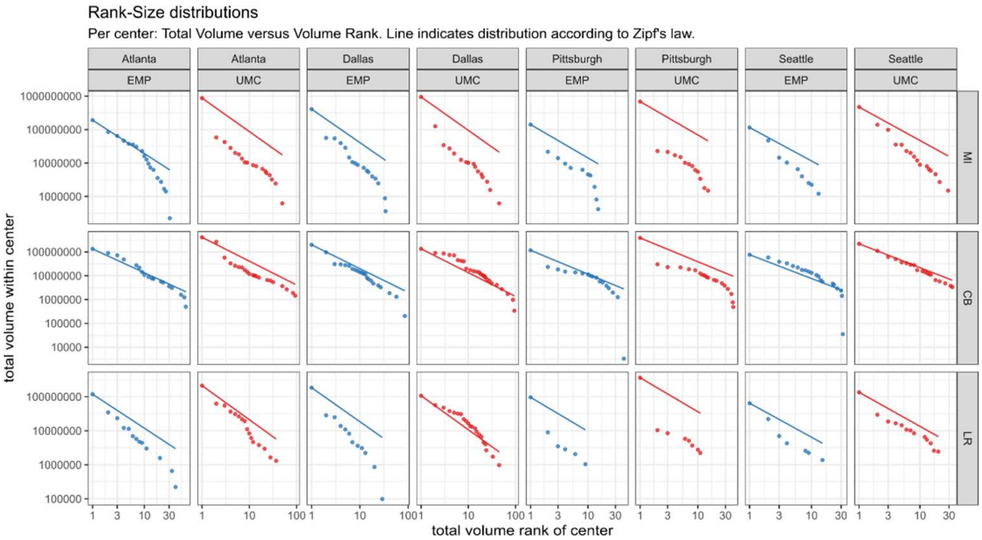
Appendix B: Site-specific distributions of cell properties

Detailed distributions by algorithmic center detection and test sites.

Appendix C: Alternative rank-size distributions



Area-based rank-size distributions of detected centers by algorithmic center detection and test sites.



Volume-based rank-size distributions of detected centers by algorithmic center detection and test sites.

Statement of Originality

Title of the Thesis:

Analyzing the relationship between urban morphology and economic subcenters with a focus on urban polycentricity using remote sensing and socioeconomic data

Submitted by (First Name, Name, Matrikel):

Johannes Mast, 2327704

Thesis topic provided by (Title, First Name, Name, Institute):

Dr. Hannes Taubenböck, Lehrstuhl für Fernerkundung

I, Johannes Mast, hereby declare that the present thesis is my own work and never was submitted for examination purposes elsewhere. I further declare that I did not use other references or tools than those mentioned within the thesis, and that citations and quotations have been indicated as such.

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig verfasst, bisher nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegeben Quellen oder Hilfsmittel benutzt sowie Zitate als solche gekennzeichnet habe.

Place, Date, Signature of the candidate