# Existing approaches to classification of urban form

To start the journey towards a taxonomy of urban form, we first need to look at the state of the art of the field. Hence, this chapter aims to understand the classification and its various models from a theoretical perspective and review how are these models applied to urban form. Following the scope of the work defined in the first chapter, the main focus is on quantitative methods of urban morphology based on the vector representation of form. However, some detours to the remote sensing field are necessary. Both to give a full picture of the possible approached and to break the existing firm barriers between the fields.

This chapter is structurally split into two main sections. The first one is outlining the general theory of classification, its principles and different models. That is directly reflected in the specification of a hypothetical, optimal model of classification of urban form, i.e. a set of principles which should a model fulfil to gain a theoretical ability to reflect the complexity of urban form. The second part of the chapter examines models of classification present in literature to this day and compares them to the specified requirements. It first focuses on the spectrum of works ranging from the metropolitan scales to the level of individual buildings (from the perspective of the unit). It then identifies three recurring concepts shared among wide groups of researchers - Urban Structural Type, Land cover classification and Local Climate Zones. The overview leads to the specific gap in the current methodology, as none of the published works fits the optimal model requirements.

In short, the following chapter answers four questions allowing this research to build on the existing knowledge:

1. What is classification?
2. How should it look like?
3. How was it done to date?
4. What is missing?

## Making sense of the world

The world is an inherently complex entity, which needs to be simplified to manageable pieces of information to ensure that the human mind will be able to understand it, i.e. to avoid combinatorial explosion [@fernbach2017knowledge]. The natural way of doing so is similar grouping features into chunks and thinking about those. We do not think about individual trees and plants; we consider them all together as a forest. We cluster them based on their functional and geographical relationships into the higher-order entity. That is the essence of classification; we join smaller things into higher-order groups and talk about groups. It is easier for the brain to manage.

### Classification as a method

Before we can talk about classification, it is necessary to declutter and define the term itself as it is used in literature in multiple meanings. In an overview of the whole concept and its meanings, @bailey1994 defines classification as “*the ordering of entities into groups or classes on the basis of their similarity*” (p.1). At the same time, he notes that “*classification is both process and an end result*” (p.2). That can cause terminological confusion as in some situation it might be unclear whether we refer to the resulting systematics or the procedure which generates one. This polysemic nature of the term itself is even more pronounced if we explore the definition offered by the Oxford English Dictionary (OED), which comes with three relevant options:

* “*A systematic distribution, allocation, or arrangement of things in a number of distinct classes, according to shared characteristics or perceived or deduced affinities.*”
* “*The action of classifying or arranging in classes, according to shared characteristics or perceived affinities*”
* “*‌A category to which something is assigned; a class.*” [@oxfordenglishdictionary]

While the first two coincide with the definition offered by @bailey1994, the last adds one more option - a class. This research will refer to classification as process and the final result, following Bailey’s approach, always trying to specify which meaning is used. The last definition offered by OED will not be used within this research.

Classification systems vary and can be systematised based on the different aspects of the resulting structure and the method used in the process [@bailey1994].

The first distinction can be made based on the number of dimensions to *unidimensional* and *multidimensional*. The scope of this work lies in multidimensional classification as a single dimension does not have the explanatory power to describe the complexity of urban form. Dimensions can be both numerical and categorical variables, where unidimensional numerical classification is also known as *binning* or a “*classification scheme*” following @marradi1990.

Looking at the structure of the classification results, we can talk about *flat* and *hierarchical* models. Flat models generally define all classes as equal and do not specify the relationship between them, either because it could not be specified or because the model does not focus on such aspect. An example is general land use classification, which is flat (e.g. residential use, commercial, industrial) and the relationship between different classes is conceptually complicated to define. Hierarchical models specify, either numerically or conceptually, the relationship between different classes and hence offer a certain level of flexibility of classification as their structure can be interpreted on multiple levels of resolution. That allows the division of elements into two macro classes or multiple micro classes within the same system. A typical example is biological taxonomy of species, where different levels of resolution are represented by domains, kingdoms, families or species. Typical hierarchical classification is, compared to flat models, computationally or conceptually more challenging, but the flexibility which it brings (where applicable) is valuable.

@marradi1990 uses the term “*typology*” for flat classification models and "*taxonomy* for hierarchical models. However, @bailey1994 suggests using the same terms to distinguish between classification models based on different criteria. Methodologically, he recognises two key approaches - one is conceptual, for which he uses the term *typology*, and the other is numerical, named *taxonomy*. Typology is then a conceptual classification, where resulting classes represent concepts, not empirical cases [@bailey1994]. This means that typologies can be seen as *qualitative classifications* because there is generally no statistics involved. The land-use case mentioned above is a good example of a typology. On the other side is a taxonomy, which is quantitative classification, with classes being empirical entities in the system proposed by @bailey1994 as well as earlier by @sneath1973.

This confusion in what *typology* and *taxonomy* mean and the criteria used to distinguish between them is present across the literature and is not easy to solve. Furthermore, as with the term classification, *taxonomy* is also used to describe both process and the final result.

Literature knows three terms for quantitative classification based on statistical analysis, to complicate the matter further. One is *numerical taxonomy* [@sneath1973], which is quantitative, algorithmic classification (more on numerical taxonomy is in section ). The other is term *cluster analysis*, describing the process of classification. As @bailey1994 points out, the methods of numerical taxonomy can be classified as clustering algorithms, making the numerical taxonomy (in the sense of a process) and cluster analysis “*virtually synonymous*” (p.7). The machine learning area, to make things more complex, uses term *unsupervised classification* for cluster analysis, but these two (together with *clustering*) are equal.

Within this research, the term *taxonomy* will be predominantly used to describe the final result of hierarchical quantitative classification, the usage which fits in the definition of both @bailey1994 and @marradi1990 but can be used for general hierarchical models if needed. *Cluster analysis* and *clustering* may be used interchangeably and describe the process. The term *typology* will be used within the relevant context to avoid potential disambiguation and its application will generally follow @marradi1990 rather than @bailey1994. Unsupervised classification is left only for occasions where it is needed.

@bailey1994 has summarised the reasons why is classification useful in ten fundamental advantages, which are all transferable to urban morphology and classification of the spatial structure of cities. In short, a classification is a tool for *description* (1), giving an overview of all classes within data based on the same criteria. The classified data has *reduced complexity* (2) to a manageable extent. We cannot deal with all individual animals on Earth, but we can work with their taxonomy, significantly reducing the amount of information to work with. Classification can be used to identify *similarities* (3) and *differences* (4) among cases. Identification of similarities allows us to treat all individuals of a single class of similar classes equally (e.g. *beware of snakes* or *go shopping to commercial district*). On the other hand, we can distinguish subtle differences between similar entities (e.g. the difference between venomous and non-venomous snakes can be helpful). A classification, if done properly, is defined by *an exhaustive list of dimensions* (5) on which different classes are based. In such a case, the resulting classification can be comprehensive, while capturing the relationships between classes and dimensions, which is useful for further analysis and profiling of classes. That allows quick, straightforward *comparison* (6) of classes from different parts of the classification structure. Complete list of classes can serve as the *inventory* (7) for management purposes and allows the study of *relationships* (8) among dimensions, relative to the structure of the classification. Moreover, classes can be used as *criteria for measurement* (9), where one class is the criterion, and other are assessed according to similarity with the criterion (e.g. *how close is snake A to a python?* or *how similar is neighbourhood A to Manhattan?*). Finally, the classification may be *versatile* (10), as it can represent both individual units under scrutiny and their location within property space, but also describes the whole sample of units. (pp.12 - 14).

### Optimal classification model

Classification model can follow a multitude of pathways and can be defined based on various aspects for a plethora of purposes. However, the aims and scope of each research should drive the classification. Since this research aims to develop a data-driven method, its methodological nature should be a quantitative. The *optimal classification model (OCM)* for this research can be then defined based on the following set of rules derived from definitions presented in the previous sections. Each of the existing models presented in literature and outlined in the next section is then related to OCM. Furthermore, the specification of the model used within this study should directly reflect the rules in its fundamental design.

The optimal classification model of the urban form should be:

1. Exhaustive
2. Mutually exclusive
3. Empirical
4. Hierarchical
5. Comprehensive
6. Detailed
7. Scalable

An *exhaustive* model covers all entities within the set, meaning that there should not be unlabelled cases in the resulting classification. While the specification of levels of uncertainty in the labelling is welcome, it should not prevail. *Mutual exclusivity* ensures that no entity is at the same time member of more than one class to minimise ambiguities. However, probabilistic clustering specifies the probability of each feature to be a member of each class, but the resulting classification typically uses a single threshold value. *Empirical* nature ensures the data-driven nature of the classification limiting the potential bias in the derivation of *concepts* and dependency on the expert knowledge to assess each entity. Structurally, the model should be *hierarchical* to allow flexibility of its reading, unavailable for flat options. With hierarchical relations between groups, the information encoded in the classification becomes adaptable and provides more straightforward interpretation. *Comprehensiveness* entails the number of dimensions, or descriptors used to cluster entities. The selection of dimensions can be biased and negatively influence the resulting taxonomy. By implementing the large number of dimensions (trying to be as inclusive as possible), such a possibility can be effectively minimised [@sneath1973]. The classification of the urban form should be *detailed* in terms of spatial granularity, meaning that labels should be assigned to individual plots or buildings rather than districts or cities. Finally, the model should be *scalable*. That is both a technical and conceptual requirement, which should ensure that the same model can be used to classify small town and large metropolitan areas.

A classification model which would adhere to these seven principles is currently not available in the published literature. However, the spectrum of existing approaches is vast, and all the points have been addressed in various works. However, not at the same time.

## Spectrum of classification models

There are two main branches of science focusing on urban form classification, which are surprisingly separated from each other - remote sensing (RS) and urban morphology (UM). The diffusion between them is minimal, even though both may conceptually focus on similar questions. Urban morphologists tend to prefer to work with the elements of the urban form directly and build the classification from the constituents parts of urban form upwards (as in @dibble2017). They need to identify buildings, plots or blocks first, identify the structural features, and then describe and classify them. On the other hand, remote sensing research usually does not dwell into the detail of individual elements, but instead tries to capture the whole pattern directly [@taubenbock2020]. The other approach is to extract the elements first, as in @dogrusoz2007, but then the methodology is essentially two-step, starting with remote sensing in the extraction part and finishing with morphological analysis of extracted elements in the second. Even though both fields approach the topic differently, there are some similar concepts, although currently unlinked.

Following sections provide an overview of the quantitative trajectory classification of urban form followed in last decades and of recurring concepts stemming from the field of remote sensing. The overview does not include concepts derived from traditional qualitative urban morphology as it is not within the scope of this research.

### The trajectory of systematic quantitative classification

The literature on quantitative urban morphology offers a wide selection of proposed methods of classification of urban form, ranging from flat typologies to hierarchical taxonomies and from city to buildings scales (as per the unit of analysis). However, if we want to focus on the methods following similar aims as this research, we find that there are not many methods, which would come close to OCM. This section presents a trajectory of relevant models, focusing here mostly on urban morphology rather than RS, between the years 2007 and 2020. The earlier work [@herold2002; @barnsley1996] is seen as too distant from the current research, likely due to the scarce data and tools availability.

The first relevant method proposed by @dogrusoz2007 attempts to classify urban form based on satellite imagery using automatic extraction of building footprints followed by a cluster analysis of the structure of minimum spanning tree between buildings. The resulting classification can determine whether neighbourhoods tend to be *organized* or *unorganized* as illustrated in figure . As such, it makes the first step towards the detection of different patterns, which is naturally relatively distant from OCM in some of the criteria.

Three steps of classification from satellite imagery, through detected building footprints to resulting classes. (Dogrusoz and Aksoy, 2007, figure 4)

@song2007 are proposing a classification of *neighbourhoods* defined as a 1/4 mile (approx. 400m) buffer around selected sites of new single-family development. Using 21 characters, factor analysis and cluster analysis (K-means), they propose flat model of six clusters. By classification of 6788 homes into neighbourhood types, they illustrate the scalability of the model and potential for detailed assessment, as each site has its own neighbourhood defined in a location-based manner (i.e. neighbourhoods are defined independently on each other and can overlap). However, 21 characters and focus on single-family housing only does not ensure the comprehensiveness of the model.

Classification method presented by @steiniger2008 uses a predefined conceptual typology of buildings and uses morphometric assessment to predict its classes. In this sense, the authors provide the first assessment of the validity of morphometrics in the context of classification of urban form, but the typology itself is not defined empirically. Of a similar nature is research presented by @neidhart2004, @wurm2016 and @hartmann2016. Even though based on a different set of indicators and data inputs, these three works illustrate the progress in labelling over the years.

Spacemate diagram by @berghauserpont2010 identifies six building types based on three measurable dimensions - Floor Space Index, Ground Space index and Open Space Ratio (figure ). They propose the same method to be applied on the scale of the building and the scale of fabric. The method has the potential to be scalable and detailed, but not comprehensive as it is based on three characters only. Moreover, the classification, even though based on empirical values, is rather conceptual than data-driven as it is based on nine pre-defined archetypes.

Spacemate diagram and building types on the scale of fabric. (Berghauser Pont and Haupt, 2010, figure 23)

@gil2012 characterise streets and block using 25 quantitative characters and cluster them based on K-means into six groups of blocks and four groups of streets. The case study area is covering two neighbourhoods of different origin, and results indicate the potential of cluster analysis based on morphometric values in urban morphology. The method is one of the first which use historical origin as a method of validation of clustering and which results reflect expected distinction (figure ).

Geographical distribution of block clusters and street clusters in two studied neighbourhoods. (Gil et al., 2012, figure 4)

@louf2014 propose hierarchical taxonomy of 131 cities based on their street network patterns, with the actual characterisation based on block area and shape. The resulting dendrogram illustrating the classification is shown in figure . The large set of cases illustrate the scalability of research, but that comes at the cost of granularity and comprehensiveness.

Dendrogram representing the structure of classification of cities by Louf and Barthelemy (2014, figure 4). Each bar represents a single case.

@schirmer2015 proposes classification on multiple levels, where the top one is a municipality, even though the unit is a building defined by centrality and accessibility characters. While their proposal is for a “multiscale typology”, methodologically purpose four flat classification models and do not relate one to the other. Their resulting “municipal typology” is illustrated on figure below. While all scales combined may be based on comprehensive information, the proposed model is not.

Four classes of Schirmer and Axhausen’s (2015, figure 11) municipal-level classification mapped in the area of Zurich.

The scale of neighbourhood @schirmer2015 shares with @serra2018a, who are classifying neighbourhoods defined as “*circular areas of 1km radius*” (p.65) characterised by 12 morphological indicators derived from street network, blocks and buildings. Resulting classification is hierarchical taxonomy of selected neighbourhoods, represented by a dendrogram on figure . In its current form classification is not exhaustive as it covers only pre-defined, yet overlapping neighbourhoods (figure ), and it is not known how would it scale to the continuous classification of whole areas.

Hierarchical classification of neighbourhoods proposed by Serra et al. (2018).

Geographical distribution of clusters of neighbourhoods proposed by Serra et al. (2018).

Similar neighbourhood scale is used by @dibble2017, where the unit of classification is Sanctuary Area [@mehaffy2010]. The resulting classification is a hierarchical taxonomy (figure ) based on the comprehensive set of morphometric characters. However, due to the selection of the basic unit, it is not detailed and exhaustive. The method itself is time-consuming [@dibble2016], and its proposed form is not scalable. However, the work of @dibble2017 is building foundations of the science of urban morphometrics and will be further examined in section 3.2.2.

Resulting dendrogram illustrating the classification of Sanctuary Areas by Dibble et al. (2017, figure 6, rotated). Notice the consistency of morphometric classification and historical origin of individual cases.

The work on urban typologies presented by scholars at Chalmers University in a series of recent publications [@berghauserpont2017; @berghauserpont2019a; @bobkova2019] proposes to use three individual typologies of morphological elements: plots, streets and buildings. Each typology is defined through a handful of morphometric characters and cluster analysis, thus making the outputs influenced by this particular selection. Compared to the optimal criteria above, their model is not hierarchical and, notably, not comprehensive (due to the limited number of morphometric characters it uses).

Spatial distribution of plot types in selected case studies developed by Berghauser Pont et al. (2019, supplementary material figure 10).

The work of @araldi2019 proposes a classification of street segments from the pedestrian point of view, based on 20+ morphometric characters derived from street networks, building footprints and digital terrain model. The model is powerful in terms of top-level classification of urban form, however, similarly to the Chalmers’, it is not hierarchical (the relationship between the types is unknown) and still far from comprehensive (compared, e.g. to others which use a more significant number of characters such as @dibble2017 with 207). The selection of street as the smallest unit is also a limitation as it assumes homogeneity of the urban form along both sides of the whole segment, which is rarely the case in urban contexts of almost all periods of development.

Urban fabric classes as a result of the MFA procedure by Araldi and Fusco (2019, figure 2).

@dong2019 proposes a classification of blocks into hierarchical taxonomy using convolutional autoencoder (CAE). The method rasterises the vector representation of block footprints to 64x64 pixels and uses a neural network for image recognition to cluster them. Due to the necessity to keep data for autoencoder of a similar size, oversized and undersized cases were excluded, drawing the method not entirely exhaustive. Resulting hierarchical clustering (figure ) identifies 16 clusters, but in a way which includes all cases, leaving some (approx 40%) unclassified. However, the application of CAE is quite unique in the context of the rest of the field.

Hierarchical clustering of sample of blocks and illustration of different types (rotated). (Dong et al., 2019, figure 6)

@li2020 focus on classification of 83 blocks into hierarchical taxonomy on the basis of 11 indicators (figure ). On top of hierarchical clustering, authors also do K-means analysis, resulting in 5 types, although the relation between K-means clusters and hierarchal one seems to be left unexplored. The sample of blocks covers only a single small case study area, leaving the question of scalability unresolved.

Hierarchical clustering of selected blocks and illustration of different types (rotated). (Li et al., 2020, figure 5)

The method proposed by @jochem2020 works on 100m grid and is based on building footprint data and 7 characters measured per grid cell. That resulted in 12 types (figure ) in the first and 5 in the consequent step. The method dependent on the arbitrary grid does not follow the natural composition of urban form and in some case might be unable to recognise certain linear patterns. Although scalable, as presented in the paper, the method is likely not comprehensive enough due to the small number of measurable character influencing the resulting classification.

Settlement types prediction in Kinshasa, Congo in the resolution of 12 types based on 100m grid. (Jochem et al., 2020, figure 3)

### Recurring concepts

The trajectory above presents instead unlinked methods and concepts as urban morphology did not agree on a specific approach yet. Likely due to the relative sub-optimality of all presented above. In remote sensing, partially overlapping with urban morphology, are identifiable three recurring partially linked concepts in the literature - Urban Structural Types (UST), Land Cover (LC) classification and Local Climate Zone (LCZ) classification.

The concept of UST (and related Urban Structural Unit) has been first developed for planning purposes in the 1960s [@lehner2019], alongside the UM concepts of the morphological region [@conzen1960; @oliveira2020] or *tessuta urbana* [@caniggia1979], with which it shares the core of the definition based on the internal homogeneity. In the RS, UST started to appear since the early 1990s [@lehner2019] and become quickly popular. However, the methods and terms are still remaining inconsistent. Defining UST is not an easy task as the literature is not consistent in the definition and a range of works reviewed by @lehner2019 agree on principles, but not on specific “definition, description and derivation” [@lehner2019, p.7]. Generally speaking, UST can be defined as a unit of a “*specific spatial characteristics, e.g., the morphology and the spatial relationships between urban artefacts such as buildings, streets, trees, lawns*” [@lehner2019, p.2]. The practical translation of this, relatively vague description, varies. Some UST classifications are flat while others are hierarchical, work on various scales and are mostly conceptual typologies, with only a few examples of data-driven models [@lehner2019]. In relation to the OCM requirements, it is complicated to assess the concept as a whole due to its internal inconsistency, but no method found in literature fulfils all the criteria.

Land cover is a related but more straightforward concept coming from Remote Sensing area. Unlike all the other, it does not focus purely on the urban environment but aims to classify all areas together. As such, its detail when it comes to urban form is generally low, with CORINE Land Cover classification [@europeanenvironmentagency1990] dividing urban form into continuous and discontinuous, plus specialist (industrial/commercial, ports, airports). Copernicus Urban Atlas refines Corine by adding density on top of continuity as a second criterion. One of the most refined is the work of @pauleit2000 which distinguish 10 types of urban form. However, due to the nature of land cover classification, which is usually done as a supervised classification (i.e. labelling), it is usually a conceptual method (which can be both flat and hierarchical).

Classification into Local Climate Zones [@stewart2012] is a mix of both urban structure and land cover into a singular conceptual typology of 10 types of the built form and 7 land cover types (figure ). It is intentionally very generic to produce inclusive classification, covering all possible types of urban development. It is a flat model with classes defined numerically, but still capturing rather conceptual divisions. In RS, LCZ are used on a large scale [@taubenbock2020] to characterise cites across the world, but the classification itself is still relatively limited by design and very top-level.

Local Climate Zones as reported by Taubenböck et al., 2020, figure 1)

### Overview

The relation of each of the concepts mentioned above to the OCM is summarised in the table below.

The majority of methods provide mutually exclusive classification, are exhaustive and empirical. Assessing scalability is not straightforward as the published research is mostly not reproducible without much effort. The scalability is then inferred from the case study area or the method description and may not be entirely precise. The least fulfilled criteria are hierarchical nature and comprehensiveness of methods, with only a few works being such. However, this seems to be changing, as the proportion is growing in recent years.

## The gap in the systematic classification

Even though literature covers fields of geography, urban morphology, remote sensing or cartography, there is no classification method in the scrutinised body of work which fulfils all seven criteria of OCM. Both exhaustivity and mutual exclusivity are generally well covered as most of the methods follow these rules. The empirical nature is less presented. We have to keep in mind the difference between classifications with clusters based on empirical data and those that use empirical data to label observations properly. The latter is not empirical in the sense of OCM as the model itself is conceptual.

Although we can find methods which produce hierarchical classification, it is not a standard, likely due to a) popularity and performance of flat clustering methods as K-means in case of empirical studies, b) additional complexity in case of conceptual studies; hence they rarely reflect inter-group relationships. The significant issue is with the comprehensiveness of presented models. Most of them are based on a handful of dimensions, making them less robust and more prone to bias induced by the careful selection of indicators. Those few examples which are based on a broad set of descriptors [@dibble2017] unfortunately fail to fulfil one or more of the other criteria.

There are studies which are detailed, i.e. their spatial unit is granular (building, plot) but again - they are sub-optimal in other aspects. The scalability is not easy to assess as the majority of methods is not readily reproducible, some do not even describe used tools and computational framework, but that is an issue by itself. In works where the method is described to a sufficient detail, the scalability varies. However, with the rapid expansion of scientific geospatial software and GIS in general, it is assumed that it should not be an issue in future unless the method depends on qualitative data.

All above lead to the conclusion that current methods are not able to provide sufficient details and complexity when it comes to the classification of urban form and hence their applicability in further research or planning practice is limited.

## Summary

Classification is a rich and polysemic term, which needs specific definitions to resolve possible ambiguities. However, even more specific terms as typology and taxonomy are used with different meanings. The ideal path is to be explicit about the context and do not rely entirely on a single definition of the term. Note that since taxonomy can stand for both procedure and the final result, the terminology also needs more refinement. Therefore, this research uses term taxonomy for the final result, whilst term cluster analysis for the procedure.

The classification itself brings a lot of valuable aspects. However, to ensure that all are present, the method for the classification of the urban form should follow the principles of the Optimal Classification Model based on seven simple requirements. It should be *exhaustive, mutually exclusive, empirical, hierarchical, comprehensive, detailed* and *scalable.* None of the models present in literature fulfils all seven criteria to date. The rest of this research will focus on developing a novel method, which embeds all principles into its design to derive the taxonomy (i.e. quantitative classification) of urban form.

Existing models for classification of urban form present in the literature vary and spread across multiple dimensions, from different basic scales (from city to building) to different methodological prepositions (from remote sensing to typo-morphology). Across the field, all OCM criteria have been met; however, not within a single model.