Classifying urban form at a national scale
The case of Great Britain

Martin Fleischmann[*†§]
Daniel Arribas-Bel[*†‡]
Geographic Data Science Lab, University of Liverpool

July 2021

ABSTRACT: There is a pressing need to monitor urban form and function in ways that can feed into better planning and management of cities. Both academic and policymaking communities have identified the need for more spatially and temporally detailed, consistent, and scalable evidence on the nature and evolution of urban form. Despite impressive progress, the literature can achieve only two of those characteristics simultaneously. Detailed and consistent studies do not scale well because they tend to rely on small-scale, ad-hoc datasets that offer limited coverage. Until recently, consistent and scalable research has only been possible by using simplified measures that inevitably miss much of the nuance and richness behind the concept of urban form. This paper outlines the notion of "spatial signatures", a characterisation of space based on form and function, and will specifically focus on its form component. Whilst spatial signature sits between the purely morphological and purely functional description of the built environment, its form-based component reflects the morphometric definition of urban tissue, the distinct structurally homogenous area of a settlement. The proposed method employs concepts of "enclosures" and "enclosed tessellation" to derive indivisible hierarchical geographies based on physical boundaries (streets, railway, rivers, coastline) and building footprints to delineate such tissues in the built fabric. Each unit is then characterised by a comprehensive set of data-driven morphometric characters feeding into an explicitly spatial contextual layer, which is used as an input of cluster analysis. The classification based on spatial signatures is applied to the entirety of Great Britain on a fine grain scale of individual tessellation cells and released as a fully reproducible open data product. The results provide a unique input for local authorities to drive planning and decision-making and for the wider research community as data input.

[*]THANKS.

[†]Geographic Data Science Lab, Department of Geography and Planning, University of Liverpool, Roxby Building , 74 Bedford St S , Liverpool , L69 7ZT, United Kingdom

[§]E-mail: D.Arribas-Bel@liverpool.ac.uk; phone: +44 (0)151 795 9727; website: http://darribas.org.

[‡]E-mail: M.Fleischmann@liverpool.ac.uk; website: https://martinfleischmann.net/.

1. Introduction

Identification of recurring patterns in the built environment is deeply embedded within all schools of urban morphology. We can discuss Conzen's Plan Unit (Conzen, 1960), which evolved into Whitehand's morphological region within the British historico-geographical tradition, aiming to capture internally homogenous areas of a shared origin and spatial character (Oliveira and Yaygin, 2020). At the same time, we may talk about *tessuta urbana* or *urban tissue*, stemming from the Italian school of typo-morphology (Caniggia and Maffei, 2001), and sharing the notion of internal homogeneity, just detected by different methods reflecting the architectural perception of space more than geographic. However, it is a complicated pursuit to scale up morphological studies of these kinds without losing too much information. Traditional schools (both historico-geographical and typo-morphological) are generally not able to scale their methods to larger areas while keeping the detail. Another well-established school of urban morphology - Space Syntax, based on the work of Hillier and Hanson (Hillier, 1996), is different and can be considered scalable to metropolitan or even national extents (Space Syntax Limited, 2018). However, Space Syntax at these scales is limited to an analysis of street networks and their configuration, completely omitting patterns formed by plots, buildings, and open spaces. The same can be said about a broader range of network-based methods like Multiple Centrality Assessment (Porta et al., 2010) - while they can be scaled, their insight is inherently limited by the limited data input.

The recent growth of purely quantitative methods of urban form analysis, often nicknamed *urban morphometrics*, and their ability to scale without losing too much detail is opening a range of opportunities to give urban morphology a toolkit to analyse recurring patterns at metropolitan and national extents. After the first explorations in the works of Gil et al. (2012) or Hamaina et al. (2012), the methods are starting to mature as illustrated by recent publications of Multiple Fabric Assessment by Araldi and Fusco (2019), a series of element-based typologies by Berghauser Pont et al. (2019), gridded classification by Jochem et al. (2020) or hierarchical model following the biological methods of taxonomy creation by Fleischmann (2021). All share a similar approach, based on the an initial set of measured characters capturing the individual aspects of form-based patterns and subsequent unsupervised classification. As all methodological steps can be expressed as computer algorithms, they can potentially scale to nation-wide analyses, as already shown by Jochem et al. in the case of Great Britain (Jochem and Tatem, 2021). Is it to be noted that each of the existing methods has its limitations, often linked either to the a spatial unit that does not ensure internally homogenous urban patterns (Jochem et al., Araldi and Fusco), dependency on rarely available data (Berghauser Pont et al.) or a limited number of measured characters, which may omit some aspects of patterns and introduce selection bias in the method (Berghauser Pont et al.).

This paper presents *spatial signatures*, a method of characterisation of space based on its form and function, and focus specifically on its form component, able to delineate internally homogenous areas of urban form based on an extensive set of morphometric characters. The method is applied in the case of Great Britain, deriving an exhaustive classification of both built and non-built environment. The remainder of the paper outlines the method (Section 2), including the introduction of the Enclosed tessellation as the spatial unit, presents resulting classification (Section 3), and discussed its implications on the analysis of urban form (Section 4).

2. Method

Urban morphology tends to consider form only when trying to understand urban environment, leaving aside its functional aspects from population and amenities to green and blue spaces. However, one can argue that only combination of form and function can provide deeper insights into the complexity of cities. The notion of *spatial signatures* combines both aspects into a single model, defined as:

A characterisation of space based on form and function designed to understand urban environments

Spatial signatures then provide a typology of space defined by both form and function together, encoding the interplay between the two. On the other hand, not all questions can be answered using both components at the same time. Each of them, form as well as function, can be studied independently when needed. This paper explores form component of spatial signatures characterisation and derives form-based signatures as a single-aspect classification of space.

Spatial signatures are conceptually defined as an aggregation of granular elements into contiguous areas based on the homogeneity of their characterisation. That poses a first methodological question - which spatial unit should be used for capturing of such characterisation? The optimal unit should be *indivisible* - when split into smaller components, none of them would be enough to encode character of a signature; *internally consistent* - each observation should reflect single signature type; and geographically *exhaustive* - every location in the area of interest, both built-up and un-built, should be covered.

Literature tends to be split into three groups when it comes to unit of analysis. One uses predefined administrative boundaries (REF) but that can be seen as a suboptimal solution as their definition follows different goals and some even argue that "administrative units obscure morphologic reality" (Taubenböck et al., 2019). The other employs uniform grids, in some cases linked to a spatial index (like hexagonal H3 grid (Brodsky, 2018)) or to other types of data commonly distributed on grids (e.g. Jochem et al. (2020)). However, grids cells are rarely indivisible and often internally inconsistent as their definition does not reflect the spatial configuration on the ground. The most common approach in urban morphology is to use structural elements as buildings (Hamaina et al., 2012), street segments (Araldi and Fusco, 2019) or plots (Berghauser Pont et al., 2019) as a unit. However, these are not exhaustive as they are not present in un-built areas. Plots would theoretically provide geographical exhaustiveness but since their conceptual definition is not stable and geometric representation varies (Kropf, 2018), they are unfit for a large scale analysis.

We propose to use an alternative spatial unit named *enclosed tessellation cell* (EC), defined as:

The portion of space that results from growing a morphological tesselation within an enclosure delineated by a series of natural or built barriers identified from the literature on urban form, function and perception.

Martin: [We should somehow cite the conceptual paper when we talk about signatures and EC]

The ECs are generated in three steps. First, a defined set of spatial features, that divide space into smaller parts, is integrated into a single set of boundaries. Such boundaries are usually formed by linear features as street network, railway or rivers. Second, these boundaries are used to subdivide space into *enclosures*, a smaller areas delimited from all sides by at least one boundary feature. Third, enclosures are combined with building footprints taking the role of anchors in space and subdivided into ECs based on proximity to each building, using a morphological tesselation algorithm (Fleischmann et al., 2020).

Martin: [Figure?]

Resulting ECs are indivisible and internally consistent, as they are very granular and linked at maximum to a single building, and exhaustive as they cover entirety of space thanks to contiguous geometry of enclosures. Their structure and scale adapts to the environment and can take a form of a small-scale granular mesh in city centres of historical origin as well as large-scale polygons encoding the vast natural open spaces.

Having ECs, building footprints and street networks, we measure a wide range of aspects of their spatial organisation, from dimensions and shapes of individual objects to their spatial distribution, intensity or connectivity reflecting configuration of streets. We call these measurements *morphometric characters*. Since we do not know a-priori which characters are the most determining the differences between types of urban development, we aim to include a relatively large set of non-collinear characters to avoid potential selection bias, believing that further steps will be able to deal with a larger amount of input data generated by a larger number of characters[1]. As the aim is to use these characters to detect contiguous homogenous areas of urban form, we are more interested in tendencies of their distributions within space. Therefore, we link all values to ECs and capture the distribution of each of them within a spatial context around each EC. Context here is defined as a topological distance (10 steps) on enclosed tessellation and each EC in reach is weighted by its distance to the original EC. This kind of aggregation is adaptive to the pattern of ECs and reflects the higher importance of close elements compared to more distant ones.

ECs characterised by contextualised morphometric characters are clustered using K-Means clustering algorithm, deriving a typology of ECs. Even though K-Means itself does not contain any contiguity constraint, the design of inherently spatially autocorrelated characters results in clusters which are spatially contiguous anyway. The clustering can be potentially applied hierarchically. When further detail is needed, individual clusters can be clustered again resulting in a hierarchy of classes. Finally, ECs are aggregated together based on their class, creating geometry of spatial signatures, where each contiguous area classified as a single cluster represent a single signature.

We test the method on the case study covering the whole Great Britain, illustrating both potential and limits of the analysis of urban form at scale.

[1]The list of measured morphometric characters and their implementation details are available in the online repository at github.com/urbangrammarai/spatial_signatures.

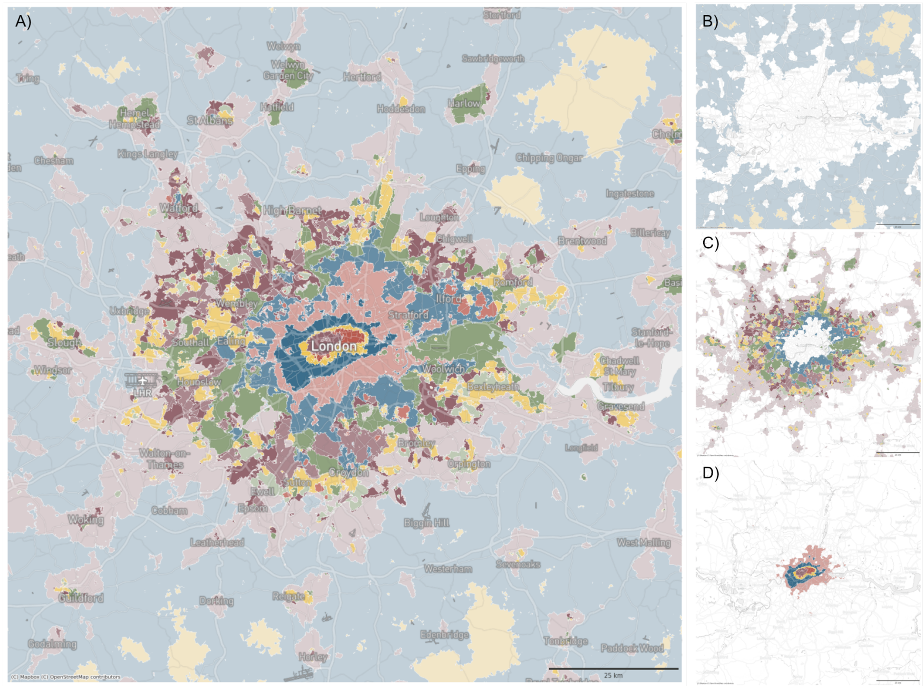

Figure 1: Form-based spatial signatures in London (A), and their classification into top-level macro groups (countryside (B), suburban development (C) and city centres (D)).

The method itself requires a relatively available data on input. First, a set of barriers encoding delimiters of enclosures is needed. In our case, we use street networks from OS OpenRoads (REF), representing street centrelines, railways from OS OpenMap Local (REF), rivers from OS OpenRivers and a coastline from OS Strategi® (REF), all released as open data under a permissive license allowing us to release resulting classification as open data product. Second input is the layer capturing building footprints, which is again retrieved from the OS OpenMap Local. However, the data reflect aggregated footprints and do not distinguish between individual buildings when they are adjacent. While that is indeed a limitation, the method is designed and tested (REF Conceptual paper??) to be robust enough to accommodate for various sub-optimal data sources.

3. Results

The resulting classification of Great Britain identified 19 types of form-based signatures that can be organised into three macro groups: countryside, suburban low density development and dense city centres. It is a result of two hierarchical steps based on K-Means clustering, where the first resulted in 7 classes, of which the two most urban have been further clustered into 6 and 8 lower-level classes (after a removal of outlier classes). An illustration of the classification in London is shown on figure 1 and in Birmingham on figure 2.

Countryside macro group is composed of four signature types covering large-scale open spaces from agricultural land in southern England to vast natural areas of Scottish Highlands. The urban

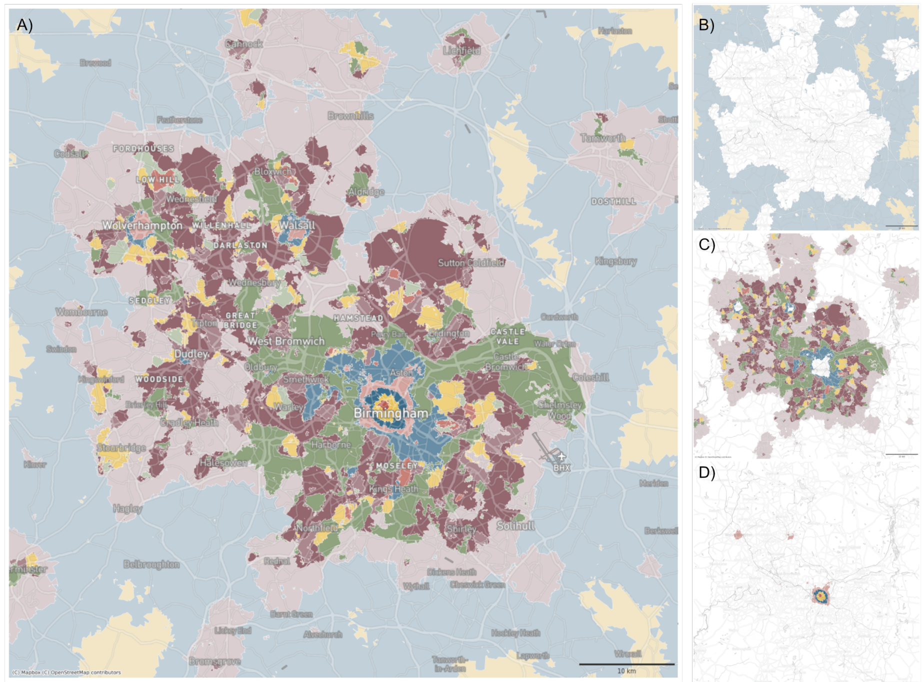

Figure 2: Form-based spatial signatures in Birmingham (A), and their classification into top-level macro groups (countryside (B), suburban development (C) and city centres (D)).

development in this group is limited to small villages or hamlets. All four classes are a result of the first clustering step.

Second macro group covers suburban low-density development areas, taking up most of the area of british cities. We can identify 9 types of signatures, originating in two classes from the first step. The range of types stretches from sparse single-family housing on the peripheries of cities, planned residential developments of 20th century to predominantly industrial areas. The types differ in many aspects, from the overall built-up density and related geometry of both enclosed tessellation and enclosures (both affecting the description via many morphometric characters) to connectivity of street networks or their solar orientation.

Final macro group comprises of dense town and city centres, all originating from the single top level cluster. These signatures reflect the main hubs of activities in each larger settlement. In some cases, they are all located in the same central areas, while in others some local district nodes show up (e.g. Liverpool). All six types can be arranged according to their level of *urbanity* and tend to form concentric rings.

Two types are exclusive to London's city centre (roughly around Soho as show on figure 3A) and are not present anywhere else in the whole country. The remaining are present in other places and their presence can encode the *urbanity* of each city or town. For example, Birmingham as the second largest city in the country contains four types of central signatures (figure 3B), compared to six in London. Scottish capital Edinburgh contains only three (like many other cities across the country (e.g. Manchester or Glasgow)) illustrated on figure 3C. Smaller cities like Southampton have only two types (figure 3D) and towns as Oxford are limited to a single central type (figure 3E). The presence of types is not accidental – smaller cities lack the most urban types, while all of

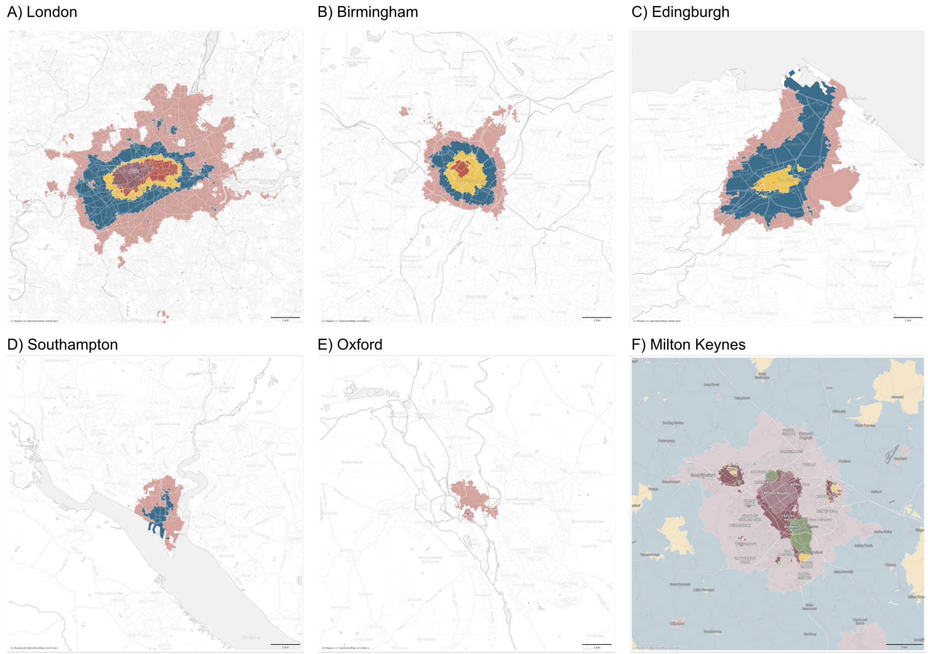

Figure 3: Hierarchy of city centres from the most urban with all six signature types belonging to city centres macro class in London (A), four types in Birmingham (B), three in Edinburgh (C), two in Southampton (D) and one in Oxford (E). Milton Keynes (F) does not have an area classified as morphological centre.

them contain the least urban type.

Notable is the case of Milton Keynes, a new town built since 1960s on the green field with a target population of 250 000 (currently at 230 000). Its development followed a carefully designed masterplan, laying out the whole city. However, the resulting structure is very different from any other city in the country as none of the signatures encoding urban centre is not present in Milton Keynes. We could say, that it does not, have a morphologic centre, as illustrated on the figure 3F.

4. Discussion

Spatial signatures, either based on a combination form and function or one components only, aim to identify and delineate recurring patterns of the built environment. As such, they are deeply embedded in the tradition of urban morphology directly related to established concepts like morphological region, urban tissue and similar. All aim to determine which areas of cities are internally more similar that they would be to any other place. At the same time, the concepts and their methods are not same and will undoubtedly produce a variation in the classification of urban form. Where historico-geographical morphological region would detect original development and its later expansion following the same pattern as two distinct areas (due to a different historical origin), form-based signatures would result in a single area as the pattern is homogenous. Which of the two is the *correct* answer then fundamentally depends on the question asked. Each of the concepts and related methods of their analysis are bound to different scales, purposes and questions and are complimentary rather than competing.

Unlike traditional urban morphology that based mostly on qualitative methods requiring a high degree of expert knowledge and interpretation, morphometric studies offer a high degree of reproducibility and replicability. The methods are often fully written as Python (as in case of this research) or R code, which given the same data on input produces the same results and given similar data for a different case study area should produce comparable results. The stability and reliability of the method is especially important in the policy-making context which needs to be consistent across years and updates of supporting morphological data.

Morphometric approaches have their limitations, related to the the dependency on the quality of input data as well as the behaviour of used algorithms. Suboptimal data input can significantly affect the performance of signature detection. Take, for example, the quality of building footprints as one of the most utilised data source. While there is a wide range of reliable, mostly governmental or municipal, sources of data, that does not apply to every place and dataset. Some, like OpenStreetMap, provide highly inconsistent data with some buildings drawn in a high detail or even subdivided into multiple parts based on variable building height, other do not distinguish between individual buildings and provide a single geometry for the whole block composed on multiple street fronts and buildings.

Algorithms used to classify features into signatures also make a difference, especially on the national scale shown in this paper. The classification and to a degree a *resolution* of clustering is done on a national level, that may result in local differences being smoothed out as insignificant from the national perspective. The zoom into individual neighbourhoods then may seem to be coarse in terms of identified signatures. Same applies to the placement of the boundaries between signature types. Trained urban morphologist or urban designer would likely question why one building belongs to a type A while the other within the same block to a type B but we have to keep in mind the algorithmic nature of the work providing a prediction more than anything else.

However, the significant difference between the presented method and traditional morphology is the scalability of the analysis. With national-scale data we can ask questions about much larger patterns than we did before. Obtaining the analysis covering the same extent as the one just presented in the same level of detail using historico-geographical or typo-morphological methods would be simply an unfeasible task. With morphometric methods we can start looking into cross-national similarities and tendencies in urban development on the neighbourhood-like scale signatures provide.

5. Conclusions

This paper identifies the scalability issue of traditional urban morphology and presents a method of characterisation of space based on its form component, one which embraces urban morphometrics and offers scalability to large extents. That is illustrated on the case study of Great Britain, which environment has been classified into 19 types of form-based spatial signatures. Those can be organised into three macro groups of countryside, suburban low density development and dense city centres. The spatial composition of city centres then tells a story of hierarchy of British cities with London being the only one containing the most *urban* of signatures.

Compared to other scalable methods, spatial signatures build on both street networks and building footprints using the inclusive set of morphometric characters aiming to capture a wide spectrum of variables, minimising a selection bias incurred by small sets. The method paves the way towards large scale urban morphology which is able to discuss urban form across nations while still retaining its most natural neighbourhood scale.

References

Araldi, A. and Fusco, G. (2019). From the street to the metropolitan region: Pedestrian perspective in urban fabric analysis:. *Environment and Planning B: Urban Analytics and City Science*, 46(7):1243–1263.

Berghauser Pont, M., Stavroulaki, G., Bobkova, E., Gil, J., Marcus, L., Olsson, J., Sun, K., Serra, M., Hausleitner, B., Dhanani, A., and Legeby, A. (2019). The spatial distribution and frequency of street, plot and building types across five European cities. *Environment and Planning B: Urban Analytics and City Science*, 46(7):1226–1242.

Brodsky, I. (2018). H3: Uber's hexagonal hierarchical spatial index. *Available from Uber Engineering website: https://eng. uber. com/h3/[22 June 2019]*.

Caniggia, G. and Maffei, G. L. (2001). *Architectural Composition and Building Typology: Interpreting Basic Building*, volume 176. Alinea Editrice, Firenze.

Conzen, M. (1960). *Alnwick, Northumberland: A Study in Town-Plan Analysis*. George Philip & Son, London.

Fleischmann, M. (2021). *The Urban Atlas: Methodological Foundation of a Morphometric Taxonomy of Urban Form*. PhD thesis, University of Strathclyde, Glasgow.

Fleischmann, M., Feliciotti, A., Romice, O., and Porta, S. (2020). Morphological tessellation as a way of partitioning space: Improving consistency in urban morphology at the plot scale. *Computers, Environment and Urban Systems*, 80:101441.

Gil, J., Montenegro, N., Beirão, J. N., and Duarte, J. P. (2012). On the Discovery of Urban Typologies: Data Mining the Multi-dimensional Character of Neighbourhoods. *Urban Morphology*, 16(1):27–40.

Hamaina, R., Leduc, T., and Moreau, G. (2012). Towards Urban Fabrics Characterization Based on Buildings Footprints. In *Bridging the Geographic Information Sciences*, volume 2, pages 327–346. Springer, Berlin, Heidelberg, Berlin, Heidelberg.

Hillier, B. (1996). *Space Is the Machine : A Configurational Theory of Architecture*. Cambridge University Press, Cambridge.

Jochem, W. C., Leasure, D. R., Pannell, O., Chamberlain, H. R., Jones, P., and Tatem, A. J. (2020). Classifying settlement types from multi-scale spatial patterns of building footprints. *Environment and Planning B: Urban Analytics and City Science*, page 239980832092120.

Jochem, W. C. and Tatem, A. J. (2021). Tools for mapping multi-scale settlement patterns of building footprints: An introduction to the R package foot. *PloS one*, 16(2):e0247535.

Kropf, K. (2018). Plots, property and behaviour. *Urban Morphology*, 22(1):1–10.

Oliveira, V. and Yaygin, M. A. (2020). The concept of the morphological region: Developments and prospects. *Urban Morphology*, 24(1):18.

Porta, S., Latora, V., and Strano, E. (2010). Networks in Urban Design. Six Years of Research in Multiple Centrality Assessment. In *Network Science*, pages 107–129. Springer London, London.

Space Syntax Limited (2018). Space Syntax OpenMapping.

Taubenböck, H., Weigand, M., Esch, T., Staab, J., Wurm, M., Mast, J., and Dech, S. (2019). A new ranking of the world's largest cities—do administrative units obscure morphological realities? *Remote Sensing of Environment*, 232:111353.