A spatially-explicit sensitivity analysis of urban definitions: Uncovering implicit assumptions in the *Degree of Urbanisation*

Van Migerode Céline[*1], *Poorthuis Ate*[1], *Derudder Ben*[1,2]

[1]KU Leuven, Leuven, Belgium
[2]Ghent University, Ghent, Belgium

Published as: Van Migerode, C., Poorthuis, A., & Derudder, B. (2024). A spatially-explicit sensitivity analysis of urban definitions: Uncovering implicit assumptions in the *Degree of Urbanisation*. *Computers, Environment and Urban Systems 112*, 102149. https://doi.org/10.1016/j.compenvurbsys.2024.102149.

Abstract: We introduce a spatially-explicit sensitivity framework to uncover potential biases in urban delineation approaches. Our starting point is that there is no broadly shared agreement on how to define or delineate urban areas, neither in terms of methods nor in terms of thresholds or criteria. Deciding on delineation criteria thus inevitably involves making certain assumptions that may unwittingly reproduce urban realities experienced by those expressing them, and have spatially unequally distributed implications. Understanding how specific criterion choices shape our understanding of 'the urban' and how, why, and – especially – *where* a definition leads to specific sensitivities is therefore key, especially when the definition is utilised beyond its intended application. Our framework to uncover these sensitivities is spatially explicit in the sense that it does not rely on aggregate statistics but instead focuses on the sensitivity of the 'urban' classification of individual spatial units at the finest spatial granularity. Applying the framework to the definition of the *Degree of Urbanisation* reveals that sensitivity is indeed not equally distributed across the world. Certain regions (e.g., areas around Dallas – Fort Worth) and specific types of urbanisation (e.g., desakota regions in Pacific Asia) exhibit higher sensitivity than others. We discuss how these sensitivities may embody certain implicit assumptions in the definition, and examine their broader theoretical implications.

Highlights:

- We developed a framework to quantify variability in urban definitions in a spatially-explicit manner.
- Results revealed that criteria tweaks in DEGURBA have spatially unequally distributed implications.
- Certain large delta regions are disproportionally sensitive to changing a contiguity rule in DEGURBA.
- DEGURBA includes a specific rule to delineate cities in North America, while no targeted rules exist for other world regions.

Keywords: Sensitivity analysis, Spatially-explicit, Urban delineation, City definition, *Degree of Urbanisation*

[*] Public Governance Institute / Division of Geography and Tourism, KU Leuven, Oude Markt 13, Leuven 3000, Belgium. Email: celine.vanmigerode@kuleuven.be

1 INTRODUCTION

In recent years, there has been renewed interest in defining cities and classifying locations into urban and rural categories. Scholars have constructed urban delineations to answer questions about urban concentration (Uchida & Nelson, 2009), economic density (Henderson et al., 2021), and population dynamics (Moreno-Monroy et al., 2021). International organisations developed definitions in the context of development policies (Asian Development Bank, 2019) and to monitor the Sustainable Development Goals (SDGs; Eurostat, 2016; Melchiorri et al., 2019). However, there is no broadly shared agreement on how to define a city, neither in terms of methods nor in terms of thresholds to be used in these definitions (Cohen, 2004). This can have profound implications: different delineation approaches lead to varying numbers of cities and variable city boundaries, which in turn affects the quantitative analysis of issues ranging from economic development (Bosker et al., 2021; Wineman et al., 2020) to urban heat island effects (Yang et al., 2023). Variability in urban definitions also influences policies that target 'urban' areas, as underestimating their presence or size may limit financial aid and political attention (Wineman et al., 2020).

Urban delineation algorithms typically group different spatial units based on their functional relationship (e.g., commuting flows) or their similarity in characteristics (e.g., population density). This often involves applying thresholds, such as enforcing a minimum amount of commuting, number of inhabitants, or night-time light emission for a location to be considered 'urban'. There are no agreed-upon ways of deciding the 'best' values for these thresholds as developing conceptual justifications with universal validity is challenging (Duranton, 2021). For example, Nigeria and Syria employ a minimum population size of 20,000 to define urban areas, while Canada and New Zealand already designate a settlement of 2,000 inhabitants as 'urban' (United Nations Population Division, 2019). Besides threshold values, there are also other implementation rules that influence the boundaries of urban areas: for instance, how to operationalise spatial contiguity (Statham et al., 2021) or how to cope with the presence of green spaces and water bodies intersecting urban areas (Eurostat, 2021).

In this paper, we use the term *criteria* to refer to all implementation rules in delineation algorithms that may be at the root of variability in urban boundaries, including but not limited to threshold values, contiguity rules, and smoothing procedures. Deciding on these delineation criteria inevitably involves making explicit and implicit assumptions about the fundamental nature of 'the urban'. These assumptions may – unwittingly – reproduce urban realities experienced and observed by those expressing them, potentially leading to spatially unequally distributed implications. For example, Statham et al. (2021) found that low-income countries are more sensitive to specific criteria in a delineation algorithm than high-income countries. This might suggest an implicit bias towards urban patterns that are abundant in certain high-income countries, and give rise to the question of whether consistent delineations are equally 'fit' to capture different types of urbanisation worldwide (Potts, 2018). It is therefore key to understand how specific criterion choices shape our understanding of 'the urban' and assess how, why, and – especially – *where* a definition leads to specific sensitivities. However, conventional sensitivity analyses typically lack insight into this spatial dimension of sensitivity: which locations are most affected by variation in delineation criteria? There is a need for new approaches to quantify sensitivities tied to urban delineations in *a spatially-explicit* manner, not just at the regional or country level, but at finer spatial granularities.

Against this backdrop, we introduce a spatially-explicit sensitivity framework that quantifies the sensitivity of a location's 'urban' classification across various criteria settings. The framework is *spatially explicit* in the sense that it does not rely on aggregate statistics such as the urban population or number of cities per country or region but instead focuses on the sensitivity of the classification of individual units at the finest possible spatial resolution (e.g., grid cells or statistical units). It is not designed to determine criterion settings in a more 'robust' manner, this would also be conceptually untenable. Rather, the framework aims to detail the spatial heterogeneities in the sensitivity. It consists of three consecutive steps. The first step involves the identification of criteria that may induce variability in a specific urban definition. Next, a set of alternative realisations of the definition is constructed by simultaneously varying the values of these criteria. Finally, the sensitivity of each spatial unit is quantified based on the dispersion in the classification across the set of alternative realisations. The framework thus quantifies the sensitivity of a spatial unit across different criteria choices, in contrast to conventional sensitivity analyses that typically quantify the total sensitivity of a specific criterion choice within a particular study area.

To illustrate the potential of our framework, we apply it to the definition of the *Degree of Urbanisation* (DEGURBA), proposed by Dijkstra et al. (2021). DEGURBA identifies urban centres (i.e. cities), urban clusters (i.e. towns), and rural areas by clustering contiguous grid cells that meet a minimum population density, and collectively contain a minimum number of inhabitants. This definition was initially developed for settlements in the European Union but later applied to the rest of the world. Since then, it has been widely used in academic circles (Moreno-Monroy et al., 2021) and policy reports (Eurostat, 2016). A handful of previous studies conducted sensitivity analyses of DEGURBA. In these analyses, a number of aggregated metrics are computed, such as the change in urban population and urban land cover per continent (Dijkstra et al., 2021) and country (Dorward et al., 2023; Statham et al., 2021). By applying our *spatially-explicit* sensitivity framework to DEGURBA, we expand on these existing analyses in three main ways. First, we explicitly focus on the sensitivity of spatial units at the finest spatial granularity – 1 km² grid cells – instead of relying on aggregate statistics. Second, our approach goes beyond the 'obvious' population density and size thresholds. Instead, we consider all criteria that may induce variability in the resulting delineations, including technical specifications such as contiguity rules and an additional built-up density threshold. Although these criteria may seem less impactful at first glance, they can have spatially unequally distributed implications. In the discussion of the results, we specifically focus on these 'other' criteria. Third, our framework facilitates adopting a critical perspective on DEGURBA. Because the definition was initially calibrated on European settlements, it is possible that the definition (unintentionally) incorporates a particular perspective on urbanisation (Dorward et al., 2023). Our analysis of the impact of criteria tweaks in DEGURBA offers insight into whether and how the urban delineations are influenced by or reflect implicit assumptions about the nature of 'the urban'. This is particularly relevant to understand when the methodology is utilised beyond its intended application of monitoring the SDGs indicators.

The remainder of this paper is structured as follows: Section 2 gives an overview of the current literature regarding urban definitions and spatially-explicit sensitivity analyses. Section 3 introduces the framework and elaborates on its empirical application on DEGURBA. Section 4 presents the results, focusing on the sensitivity in the Red River Delta (Vietnam), the urbanised area around Dallas, Fort Worth, and Arlington (Texas, USA), and Vienna (Austria). Section 5 raises the main discussion points, after which Section 6 concludes the paper.

2 BACKGROUND

2.1 Inconsistency in urban definitions

There is no definite viewpoint or definition of what constitutes 'the urban' or 'a city.' Growing levels of urbanisation have gone hand in hand with new urban phenomena, implying that cities and their urban fabric are increasingly complex. Compare, for instance, the polycentric megalopolis around

Mexico City (Cohen, 2004) with the extended *desakota* zone[†] around Bangkok (McGee, 2013) and the suburban development in Los Angeles. All three regions can, and will likely, be considered 'urban'. However, the urbanisation in and of these three regions reflects different processes, timings, and geographical circumstances. Differences that are subsequently manifested in heterogeneity in terms of function, shape, and size. The 'urban' nature of settlements of a certain size is also context-dependent (Cohen, 2004). In some countries, settlements of a few thousand inhabitants may encompass typically 'urban' characteristics such as central place functions, while in other countries a similarly-sized settlement could still be perceived as 'rural' due to a limited employment rate in the non-agricultural sector (Wineman et al., 2020).

There are, therefore, fundamental reasons for assuming that different definitions may apply in other parts of the world. The question then arises whether a globally consistent definition of urbanisation – that is, relying on globally consistent criteria – is feasible and even desirable (Potts, 2018; Statham et al., 2021). No single definition can capture the worldwide complexity of local urban conditions. For example, DEGURBA seems to be an accepted measure of global urbanisation as it is approved by all member states of the United Nations (Dijkstra et al., 2021). However, there is some controversy around its criteria. Urban areas are defined based on a minimum density threshold of 300 inhabitants per km² and a minimum total population of 5,000 (Dijkstra et al., 2021). Angel et al. (2018) argue that the density threshold is too low, as it classifies many predominantly agrarian regions (including large parts of Java, Indonesia) as 'urban'. Similarly, Henderson et al. (2021) state that the density threshold should be increased to adequately capture urban areas in Sub-Saharan Africa. Potts (2018) argues that in the past, a minimum of 5,000 inhabitants for urban areas was appropriate to exclude settlements in Africa that are 'rural' in terms of their employment profile. However, these 'rural' settlements now exceed the population threshold of 5,000 due to recent population growth. They are classified as 'urban', although they did not experience a shift from agricultural to industrial activities, as is often presumed in theories on urbanisation (Potts, 2018). More fundamentally, Dorward et al. (2023) raise concerns regarding the population thresholds in DEGURBA in African contexts, given that the thresholds were initially designed and calibrated on settlements in the European Union and only later applied to the rest of the world.

These debates regarding the global suitability of criteria are not specific to population-based definitions such as DEGURBA, but arise in any urban delineation approach. There are, for example, similar disagreements about the minimum emission threshold when delineating urban areas based on night-time lights. Urban categories are generally constructed as contiguous areas with sufficiently large light emission, but some scholars rely on a single global emission threshold (Ch et al., 2021; Florida et al., 2008), while others argue that different types of urbanisation across the world require different threshold values (Henderson et al., 2003).

Nonetheless, even though globally consistent urban definitions entail complexities and face arbitrariness, it is clear that scientific and policy praxis requires such definitions as part of its analytical toolkit (Duranton, 2021; United Nations Statistics Division, 2019). They are necessary to monitor, for example, the SDGs on a global scale (Melchiorri et al., 2019) and enable globally comparable urban research on pressing issues such as segregation and urban sprawl. It is nevertheless crucial to recognise the broader policy and theoretical implications that arise from sensitivities to their criterion settings. An example of a policy implication is that policy programmes often use urban delineations to produce broader sets of statistics (Eurostat, 2016) and/or determine eligibility for funding. As slightly

[†] *Desakota* zones are extensive densely populated regions, often situated along the corridor between major cities in Pacific Asia, characterised by a mixture of both agricultural and non-agricultural activities (McGee, 2013).

changing a threshold value can determine whether a settlement is classified as 'urban' or not, this can affect its eligibility for financial aid and consequently have a substantial impact on its inhabitants.

An example of a theoretical implication is rooted in the observation that deciding on operational criteria inevitably involves making certain explicit or implicit assumptions about what constitutes 'the urban'. However, urban definitions are – to some degree – social constructs that part-represent realities that have been experienced, observed, and embodied by those articulating them (Cottineau et al., 2024). This leads to what in philosophy is called the analytical or conceptual circularity problem (Humberstone, 1997). This problem arises when a loop exists where a concept's operational and theoretical definitions co-constitute each other. One aspect of this circularity in urban studies is that urban globalisation research is mainly dominated and subsequently shaped by specific researchers, institutions, and cities (Kanai et al., 2018)[‡]. In light of DEGURBA's roots, it cannot be ruled out that, say, a Eurocentric bias is inserted – even unwittingly – in the theoretical-operational loop of what is purportedly a global exercise. Taking these outcomes at face value would thus reinforce our skewed understanding of the nature of urbanisation itself (Roy, 2009). Another theoretical implication stems from the consideration that a definition's circulation across academia and in policymaking may be affected by its contextual setting. In case of DEGURBA, the definition is utilised in diverse applications, in part because it is developed by the European Commission and endorsed by the United Nations (Dijkstra et al., 2021). However, when applying a definition to other use cases, it is important to consider the underlying assumptions stemming from the definition's original purpose (Cottineau et al., 2024). The DEGURBA delineations were, for instance, constructed to facilitate international comparison of the SDG indicators, which explains why urban areas are predominantly operationalised based on population data (Dijkstra et al., 2021). However, the assumption that population concentration is *the* defining feature of urbanity might be suboptimal in applications beyond this particular use case. A spatially-explicit sensitivity analysis of choices in delineation algorithms may help grasp whether, and if so, how urban definitions shape, incorporate, reflect, and reproduce policy frameworks and implicit theoretical assumptions.

2.2 Spatially-explicit sensitivity analysis

Although *spatial* sensitivity analyses have been applied in myriad geographical applications (e.g., Chen et al., 2013; Kocabas & Dragicevic, 2006; Lilburne & Tarantola, 2009), there are only few studies conducting *spatially-explicit* analyses in which a sensitivity index is calculated for each individual spatial unit. Şalap-Ayça et al. (2018) performed a spatially-explicit sensitivity analysis of a cellular-automata-based urban growth model, and employed a meta-modelling technique to approximate the model's response to changes in input values. This meta-model was applied at the local level to estimate the sensitivity for each spatial unit independently. Ligmann-Zielinska & Jankowski (2014) and Tang et al. (2018) both used a similar approach to quantify spatially-explicit sensitivity of input weights in multi-criteria analyses (MCAs). More specifically, Ligmann-Zielinska & Jankowski (2014) developed an integrated spatially-explicit uncertainty and sensitivity analysis (iUSA) to (1) quantify the variability in the spatial outcome of MCAs and (2) identify which input weights contribute most to this variability. Their framework consisted of three different steps. First, Monte Carlo simulations are used to sample values from the probability density functions of the input weights. Each combination of weight values represents a possible realisation of the MCA. Second, the different realisations are summarised geographically by calculating the average and the standard deviation of the values for each spatial unit

[‡] Bunnell and Maringanti (2010) discussed an example of this, which they call 'metrocentricity': the often-unwitting assumption that large cities in the Global North represent a norm against which other cities can or should be benchmarked.

in the study area. Third and finally, a variance-based global sensitivity analysis is conducted to determine the partial contribution of each input weight to the total output variability (Ligmann-Zielinska & Jankowski, 2014). This involves determining both first-order and total effect sensitivity indices by integrating the model equation (Saltelli et al., 2010).

The iUSA framework has proven valuable in several spatial MCAs, including land suitability evaluation (Ligmann-Zielinska & Jankowski, 2014; Şalap-Ayça & Jankowski, 2016), landslide susceptibility modelling (Feizizadeh et al., 2014), and flood vulnerability evaluation (Feizizadeh & Kienberger, 2017). However, it cannot readily be applied to quantify spatially-explicit sensitivities in urban delineations due to inherent differences between spatial MCAs and urban delineation algorithms. Urban delineation algorithms are typically rule-based and include topological relationships or complex procedures relying on the value of surrounding units. It is consequently not feasible to perform a variance-based global sensitivity analysis, as this requires algorithms that allow integral calculus. In addition, the sources of variability in urban delineations are more diverse than in spatial MCAs. They often include a large variety of criteria (e.g., density threshold, contiguity rule, smoothing rule, etc.) with different levels of measurement (e.g., logical, continuous, discrete, etc.). Therefore, a spatially-explicit sensitivity framework for urban definitions needs to be revised to allow greater flexibility to accommodate more complex criteria beyond continuous weight values.

3 FRAMEWORK

3.1 Spatially-explicit sensitivity framework

We propose a spatially-explicit sensitivity framework specifically designed to quantify spatial variability in urban definitions. It is developed to be generally applicable to a wide range of delineation methodologies, regardless of whether it is a morphological or functional approach or resulting in an urban-rural dichotomy or multi-level classification (e.g., city, town, rural). The framework examines variability in the classification of individual spatial units at the finest granularity and consists of three different steps (see Figure 1):

STEP 1 – The first step of the framework involves identifying the main criteria in a specific delineation algorithm that may affect the resulting urban boundaries. These criteria can comprise threshold values, contiguity rules, or any other implementation rule for which it is hard to find a clear-cut conceptual justification.

STEP 2 – Next, alternative values are determined for each criterion. These alternative values are systematically combined to create a large set of alternative realisations of the definition. For example, if there are four different criteria with five alternative values each, 5^4 or 625 alternative realisations can be constructed. The number of alternative values and their range should be chosen in the context of the specificities of the definition, its complexity, and the available computing resources, while ensuring compliance with theoretical considerations (e.g., a minimum population threshold of 5 inhabitants is, of course, theoretically untenable).

STEP 3 – Afterwards, the sensitivity S of each spatial unit is computed based on a measure of dispersion. Consistent unit classification across all realisations results in a sensitivity index of zero. When there is high dispersion in the classification – for example, covering both urban and rural categories – then the sensitivity index should also reach a larger value. Many urban delineation algorithms result in an ordinal outcome: a classification in city, town, and rural implies that 'town' is the category between 'city' and 'rural.' A potential measure of dispersion of ordered categorical data is the inverse of l^2 proposed by Blair & Lacy (2000), although other measures might be chosen as well. The sensitivity index S can consequently be computed as:

$$S = 1 - l^2 \quad with \quad l^2 = \sqrt{\frac{d^2}{d^2{}_{max}}} \tag{1}$$

In Equation 1, d^2 is a measure of ordinal concentration calculated by Equation 2:

$$d^2 = \sum_{i=1}^{k-1} (F_i - 0.5)^2 \tag{2}$$

with k the number of categories and F_i the cumulative frequency of the i^{th} category. $d^2{}_{max}$ represents the maximum concentration for any given number of categories and is calculated by Equation 3 (Blair & Lacy, 2000) [§]:

$$d^2{}_{max} = (k - 1) * 0.25 \tag{3}$$

The resulting sensitivity index S is a continuous variable ranging from 0 to 1. A value of 0 represents a minimum dispersion and occurs when a location is consistently classified across all realisations. For example, New York is likely classified as 'urban' and the center of Greenland as 'rural' regardless of the criterion settings. On the contrary, $S = 1$ indicates extreme polarisation in the realisations. If a particular location is classified as 'rural' in half the realisations and as 'urban' in the other half, then S will reach its peak value of 1. Values between 0 and 1 represent the relative dispersion of a set of realisations, proportional to the maximum dispersion with a fixed number of ordinal categories. For example, a value of 0.4 signifies that the set of realisations has 40% of the maximal possible dispersion $d^2{}_{max}$, given the number of possible categories. However, it is crucial to note that a single sensitivity value can correspond to various class distributions. For example, a unit's classification varying between 'city' and 'town' or between 'town' and 'rural' may yield similar sensitivity values. To determine how the classification of a unit varies exactly, one should look at the separate realisations.

[§] The maximum concentration is achieved when all observations are in one class. The cumulative frequency of any class (F_i) is then either 0 or 1. Both $F_i = 0$ and $F_i = 1$ result in $(F_i - 0.5)^2 = 0.25$. The maximum value of d^2 for k classes is accordingly $(k - 1) * 0.25$. For more information, see Blair & Lacy (2000).

The final output of the spatially-explicit sensitivity framework is a map with S calculated for each spatial unit. This map serves as a starting point for an in-depth examination of spatial sensitivity and geographical biases within the definition. The user/researcher can zoom in on specific areas with high levels of sensitivity and study the implications of changing a specific criterion in this area by reviewing a selection of alternative realisations[**].

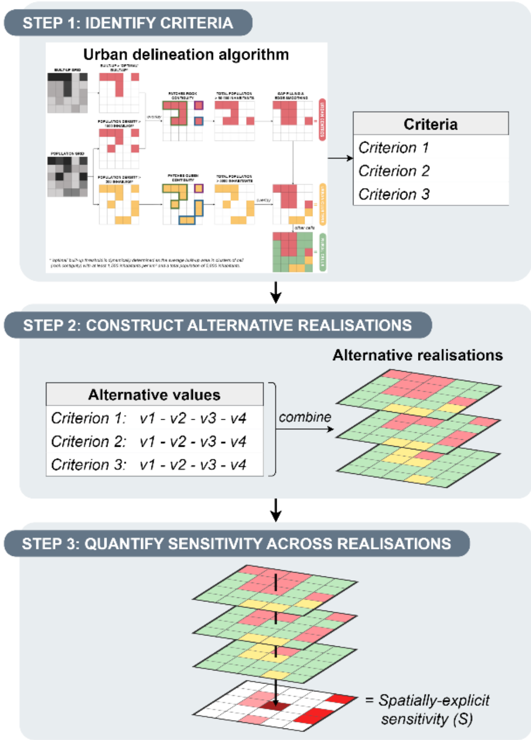

Figure 1: Schematic representation of the spatially-explicit sensitivity framework.

While we take inspiration from the iUSA proposed by Ligmann-Zielinska & Jankowski (2014), there are three crucial differences to ensure compatibility with urban definitions. First, the iUSA framework relies on Monte Carlo simulations to vary the input criteria, while the framework proposed in this paper uses a systematic combination of user-determined alternative values. We choose this approach because it shares strong methodological parallels with other sensitivity analyses in the field of urban studies (see, for example, Cottineau et al., 2017; Dijkstra et al., 2021; Statham et al., 2021; Taubenböck et al., 2022). Second, we measure sensitivity using an indicator of ordinal dispersion as the outcomes of urban delineations are typically ordered categories. The iUSA framework uses the standard deviation since MCAs generate continuous outcome values. The third and most significant difference is the absence of decomposing the total variability into the partial contribution of each input criterion. The proposed framework aims to uncover potential regional biases in urban definitions. This requires identifying the spatial units that are most sensitive to criteria tweaks, but does not demand the exact calculation of how much each criterion contributes to the outcome variability. The individual contribution of variation in each criterion to the sensitivity of a region is instead assessed by exploring a set of individual realisations (see Section 4 for concrete examples).

3.2 Empirical application to DEGURBA

To illustrate the potential of the spatially-explicit sensitivity framework, we apply it to the grid cell classification of DEGURBA (Dijkstra et al., 2021). It is important to emphasise that we do not want to criticise the specific criterion choices in this definition, but rather demonstrate how our proposed framework can uncover consequences of inherent assumptions and technical choices. The DEGURBA methodology classifies the cells of a 1 km² grid into three distinct categories based on the rules

[**] Sensitivity analyses generally determine how variability in the output of the model can be apportioned to variation in model inputs (Saltelli et al., 2010). We do not aim to exactly apportion the variability in urban delineations (cf. output) to specific criteria (cf. model inputs). Instead, we explore the contribution of each criterion by exploring a set of alternative realisations. This type of analysis where the total variability is quantified, without exact determination of partial contribution of individual inputs is sometimes referred to as *uncertainty analysis* instead of *sensitivity analysis* (Saltelli et al., 2010). However, we opted for term *sensitivity analysis* in this context, as we *are* examining the contribution of individual criteria, just not by quantifying it, but by exploring a set of alternative realisations.

summarised below (see Figure 2). For more details about the DEGURBA methodology, readers can consult the official documentation (Dijkstra et al., 2021; European Commission, 2023; Eurostat, 2021).

- **Urban centres** are clusters of cells (rook contiguity) with a minimum population density of 1,500 inhabitants per km² of permanent land, and cells with a built-up density above a certain threshold. In addition, the total population in these clusters should be at least 50,000. Gaps in the urban centres are filled, and edges are smoothed with a majority rule.
- **Urban clusters** are clusters of cells (queen contiguity) with a minimum population density of 300 inhabitants per km² of permanent land and a minimum total population of 5,000 inhabitants. Cells that belong to urban centres are removed from urban clusters.
- **Rural grid cells** neither belong to an urban centre nor an urban cluster.

Three different data sources are required to construct the grid cell classification: (1) a population grid, (2) a grid with built-up density, and (3) a grid representing the proportion of permanent land (European Commission, 2023). For the three data sources, we use the data products GHS-POP, GHS-BUILT-S, and GHS-LAND with the estimates for 2020, respectively[††] (European Commission, 2023). In this sense, we create a reconstruction of the level 1 grid classification, available under the SMOD layer on website of the Global Human Settlement Layer.

3.2.1 Step 1: Identifying the criteria that may induce variability

We identified five criteria in DEGURBA's definition that are (to some extent) difficult to justify: (1) the minimum population density thresholds, (2) the minimum population size thresholds, (3) the built-up density threshold for urban centres, (4) the contiguity rules and (5) the smoothing procedure for urban centres. The minimum population density and size thresholds are the most apparent criteria in DEGURBA. Although these thresholds were determined as a trade-off between national urban definitions (Dijkstra et al., 2021), they are hard to conceptually justify from a global point of view, as exemplified by the controversies in the recent literature (see Section 2.1). There is also room for discussion regarding the urban centres' built-up density threshold (Balk et al., 2021). The logic behind this threshold has, for example, been altered in the most recent versions of the DEGURBA methodology[‡‡]. Furthermore, urban centres are identified with rook contiguity, while urban clusters are identified with queen contiguity. The reason for this distinction is not clear (Statham et al., 2021). Gaps in urban centres are filled, and edges are smoothed using an iterative majority rule (Eurostat, 2021). This smoothing procedure is hard to justify, as other algorithms exist to achieve similar objectives.

[††] We employed the most recent version of this data, published under the release 'GHSL Data Package 2023'. All data sets are open-source and freely available at https://ghsl.jrc.ec.europa.eu/. For the land grid, data of 2018 is employed, as this is the only available data on the GHSL data website.

[‡‡] In GHSL Data Package 2022, DEGURBA included a fixed minimum built-up density threshold of 50% (Schiavina et al., 2022). However, in GHSL Data Package 2023, a dynamically identified 'optimal' built-up density threshold is employed. This 'optimal' threshold corresponds to a minimum of 20% built-up density when employing the data products of epoch 2020 (Van Migerode et al., 2024). The 'optimal' built-up density threshold, as defined in GHSL Data Package 2023, is determined as the global average built-up density in clusters of cells (rooks contiguity) with at least 1,500 inhabitants per km² of permanent land and a minimum total population of 5,000 inhabitants (for more information, see GHSL Data Package 2023, footnote 30 on page 51; European Commission, 2023).

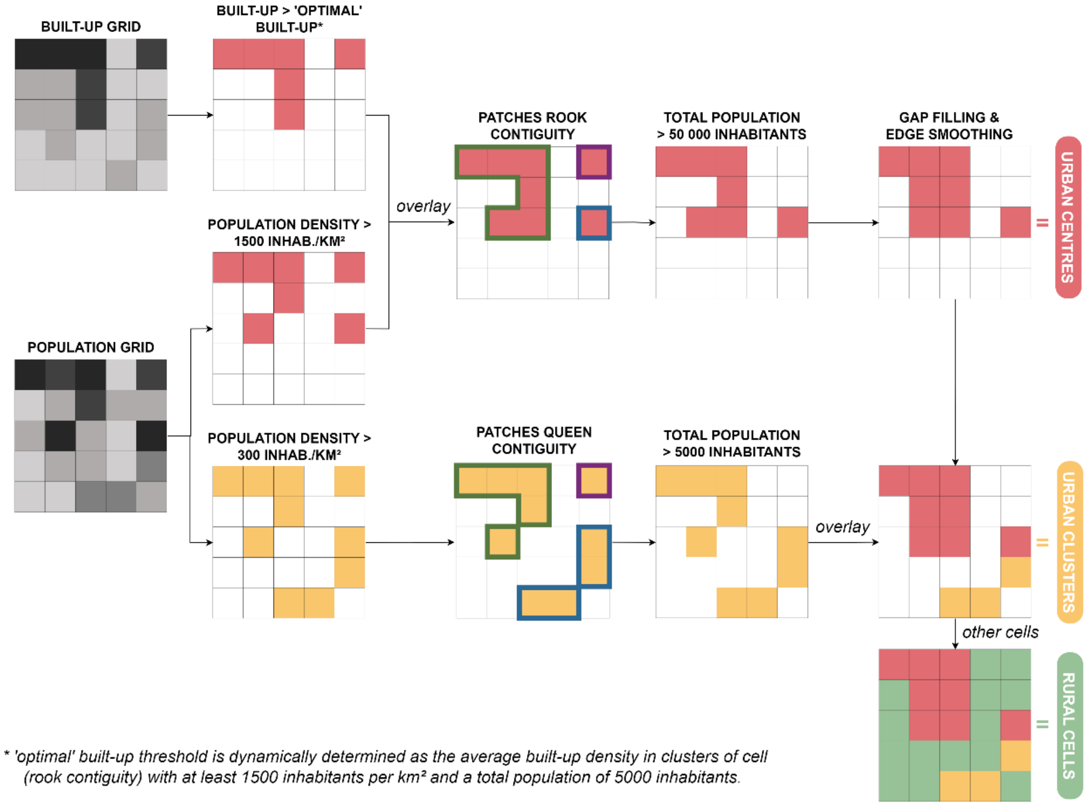

Figure 2: Workflow of DEGURBA's grid classification (based upon the implementation rules in GHSL Data Package 2023; European Commission, 2023).

3.2.2 Step 2: Constructing a set of alternative realisations

For each criterion identified in Step 1, different alternative values are determined (see Table 1). The alternative values for the population thresholds are generated by multiplying the standard values with the following range $[\frac{1}{2}, \frac{1}{1.75}, \frac{1}{1.5}, \frac{1}{1.25}, 1, 1.25, 1.5, 1.75, 2]$. For the built-up density threshold, three values are considered: (1) the dynamically identified 'optimal' threshold[§§], (2) a fixed threshold of 20% ('optimal' threshold in GHSL Data Package 2023), and (3) a fixed value of 50% (threshold used in GHSL Data Package 2022). All combinations of rook and queen contiguity are considered for the contiguity rules, ensuring that the rule for urban centres is at least as strict as the rule for urban clusters. In addition to the standard edge smoothing algorithm, one alternative smoothing rule is included, which entails applying an average moving window on the population grid (Henderson et al., 2021).

Systematically combining the alternative values for the criteria yields 9 x 9 x 3 x 3 x 2 = 1458 alternative realisations of DEGURBA. The alternative realisations are computed on a global scale by using the

[§§] The 'optimal' built-up density threshold, as defined in GHSL Data Package 2023, is determined as the global average built-up density in clusters of cells (rooks contiguity) with at least 1,500 inhabitants per km² of permanent land and a minimum total population of 5,000 inhabitants (see GHSL Data Package 2023, footnote 30 on page 51; European Commission, 2023). According to this official definition, the threshold is identified based on the population density of urban centres (*UC_DEN* = 1,500 inhabitants/km²), the population size of urban clusters (*UCL_SIZ* = 5,000) and the contiguity rule for urban centres (*UC_CONT* = rook). In the sensitivity calculations, we employ the same implementation, but dynamically identify the built-up threshold based on the combination of alternative *UC_DEN, UCL_SIZ and UC_CONT* values. The 'optimal' threshold is thus calculated as the global average built-up density in clusters of cells (with *UC_CONT* contiguity) with at least *UC_DEN* inhabitants per km² of permanent land and a minimum total population of *UCL_SIZ* inhabitants.

flexurba package, an open-source R package that allows reconstructing the classification of DEGURBA with customised criteria (Van Migerode et al., 2024). The computations are performed on a server with 32 cores and 256 GB RAM. The code for the analyses in this paper is available under: https://doi.org/10.48804/VFAGOQ.

Table 1: Alternative values for the criteria in DEGURBA.

Criterion		Alternative values									#
DENSITY THRESHOLD	***Urban centre***	750	857	1000	1200	1500	1875	2250	2625	3000	9
	Urban cluster	150	171	200	240	300	375	450	525	600	
SIZE THRESHOLD	***Urban centre***	25000	28571	33333	40000	50000	62500	75000	87500	100000	9
	Urban cluster	2500	2857	33333	4000	5000	6250	7500	8750	10000	
BUILT-UP THRESHOLD	dynamically identified "optimal" threshold			fixed value of 20%			fixed value of 50%				3
CONTIGUITY RULES	*rook* contiguity for both urban centres and urban clusters			*rook* contiguity for urban centres and *queen* contiguity for urban clusters			*queen* contiguity for both urban centres and urban clusters				3
SMOOTHING PROCEDURES	After classification: gaps > 15 km² are filled, edges are smoothed by a 3x3 majority rule					Before classification: smooth population grid with an 7x7km average moving window					2

The standard criterion values in DEGURBA are underlined.

3.2.3 Step 3: Quantifying the sensitivity of each spatial unit

The last step of the framework involves calculating the sensitivity for each spatial unit, which in DEGURBA's grid classification implies the 1 km² grid cells. The sensitivity of a grid cell is quantified by the indicator of ordinal dispersion (as described in Section 3.1). Figure 3 illustrates the value of S for several class distributions of DEGURBA. For example, if a cell is classified as urban centre in 60% of the realisations, as urban cluster in 30% of the realisations, and as rural cell in the remaining 10% (example 4 in Figure 3), then S is calculated as follows.

Figure 3: Illustration of sensitivity value for different class distributions.

$$d^2 = \sum_{i=1}^{3-1} (F_i - 0.5)^2 = (0.60 - 0.50)^2 + \big((0.60 + 0.30) - 0.50\big)^2 = 0.17 \qquad (4)$$

$$d^2{}_{max} = (3 - 1) * 0.25 = 0.5 \qquad (5)$$

$$S = 1 - \sqrt{\frac{0.17}{0.5}} = 0.4169 \quad (6)$$

The value of 0.4169 signifies that the dispersion of the grid cell in the set of alternative realisations is 41.69% of the maximal dispersion possible with three ordinal categories. The maximum dispersion is achieved when a cell is classified as urban centre in 50% of the classifications, and as rural cell in 50% of the classification (example 7 in Figure 3). The sensitivity index should thus be interpreted in relative terms. Example 2 and 3 in Figure 3 illustrate the fact that different class distributions can have the same sensitivity value. In interpreting the results, it is thus relevant to complement the final sensitivity map with a set of alternative realisations, and class distribution histograms, as shown in Figure 3.

4 RESULTS

We applied the spatially-explicit sensitivity framework to the definition of DEGURBA and computed the sensitivity of each grid cell on a global scale based on the 1458 different realisations. More than 93% of informative cells*** have $S \leq 0.1$. These relatively low-sensitive cells are mainly located in sparsely populated areas or the core of large cities with high population densities. Unsurprisingly, cells with larger sensitivities are generally located at the urban fringe and in regions with average population densities. These more highly sensitive cells are, however, not equally distributed across the world (see Figure 4). Sensitive cells are clustered in certain regions and specific types of urbanisation patterns. For a more detailed exploration of the sensitivity map, the 1458 alternative realisations, and the class distribution histograms of all individual grid cells at a world scale: https://platform-dou.snl.ees.kuleuven.be.

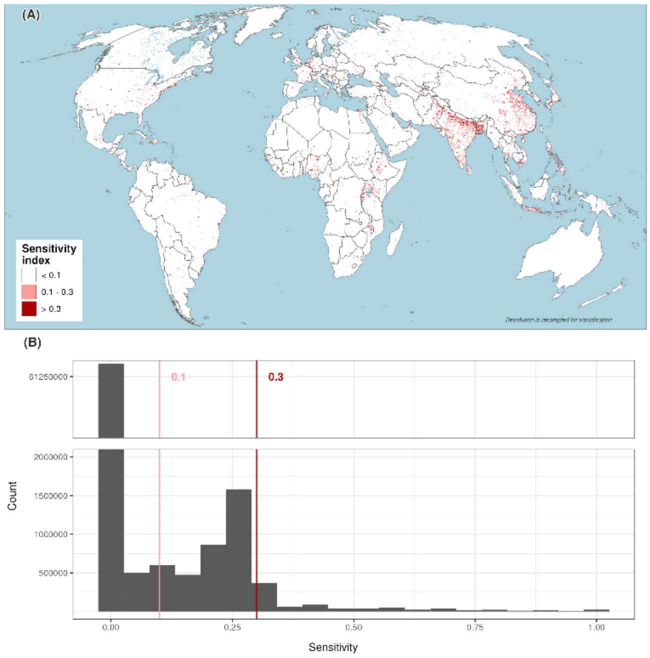

Figure 4: (A) Global distribution of sensitivity and (B) Histogram of sensitivity values of informative cells.

In the following sections, we focus on three specific case studies to illustrate how the observed sensitivity in a region relates to underlying urbanisation processes. We also reveal what specific criteria might be at the root of this sensitivity by exploring a set of alternative realisations. In Section 4.1 and 4.2, we discuss two regions with a relatively high sensitivity: the Red River Delta (RRD; Vietnam) and the urbanised area around Dallas, Fort Worth, and Arlington (DFWA; Texas, USA), respectively. In Section 4.3, we focus on Vienna (Austria), a city with comparatively less sensitivity. Figure 5 visualises the sensitivity maps and three alternative realisations of DEGURBA in the three case studies. There are notable differences in the urban delineations in the RRD and the area around DFWA when varying criterion settings in DEGURBA, while delineations around Vienna are more robust to small changes.

*** Non-informative cells are cells with no population in a window of two cells around it and cells that are classified as water in all realisations. This ensures oceans and uninhabited mountainous areas, large forest, deserts, etc. are not included in our results.

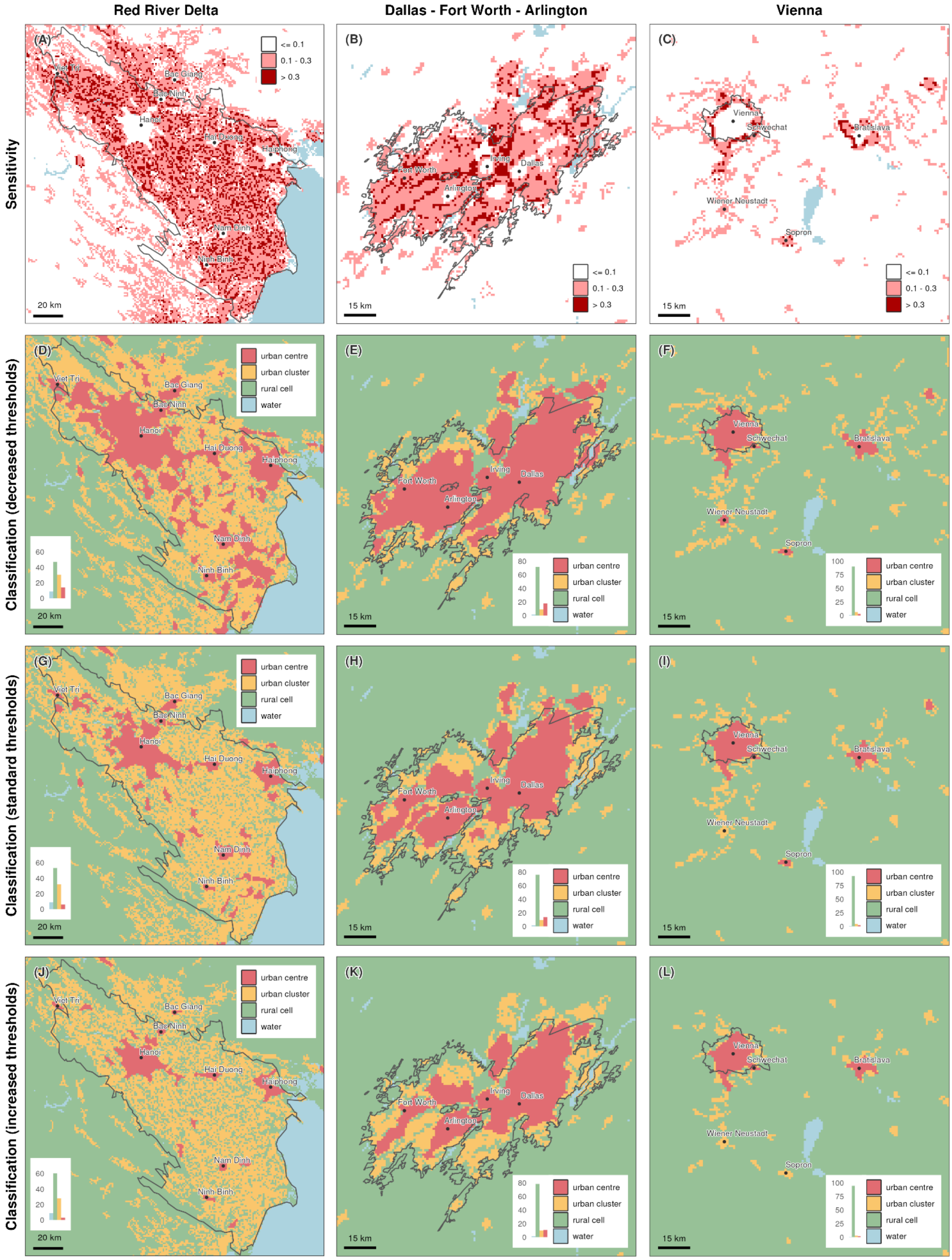

Figure 5: Sensitivity and three grid classifications in the Red River Delta, the urbanised area of Dallas – Fort Worth – Arlington and the city of Vienna. Grid classification D-E-F with density thresholds 1,000 and 200, and size thresholds 33,333 and 3,333, G-H-I with density thresholds 1,500 and 300, and size thresholds 50,000 and 5,000, and J-K-L with density thresholds 2,250 and 450, and size thresholds 75,000 and 7,500 for urban centres and urban clusters, respectively. For grid classifications D to L, other criteria were kept constant as follows: rook contiguity for urban centres and queen contiguity for urban clusters, the standard smoothing procedure of DEGURBA, and a fixed built-up threshold of 20%.

4.1 Sensitivity in the Red River Delta (Vietnam)

The RRD, consisting of the low-lying fertile plains around Hanoi, is among the regions with the highest sensitivity across the world: the average and median value of S in this area[†††] is 0.26 and 0.25, respectively. More than 22% of the cells have $S > 0.3$, housing approximately 1 million people or 4% of the total population in the entire RRD. The classification of the grid cells varies across all categories, with some cells changing directly between urban centre and rural cell, thus skipping the middle category of urban cluster. This sensitivity pattern can be understood based on the morphological character and geohistorical context of this region. After the introduction of *Doi Moi* policies in Vietnam in 1986, the RRD transformed from predominantly agriculturally-oriented into densely populated with mixed-use development (van Horen, 2005). These policies aimed to transform Vietnam into an open, market-oriented economy and had significant effects on the RRD – especially along the Hanoi-Haiphong axis – with the emergence of small industries and craft villages, diversification in employment, and seasonal migration to the major cities (Labbé, 2016). The *in situ* urbanisation of the previous agrarian society led to a *desakota* landscape of heavily populated cores surrounded by intensively cultivated agricultural land (McGee, 2000, 2013). This landscape, containing both urban and rural elements (Cohen, 2004), is sensitive to minor variations in population thresholds, as can be seen in Figure 5D, G and J.

Apart from the population thresholds, the classification of the *desakota* landscape is also sensitive to other criteria. For instance, adapting the contiguity rule for urban centres significantly increases the extent of and population in urban centres (see Figure 6A and B). With rook contiguity, approximately 10 million people live in urban centres, while with queen contiguity, this becomes more than 12.5 million. This sensitivity can again be explained by the morphological patterns in the RRD. The region has significant local variations in population density with densely populated cores and sparsely populated agricultural lands. As a result, neighbouring high-density cells might only share a corner and no full edge and are consequently only considered contiguous when using the queen's contiguity rule. Interestingly, this sensitivity to the urban centres' contiguity rule occurs in the RRD *and* other large delta areas, such as the Ganges River Delta (India/Bangladesh) and the Nile Delta (Egypt; see Figure 6C). It appears that the specific morphological pattern that is abundant in river deltas with fertile lands in combination with recent population growth is disproportionally affected by changing this seemingly minor criterion in DEGURBA.

[†††] Boundary of the RRD according to General Statistics Office Vietnam (2021). Quảng Ninh Province and Cát Bà Island were excluded because these areas are occasionally considered to be part of the North-eastern region of Vietnam.

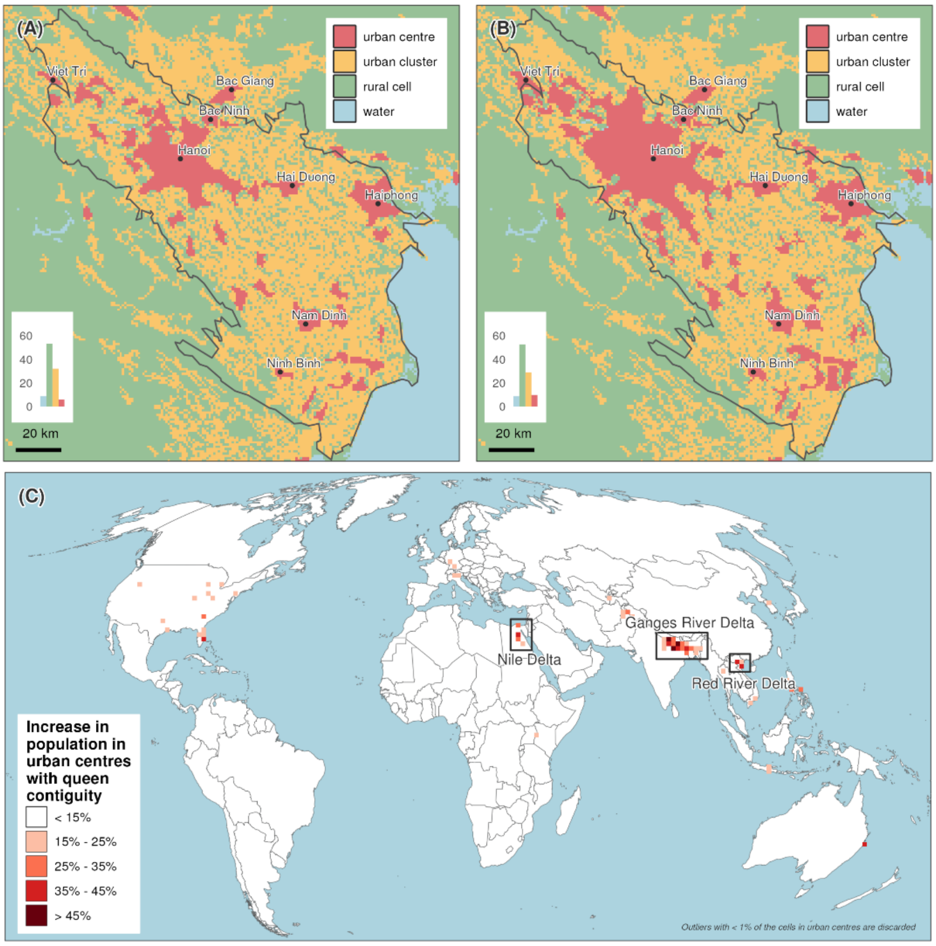

Figure 6: Grid classification in the RRD with (A) rook contiguity for urban centres and queen contiguity for urban clusters and (B) queen contiguity for both urban centres and urban clusters. Other criteria were kept constant as follows: density thresholds 1,500 and 300, and size thresholds 50,000 and 5,000 for urban centres and urban clusters, respectively, the standard smoothing procedure of DEGURBA, and a fixed built-up threshold of 20%. (C) Global sensitivity to changing the contiguity rule for urban centres from rook to queen.

4.2 Sensitivity in the urbanised area of DFWA (USA)

There is also significant variation in urban delineations in the urbanised area of DFWA[‡‡‡], with an average and median S of 0.24 and 0.23. More than 16% of the cells have $S > 0.3$, collectively containing more than 200,000 inhabitants or 3% of the total population in the DFWA area. The region's morphology is mainly dominated by highways and suburbanisation (Liu et al., 2019) and has a strong decentralised and polycentric character (McMillen, 2001). Apart from the Central Business Districts of Dallas and Fort Worth, there are numerous other high-density centres, including the technological business district of Richardson and the Stemmons industrial corridor (Shukla & Waddell, 1991). This polycentric development results in significant fragmentation of the urban landscape when employing high population thresholds in DEGURBA. For example, Arlington and Fort Worth are considered part of a single urban centre in Figure 5E, while they are separated in H and K.

Not only the population thresholds but also the built-up density threshold in DEGURBA is critical in the DFWA area. Its value determines the degree to which urban centres are fragmented. For example, the large commercial and industrial area between Irving and Dallas contains many buildings but almost no residential population. With a relatively low built-up density threshold of 20% (Figure 7A), the site

[‡‡‡] Boundary according to the US Census Bureau (2022).

is included when identifying urban centres, thus connecting the cores of Irving and Dallas. However, an increased built-up density threshold of 50% (Figure 7B) no longer sees this site added to urban centres and generates smaller and more fragmented urban centres. This effect of the built-up density threshold occurs not only in the DFWA area but in most cities in the United States (see Figure 7C). In Tulsa and Jacksonville, the population in urban centres decreases with more than 60% when employing a built-up density threshold of 50% in comparison to 20%. Nevertheless, the effect of changing the threshold is almost negligible in other parts of the world. North American cities – generally containing more low-density development – are thus disproportionally sensitive to this criterion in DEGURBA.

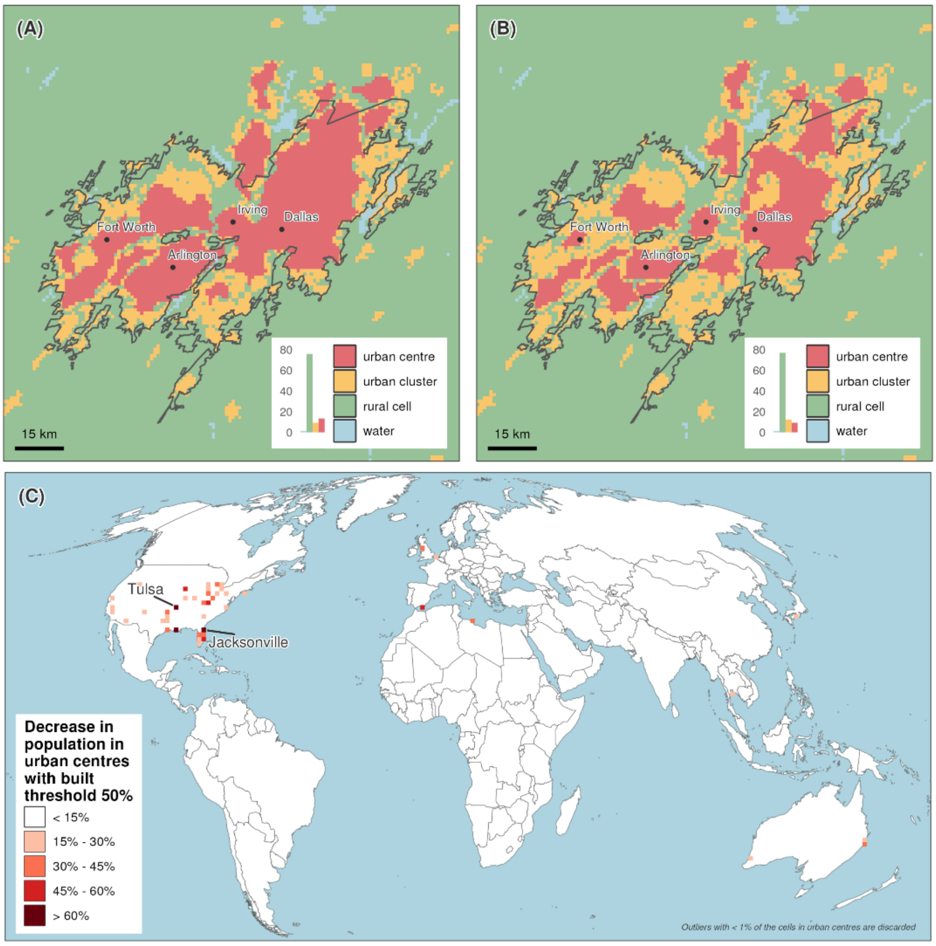

Figure 7: Grid classification in the DFWA area with (A) a fixed built-up density threshold of 20% and (B) a fixed built-up density threshold of 50% for urban centres. Other criteria were kept constant as follows: density thresholds 1,500 and 300, and size thresholds 50,000 and 5,000 for urban centres and urban clusters, respectively, the standard smoothing procedure of DEGURBA, and rook contiguity for urban centres and queen contiguity for urban clusters. (C) Global sensitivity to changing the built-up density threshold from 20% to 50%.

4.3 Sensitivity in the city of Vienna

The city of Vienna, with a mean and median S of 0.11 and 0, respectively, exhibits less sensitivity to criteria variation in DEGURBA compared to the RRD and the area of DFWA. Despite the remarkable median value of S equal to zero, 12% of the cells in the city proper have $S > 0.3$. These cells are mainly situated along the green belt at the edge of the city and house only 0.6% of Vienna's population. This implies that few inhabitants would be affected by potential variation in the delineation. The sensitivity of these cells can be attributed to the variation in the smoothing procedure: the cells may not belong to an urban centre with standard edge smoothing, while they are included with the alternative population smoothing procedure.

The relatively low sensitivity of Vienna can be understood through the physical surroundings of the city – the west side of the city is delimited by green and hilly areas, serving as a natural frontier –, and its historical urban development. With the fall of the Iron Curtain in 1989, Vienna experienced a population growth and drastic shift in planning culture. There was a strong focus on inner-city development and reutilisation of abandoned land, and new office complexes were constructed at the edge of the city – but within the city proper. The efficiency of the comprehensive inner-city development policy was eased by the fact that Vienna has its own politically autonomous region (European Investment Bank, 2018; Hatz, 2008).

Vienna is not the only city exhibiting relatively low sensitivity. Other examples are the agglomeration of São Paulo (Brazil), Tehran (Iran) and Kinshasa–Brazzaville (DRC – Congo Brazzaville). The periphery of these urban areas is sensitive to changes, which is logical and perhaps almost inevitable. However, apart from this, the DEGURBA methodology performs generally well in these regions (see Figure 8).

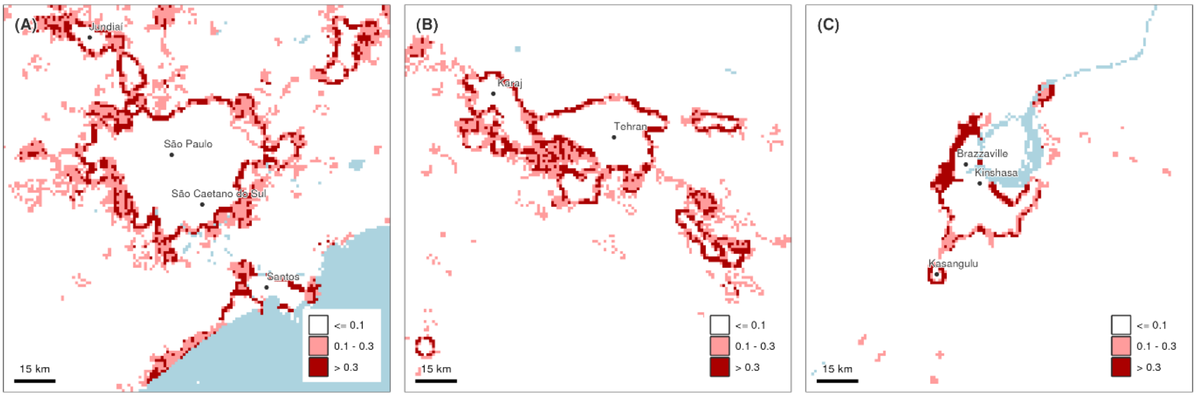

Figure 8: Sensitivity map in the agglomeration of (A) São Paulo, (B) Tehran and (C) Kinshasa–Brazzaville.

5 DISCUSSION

We proposed a spatially-explicit sensitivity framework to examine potential regional bias in urban delineation algorithms and employed it to analyse the implications tied to specific criterion choices in DEGURBA at a fine spatial granularity. The resulting sensitivity map gives an overview of the set of alternative realisations and serves as exploration tool; it guides us to potentially interesting regions. By zooming in on these regions and exploring a set of alternative realisations, we can assess the link between sensitivity and the underlying urban morphological pattern and its drivers. In that way, we gain a deeper understanding of the implications of (explicit and implicit) assumptions tied to certain criterion choices. Many of these assumptions in DEGURBA are inherently linked to its intended purpose to facilitate international comparison of the SDG indicators. Nonetheless, it is crucial to bear in mind these assumptions when employing the DEGURBA delineations for other applications, for example when computing accessibility to ‘urban’ services (Weiss et al., 2018).

The results demonstrated that different regions in the world are indeed sensitive to (different types of) small changes in DEGURBA’s algorithm and that the distribution of highly sensitive cells is not equally distributed across the world. Certain regions and specific types of urbanisation exhibit higher sensitivity than others, potentially pointing to geographical skewness in the definition. For instance, the results showed that changing the contiguity rule for urban centres in DEGURBA disproportionally affects certain regions, specifically large delta regions including the Ganges River Delta, Nile Delta, and

Red River Delta. In other parts of the world – including Europe – the effect of changing the contiguity rule is rather limited. However, this contiguity rule is actually a practical consequence of working with rectangular grid cells rather than an inherent aspect of urbanisation. The developers of DEGURBA might not have been aware of the implications of this rather technical criterion, because the definition was initially developed in the European context – where the choice of contiguity is less impactful. This, in a way, demonstrates that knowledge production about 'the urban' may be partly shaped by specific urban realities experienced by those producing the knowledge.

The results also revealed that cities in North America are disproportionally influenced by changing the built-up density threshold for urban centres. According to DEGURBA's documentation, the built-up criterion is introduced to reduce fragmentation and avoid generating multiple urban centres for a single 'city'. The European Commission (2023, p. 51, footnote 30) specify that the rule is explicitly established for "a few countries with relatively low-density urban development and a strong separation of land use functions". It thus appears that DEGURBA incorporates the specific urbanisation pattern that is abundant in North American cities, as the definition contains a rule to 'better' delineate these types of urban agglomerations. The fact that DEGURBA does not include such a targeted rule for urbanisation patterns in other parts of the world – although this would be possible theoretically speaking[§§§] – might point to an implicit bias towards North American urbanisation patterns.

The point here is not to criticise the specific definition of DEGURBA. Instead, DEGURBA serves as an example to illustrate that, because there is no single 'true' definition of urbanisation, every urban definition inevitably requires making specific choices and assumptions about the outlook of 'the urban', which may, unwittingly, reflect specific types of urban patterns more effectively than others. Given the challenging task of developing globally consistent urban delineations, the DEGURBA method is, in fact, well-conceived and contributed significantly to the field of urban studies. There are also strong theoretical arguments to be made for the implementation choices in DEGURBA in light of its objective of monitoring the SDGs. However, it is important to keep in mind the spatial sensitivities tied to its delineation criteria, especially when employing DEGURBA beyond its intended application.

Applying the spatially-explicit sensitivity framework to other urban definitions is an interesting avenue for future research. The framework is designed generically and can thus be applied to reveal implicit assumptions in other definitions and their potential implications. Furthermore, future studies might extend the framework by integrating the decomposition of sensitivity into the partial contribution of each criterion. Currently, we do not quantify the exact individual contribution of each, as is often done in other sensitivity analyses. Variance-based global sensitivity analyses can be employed to achieve this (Saltelli et al., 2010), but these generally require integral calculus, which is not feasible for rule-based urban delineation approaches that include topological relationships or complex smoothing procedures. However, future work may develop innovative and clever ways to tackle this issue.

6 CONCLUSION

There is no broadly shared agreement on how to define or delineate urban areas. As a consequence, every urban definition inevitably requires making certain decisions and assumptions that are challenging to justify from a global point of view. These may lead to spatially heterogeneous sensitivities that embody regional skewness in the definition. Against this backdrop, we developed a spatially-explicit sensitivity framework to make the implications of certain criterion choices both

[§§§] We do not suggest that DEGURBA should be altered in any way. We realise that introducing other rules such as a maximum employment in the agricultural sector may not be feasible due to data availability issues, neither desirable as it can lead to interference when monitoring certain SDG indicators (Dijkstra et al., 2021).

legible and visible, and uncover potential geographical biases in urban delineation approaches. Applying the framework to the definition of DEGURBA revealed two specific points. First, implications tied to the contiguity criterion for urban centres are most pronounced in specific delta regions including the Ganges River Delta, Nile Delta, and Red River Delta. The developers of DEGURBA may not have been aware of these implications as the definition is initially designed for settlements in the European Union, where the influence of varying the contiguity rule is rather limited. Second, DEGURBA includes a specific rule to reduce fragmentation in regions with low-density urban development, a pattern that is abundant in the North American context. As such, the definition explicitly incorporates North American urbanisation pattern, while no targeted rules exist for urbanisation patterns that are common in other world regions. Both findings demonstrate, in a way, that urban definitions inevitably reflect and are shaped by the urban context and situated knowledge of those articulating them, even though this is likely not intentional. Essentially, this reiterates the broader conceptual circularity problem, where the definition of a theoretical concept (the 'urban') and its operationalisation (DEGURBA) are intertwined and might even reenforce one another.

7 DATA AVAILABILITY

The code and data needed to replicate the analysis are published under the following doi: https://doi.org/10.48804/VFAGOQ. Interactive visualisation of the results is provided on the following platform: https://platform-dou.snl.ees.kuleuven.be.

8 FUNDING ACKNOWLEDGEMENT

The research for this paper was supported by Research Foundation Flanders (FWO) under PhD number 11P4224N; by patronage funding provided to KU Leuven for carrying out scientific research into urban processes and change; and partly by Internal Funds KU Leuven grant number STG/20/021.

9 REFERENCES

Angel, S., Lamson-Hall, P., Galarza, N., & Blei, A. (2018). *Our Not-So-Urban World*. The Marron Institute of Urban Management. https://marroninstitute.nyu.edu/papers/our-not-so-urban-world

Asian Development Bank. (2019). *Fostering Growth and Inclusion in Asia's Cities*. https://doi.org/10.22617/FLS190070-3

Balk, D., Leyk, S., Montgomery, M. R., & Engin, H. (2021). Global Harmonization of Urbanization Measures: Proceed with Care. *Remote Sensing*, *13*(24), 4973. https://doi.org/10.3390/rs13244973

Blair, J., & Lacy, M. G. (2000). Statistics of Ordinal Variation. *Sociological Methods & Research*, *28*(3), 251–280. https://doi.org/10.1177/0049124100028003001

Bosker, M., Park, J., & Roberts, M. (2021). Definition matters. Metropolitan areas and agglomeration economies in a large-developing country. *Journal of Urban Economics*, *125*, 103275. https://doi.org/10.1016/j.jue.2020.103275

Bunnell, T., & Maringanti, A. (2010). Practising Urban and Regional Research beyond Metrocentricity. *International Journal of Urban and Regional Research*, *34*(2), 415–420. https://doi.org/10.1111/j.1468-2427.2010.00988.x

Ch, R., Martin, D. A., & Vargas, J. F. (2021). Measuring the size and growth of cities using nighttime light. *Journal of Urban Economics*, *125*, 103254. https://doi.org/10.1016/j.jue.2020.103254

Chen, Y., Yu, J., & Khan, S. (2013). The spatial framework for weight sensitivity analysis in AHP-based multi-criteria decision making. *Environmental Modelling & Software*, *48*, 129–140. https://doi.org/10.1016/j.envsoft.2013.06.010

Cohen, B. (2004). Urban Growth in Developing Countries: A Review of Current Trends and a Caution Regarding Existing Forecasts. *World Development*, *32*(1), 23–51. https://doi.org/10.1016/j.worlddev.2003.04.008

Cottineau, C., Batty, M., Benenson, I., Delloye, J., Hatna, E., Pumain, D., Sarkar, S., Tannier, C., & Ubarevičienė, R. (2024). The role of analytical models and their circulation in urban studies and policy. *Urban Studies*, 00420980241237410. https://doi.org/10.1177/00420980241237410

Cottineau, C., Hatna, E., Arcaute, E., & Batty, M. (2017). Diverse cities or the systematic paradox of Urban Scaling Laws. *Computers, Environment and Urban Systems*, *63*, 80–94. https://doi.org/10.1016/j.compenvurbsys.2016.04.006

Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M., & Schiavina, M. (2021). Applying the Degree of Urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*, *125*, 103312. https://doi.org/10.1016/j.jue.2020.103312

Dorward, N., Fox, S., Statham, T., & Wolf, L. J. (2023). A spatial-demographic analysis of Africa’s emerging urban geography. *Environment and Urbanization*, *35*(2), 310–327. https://doi.org/10.1177/09562478231190735

Duranton, G. (2021). Classifying locations and delineating space: An introduction. *Journal of Urban Economics*, *125*, 103353. https://doi.org/10.1016/j.jue.2021.103353

European Commission. (2023). *GHSL Data Package 2023* (JRC133256; Publications Office of the European Union). https://ghsl.jrc.ec.europa.eu/documents/GHSL_Data_Package_2023.pdf

European Investment Bank. (2018). *Why Vienna gets high marks*. https://doi.org/10.2867/9448

Eurostat. (2016). *Urban Europe—Statistics on cities, towns and suburbs*. https://ec.europa.eu/eurostat/web/products-statistical-books/-/ks-01-16-691

Eurostat. (2021). *Applying the Degree of Urbanisation—A methodological manual to define cities, towns and rural areas for international comparisons—2021 edition*. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-02-20-499

Feizizadeh, B., Jankowski, P., & Blaschke, T. (2014). A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. *Computers & Geosciences*, *64*, 81–95. https://doi.org/10.1016/j.cageo.2013.11.009

Feizizadeh, B., & Kienberger, S. (2017). Spatially explicit sensitivity and uncertainty analysis for multicriteria-based vulnerability assessment. *Journal of Environmental Planning and Management*, *60*(11), 2013–2035. https://doi.org/10.1080/09640568.2016.1269643

Florida, R., Gulden, T., & Mellander, C. (2008). The rise of the mega-region. *Cambridge Journal of Regions, Economy and Society*, *1*(3), 459–476. https://doi.org/10.1093/cjres/rsn018

General Statistics Office Vietnam. (2021). *Statistical Yearbook of 2020*.

Hatz, G. (2008). Vienna. *Cities*, *25*(5), 310–322. https://doi.org/10.1016/j.cities.2008.02.002

Henderson, J. V., Nigmatulina, D., & Kriticos, S. (2021). Measuring urban economic density. *Journal of Urban Economics*, *125*, 103188. https://doi.org/10.1016/j.jue.2019.103188

Henderson, M., Yeh, E. T., Gong, P., Elvidge, C., & Baugh, K. (2003). Validation of urban boundaries derived from global night-time satellite imagery. *International Journal of Remote Sensing*, *24*(3), 595–609. https://doi.org/10.1080/01431160304982

Humberstone, I. L. (1997). Two Types of Circularity. *Philosophy and Phenomenological Research*, *57*(2), 249–280. https://doi.org/10.2307/2953718

Kanai, J. M., Grant, R., & Jianu, R. (2018). Cities on and off the map: A bibliometric assessment of urban globalisation research. *Urban Studies*, *55*(12), 2569–2585. https://doi.org/10.1177/0042098017720385

Kocabas, V., & Dragicevic, S. (2006). Assessing cellular automata model behaviour using a sensitivity analysis approach. *Computers, Environment and Urban Systems*, *30*(6), 921–953. https://doi.org/10.1016/j.compenvurbsys.2006.01.001

Labbé, D. (2016). Critical reflections on land appropriation and alternative urbanization trajectories in periurban Vietnam. *Cities*, *53*, 150–155. https://doi.org/10.1016/j.cities.2015.11.003

Ligmann-Zielinska, A., & Jankowski, P. (2014). Spatially-explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation. *Environmental Modelling & Software*, *57*, 235–247. https://doi.org/10.1016/j.envsoft.2014.03.007

Lilburne, L., & Tarantola, S. (2009). Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, *23*(2), 151–168. https://doi.org/10.1080/13658810802094995

Liu, X., Pan, Q., King, L., & Jin, Z. (2019). Analysing the changes of employment subcentres: A comparison study of Houston and Dallas. *Urban Studies*, *56*(12), 2532–2548. https://doi.org/10.1177/0042098018789554

McGee, T. G. (2000). The urban future of Vietnam reconsidered. *J Ritsumeikan Geogr Soc*, *12*, 1–18.

McGee, T. G. (2013). The Emergence of Desakota Regions in Asia: Expanding a Hypothesis. In N. Brenner (Ed.), *Implosions /Explosions: Towards a Study of Planetary Urbanization* (pp. 121–137). JOVIS. https://doi.org/10.1515/9783868598933-010

McMillen, D. P. (2001). Nonparametric Employment Subcenter Identification. *Journal of Urban Economics*, *50*(3), 448–473. https://doi.org/10.1006/juec.2001.2228

Melchiorri, M., Pesaresi, M., Florczyk, A. J., Corbane, C., & Kemper, T. (2019). Principles and Applications of the Global Human Settlement Layer as Baseline for the Land Use Efficiency Indicator—SDG 11.3.1. *ISPRS International Journal of Geo-Information*, *8*(2), 96. https://doi.org/10.3390/ijgi8020096

Moreno-Monroy, A. I., Schiavina, M., & Veneri, P. (2021). Metropolitan areas in the world. Delineation and population trends. *Journal of Urban Economics*, *125*, 103242. https://doi.org/10.1016/j.jue.2020.103242

Potts, D. (2018). Urban data and definitions in sub-Saharan Africa: Mismatches between the pace of urbanisation and employment and livelihood change. *Urban Studies*, *55*(5), 965–986. https://doi.org/10.1177/0042098017712689

Roy, A. (2009). The 21st-Century Metropolis: New Geographies of Theory. *Regional Studies*, *43*(6), 819–830. https://doi.org/10.1080/00343400701809665

Şalap-Ayça, S., & Jankowski, P. (2016). Integrating local multi-criteria evaluation with spatially explicit uncertainty-sensitivity analysis. *Spatial Cognition & Computation*, *16*(2), 106–132. https://doi.org/10.1080/13875868.2015.1137578

Şalap-Ayça, S., Jankowski, P., Clarke, K. C., Kyriakidis, P. C., & Nara, A. (2018). A meta-modeling approach for spatio-temporal uncertainty and sensitivity analysis: An application for a cellular automata-based Urban growth and land-use change model. *International Journal of Geographical Information Science*, *32*(4), 637–662. https://doi.org/10.1080/13658816.2017.1406944

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, *181*(2), 259–270. https://doi.org/10.1016/j.cpc.2009.09.018

Schiavina, M., Melchiorri, M., Pesaresi, M., Politis, P., Freire, S., Maffenini, L., Florio, P., Ehrlich, D., Goch, K., Tommasi, P., & Kemper, T. (2022). *GHSL Data Package 2022* (JRC129516; Publications Office of the European Union). https://ghsl.jrc.ec.europa.eu/documents/GHSL_Data_Package_2022.pdf

Shukla, V., & Waddell, P. (1991). Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort worth. *Regional Science and Urban Economics*, *21*(2), 225–253. https://doi.org/10.1016/0166-0462(91)90035-L

Statham, T. A., Wolf, L. J., & Fox, S. (2021). *Identifying urban areas: A new approach and comparison of national urban metrics with gridded population data*. https://doi.org/10.31235/osf.io/abvc6

Tang, Z., Zhang, H., Yi, S., & Xiao, Y. (2018). Assessment of flood susceptible areas using spatially explicit, probabilistic multi-criteria decision analysis. *Journal of Hydrology*, *558*, 144–158. https://doi.org/10.1016/j.jhydrol.2018.01.033

Taubenböck, H., Droin, A., Standfuß, I., Dosch, F., Sander, N., Milbert, A., Eichfuss, S., & Wurm, M. (2022). To be, or not to be 'urban'? A multi-modal method for the differentiated measurement of the degree of urbanization. *Computers, Environment and Urban Systems*, *95*, 101830. https://doi.org/10.1016/j.compenvurbsys.2022.101830

Uchida, H., & Nelson, A. (2009). *Agglomeration Index: Towards a New Measure of Urban Concentration*. World Bank. https://openknowledge.worldbank.org/handle/10986/9039

United Nations Population Division. (2019). *World Urbanization Prospects. The 2018 Revision.* https://population.un.org/wup/Publications/

United Nations Statistics Division. (2019). *United Nations Expert Group Meeting on Statistical Methodology for Delineating Cities and Rural Areas*. https://unstats.un.org/unsd/demographic-social/meetings/2019/newyork-egm-statmeth.cshtml

US Census Bureau. (2022). *TIGER/Line Shapefile of Urbanized Areas 2020* [Dataset].

van Horen, B. (2005). Hanoi. *Cities*, *22*(2), 161–173. https://doi.org/10.1016/j.cities.2005.01.006

Van Migerode, C., Poorthuis, A., & Derudder, B. (2024). Flexurba: An open-source R package to flexibly reconstruct the Degree of Urbanisation classification. *Environment and Planning B: Urban Analytics and City Science*, 23998083241262545. https://doi.org/10.1177/23998083241262545

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., … Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, *553*(7688), Article 7688. https://doi.org/10.1038/nature25181

Wineman, A., Alia, D. Y., & Anderson, C. L. (2020). Definitions of "rural" and "urban" and understandings of economic transformation: Evidence from Tanzania. *Journal of Rural Studies*, *79*, 254–268. https://doi.org/10.1016/j.jrurstud.2020.08.014

Yang, Q., Xu, Y., Tong, X., Hu, T., Liu, Y., Chakraborty, T. C., Yao, R., Xiao, C., Chen, S., & Ma, Z. (2023). Influence of urban extent discrepancy on the estimation of surface urban heat island intensity: A global-scale assessment in 892 cities. *Journal of Cleaner Production*, *426*, 139032. https://doi.org/10.1016/j.jclepro.2023.139032