Classifying urban form at national scale
The British morphosignatures

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

July 2021

ABSTRACT: Identification of recurring patterns in the built environment is deeply embedded within all schools of urban morphology. However, a constant challenge in advancing the systematic study of urban morphology has been that of scaling studies up to consistently cover large regions with enough degree of detail. The recent growth of morphometric methods, 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 consistently at metropolitan and national extents. In this paper, we use the spatial signatures, a recent quantitative framework to characterise space based on form and function, and focus on the form component - morphosignatures. Morphosignatures are conceptually defined as an aggregation of granular elements into contiguous areas based on the homogeneity of their characterisation. We adopt an enclosed tessellation cell (ETC) as a spatial unit onto which the characterisation is projected. Having ETCs, 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. ETCs described by these characters are grouped using the K-Means algorithm, effectively deriving a typology of urban form. We deploy this approach on the case of Great Britain, illustrating both potential and limits of the analysis of urban form at scale. The resulting classification identified 19 types of morphosignatures that can be organised into three macro groups: countryside, suburban low density development and dense city centres. The main contribution of the proposed method, with respect to traditional morphological studies, resides in the scalability of the analysis. With national-scale data we can ask fundamentally new questions about the emerging patterns than is possible with smaller scales.

*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 start from 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.

A constant challenge in advancing the systematic study of urban morphology has been that of scaling studies up to consistently cover large regions with enough degree of detail. 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 seminal work of Hillier and Hanson (Hillier, 1996), is different as it is predominantly quantitative 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 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 consistently 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 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. (2020) in the case of Great Britain. It is to be noted that each of the existing methods has its own limitations, often linked either to the choice of a spatial unit that does not ensure internally homogenous urban patterns (e.g. Jochem et al., 2020 or Araldi and Fusco, 2019), the dependency on rarely available and/or scalable data (e.g. Berghauser Pont et al., 2019), or to a limited number of measured characters, which may omit some aspects of patterns and introduce selection bias in the method (e.g. Berghauser Pont et al., 2019).

In this paper, we use the *spatial signatures* (Arribas-Bel and Fleischmann, 2021), a recent framework to characterise space based on form and function, and focus on the form component. We apply these concepts to 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 description of the Enclosed Tessellation cell as the spatial unit; presents resulting classification (Section 3); and discusses its implications on the analysis of urban form (Section 4).

2. Method

We follow the notion of spatial signatures as a framework to develop our classification. This approach was developed by Arribas-Bel and Fleischmann (2021), who define it as:

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

Spatial signatures thus provide a typology of space defined by both form and function together, encoding both patterns of each aspect as well as the interplay between the two. The field of urban morphology tends to consider form only when studying urban environments, leaving aside functional aspects such as population, amenities, or green and blue spaces. This paper explores the form component of the spatial signatures. Our characterisation focuses only on morphological aspects of urban environments, thus better aligning itself with the urban morphology tradition. Because we follow the spatial signatures framework but focus on form-based characters only, we call the results of our analysis, both the types we derive and delineations we obtain, *morphosignatures*.

Spatial signatures, and morphosignatures by extension, are conceptually defined as an aggregation of granular elements into contiguous areas based on the homogeneity of their characterisation. That poses the subsequent methodological question: which spatial unit should be used for developing of such characterisation? We argue the optimal unit should be: *indivisible*, so that if split into smaller components, none of them would be enough to encode character of a (morpho)signature; *internally consistent*, in a way that each observation reflects a single (morpho)signature type; and geographically *exhaustive*, with all the space in the area of interest assigned into one and only one (morpho)signature.

The literature tends to be split into three groups when it comes to unit of analysis. The first relies on predefined administrative units (REF) which, although they are convenient to source, can be counter-productive and "obscure morphologic reality" (Taubenböck et al., 2019). The second employs uniform grids, either linked to a spatial index (like hexagonal H3 grid from Brodsky, 2018) or to ancillary data commonly distributed on grids (e.g. WorldPop grids as in Jochem et al., 2020). Regular geometries are often internally inconsistent as their definition does not reflect the spatial configuration on the ground, thus commonly prone to the modifiable areal unit problem (MAUP, Openshaw, 1981). The third and most common approach in urban morphology is to use structural elements such as buildings (Hamaina et al., 2012), street segments (Araldi and Fusco, 2019) or plots (Berghauser Pont et al., 2019) as a unit. These are closer to the underlying entity of interest, but are not exhaustive of space 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 large scale analysis.

We follow Arribas-Bel and Fleischmann (2021) in adopting an alternative spatial unit named *enclosed tessellation cell* (ETC) which is defined as:

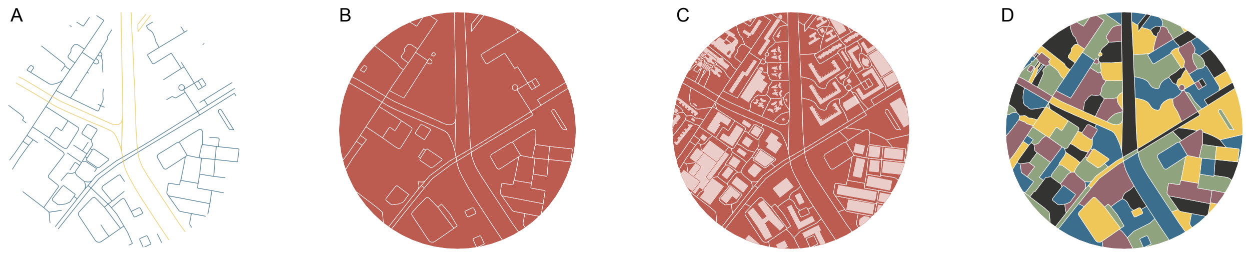

Figure 1: Diagram illustrating the sequential steps leading to the delineation of enclosed tessellation. From a series of enclosing components, where blue are streets and yellow river banks (A), to enclosures (B), incorporation of buildings as anchors (C) to final tessellation cells (D).

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.

The ETCs are generated in three steps illustrated on a Figure 1. First, a defined set of spatial features that divide space into smaller parts (1A) is integrated into a single set of boundaries (1B). 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 (1C). Third, enclosures are combined with building footprints taking the role of anchors in space and subdivided into ETCs using the morphological tesselation algorithm (Fleischmann et al., 2020) (1D).

Resulting ETCs are indivisible and internally consistent, as they are linked at maximum to a single building, and exhaustive as they cover the 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 ETCs, 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 a-priori know which characters are determining the differences between types of urban development the most, we aim to include a relatively large set of 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 ETCs and capture the distribution of each within a spatial context around each ETC. Context here is defined as in terms of topological distance (10 steps) between ETCs, with each one in reach being weighted by its distance to the original ETC. This kind of aggregation is adaptive to the pattern of ETCs and reflects the higher importance of close elements compared to more distant ones.

ETCs characterised by contextualised morphometric characters are grouped using the K-Means algorithm, effectively deriving a typology of ETCs. Even though K-Means itself does not contain

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

any contiguity constraint, the design of inherently spatially autocorrelated characters results in larger clusters that are spatially contiguous. Moreover, this approach can be potentially applied hierarchically. When further detail is needed, the algorithm can be run within individual clusters. Finally, ETCs are aggregated together based on their class, resulting in the set of spatial morphosignature geometries, where each contiguous area assigned the same cluster represents a single morphosignature.

We deploy this approach on the case of Great Britain, illustrating both potential and limits of the analysis of urban form at scale. All the datasets we rely on are available openly. First, for barriers encoding delimiters of enclosures, we use the following: street networks from the "OS Open Roads" data product, representing street centrelines; railways from the "OS OpenMap - Local"; rivers from "OS OpenRivers"; and a coastline from OS Strategi® (REF). All these products are released by the Ordnance Survey under an Open Government License, allowing us to release the resulting classification as an open data product. The 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 (Arribas-Bel and Fleischmann, 2021) to be robust enough to accommodate for various sub-optimal data sources.

3. Results

The resulting classification of Great Britain identified 19 types of morphosignatures 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 seven classes, of which the two most urban have been further clustered into six and eight lower-level classes (after a removal of outlier classes). An illustration of the classification in London is shown on Figure 2 and in Birmingham on Figure 3.

The countryside macro group is composed of four morphosignature types covering large-scale open spaces from agricultural land in southern England to vast natural areas of the Scottish Highlands. The urban development in this group is limited to small villages or hamlets. All four classes are a result of the first clustering step.

The second macro group covers suburban low-density development areas, taking up most of the area of british cities. We identify nine types of morphosignatures, 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 morphosignatures 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

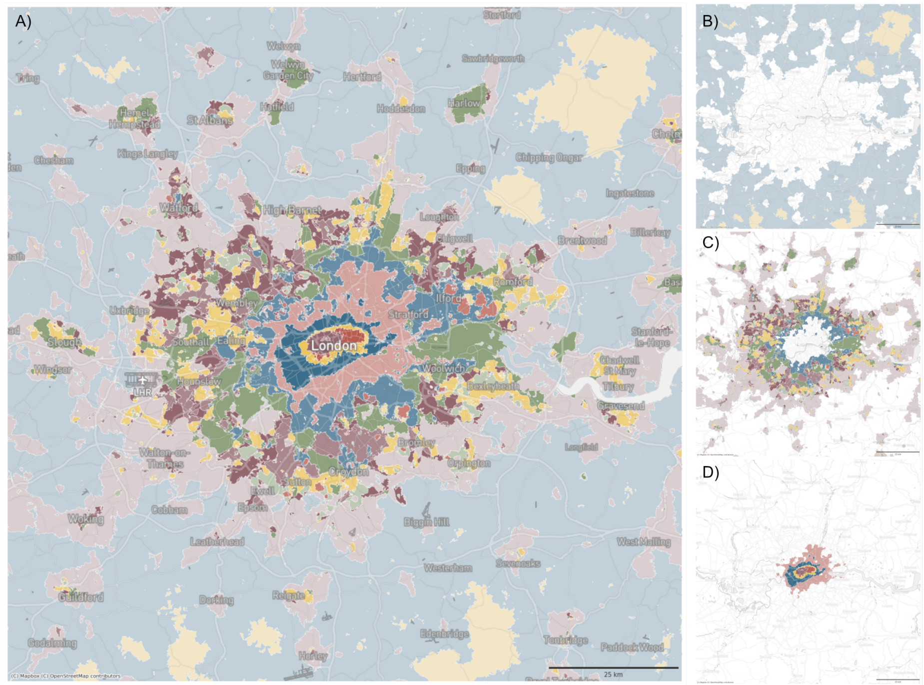

Figure 2: Morphosignatures in London (A), and their classification into top-level macro groups (countryside (B), suburban development (C) and city centres (D)). Contains OS data © Crown copyright and database right 2021.

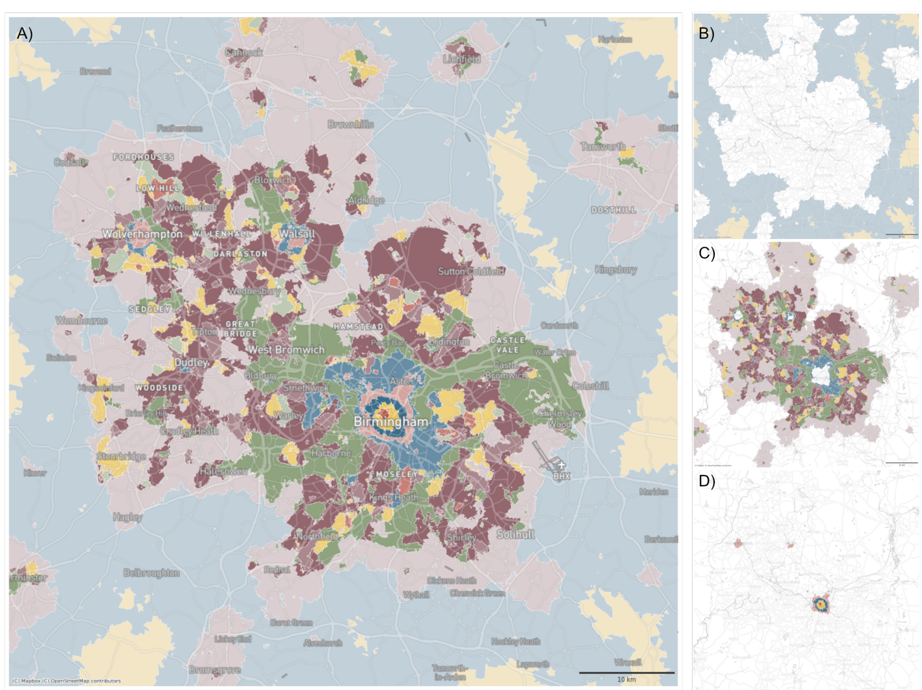

Figure 3: Morphosignatures in Birmingham (A), and their classification into top-level macro groups (countryside (B), suburban development (C) and city centres (D)). Contains OS data © Crown copyright and database right 2021.

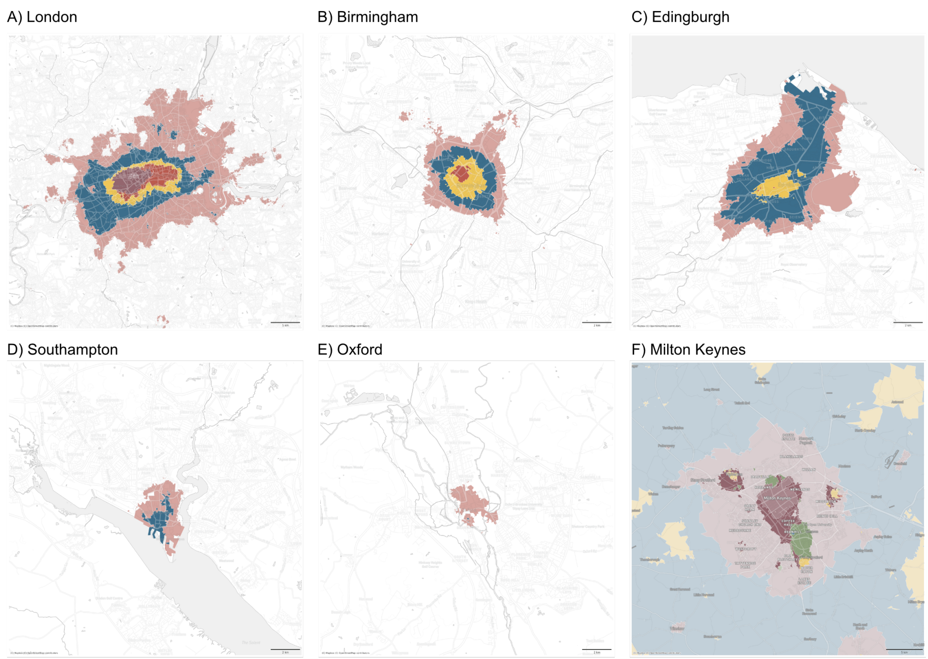

Figure 4: Hierarchy of city centres from the most urban with all six morphosignature types belonging to city centres macro class in London (A), four types in Birmingham (B), three in Edinburgh (C), two in Southampton (D) and one in Oxford (E). Milton Keynes (F) does not have an area classified as morphological centre. Contains OS data © Crown copyright and database right 2021.

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 4A) 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 morphosignatures (Figure 4B), 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 4C. Smaller cities like Southampton have only two types (Figure 4D) and towns as Oxford are limited to a single central type (Figure 4E). The presence of types is not accidental – smaller cities lack the most urban types, while all of 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 morphosignatures 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 4F.

4. Discussion

The spatial signatures framework, either based on a combination form and function or a single component, aims to identify and delineate recurring patterns of urban spaces. As such, it is deeply embedded in the tradition of urban morphology, directly related to established concepts like morphological region, urban tissue and similar. All of them aim to determine which areas of cities are internally consistent. The differences between these perspectives stem from the methods they rely on, which undoubtedly results in variation in the classifications of urban form they produce. Where historico-geographical morphological region would detect the original development and its later expansion following a similar pattern as two distinct areas (due to a different historical origin), the morphosignatures would class them 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 is based mostly on qualitative methods requiring expert knowledge and interpretation, morphometric studies offer a high degree of reproducibility and scalability. Methods are often embedded in computer code release under open licenses which, given the same input data, produces the same results. This 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 input data can significantly affect the performance of morphosignature 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 for some regions of the world, this is not the case everywhere. Some sources, like OpenStreetMap, provide highly inconsistent data, with some buildings drawn in high detail and others in very coarse ways. Fortunately, technology is moving very rapidly and advances in data collection and creation (e.g. building footprints derived from satellite imagery) may change this landscape for the better in the coming years.

The algorithms used to classify features into morphosignatures also make a difference, as does the scale at which an analysis is run. Our classification and, to a degree, the *resolution* at which the clustering is done is national. This choice may result in local differences being smoothed out if they are less relevant from the national perspective. Zooming into individual neighbourhoods, one may find some of the decisions made by the algorithm too coarse. The same applies to the placement of the boundaries between morphosignatures. The trained urban morphologist or urban designer might question why one building belongs to a type A, while the other one within the same block is assigned to type B. It is important to keep in mind the scale and goal of the classification, which is then embedded in the algorithmic choices.

Our main contribution with respect to traditional morphological studies resides in the scalability of the analysis. With national-scale data we can ask fundamentally new questions about the emerging patterns than is possible with smaller scales. Obtaining the analysis covering the same extent as the one just presented at the same level of detail using historico-geographical or typo-morphological methods would be simply an unfeasible task. Morphometrics allows us to begin considering cross-national comparisons that consider detail at neighbourhood-like scale. As we develop more fine grained and scalable classifications, an important challenge will be to devise ways that allow us to blend the insights we can gain at this scale with the body of knowledge developed over time from more qualitative approaches.

5. Data and code availability

The repository, containing reproducible code in Jupyter notebooks is available at `https://urbangrammarai.github.io`. Guidelines on how to access and download the resulting classification are available in the same website.

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.

Arribas-Bel, D. and Fleischmann, M. (2021). Spatial Signatures - Understanding (urban) spaces through form and function. Mimeo.

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.

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.

Openshaw, S. (1981). The modifiable areal unit problem. *Quantitative geography: A British view*, pages 60–69.

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.