Land Use Policy 112 (2022) 105822

Contents lists available at ScienceDirect

Land Use Policy

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

Exploring the spectrum of urban area key figures using data from Finland and proposing guidelines for delineation of urban areas

Jaana Vanhatalo [a,*], Jenni Partanen [b]

[a] Faculty of Built Environment, School of Architecture, Tampere University, 33014 Tampere, Finland
[b] Academy of Architecture and Urban Studies, Tallinn University of Technology, Estonia

ARTICLE INFO

Keywords:
Urban population
Urban area definition
Delineation of urban area
Degree of urbanisation
Spatial analyses

ABSTRACT

Urban area definitions are commonly used worldwide to reflect countries' urban population percentages. The measurements are based on local factors and differ widely across countries, making them non-comparable. This is well acknowledged among specialists; however, they are commonly used in everyday practice as universal measurements, and even compared with each other. The problem is that we do not know the degree of error in such comparisons. For this purpose, in the study presented here we analysed and categorised different European national urban area definitions, testing them in the case of Finland. Definitions from 27 European countries were divided into two main categories according to the end result areal unit of the definition and further into seven subcategories based on the criteria used. Thirteen different definitions in the case of Finland were tested using spatial analyses with GIS. The results indicate that urban population percentage varies widely according to different definitions, making their comparability infeasible. The difference is even greater in the case of urban area ratio and population densities of urban areas between the two main categories. The results prove that definitions based on LAU areas cannot illustrate urban areas and their densities coherently. In light of a literature review on certain relational urban area delineation methods and the case study, desirable characteristics for the delineation of urban areas were highlighted. Consequently, a constant, structured evaluation of urbanity measurements and their underlying logic is necessary to enable unambiguous discourse on urban area in urban sciences. The results could help in formulating the applicability of the concept of urban area in scientific and popular communication and media.

1. Introduction

Extensive urbanisation is a key challenge globally. Urban areas are responsible for the majority of pollution, services and industrial production and the key players in world trade and global economics. It has been estimated that by 2050 68% of the world's population will be living in cities (UN, 2019e, p. 1), increasing the importance of urban areas excessively. A reasonably coherent definition of urban area is indispensable to enable comparison of certain urban indicators and to achieve better planning and policymaking.

Urban area definitions commonly applied, however, are determined separately by the respective countries. The United Nations (UN) has stated that each country is responsible for establishing measures for the local level of urbanisation (UN, 2017b, p. 188). The indicators for urban area and applied areal units are tacitly agreed and vary widely across countries (see e.g. Cohen, 2004, p. 25; UN, 2019b, pp. 120–125). The areal unit may consist of administrative units, grid, or building-based areal units, with other embedded variables (e.g. population, workplaces, or ratio of agricultural workers). The figures based on national definitions, such as those presenting national urban population percentages, are commonly applied in everyday use in a comparative manner. Such a practice disregards the professionals' guidelines to refrain from comparison (Cohen, 2004, pp. 24–25; Satterthwaite, 2010, pp. 83–84). While accepting and promoting the use of national definitions, attempts have been made simultaneously to align European and global measures to achieve more coherent urban area definitions. These include, for example, the OECD Regional Typology (OECD, 2008, p. 4), the OECD Extended Regional Typology (Brezzi et al., 2011), the Urban-rural Typology (Eurostat, 2010, p. 240) and the DEGURBA-method (Degree of Urbanisation) (EC et al., 2020a, p. 6, 9; EC et al., 2020b, p. 2, 5; Eurostat, n.d.a).

Furthermore, the national definitions produce by default a dual

* Corresponding author.
E-mail addresses: jaana.vanhatalo@tuni.fi (J. Vanhatalo), jenni.partanen@taltech.ee (J. Partanen).

https://doi.org/10.1016/j.landusepol.2021.105822
Received 16 April 2021; Received in revised form 27 September 2021; Accepted 15 October 2021
Available online 27 October 2021

representation of reality: urban-rural. However, many urban scholars have recently suggested that such a dichotomy no longer reflects the real-world spatio-functional configurations. Instead, the boundary between urban and rural is considered obscure or fuzzy, a gradient, or the urban and rural form a continuum (Champion and Hugo, 2004; Dymitrow and Stenseke, 2016; Haase and Tötzer, 2012; López-Goyburu and Gonzaga García-Montero, 2018; Van Vliet et al., 2020; Wandl et al., 2014). Yet the dichotomous reading of urbanity is commonly used, especially in demography (Hugo et al., 2003). For example, the UN population tables are based on this dichotomy and the U.S. Census Bureau defines everything that is not urban as rural (Ratcliffe et al., 2016, p. 1; UN, 2019a, 2019c).

The fundamental question concerns the degree of error in comparing these national figures reflecting urbanisation, such as urban population percentages, resulting from the various measures. For the discourse concerning cities and their challenges, it would be crucial to arrive at a common understanding of what is meant by urban area in a quantitative sense along with the variety of the spectrum of measures. It is necessary to outline the spatial implications of these measurements, which indicators yield a certain percentage of "urban" environment: the concept of urban area needs clarification through a revision of the logic underlying the myriads of measurements. Further it is important to ponder what characteristics the delineation methods for urban areas should possess.

We therefore ask what kind of measurements European countries use in defining their urban areas as manifest in urban population percentages, and to what extent their interpretations of urban areas differ spatially and quantitatively. For this purpose, we compared urban area definitions across Europe against data from Finland to compare their variance. In addition, we ask what other possible approaches could be used to delineate urban areas.

We first categorised the definitions of urban used in 27 European countries according to their measurement idiosyncrasies. Secondly, for a spatial comparison between these we applied the different measurements using data from Finland. We produced 13 different interpretations of urban areas which were scrutinised using spatial analytical methods along with statistical measurements. Urban population percentage, urban area ratio and population density of urban areas were calculated for all cases, and similarities, regularities and distinguishing features between the resulting patterns were compared. Finally, based on other explored urban area delineation methods and this case study, desirable characteristics for future guidelines are highlighted.

2. Delineating urban areas and measuring urban population

2.1. A framework of measuring urban population in Europe

The quest to find universal measures for the level of urbanisation or to define urban areas is neither new nor contemporary, but dates back to the first half of the 20th century (see e.g. Arriaga, 1970; Bergel, 1955, pp. 3–14; Bogue, 1961; Gibbs, 1966; Wirth, 1938). Although cities and urban form have changed radically, breaking the traditional urban-rural dichotomy (see e.g. Brenner and Schmid, 2014; Champion and Hugo, 2004; Price according to Shane, 2006, p. 56; Sieverts, 2003), the discourse and expectations have remained surprisingly similar.

Soon after the founding of the United Nations, the Economic and Social Council established a Statistical Commission in 1946 (UN, 1946a). The aim was to support the work of the Economic and Social Council and to promote 'the development of national statistics and the improvement of their comparability' and 'the improvement of statistics and statistical methods generally' (UN, 1946b, p. 398). As the only global organisation with a mandate from almost all the nations of the world, the UN is authorised to assess and evaluate how urban area definitions, urban population statistics and recommendations have been contemplated in UN documents.

The first UN urban-rural population figures were published in the first United Nations Demographic Yearbook 1948 and annually from 1967 (UN, 1949 [...] 2018). Right from the beginning these figures were based on the national definitions and the non-comparability of the figures has been clearly stated. The UN considers that the respective countries have the best available knowledge of the local circumstances, making them the most reliable source of information regarding their urban area definitions (UN, 1973, p. 8, 1980a, p. 9, 2017b, p. 188).

To improve international comparison, while avoiding pursuing a universal measure for urban and rural areas, the UN (1949a, pp. 170–171, 1980b, p. 30) suggested that countries should report the tabulated population figures by size classes of *localities*. Locality should be defined 'as a distinct population cluster (also designated as inhabited place, populated centre, settlement and so forth) in which the inhabitants live in neighbouring or contiguous sets of living quarters, and that has a name or a locally recognised status' (UN, 1958, p. 11, 2017b, p. 187). This indicator was never established in the annual procedure and was used only occasionally (UN, 1949 [...] 2018). Capitals and cities with populations over 100,000 were first published in the Demographic Yearbook for 1952 and annually since 1965 (UN, 1949 [...] 2018). Like the urban area delineation, they are dependent on the national definitions.

Despite the variation between countries indicating that a single comprehensive urban area measure probably cannot provide a solution (UN, 1980b, p. 11, 2017b, p. 188), the discussion about an international definition continues (see e.g. EC et al., 2020a; EC et al., 2020b; UN, 1949b, p. 6, 1980b, 2017a, p. 39). However, the UN refrains from proposing any universal criteria for definitions of urban areas for countries, providing only recommendations.

In addition, the EU and the OECD have endeavoured to produce universal measures for urban areas. In the 1990s the OECD developed the OECD Regional Typology (OECD, 2008, p. 4). After the EU suggested some modifications to this, the method was further adopted by the EU in 2010 as Urban-rural Typology (Eurostat, 2010, p. 240). In 2011 the OECD redefined their typology by adding a remoteness factor and calling it the OECD Extended Regional Typology (Brezzi et al., 2011). The EU and OECD also use the Degree of Urbanisation classification (DEGURBA) method (Eurostat, n.d.a). However, instead of classifying population data, this method has been used for other statistical purposes, such as economic indicators.

These OECD and EU methods classify areas into urban, intermediate and rural. The problem with typologies producing three or more classes is their applicability for calculating the urban population percentage, which requires an earlier classification to urban and rural areas. Their scale of observation also determines their usage.[1] Since the Regional Typology and the Urban-rural Typology are regional definitions, the results are inaccurate on a national level. Consequently, they have also been criticised for the omission of essential information on smaller areas and for non-comparable basic areal units: their sizes (km2) vary widely across countries (Dijkstra and Poelman, 2008, p. 3; Eurostat, 2010, p. 240; Fertner, 2012). The DEGURBA method appears more suitable for application to smaller, municipality-level units.

In 2016 the UN started co-operation with the EU Commission, OECD, ILO, FAO and the World Bank to evaluate the suitability of DEGURBA for global use. After testing, consultation and adjustment (EC et al., 2020b, pp. 23–24), DEGURBA is considered to provide a global definition of urban areas to enable global comparisons of certain indicators and to ensure higher-quality urban and rural statistics (EC et al., 2020a, p. 6, 9; EC et al., 2020b, p. 2, 5). However, it is intended to supplement the

[1] The original OECD Regional Typology is based on Local Administrative Unit 2 -areas (LAU2) corresponding to municipalities or similar units (Eurostat, n.d.b), and further on Territorial Level 3 (TL3) (OECD, 2008, p. 4), which mostly corresponds, for example with provinces or regional councils (in Europe NUTS3 level) (OECD, 2020, p. 3). The EU Urban-rural Typology is based on a 1 km grid and further on NUTS3 units, whereas DEGURBA is based on a 1 km grid and further on LAU areas (Eurostat, 2019, p. 2).

national definitions rather than replacing them (EC et al., 2020a, p. 9; EC et al., 2020b, p. 5). Regarding urban population, DEGURBA follows its predecessors, problematically comprehending three classes.

The DEGURBA document states that the strength of the method is its reliance on the population grid (EC et al., 2020b, p. 2). This is correct if we settle for population residency as the only, and sufficient, measure for urban areas. However, we argue that urban is essentially a 'place-based characteristic', seeing urban areas as 'a spatial concentration of people whose lives are organised around non-agricultural activities' (Weeks, 2010, p. 34). People's lives are also multilocal and mobile for various reasons (see e.g. Hugo, 1982; Hugo et al., 2003; Weeks, 2010, p. 38). Therefore, instead of population alone, we should also pay attention to the characteristics of places (see e.g. Hugo et al., 2003; Wirth, 1938).

The variables related to urban area are applied for global comparisons between different countries. One example is the comparison of urban population percentage and GDP. Such comparisons involving national urban variables, however, pose a fundamental problem for being state-istic.[2] The figures are not comparable as their underlying definitions are different. The figures may only indicate national situations.

Another problem with statistical comparisons is that the providers may use various national definitions or countries may have several concurrent official definitions (see e.g. Satterthwaite, 2010, p. 84). This, and the state-istic nature of the definitions, becomes problematic when the end-users interpret the statistics but are unaware of these (see e.g. Cohen, 2004, pp. 24–25; Satterthwaite, 2010, pp. 83–84). Here, scrutiny of the urban population percentages of 27 European countries using the United Nations Demographic Yearbook 2018 (UN, 2019b, pp. 159–168) and the World Urban Prospects 2018 (UN, 2019c) as a source, reveals variations in percentages and in national definitions used (Table 1). As these are both UN documents, the figures and definitions in them are all from national sources.

Only seven cases share the same definitions (Table 1). In nine cases the difference is unclear (e.g. the terms used could be different[3]) and seven cases were markedly different. The percentages differ widely: the greatest difference is 24.7 percentage points, which included four cases with a difference above 8.9pp. All four cases used a different definition in the documents. In two cases, Croatia and Germany, the urban population percentage is provided in both documents, but the definition is missing from the Demographic Yearbook 2018 (UN, 2019b, pp. 123–124). Presumably, Croatia used the same definition, in contrast to Germany, as their figure is higher in 2011 than in 2018. This example highlights the problematic nature of these statistical compilations, particularly when used without knowing the background of state-istic definitions[4] or the definitions used.

However, as urbanisation proceeds, it impacts population, economics, nature and resources (see e.g. Van Vliet et al., 2019). Many issues require global and regional comparisons of certain indicators (EC et al., 2020b, p. 2, 5). In addition, in the Global South the urban-rural division is important for efficient policies (Potts, 2018; Wineman et al., 2020, p. 254).[5] Consequently, along with statistical and modelling needs, classification implying spatial delineation is necessary, often leading to dichotomous solutions. For these, and as the state-istic urban population figures are constantly compared, it is crucial to explore the urban area definitions applied for measuring urban population: the types, the principles and their threshold values.

In parts 3, 4 and 5 we evaluate these in the European context. The testing of the impact is carried out in the case of Finland, asking: *How much will the urban population percentage, the urban area ratio and the population density of urban areas vary when applying different European national definitions*? The aim is to elucidate the spectrum of European definitions and discover h*ow much the differentiation between the definitions actually affects the figures.* But first, in Section 2.2, we explore the relational models for delineating urban areas. With these and the case study, we want to highlight certain desirable characteristics for urban area delineation approaches.

2.2. Recent progress in methodology: relational models for delineating urban areas

There has for decades been a tremendous number of new openings for methodology to delineate urban areas. In the 60s and 70s, for instance, Jack Gibbs (1966) and Eduardo Arriaga (1970) developed mathematical models to measure urbanisation. The basic limitation with these was that they were based on localities and cities defined by the countries themselves, which made them non-comparable in the first place. They were also very often based on administrative units (see e.g. UN, 1971, pp. 159–165). More recent proposals for a more universal method are likewise based on administrative units, for example on regions, municipalities, or postal code areas.[6] However, as stated in 2.1, the sizes of administrative units very often vary across countries. Therefore, the methods using administrative end result units may be incapable of portraying urban areas universally. Nevertheless, it is noteworthy that nowadays most of the methods for the delineation of urban areas apply GIS in some form. In addition to the location-based data and analyses, however, more use should be made of delineation of urban areas based on relational approaches as these reflect more accurately the dynamic and relational nature of cities. They include such approaches using, for example, remote sensing, use of artificial intelligence (AI) and computing.

For these latest applications, new types of datasets, such as telecommunication data, can become essential and open novel possibilities for delineating urban areas. The remote sensing-based methods and data

Table 1
The Demographic Yearbook 2018 versus World Urban Prospects 2018 data on urban population percentage of certain European countries (UN, 2019b, pp. 123–124, 159–168, 2019c, 2019d). All definitions and percentages the UN presents in both documents are from national sources. All missing information marked with '-'. In WUP 2018 column the italicised year indicates that the figure is an estimate (see UN, 2019d).

(*continued on next page*)

[2] *State-istic* is a term used by Brenner and Schmid (2014, p. 740) to illustrate the national characteristic of the urban area definitions.

[3] E.g. in the case of the Czech Republic one source uses the term localities and the other municipalities (UN, 2019b, pp. 124, 159–168, 2019d).

[4] The UN states in their document that the definitions are national and should be treated as such (see e.g. UN, 2019b, p. 5, 120–124), however, it is only possible to download the tables from the UN web pages (see. e.g. UN, 2019a).

[5] In Western societies the differences between urban and rural have levelled out (see e.g. Andersen et al., 2011; Potts, 2018; Schaeffer et al., 2013), i.e. the policymaking is less dependent on the urban-rural area distinction per se, than on more nuanced differences, how these are discussed, and also the existence of urban-rural interface or continuum, or territories-in-between (see e.g. Champion and Hugo, 2004; Dymitrow and Stenseke, 2016; López-Goyburu and Gonzaga García-Montero, 2018; Laurin et al., 2020; Urso, 2020; Van Vliet et al., 2020; Wandl et al., 2014; Woods, 2009).

[6] Very often the testing of the methods was done on administrative units, for example, because of data availability, and therefore the methods could most often also be adapted to other analysis units. See e.g.: a methodology for identifying urban areas combining subjective assessments with machine learning (Galdo et al., 2019), the fuzzy rurality indicator (Pagliacci, 2017), the Urbanity Index (Niklas et al., 2020), the OECD Regional Typology (OECD, 2008, p. 4), the OECD Extended Regional Typology (Brezzi et al., 2011), the Urban-rural Typology (Eurostat, 2010, p. 240) and the DEGURBA-method (Degree of Urbanisation) (EC et al., 2020a, p. 6, 9; EC et al., 2020b, p. 2, 5; Eurostat, n.d.a).

Table 1 (*continued*)

	Dem. Year-book 2018	WUP 2018	Dif-ference *	Dif-ferent def. **
Austria	67.2 (2011)	58.3 (2017)	-8.9	Y
Belgium	98.6 (2011)	98.0 (2017)	-0.6	Y
Bulgaria	73.5 (2018)	75.0 (2011)	1.5	?
Croatia	55.3 (2011)	56.9 (2011)	1.6	?
Czech Republic	73.1 (2018)	73.8 (2016)	0.7	?
Denmark	-	87.9 (2017)	-	-
Estonia	69.4 (2018)	68.9 (2016)	-0.5	N
Finland	71.3 (2017)	85.4 (2016)	14.1	Y
France	77.4 (2015)	80.4 (2007)	3.0	N
Germany	80.3 (2011)	77.3 (2015)	-3.0	?
Greece	76.6 (2011)	79.1 (2011)	2.5	Y
Hungary	70.5 (2017)	71.4 (2015)	0.9	N
Ireland	62 (2011)	63.2 (2016)	1.2	N
Italy	-	70.4 (2016)	-	-
Latvia	68.5 (2018)	68.1 (2015)	-0.4	N
Lithuania	67.1 (2017)	67.7 (2015)	0.6	?
Nether-lands	66.8 (2011)	91.5 (2017)	24.7	Y
Norway	79.3 (2011)	82.2 (2017)	2.9	Y
Poland	60.2 (2017)	60.1 (2015)	-0.1	N
Portugal	61.1 (2011)	65.2 (2011)	4.1	N
Romania	53.7 (2017)	54.0 (2014)	0.3	?
Slovakia	53.7 (2017)	53.7 (2014)	0.0	?
Slovenia	55.2 (2017)	54.5 (2015)	-0.7	?
Spain	-	80.3 (2011)	-	-
Sweden	-	87.4 (2016)	-	-
Switzer-land	84.8 (2018)	73.8 (2016)	-11.0	Y
U.K.	81.1 (2011)	83.4 (2011)	2.3	?

*The difference between WUP 2018 and Demographic Yearbook 2018 (percentage points). Negative value indicating the value from the Demographic Yearbook being larger than from WUP.
**The definitions provided by the Demographic Yearbook and WUP are different (Y), the same (N) or difference is unknown (?).

are nowadays often scrutinised. Satellites provide continuous images of such as landcover, morphology, or night-time illumination (Guo et al., 2019). This is beneficial in the cases where no analyses-ready datasets are available. This occurs frequently, especially in the case of developing countries or, for example, in the case of informal settlements: official data does not exist, or its accuracy is poor (see e.g. Giaccom Ribeiro, 2015). In these cases, remote sensing provides a way to create these datasets with reasonable built environment variables thereby enabling analyses of different indicators, development etc. (Weeks, 2010, p. 38). The remote sensing-based data is often accused of 'time-consuming interpretation steps and poor image resolution' (Long et al., 2016). However, the work processes, as well as the image resolution, are constantly being enhanced.

Telecommunication data has great potential to reveal cyclical, e.g. daily, weekly or yearly choreographies in inhabitants' mobility and locations. This can reveal unexpected patterns and assist in building more diverse delineations of urban areas that are based on both work and resident populations, their mobility and their volumes (see e.g. Becker et al., 2011; Calabrese et al., 2014; Grauwin et al., 2015; Lenormand et al., 2015; Makhrova and Babkin, 2020; Ratti et al., 2006; Reades et al., 2007). The main limitation in using the telecommunication data is the national regulations and laws and ensuring the privacy of users (see e.g. Calabrese et al., 2014). To utilise telecommunication data in delineating urban areas, the data must be on an adequate scale. Naturally, a sufficient number of mobile device users as well as service area coverage are also necessary.

Computing and mathematical modelling also provide promising methods. Enhanced computing capacity enables highly complex analysis, larger datasets and new approaches. For example, scaling laws are a well-known phenomenon common in complex systems not only in nature but also e.g. in cities, and their research forms an established tradition in urban sciences. The sizes of natural cities in the many countries appear to follow Zipf's rank-size rule, but more importantly, their attributes (salaries, crime, patents etc.) scale against their population (see e.g. Bettencourt et al., 2007; Chen and Jiang, 2018; Duranton, 2021; Pumain, 2004). Jiang and Liu (2012) discovered that scaling can be used to delineate natural cities while observing urban blocks and their sizes. With the OpenStreetMap (OSM) street data they produced a dataset with individual blocks. By applying the head/tail division rule with the heavy-tailed distribution to them, they distinguished the city and urban blocks from the data and clustered them to delineate cities following the spatial autocorrelation (Jiang and Liu, 2012). Jiang and Liu (2012) argue that the side effect of spatial autocorrelation is that the smaller urban edge blocks might erroneously be classified as rural, but they believe this produces only a very minor error in total, since these blocks usually have relatively small population.

De Bellefon et al. (2019) also rely on computing and statistical distribution. They created a dartboard methodology utilising the density of building volume in a 200 m grid. After calculating the densities in each grid cell, they used kernel to smooth the values, and they repeated the procedure for generated counterfactual densities for randomly distributed buildings in buildable grid cells. When comparing these values, if the real smoothed value is over the threshold, i.e. 'above the 95th percentile of the distribution of counterfactual smoothed densities computed for that pixel', the grid cell is categorised as urban and the adjacent grid cells form the urban areas (De Bellefon et al., 2019). Moreover, Tellier (2020) and Tellier and Gelb (2018) has developed the Urban Metric System (USM). USM is a space-economy based model suitable for different scales and based on vector fields and on attractive and repulsive forces. The method stresses the functionality of the city, concerning residents but also employees and economic actors, i.e. their attractive and repulsive forces and their magnitudes. The vector-resultants based on gravity models estimating the 'probabilities of interaction between pairs of points in the space' (Tellier and Gelb, 2018, p. 150) are able to reveal urban structures on different scales, from districts to global systems. Regarding the urban-rural dichotomy, UMS entails that a critical threshold value should be determined (Tellier, 2020).

All these three approaches using different mathematical methods (statistical distributions or probabilities) are promising for delineating urban areas for their capability to reflect the autonomously emerging characteristics of the city. What is important is that these methods consider regional variation and volumes, recognising the local systemic fluctuation. Hence, they could reduce the global non-comparability of varying urbanity and measures. Furthermore, the temporal variation is also considered as the comparison is always made to the variables of the time as they produce the reference values. The question is to what extent these mathematical methods and statistical values can capture the real urban environment overall. This would require further validation.

AI is an additional tool that utilises other methods. It expands the possibilities for delineating urban areas, as the use of AI enables processing even larger data masses while enhancing the processes themselves. It can be utilised to classify for example building densities, landcovers or urban fabric based on subset algorithms, decision trees or computational models (see e.g. Abarca-Alvarez et al., 2019; Arribas-Bel et al., 2019; Guo et al., 2019). It can also be combined, for example, with human estimation (Duranton, 2021; Galdo et al., 2019). In the method proposed by Galdo et al. (2019) the set of images was first categorised by human actors. Based on that, AI was trained to classify any set of images. Involving human estimation in the process could improve the observations in the urban environment that is often limited in qualitative methods. However, the use of AI initially entails a large amount of work and the processes can be complex. In addition, for example neural networks have been criticised as the process and the choices may remain hidden (Abarca-Alvarez et al., 2019). Nevertheless, swiftly developing AI will certainly be applied widely in the future in many fields in urban research.

3. Research structure, methods and data

In this study, we explored urban area definitions yielding urban population percentages from 24 EU countries and three non-EU Schengen countries. The small and certain island countries, for example Luxembourg, Malta and Iceland were omitted. While some countries use the same definition, 22 different urban area definitions were adopted for the study.

The empirical study consisted of two parts: a categorisation of the definitions of the 27 European countries and testing of these in the case of Finland. In the first part all 22 definitions were classified into categories, implying that the patterns observed were organised into relevant categories (Chenail, 2008).

In the second part these definitions and the resulting categories were tested in the case of Finland as accurately as possible considering the applicability of data and areal units. The urban population percentage, urban area ratio and population densities of urban areas were calculated in all cases. The 15 countries included in the testing were Belgium, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, Latvia, Norway, Poland, Spain, Sweden and Switzerland. The resulting 13 definitions consisted of ten official and three unofficial definitions (urban areas resulting from the official definitions before aggregating them into LAU areas). The countries for which the definitions could not be tested in the case of Finland were: Austria (no comparable data available); Bulgaria, Croatia, Lithuania, Portugal, Slovenia and the United Kingdom (no comparable areal unit available) and Greece, Ireland, the Netherlands, the Slovak Republic and Romania (methodology not described in adequate detail).

The testing was performed as spatial analyses using GIS tools. The urban areas were formed according to the national definitions selected and data available. The Finnish datasets used were:

- Population data, 2018 (Population Information system of Digital and Population Data Services Agency (Finland))
- Buildings, 2018 (Topographic Database of National Land Survey of Finland)
- Administrative borders, 2018 (National Land Survey of Finland)
- 500 m, 1 km grid files (YKR (Monitoring System of Spatial Structure and Urban Form) by the Statistics Finland and the Finnish Environment Institute)
- Water elements, 2015 (General map 1:1 000 000, National Land Survey of Finland)
- The local functional areas by services and grocery shopping, 2017 (Finnish Environment Institute)
- Corine Land Cover 2018 25 ha, 2018 (Finnish Environment Institute)

The Finnish population dataset used in this study had 5421122 records (sampling date 27 December 2018). From the original database approximately 2% of the records were eliminated by the Digital and Population Data Services Agency (Finland). These included population without permanent place of residence (~1.7%), orders of non-disclosures (~0.2%) and population temporarily in medical care institutions (~0.05%) (Liimatainen, 2019, personal communication). All analyses were conducted using the same population dataset to eliminate possible differences between the dataset used and the official municipality population information. In urban area calculations sea area and lakes (=< 20 km^2) were eliminated from the category A results.

4. Research operations

4.1. Country-specific definitions and their categorisation

Definitions from 27 European countries were divided into two main categories according to the basic areal unit, i.e. the end result unit of the definition. The main resulting categories were those with (A) LAU areas and (B) settlements/localities as end result units. Accordingly, in category A, despite the use of a grid or buildings in defining urban areas, the results are aggregated into the LAU areas indicating that the whole LAU area was considered urban. In category B, either no or only smaller unit aggregations were carried out (for example, settlements).

The two main categories were further subdivided into seven subcategories according to how the urban areas were defined and formed. This resulted in two main categories and seven subcategories:

A) LAU areas as end result units (1 un-subcategorised country)
   1. Official cities (2 countries)
   2. Municipalities with a certain number of inhabitants (possibly combined with other characteristics) (4)
   3. Grid-based population or address density (7)
   4. Localities (built-up area or similar) and population (3)

B) Settlements/localities as end result units
   1. Buildings and population (5)
   2. Buildings, minimum area and population aggregated to non-LAU-unit (1)
   3. Settlements (as administrative units and/or with other formation criteria) with certain population and/or socio-economic structure (4)

Details for these subcategories are presented in Table 2 and below. The most common formation criteria of the localities were the administrative units, grid or built-up areas (Table 2). The most commonly used defining criterion for urbanity was population.

4.1.1. LAU areas as end result units

For the officially designated cities and towns (category A1), those with official city status are considered urban. This category has no minimum population requirements or other indicators. Hungary and Poland apply such a classification.

In category A2 the country-specific population threshold value and possible other characteristics are applied to define urban areas. Belgium and the Czech Republic use only the population attribute. In the case of Latvia, the LAU area must additionally involve a local centre and in the Slovak Republic also 'signs of urban development'. The definition for Romania remained somewhat unclear and it was not possible to determine whether Romania belongs to subcategory A1 or A2.

Category A3 countries' urban area definitions are based on grids noting population or population and commuters. For these, grid cell size varies from 500 m x 500 m to 1 km x 1 km. After grid-based calculations, the results were aggregated into LAU level. Countries using a grid-based population are Austria (also including commuters), Germany, Italy, Spain, Switzerland and Estonia. The Netherlands uses address density instead of population attribute and hence, by adopting a grid-based measurement, the Netherlands appears to simulate a 1 km buffer around each address point.

Category A4 is based on localities and minimum population. In France the localities are based on maximum distance between buildings (see also category B1) and in Portugal on territorial units of smallest homogenous built areas. Other attributes are also used in both: in France population weighting and certain areal exceptions and in Portugal area weighting, administrative position and existing urban planning. The explicit criteria for defining localities in Greece were not available.

4.1.2. Settlements/localities as end result units

The urban areas in category B1 are based on the built-up areas formed using a combination of maximum distance between buildings and minimum number of population. These are used in Denmark, Finland, Sweden, Norway and Ireland. The latter two also use more detailed measures for areas that make an exception to the maximum distance rule.

The United Kingdom forms the B2 class. It uses output areas as an end result unit. These are formed on the basis of postcode units. The output areas located on built-up areas formed using a maximum distance between buildings, a minimum area size and a minimum population threshold, are defined as urban.

Category B3 uses settlements as end result units. Countries belonging to this category are Bulgaria, Croatia, Lithuania and Slovenia. In these countries the settlements are administrative units. Whether they are urban or not is determined by the number of population and/or workplaces. In some cases, type of workplaces and historic town status are also taken into account, depending on the country.

4.2. Testing the definitions: spatial analyses

In the spatial analyses phase, the categories formed were tested using the criteria categorised above applying data from Finland. The steps of the analyses are described below. From categories A3 and A4 the urban population percentage of Finland was also calculated before aggregation into LAU areas. These unofficial figures are expected to yield additional information on the variance in B subcategories as well as on the possible effect of the aggregation to the upper scale unit. In these cases, the urban areas were treated similarly to subcategories of the B category. In the results, these were treated as B+ subcategories in the following manner: A4 (France) forms B1+ and A3 (Germany, Italy, Spain & Switzerland and Estonia) B+ subcategories.

4.2.1. Category A: LAU areas as end result units

After all the steps described below, in each case urban population percentage, urban area ratio and density of urban areas of the whole of Finland were calculated.

A1. In the case of category A1, Finnish municipalities with official city status were identified and their populations calculated.

A2. In category A2 the total population of all municipalities was calculated. For the Czech Republic, all municipalities with populations over 2000 and in Belgium over 5000 were chosen. In the case of Latvia, to identify the local centres, the Finnish municipalities with local functional areas by services and grocery shopping were selected. From these the municipalities with a minimum of 2000 inhabitants were chosen.

A3. In category A3 Germany, Italy, Spain and Switzerland all use the EU DEGURBA classification. They use the 1 km grid and total population was calculated for each cell. Next, clusters of populated cells were formed (i.e. adjacent cells) and their population densities (km2) calculated. Clusters with over 300 inhabitants per km2 and a minimum of 500 inhabitants were defined as urban. The population of their municipalities was calculated. The whole municipality was considered urban if the number of inhabitants in these urban clusters exceeded 50%.

Estonia uses a 500 m grid. Population density (km2) was calculated for each cell and those smaller than 200 inh./km2 were deleted. After this, clusters were formed (adjacent cells, and the cells one step further) and the population figure of each cluster was calculated. For clusters of over 5000 inhabitants the cluster cells were designated as urban. The whole municipality was considered urban if more than 50% of the cells fulfilled this condition.

A4. From category A4 France was tested. First the built-up areas were formed using a 100-metre buffer (implying 200 m distance between buildings). Next, the classes '121 - Industrial or commercial units', '122 - Transport areas', '124 - Airport areas', '141 - Green urban areas' and '142 - Sport and leisure facilities' from Corine Landcover 2018 data were combined with these. After this the population in the buffers in municipalities with over 2000 inhabitants was calculated. The municipality was considered urban if 50% or more of the population was located in a single buffer (contiguous built-up area).

4.2.2. Category B: settlements/localities as end result units

After all steps described below, the urban population percentage, the urban area ratio and the density of urban areas of the whole of Finland were calculated on the basis of the remaining buffers.

B1. Here the urban areas are based on the distances between buildings. For the Finnish and Swedish definitions, the buffer around the buildings was 100 m. Population information was merged with these buffers. Those with less than 200 inhabitants were deleted.

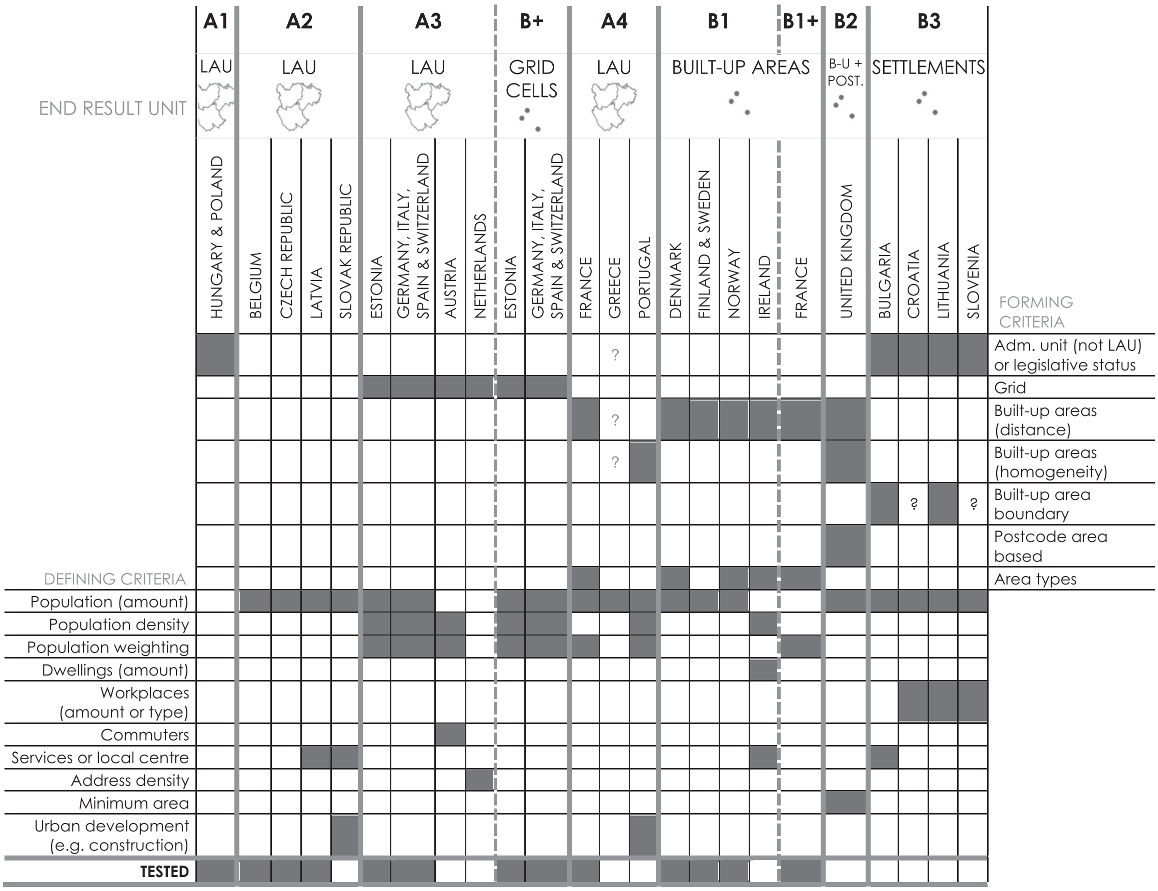
For Norway, using a 25-metre buffer and in the case of Denmark 100 m, the classes '121 - Industrial or commercial units', '141 - Green urban areas' and '142 - Sport and leisure facilities' from the Corine Landcover 2018 data were combined with these. After this the population information was merged with the built-up areas formed and those with below 200 inhabitants were deleted. In the case of Norway, other 25-metre buffers inside 400 m from the centre point of these urban areas formed were also included.

5. Discussing the resulting urban population percentage and urban areas of Finland

In the study reported in this article, we wanted to explore the degree of variation in the national urban area figures. These result from significant differences in national urban delineation methods, causing, for example, confusion in discourse on urban areas. The results yielded new knowledge about the scale of this variation. The differences between the two main categories of European urban delineations presented in Section 4 were even greater than expected. The results of this study are illustrated in Table 3 and the more detailed statistics in Appendix A. The figures also include the non-aggregated, unofficial figures (tagged as category B1+ and B+ in tables/graphs).

The calculations indicated that the urban population percentage of Finland varied between 63.6 and 98.9 depending on the criteria applied

Table 2
Countries by subcategorisation and criteria of how urban areas are formed and defined. The B+ and B1+ categories include tested A3 and A4 countries before the aggregation into LAU areas (i.e. the urban figures were also calculated based on the values in the grids or built-up areas before the aggregation to LAU area level was carried out).



(mean 79.8%, median 79.4%). The official population percentage of Finland in 2018 was 85.4% (UN, 2019c). We perceived considerable variance in urban population percentages when applying certain European urban area definitions in Finland. According to category A definitions, the urban population percentage was between 72.0 and 98.9 (mean 84.4%, median 81.6%). Category B resulted in variation percentages between 63.6 and 86.8 (mean 74.5%, median 71.8%). The greatest differences emerged in the cases of the Czech Republic + 13.5 percentage points and Belgium + 7.9pp (subcategory A2), Germany, Italy, Spain and Switzerland − 13.4pp and Estonia − 7.5pp (A3), Norway − 9.0pp (B1), unofficial France − 19.3pp (B1+) and unofficial Germany, Italy, Spain and Switzerland − 21.8pp and unofficial Estonia − 18.3pp (B+). When observing all subcategories, the most notable differences occurred in category A2 (mean 93.3%) and B+ (mean 65.4%). No such clear pattern between the main categories could otherwise be observed.

For A3, B1+ and B+ countries, one explanation for the considerable deviation may have been the population weighting criterion applied. For A2 countries, the difference could be seen for those using only population threshold value as a criterion. The difference could be explained by inappropriate population threshold values for Finland, since both the Czech Republic and Belgium have substantially smaller LAU area unit sizes than Finland.[7] Although the EU aims at a coherent system of administrative units,[8] their actual size differs considerably across countries, complicating analyses and their interpretations. Criteria applied to the smaller LAU areas may produce erroneous interpretations overall with category A. The different biogeographical characteristics[9] may also play a role.

In the case of urban area ratio, the results revealed notable differences in urban Finland with different criteria. Examples are illustrated in Fig. 1. The variation in the total urban area ratio indicates considerable variation in delineations (0.7–88.9%). The mean of urbanised land was 28.0% and the median 19.3%. Comparison between categories A and B indicated a substantial difference: while the mean of urban area of category A percentage of Finland was 50.9%, the mean of category B is only 1.4%. It is also noteworthy that in category A the variation is extremely high: 19.3–88.9%, as in category B, all figures settle between 0.7% and 2.5%, i.e. 1.8pp. The wide variation in urban area ratio probably results from the different end result units of the main categories, i.e. municipality vs. localities. Although category B had a

[7] Mean area of LAU (km2) in Finland 977, in Chech Republic 12 and Belgium 51 (Eurostat, 2018; The World Bank, 2018).

[8] To tackle the challenge resulting from the use of local-level statistics, Eurostat created two-level local administrative units (LAU1&2), which were changed to one-level units from the beginning of 2017. LAU areas (former LAU2) correspond to municipalities or equivalent units. (Eurostat, 2019.)

[9] Finland is one of the most sparsely populated, heavily forested countries in the EU.

decidedly low urban area ratio, below 2.5%, the urban population percentage remained high: 63.6–86.8%.

Regarding the population density of urban areas produced by the different definitions, the results indicate that in the case of Finland the density varied between 19.2 and 1853.5 inhabitants per km2 (mean 561.3 per km2, median 64.4 per km2). Similarly to the urban area ratio, the density of urban areas followed a similar pattern perceived when scrutinising the categories: category A and B differ from each other considerably. The mean density of category A is 35.6 per km2 and of category B 1174.7 per km2.

Thus, the figures resulting from the definitions in category A do not directly correlate with the actual urban populations of the respective municipalities, nor with the urban area or density, because category A classifies the population and the area of the municipality as a whole homogenously as either urban or rural regardless of local variations. These classifications may increase the total degree of agglomeration of the country. Nevertheless, the LAU area borders are always somewhat arbitrary, not necessarily correlating with spatio-functional urban configurations (Cohen, 2004, p. 25) - and conversely, the urban dynamic, functional regions and agglomerations rarely follow municipal borders. Category B appears more appropriate for ascertaining the actual highly populated areas. However, definitions based on the local unit (e.g. municipality) are necessary as well, for example for evaluating national and wider economic indicators.

The limitations of this study are mainly related to methodology. First, translations from original languages imply a possibility for translation errors. Interpretations of criteria and areal units were necessary to conduct the analyses causing potential inaccuracy. The data was chosen carefully and in some cases we refrained from testing certain countries due to the excessive adjustments required, potentially impairing the reliability. For some of the variables the correspondence between the applied and original data was moderate. Since the analyses were conducted only in the case of Finland, with its regionally dispersed and small population, the generalisability is limited. However, these results are indicative, allowing certain universal interpretations.

6. Guidelines for selecting an approach to urban area delineations

Based on the literature and the empirical study presented in this article we here discuss the essential challenges along with key characteristics for an applicable urban area delineation approach[10]:

Avoid administrative units as a basis. As our case study indicated, the administrative unit-based definitions also generally include non-urban populations and places. In addition, these units differ from country to country. Therefore, they are incapable of generating a universally commensurate delineation method for urban areas. In other words, instead, the approach should be based on grids, buffers or other small units, and it ought to be scalable. (see e.g. Arribas-Bel et al., 2019; De Bellefon et al., 2019; Long et al., 2016; Tellier, 2020; Tellier and Gelb, 2018; Weeks, 2010, p. 39).

Avoid using residential population only as a basis. The approach could be built on only one, well-justified variable (other than merely population). However, a multi-criteria method would be preferable. This GIS-based method should consider features of built environment and urban lifestyle such as density or functionality, e.g. work, points of interest etc. (see e.g. Abarca-Alvarez et al., 2019; Ban et al., 2021; De Bellefon et al., 2019; Jiang and Liu, 2012; Long et al., 2016; Weeks, 2010, pp. 38–39) In addition natural elements can be used as well (see e.g. Benza et al., 2016).

Avoid using datasets based on methodology using larger aggregations, e.g. in administrative units *or artificial distributions*. Their problems may increase proportionately in the urban area delineation method. For example, LAU area-based or regional statistics cannot be properly disaggregated to a smaller scale for a detailed delineation of urban areas. Also, avoid datasets resulting from complex, and especially unknown, processes, as their problems are hard to trace and they, too may increase proportionately in the delineation method. Instead, favour datasets which are as detailed as possible, as little processed as possible and which are based on authentic location-based information.

Enable both gradients and dichotomies. The approach should be able to reflect the urban area as zones, or gradients, acknowledging the continuous and entangled nature of urban and rural space, still enabling the representation of dichotomic classification when needed. (see e.g. Tellier, 2020; Tellier and Gelb, 2018; Weeks, 2010, p. 39) This requires that with gradient-based methods, it is clearly stated how they can be used as dichotomous, i.e. which classes belong to which one. Mathematical models, such as scaling/power laws, could be useful in building the gradients, applying efficient computing power.

The most challenging task in developing a new method is *setting the proper threshold values* for different zones or classes (see e.g. De Bellefon et al., 2019; Duranton, 2021). With a universal method, it must be considered whether a single threshold value can pertain for urbanity, as it can be seen as a relative phenomenon in both global and history wise (Duranton, 2021). For this reason, applying mathematical and statistical methods for these purposes should be carefully considered as they take into account the global fluctuations and local urban characteristics.

The ultimate guideline is: *Keep it as simple as possible.* Whatever approach is chosen, it should be simple enough, from both the analysis and data point of view, to be utilised universally (see e.g. Duranton, 2021). Overly convoluted structure or nested equations may appear as a 'black box', which is allegedly also the problem of AI. These complicate both the overall use and the adaptation of the method. Consequently, the simpler the method and the dataset used, the more widely applicable it generally becomes.

7. Conclusion

The ongoing planetary urbanisation calls for thorough scrutiny of figures reflecting urbanisation across countries in a comparative manner, both in the media and in professional discourse. These figures provide only a lens for the individual, temporal development of each country. Surprisingly, no study has so far been presented highlighting their non-comparability by testing the factual impact of these definitions on the resulting urban areas and their populations. We took the first step towards bridging this gap by presenting a systematic study of the majority of indicators across European countries reflected against the data from one country, Finland.

This pioneering study corroborates the UN guidelines recommending caution in comparing the national urban population percentages as the urban area definitions have been adjusted to reflect local requirements and features. It opens a necessary discussion about the transparency of urbanisation measures, stressing the necessity to explicate the national differences. We would like to draw attention to the way these figures and concepts are, and should be, used in scientific and popular communication and media. Particularly, we stress that while comprehensive definitions and indicators for urbanity are hardly feasible, one should refrain from creating a straightforward juxtaposition of national degrees of urbanisation uninformed of the measuring methods they are based on.

Building consistent definitions while retaining the national ones poses another challenge. It is evident that these are used for different purposes, but it would be necessary to explicate more clearly their premises, purpose of use and underlying assumptions, along with their relationship to each other. This is also necessary for establishing relevant and properly targeted policies from national to the EU-wide and even to the global scale, as it is noteworthy that the urban area forms a

[10] We focus here on the urban areas per se, and exclude, for example, functional urban areas and metropolitan areas, which also include non-urban areas or population when including commuting zones (see e.g. Moreno-Monray et al., 2020).

Table 3
Results of testing different European urban area definitions in the case of Finland.

	A1	A2			A3		A4	B1			B1+	B+	
	HUNGARY & POLAND	BELGIUM	CZECH REPUBLIC	LATVIA	ESTONIA	GERMANY, ITALY, SPAIN & SWITZERLAND	FRANCE	DENMARK	FINLAND & SWEDEN	NORWAY	FRANCE	ESTONIA	GERMANY, ITALY, SPAIN & SWITZERLAND
URBAN POPULATION %	79.4	93.3	98.9	87.9	77.9	72.0	81.6	86.8	86.7	76.4	66.1	67.1	63.6
Urban pop. % by cat. (AM)		93.3			74.9			79 (inc. B1+)				65.4	
Urban pop. % A / B (AM)	84.4							74.5					
URBAN AREA %	38.4	64.0	88.9	74.0	24.5	19.3	46.9	2.5	2.5	0.7	1.1	0.8	0.8
Urban area % by cat. (AM)		75.6			21.9			1.7 (inc. B1+)				0.8	
Urban area % A / B (AM)	50.9							1.4					
URBAN AREA DENSITY	36	25	19	20	55	64	30	595	610	1854	1017	1546	1428
Urban area density by cat. (AM)		22			60			1018 (inc. B1+)				1487	
Urban area density A / B (AM)	36							1175					

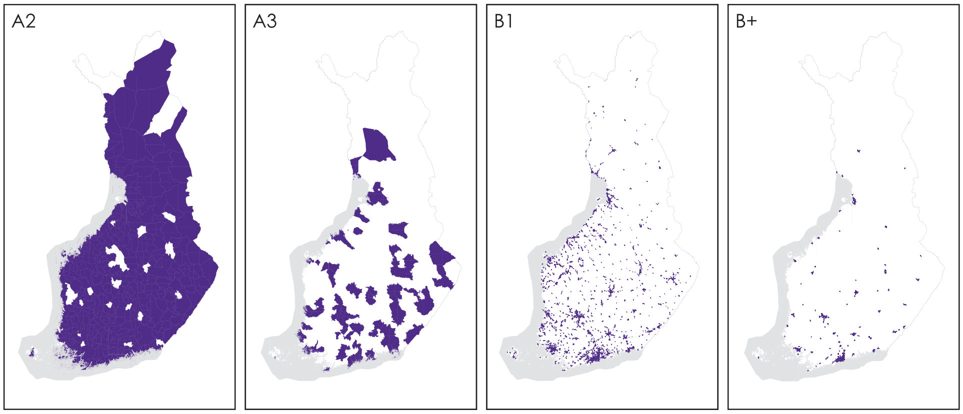


Fig. 1. Urban areas of Finland applying the definitions of the Czech Republic (category A2, share of urban area 88.9% / urban population 98.9% / urban population density 19/km2), Germany, Italy, Spain and Switzerland (A3, 19.3% / 72.0% / 64/km2), Finland and Sweden (B1, 2.5% / 86.7% / 610/km2) and Germany, Italy, Spain and Switzerland before aggregation to LAU areas (B+, 0.8% / 63.6% / 1428/km2).

politically and economically essential unit, and state subsidies and other interventions are often dependent on these measures. Appropriate policy proposals would call for an exhaustive and critical review on a larger scale. Such a systematic review could outline good practices and help to adjust the existing definitions. This would produce a more generalisable understanding of these definitions and their implications.

Undoubtedly this type of straightforward delineation necessarily leads to a dichotomy (urban/rural) that poorly reflects the late modern urban landscape. Similarly, emphasising population as the sole measure for urbanity is questionable, since it ignores many essential spatio-functional, corporeal characteristics of urbanity. For example, cases such as office parks, CBDs, or dense rural populations easily produce flawed results. Furthermore, residence-based definitions have a limited capacity to reflect current lifestyle and 'new work' implying a high level of mobility, multi-place or non-place residence or work, urban nomadism, and, generally speaking, people's rapid movements across time and space. To overcome this challenge, we explored a variety of possible data-driven approaches that had been applied to delineate urban areas. We avoided proposing a single solution since all the options scrutinised would require further validation. Hence our proposal would be more of a set of guidelines for future approaches.

We argue that a new urban area delineation method would call for more relational and dynamic approaches. These methods could use artificial intelligence cautiously yet more efficiently, applying novel datasets, such as telecommunication or GPS data, or utilise mathematical methods, or even revisit old methods with new viewpoints. Yet the most important thing is to keep the delineation method as simple as possible. In addition, the concept of *border* in the context of urban areas would require scrutiny, indicating that future research could delve into the question of the capability of current delineations to reflect reality when comparing different indicators.

Overall, the inherent complexity of urban features makes it difficult to discuss cities, leading to several parallel discourses. Here the 'two worlds' often collide (Snow, 2001 [1959]). From the city life or social scientific point of view, the city consists of human, social and cultural processes taking place in urban space. On the other hand, a city is a quantifiable entity, constructed from decidedly concrete elements: buildings, roads and infrastructure, occupied and produced by humans (see e.g. Bergel, 1955, pp. 5–6). Echoing Bergel (1955, p. 5) saying 'Everybody seems to know what a city is but no one has given a satisfactory definition' (Bergel, 1955, p. 5), exhaustive measurement for urbanity is hardly possible. However, urban delineations are still necessary for statistical, political and economic purposes: With the contribution presented in this article we would encourage urban scholars to engage in a comprehensive systematic comparative exploration of these measurements, resulting spatial delineations, and their local consequences. Such iterative, accumulative work would gradually build robust knowledge of the ways to define urban areas in varying context.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by the Academy of Finland, Strategic Research Council [grant number 303618].

Declaration of conflicting interests

None.

Acknowledgements

The authors would like to sincerely thank the anonymous reviewers for their thorough work and valuable comments.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.landusepol.2021.105822.

References

Abarca-Alvarez, F., Campos-Sánchez, F., Osuna-Pérez, F., 2019. Urban shape and built density metrics through the analysis of European urban fabrics using artificial intelligence. Sustainability 11. https://doi.org/10.3390/su11236622.

Andersen, H.T., Møller-Jensen, L., Engelstoft, S., 2011. The end of urbanization? Towards a new urban concept or rethinking urbanization. Eur. Plan. Stud. 19 (4), 595–611. https://doi.org/10.1080/09654313.2011.548472.

Arriaga, E.J., 1970. A new approach to the measurements of urbanization. Econ. Dev. Cult. Chang. 18 (2), 206–218. https://doi.org/10.1086/450419.

Arribas-Bel, D., Garcia-López, M.-A., Viladecans-Marsal, E., 2019. Building(s and) cities: delineating urban areas with a machine learning algorithm. J. Urban Econ. 19, 9. https://doi.org/10.1016/j.jue.2019.103217.

Ban, P., Zhan, W., Yuan, Q., Li, X., 2021. Delineating the urban areas of a cross-boundary city with open-access data: Guangzhou–Foshan, South China. Sustainability 13, 2930. https://doi.org/10.3390/su13052930.

Becker, R., Cáceres, R., Hanson, K., Loh, J.M., Urbanek, S., Varshavsky, A., Volinsky, C., 2011. The tale of one city: using cellular network data for urban planning. IEEE Pervasive Comput. 10 (4), 18–26. https://doi.org/10.1109/MPRV.2011.44.

Benza, M., Weeks, J.R., Stow, D.A., López-Carr, D., Clarke, K.C., 2016. A pattern-based definition of urban context using remote sensing and GIS. Remote Sens. Environ. 183, 250–264. https://doi.org/10.1016/j.rse.2016.06.011.

Bergel, E.E., 1955. Urban Sociology. McGraw-Hill, New York.

Bettencourt, L.M.A., Lobo, J., Helbing, D., Kühnert, C., West, G.B., 2007. Growth, innovation, scaling, and the pace of life in cities. Proc. Natl. Acad. Sci. USA 104 (17), 7301–7306. https://doi.org/10.1073/pnas.0610172104.

Bogue, D.J., 1961. Population ecology: studies of population distribution and density, measures of urbanization and problems of urbanization. Ekistics 11 (65), 197–202.

Brenner, N., Schmid, C., 2014. The 'urban age' in question. Int. J. Urban Reg. Res. 38 (3), 731–755. https://doi.org/10.1111/1468-2427.12115.

Brezzi, M., Dijkstra, L., Ruiz, V., 2011. OECD Extended Regional Typology: The Economic Performance of Remote Rural Regions. OECD Regional Development Working Papers 2011/06, OECD Publishing. ⟨https://www.oecd-ilibrary.org/governance/oecd-extended-regional-typology_5kg6z83tw7f4-en⟩. (Accessed 29 June 2020).

Calabrese, F., Ferrari, L., Blondel, V., 2014. Urban sensing using mobile phone network data: a survey of research. ACM Comput. Surv. 47 (2) https://doi.org/10.1145/2655691. Article 25.

Champion, T., Hugo, G., 2004. Introduction. In: Champion, T., Hugo, G. (Eds.), New Forms of Urbanization. Beyond the Urban-rural Dichotomy. Routledge, New York, pp. 3–24.

Chen, Y., Jiang, B., 2018. Hierarchical scaling in systems of natural cities. Entropy 20 (6), 432–449. https://doi.org/10.3390/e20060432.

Chenail, R., 2008. Categorization. In: Given, L.M. (Ed.), The Sage Encyclopedia of Qualitative Research Methods. ⟨http://methods.sagepub.com/Reference//sage-encyc-qualitative-research-methods/n41.xml⟩. (Accessed 17 April 2020).

Cohen, B., 2004. Urban growth in developing countries: a review of current trends and a caution regarding existing forecasts. World Dev. 32 (1), 23–51. https://doi.org/10.1016/j.worlddev.2003.04.008.

De Bellefon, M.P., Combes, P.P., Duranton, G., Gobillon, L., 2019. Delineating urban areas using building density. J. Urban Econ. https://doi.org/10.1016/j.jue.2019.103226.

Dijkstra, L., Poelman, H., 2008. Remote rural regions. How proximity to a city influences the performance of rural regions. Regional Focus 01/2008. European Union. ⟨https://ec.europa.eu/regional_policy/sources/docgener/focus/2008_01_rural.pdf⟩. (Accessed 3 July 2020).

Duranton, G., 2021. Classifying locations and delineating space: an introduction. J. Urban Econ. https://doi.org/10.1016/j.jue.2021.103353.

Dymitrow, M., Stenseke, M., 2016. Rural-urban blurring and the subjectivity within. Rural Landsc. Soc. Environ. Hist. 3 (1), 1–13. https://doi.org/10.16993/rl.1.

EC (European Commission), FAO, UN-Habitat, OECD and World Bank, 2020a. Applying the Degree of Urbanisation: A methodological manual to define cities, towns and rural areas for international comparisons. November 2020 edition. ⟨https://ec.europa.eu/eurostat/documents/10186/11395216/DEGURBA-manual.pdf/3a6bab6a-3fb1–4261-ad5b-e604cb67dc0d⟩. (Accessed 21 January 2021).

EC (European Commission) (inc. Eurostat and DG Regio), ILO, FAO, OECD, UN-Habitat and World Bank, 2020b. A recommendation on the method to delineate cities, urban and rural areas for international statistical comparisons. UN Statistical Commission, 51st session, 3–6 March 2020, Background document. ⟨https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Recommendation-E.pdf⟩. (Accessed 3 July 2020).

Eurostat, 2010. Eurostat regional yearbook 2010. Eurostat Statistical books, European Commission. ⟨https://ec.europa.eu/eurostat/documents/3217494/5727301/KS-HA-10–001-EN.PDF/1ba3cf6a-5e25–44c1–99f9-fada17625212⟩. (Accessed 1 July 2020).

Eurostat, 2018. Local administrative units. Correspondence table 2018. ⟨https://ec.europa.eu/eurostat/web/nuts/local-administrative-units⟩. (Accessed 26 March 2019).

Eurostat, 2019. Territorial typologies manual - urban-rural typology. Statistics Explained. ⟨https://ec.europa.eu/eurostat/statistics-explained/pdfscache/72656.pdf⟩. (Accessed 2 July 2020).

Eurostat, n.d.a. Degree of urbanisation – methodology. Eurostat webpage. ⟨https://ec.europa.eu/eurostat/web/degree-of-urbanisation/methodology⟩. (Accessed 2 July 2020).

Eurostat, n.d.b. Local administrative units (LAU). Eurostat -webpage on NUTS. ⟨https://ec.europa.eu/eurostat/web/nuts/local-administrative-units⟩. (Accessed 2 July 2020).

Fertner, C., 2012. Downscaling European urban-rural typologies. Dan. J. Geogr. 112 (1), 77–83. https://doi.org/10.1080/00167223.2012.707805.

Galdo, V., Li, Y., Rama, M., 2019. Identifying urban areas by combining human judgement and machine learning: an application to India. J. Urban Econ. https://doi.org/10.1016/j.jue.2019.103229.

Giaccom Ribeiro, B.M., 2015. Mapping informal settlements using WorldView-2 imagery and C4.5 decision tree classifier. In: Joint Urban Remote Sensing Event (JURSE), 2015–03, pp. 1–4. DOI:10.1109/JURSE.2015.7120470.

Gibbs, J.P., 1966. Measures of urbanization. Soc. Forces 45 (2), 170–177. https://doi.org/10.2307/2574387.

Grauwin, S., Sobolevsky, S., Moritz, S., Gódor, I., Ratti, C., 2015. Towards a comparative science of cities: using mobile traffic records in New York, London, and Hong Kong. In: Helbich, M., Jokar Arsanjani, J., Leitner, M. (Eds.), Computational Approaches for Urban Environments. Geotechnologies and the Environment, vol. 13. Springer, Cham.

Guo, J., Ren, H., Zheng, Y., Nie, J., Chen, S., Sun, Y., Qin, Q., 2019. Identify urban area from remote sensing image using deep learning method. In: IEEE International Geoscience and Remote Sensing Symposium 2019, 2019–7/8. DOI:10.1109/IGARSS.2019.8898874.

Haase, D., Tötzer, T., 2012. Guest editorial: urban-rural linkages – analysing, modelling, and understanding drivers, pressures, and impacts of land use changes along the rural-to-urban gradient. Environ. Plan. B Plan. Des. 39, 194–197. https://doi.org/10.1068/b3902ge.

Hugo, G.J., 1982. Circular migration in Indonesia. Popul. Dev. Rev. 8 (1), 59–83. https://doi.org/10.2307/1972690.

Hugo, G.J., Champion, A., Lattes, A., 2003. Toward a new conceptualization of settlements for geography. Popul. Dev. Rev. 29 (2), 277–297. https://doi.org/10.1111/j.1728-4457.2003.00277.x.

Jiang, B., Liu, X., 2012. Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information. Int. J. Geogr. Inf. Sci. 26 (2), 215–229. https://doi.org/10.1080/13658816.2011.575074.

Laurin, F., Pronovost, S., Carrier, M., 2020. The end of the urban-rural dichotomy? Towards a new regional typology for SME performance. J. Rural Stud. 80, 53–75. https://doi.org/10.1016/j.jrurstud.2020.07.009.

Lenormand, M., Picornell, M., Cantú-Ros, O., Louail, T., Herranz, R., Barthelemy, M., Frías-Martínez, E., San Miguel, M., Ramasco, J., 2015. Comparing and modelling land use organization in cities. article 150449 R. Soc. Open Sci. 2 (12). https://doi.org/10.1098/rsos.150449.

Long, Y., Shen, Y., Jin, X., 2016. Mapping block-level urban areas for all Chinese cities. Ann. Am. Assoc. Geogr. 106 (1), 96–113. https://doi.org/10.1080/00045608.2015.1095062.

López-Goyburu, P., Gonzaga García-Montero, L., 2018. The urban-rural interface as an area with characteristics of its own in urban planning: a review. Sustain. Cities Soc. 43, 147–165. https://doi.org/10.1016/j.scs.2018.07.010.

Makhrova, A.G., Babkin, R.A., 2020. Methodological approaches to the delimitation of the boundaries of the Moscow agglomeration based on data from mobile network operators. Reg. Res. Russ. 10 (3), 373–380. https://doi.org/10.1134/S2079970520030090.

Moreno-Monray, A., Schiavina, M., Veneri, P., 2020. Metropolitan areas in the world. Delineation and population trends. J. Urban Econ. https://doi.org/10.1016/j.jue.2020.103242.

Niklas, U., von Behren, S., Soylu, T., Kopp, J., Chlond, B., Vortisch, P., 2020. Spatial factor – using a random forest classification model to measure an internationally comparable Urbanity Index. Urban Sci. 4 (36) https://doi.org/10.3390/urbansci4030036.

OECD, 2008. OECD regions at a glance 2009. Final draft. Public Governance and Territorial Development Directorate, Territorial Development Policy Committee. GOV/TDPC/TI(2008)8. ⟨http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=GOV/TDPC/TI(2008)8&docLanguage=En⟩. (Accessed 1 July 2020).

OECD, 2020. OECD Territorial grids. OECD Centre for Entrepreneurship, SMEs, Regions and Cities. 3/2020. ⟨http://www.oecd.org/regional/regional-policy/42392313.pdf⟩. (Accessed 1 July 2020).

Pagliacci, F., 2017. Measuring EU urban-rural continuum through fuzzy logic. Tijdschr. Econ. Soc. Geogr. 108 (2), 157–174. https://doi.org/10.1111/tesg.12201.

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

Pumain, D., 2004. Scaling laws and urban systems. Santa Fe Institute Working Paper 2004-02-002. ⟨https://www.santafe.edu/research/results/working-papers/scaling-laws-and-urban-systems⟩. (Accessed 5 July 2021).

Ratcliffe, M., Burd, C., Holder, K., Fields, A., 2016. Defining rural at the U.S. Census Bureau. American Community Survey and Geography Brief. ⟨https://www.researchgate.net/publication/311533270_Defining_Rural_at_the_US_Census_Bureau⟩. (Accessed 8 January 2021).

Ratti, C., Pulselli, R., Williams, S., Frenchman, D., 2006. Mobile Landscapes: using location data from cell phones for urban analysis. Environ. Plan. B Plan. Des. 33, 727–748. https://doi.org/10.1068/b32047.

Reades, J., Calabrese, F., Sevtsuk, A., Ratti, C., 2007. Cellular census: explorations in urban data collection. IEEE Pervasive Comput. 6 (3), 30–38. https://doi.org/10.1109/MPRV.2007.53.

Satterthwaite, D., 2010. Urban myths and the mis-use of data that underpin them. In: Beall, J., Guha-Khasnobis, B., Kanbur, R. (Eds.), Urbanization and Development: Multidisciplinary Perspectives. Oxford Scholarship Online, pp. 83–100.

Schaeffer, P.V., Kahsai, M.S., Jackson, R.W., 2013. Beyond the rural–urban dichotomy: essay in honor of Professor A. M. Isserman. Int. Reg. Sci. Rev. 36 (1), 81–96. https://doi.org/10.1177/0160017612449981.

Shane, G., 2006. The emergence of landscape urbanism. In: Wandheim, C. (Ed.), The Landscape Urbanism Reader. Princeton Architectural Press, New York, pp. 55–67.

Sieverts, T., 2003. Cities Without Cities: An Interpretation of the Zwischenstadt. Routledge, New York.

Snow, C.P., 2001 [1959]. The Two Cultures. Cambridge University Press, London.

The World Bank, 2018. DataBank – World Development Indicators: Land area (sq. km). ⟨https://databank.worldbank.org/data/home⟩. (Accessed 26 March 2019).

Tellier, L.-N., Gelb, J., 2018. An urban metric system based in space-economy: foundations and implementation. Reg. Sci. Policy Pract. 10, 145–160. https://doi.org/10.1111/rsp3.12141.

Tellier, L.-N., 2020. Characterizing urban form by means of Urban Metric System. Land Use Policy. https://doi.org/10.1016/j.landusepol.2020.104672.

UN (United Nations), 1946a. Resolutions adopted by the Economic and Social Council during its 1st session held in London from 23 January to 18 February 1946. ⟨https://digitallibrary.un.org/record/3843513?ln=en⟩. (Accessed 26 June 2020).

UN (United Nations), 1946b. Resolutions adopted by the Economic and Social Council during its 2nd session held in New York from 25 May to 21 June 1946. ⟨https://digitallibrary.un.org/record/3845514?ln=en⟩. (Accessed 26 June 2020).

UN (United Nations), 1949a. Population census methods. Population studies, No. 4. Department of Social Affairs (Population Division) and Department of Economic Affairs (Statistical Office of the United Nations). ⟨https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Study_No.4-E.pdf⟩. (Accessed 18 February 2020).

UN (United Nations), 1949b. Report of the Population Commission (fourth session). Supplement no. 7, E/1313, 21 Apr 1949. Economic and Social council. ⟨https://digitallibrary.un.org/record/824932?ln=en⟩. (Accessed 11 February 2020).
UN (United Nations), 1949 [...] 2018. United Nations Demographic yearbooks 1948–2017. From first to sixty-eight issue. Statistical office of the United Nations and/or Department of Economic and Social Affairs. ⟨https://www.un-ilibrary.org/population-and-demography/demographic-yearbook-ser-r_406df598-en-fr⟩. (Accessed 11 February 2020).
UN (United Nations), 1958. Principles and recommendations for national population censuses. Statistical papers, series M, no. 27. Statistical Office of the United Nations. ⟨https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M27-E.pdf⟩. (Accessed 1 July 2020).
UN (United Nations), 1971. United Nations Demographic yearbook 1970. 22nd issue. Department of Economic and Social Affairs, Statistical Office. ⟨https://www.un-ilibrary.org/content/books/9789210581998/read⟩. (Accessed 6 July 2021).
UN (United Nations), 1973. United Nations Demographic yearbook 1972. 24th issue. Department of Economic and Social Affairs, Statistical Office. ⟨https://read.un-ilibrary.org/population-and-demography/united-nations-demographic-yearbook-1972_b85a1f97-en-fr#page7⟩. (Accessed 24 February 2020).
UN (United Nations), 1980a. Patterns of urban and rural population growth. Population Studies, no. 68, ST/ESA/SER.A/68. Department of International Economic and Social Affairs. ⟨https://population.un.org/wup/Archive/Files/studies/United%20Nations%20(1980)%20-%20Patterns%20of%20Urban%20and%20Rural%20Population%20Growth.pdf⟩. (Accessed 29 June 2020).
UN (United Nations), 1980b. Social and demographic statistics: classifications of size and type of locality and urban/rural areas. National practices in classifications of size and type of locality and urban/rural areas. E/CN.3/551, Report of the Secretary-General. Economic and Social Council, Statistical Commission. ⟨https://unstats.un.org/unsd/statcom/21st-session/documents/doc81/1981–551-LocalityClassification-E.pdf⟩. (Accessed 26 June 2020).
UN (United Nations), 2017a. New Urban Agenda. UN-Habitat. A/RES/71/256. ⟨http://uploads.habitat3.org/hb3/NUA-English.pdf⟩. (Accessed 3 July 2020).
UN (United Nations), 2017b. Principles and Recommendations for Population and Housing Censuses, revision 3. ST/ESA/STAT/SER.M/67/Rev.3. Department of Economic and Social Affairs, Population Division. ⟨https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Principles_and_Recommendations/Population-and-Housing-Censuses/Series_M67rev3-E.pdf⟩. (Accessed 30 June 2019).
UN (United Nations), 2019a. Table 6: Total and urban population by sex: 2009 - 2018. United Nations Demographic Yearbook 2018. ⟨https://unstats.un.org/unsd/demographic-social/products/dyb/documents/DYB2018/table06.pdf⟩. (Accessed 15 December 2020).
UN (United Nations), 2019b. United Nations Demographic Yearbook 2018. ST/ESA/STAT/SER.R/48, 69th issue. Department of Economic and Social Affairs. New York: United Nations. ⟨https://unstats.un.org/unsd/demographic-social/products/dyb/dybsets/2018.pdf⟩. (Accessed 4 December 2020).
UN (United Nations), 2019c. World Urbanization Prospects 2018: Data download center: Urban and rural populations, File1: Population of Urban and Rural Areas at Mid-Year (thousands) and Percentage Urban, 2018. ⟨https://population.un.org/wup/Download/⟩. (Accessed 23 January 2019).
UN (United Nations), 2019d. World Urbanization Prospects 2018: Data download center: Sources and documentation, File: Data sources and statistical concepts for estimating the urban population (WUP2018-DataSource-UrbanPopulation.xls). ⟨https://population.un.org/wup/Download/⟩. (Accessed 23 January 2019).
UN (United Nations), 2019e. World Urban Prospects: the 2018 revision (ST/ESA/SER.A/420). Department of Economic and Social Affairs, Population Division. New York: United Nations. ⟨https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf⟩. (Accessed 8 January 2020).
Urso, G., 2020. Metropolisation and the challenge of rural-urban dichotomies. Urban Geogr. https://doi.org/10.1080/02723638.2020.1760536.
Van Vliet, J., Birch-Thomsen, T., Gallardo, M., Hemerijckx, L.M., Hersberger, A.M., Li, M., Tumwesigye, S., Twongyirwe, R., Van Rompaey, A., 2020. Bridging the rural-urban dichotomy in land use science. J. Land Use Sci. 15 (5), 585–591. https://doi.org/10.1080/1747423X.2020.1829120.
Van Vliet, J., Verburg, P.H., Grădinaru, S.R., Hersperger, A.M., 2019. Beyond the urban-rural dichotomy: towards a more nuanced analysis of changes in built-up land. Comput. Environ. Urban Syst. 74, 41–49. https://doi.org/10.1016/j.compenvurbsys.2018.12.002.
Wandl, D.I.A., Nadin, V., Zonneveld, W., Rooij, R., 2014. Beyond urban–rural classifications: characterising and mapping territories-in-between across Europe. Landsc. Urban Plan. 130, 50–63. https://doi.org/10.1016/j.landurbplan.2014.06.010.
Weeks, J.R., 2010. Defining urban areas. In: Rashed, T., Jürgens, C. (Eds.), Remote Sensing of Urban and Suburban Areas. Remote Sensing and Digital Image Processing, vol. 10. Springer, Dordrecht, pp. 33–45.
Wineman, A., Yélognissé Alia, D., Andersson, C.L., 2020. Definitions of “rural” and “urban” and understandings of economic transformation: evidence from Tanzania. J. Rural Stud. 79, 254–268. https://doi.org/10.1016/j.jrurstud.2020.08.014.
Wirth, L., 1938. Urbanism as a way of life. Am. J. Sociol. 44 (1), 1–24. https://doi.org/10.1086/217913.
Woods, M., 2009. Rural geography: blurring boundaries and making connections. Prog. Hum. Geogr. 33 (6), 849–858. https://doi.org/10.1086/217913.

Glossary

Country-specific definitions:

Austria: Statistics Austria
Belgium: World Urbanization Prospects (2018 Revision) of the United Nations
Bulgaria: National Statistical Institute of Bulgaria
Croatia: Croatian Bureau of Statistics
Czech Republic: Czech Statistical Office
Denmark: Statistics Denmark
Estonia: Statistics Estonia
Finland: Statistics Finland; Helminen V (Finnish Environment Institute)
France: The French National Institute of Statistics and Economic Studies (INSEE)
Germany: Statistisches Bundesamt
Greece: Hellenic Statistical Authority
Hungary: Hungarian Central Statistical Office
Ireland: Central Statistics Office of Ireland
Italy: Instituto Nazionale di Statistica (Istat)
Latvia: The Saeima Law on Administrative Territories and Populated Areas (2010)
Lithuania: Statistics Lithuania
Netherlands: Statistics Netherlands Infoservice
Norway: Statistics Norway
Poland: Central Statistical Office of Poland
Portugal: Statistics Portugal
Romania: World Urbanization Prospects (2018 Revision) of the UN and UN Demographic Yearbook 2018
Slovak Republic: Statistical Office of the Slovak Republic
Slovenia: Statistical Office of the Republic of Slovenia
Spain: National Statistical Institute of Spain (INE)
Sweden: Statistics Sweden
Switzerland: Swiss Federal Statistical Office
The United Kingdom: Office for National Statistics (ONS)
DEGURBA (EU): Eurostat. 2018. Degree of urbanisation classification – 2011 revision. Statistics Explained. [http://ec.europa.eu/eurostat/statisticsexplained/], accessed 15 March 2019.
UN/WUP: United Nations. 2018. World Urbanization Prospects 2018. Data sources and statistical concepts for estimating the urban population. Internet page, available: https://population.un.org/wup/Download/. Accessed 19 September 2018.