Quantitative comparison of the distribution of densities in three Swedish cities

Meta Berghauser Pont, Gianna Stavroulaki, Kailun Sun, Ehsan Abshirini, Jesper Olsson, Lars Marcus

Department of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg, Sweden
E-mail: meta.berghauserpont@chalmers.se, gianna.stavroulaki@chalmers.se, kailun@chalmers.se, ehsan.abshirini@abe.kth.se, gusjespol@student.gu.se, lars.marcus@chalmers.se

Abstract. *Typologies play a role in urban studies since a long time, but definitions are often rather abstract, ill-defined and at worst end in fixed stereotypes hiding underlying spatial complexity. Traditional typologies are focussing on separate elements, which allow for understanding crucial differences of one spatial feature in greater detail, but lack the capacity to capture the interrelation between elements. Further, they often focus on one scale level and therefore lack to acknowledge for interscalarity. Recent publications define morphological typologies based on quantitative variables, building on the seminal book ´Urban Space and Structures´ by Martin and March, published in 1972, but using more advanced spatial analysis and statistics. These approaches contribute to the discussion of types in two ways: firstly, they define types in a precise and repeatable manner allowing for city-scale comparisons; secondly, they acknowledge cross-scale dynamics important for e.g. living qualities and economic processes where not only the local conditions are important, but also the qualities in proximity. This paper focuses on the comparison of building types in three Swedish cities, using the multi-variable and multi-scalar density definition. A statistical clustering method is used to classify cases according to their measured similarity across the scales. The results show that working with types is a fruitful way to reveal the individual identity of these types, compare cities and highlight some differences in the way the three cities are structured.*

Keywords: Density, cluster analysis, building types, city comparison, interscalarity.

Introduction

Density

Density is a crucial variable in urban studies for various reasons. It is often argued that cities with higher density have a higher level of ´urbanity´(e.g. Lozano 2007). Further, cities with higher population densities are generally found to have lower CO2-emissions due to lower private car use. This is why high density often is promoted such as in several of the UN reports on sustainable development (e.g. Sustainable Development Goals, UN 2014 and UN Habitat's 5 principles for sustainable neighbourhood planning, 2015).

Berghauser Pont and Haupt (2010) have, driven by these factors above, developed a method to relate density to urban form, and more specifically building types, to make this numeric variable a more relevant tool for architects, urban planners and designers. Their method is developed for the very local scale (i.e. urban fabric or neighbourhood), but we all know that density changes depending on the scale it is measured. Amsterdam and Stockholm have for instance similar population densities at the city level, but Stockholm offers three

times more urban green space per city dweller, because the density of its' neighbourhoods is higher (Berghauser Pont and Haupt, 2010). This offers Stockholm the advantage that 40% of all transport dependent CO_2-emissions is cut thanks to these green spaces (Jansson and Colding 2007). This scalar dimension is missed by most current density measures.

What this paper sets out to do is combining the multi-variable method as developed by Berghauser Pont and Haupt (2010) with the multi-scalar approach as proposed by Serra for street networks (2013a). When the first reveals building types on the local scale, the second provides us with an understanding of cross-scale dynamics (Barbano and Egusquiza 2015). It is widely recognized that the study of the built environment ideally requires both fine and large scale analysis to understand urban processes (ibid.).

The aim is to provide a more precise method of measuring density and in a next step a better understanding how density in all its dimensions influences urban life. The means to do so is to develop a typology that includes the scalar dimension as was successfully implemented for networks by Berghauser Pont et al. (2017). They were able to reveal the structure of the city distinguishing four street types. The same kind of analysis, but now using density variables, will hopefully reveal meaningful types where the local density for instance is high and drops when the scale of analysis increases or types

with a stable density throughout all scales. Such a multi-scalar understanding of density types not only captures local qualities, but simultaneously contextual qualities, important for living qualities, economic and ecological processes where not only the local density is important, but certainly also the densities in proximity.

This paper will develop such a typology of density including cross-scale dynamics that then later will be used to study the relation between these density types and for instance economic activities and housing prices. Before introducing the cases, describing the methodology and discussing the results, we will first discuss the concept of typologies in urban analysis.

Typologies

In urban morphology types are a central concept and often used to describe singular urban elements (e.g. types of streets, urban blocks, parcels, buildings) or aimed for a description of complete systems (e.g. Caniggia and Maffei, 2001; Conzen, 1960, Whitehand, 2001; Panerai et al., 1977, Panerai et al., 1999). Focussing on separate elements allows for the understanding of differences of one spatial feature, but lack the interrelation between the elements. The understanding of the whole system on the other hand requires a reduction of detail and therefore often lacks precision. This results in often ill-defined types, at

Data Sources	Steps	Description
Laser data		Create Digital Height Model (DHM)
Building footprints	1. Edit map	Calculate average building height
Road-Centre Line map		Translate to amount of floors
		Calculate accessible gross floor area (GFA) and footprint
		Calculate density:
	2. Spatial analysis	1. Floor Space Index
		2. Ground Space Index
		(radius 500m - 5000m)
		Standardize results
	3. Statistical analysis	PCA analysis for GSI and FSI
		Cluster analysis using results PCA

Table 1.
Summary of the methodology.

worst stereotypes, hiding underlying spatial complexity (Serra 2013a).

Further, classic studies in urban morphology focused mainly on qualitative methods. Only recently, studies of urban morphology have (again) been aiming for classifications or typologies based on quantitative description of spatial elements following the pioneers of this kind of research, Martin and March (1972). They elaborated that a quantitative approach towards urban form and structure is important when the aim is to understand entire spatial systems and specifically the relation between its elements, and an important step towards a comparative assessment of urban environments.

Recently we have seen a revival of this quantitative approach (e.g. Berghauser Pont and Haupt, 2010; Barthelemy, 2015 and Serra, 2013a; 2013b; Serra et al., 2016). In this paper we will draw from these experiences and compare cities using a multi-variable and multi-scalar approach for the development of a typology of density.

Study area

Three Swedish cities are selected for comparison in this study because they, on the one hand, carry certain socio-economic and historical similarities, while on the other hand, vary in size: Stockholm has 935 619 inhabitants[1,] Gothenburg 556 640 and Eskilstuna 103 684 inhabitants. The study areas aim to include the whole urbanized part of the cities which not necessarily coincide with the municipal borders. For this reason, we used the Urban Morphological Zone (UMZ) boundaries, as they are defined by the European Environment Agency (EEA) and Eurostat[2]. Because of the highly irregular boundaries of the UMZs, a convex hull[3] of each UMZ was made and this

area is used for the study. Finally, to address the possible 'boundary effect', the area which was analysed was 5km larger than the study area.

Methodology

The central methodology used in this paper can be divided into methods for editing the main datasets, spatial analysis and statistical analysis which will be discussed in the section below. The sequence of the methods and how they relate to each is summarized in Table 1.

Density model

The density model is based on laser dataset[4], including coordination and elevation values for each point collected from LIDAR (Light Detection and Ranging). A Digital elevation model (DEM) and Digital surface model (DSM) was extracted from the laser dataset with an average resolution of 2m. Then, DEM was subtracted from DSM to make a new surface model called Digital height model (DHM) which contains the real height values of the features on the ground. Finally, building footprints were added and the average height value of the area covered by each footprint was considered as the height of each building[5]. The building heights were then translated into average amount of floors per building using the average floor height for different functions as given in table 2. These derived floor heights were compared with real count of floors for a randomly selected set of buildings in each city to see whether the deviations are acceptable. On average the error was between 10-15% except for the industrial buildings with a larger error of 23% which, expressed as deviation in amount of floors, is not more than 0,4.

Function types	Storey height in metres		Error rate				Amount of floors	
			Stockholm median	Gothenburg median	Eskilstuna median	*average error*	*real average (RA)*	*deviation from RA*
Residential	3,5	Villa	7%	9%	14%	*10%*	*1,8*	*0,2*
		Block	8%	9%	13%	*10%*	*4,9*	*0,5*
		Modernist	9%	20%	9%	*13%*	*4,3*	*0,6*
Industrial	6,0		23%	28%	19%	*23%*	*1,8*	*0,4*
Commerce and offices	4,5		15%	14%	14%	*14%*	*3,9*	*0,5*
Public	5,0		10%	18%	15%	*14%*	*3,1*	*0,4*

Table 2.
Method to translate building heights to amount of floors and summary of error and deviation.
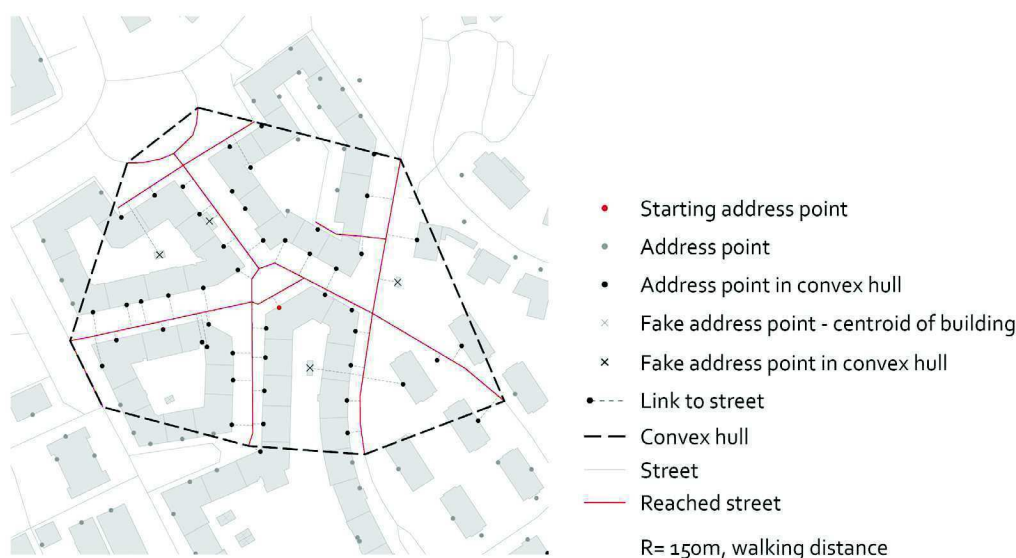
Figure 1.
Method of measuring accessible density including the definition of the area used as the denominator B in the density fraction A/B (i.e. convex hull).

The average amount of floors of the buildings was multiplied with the building footprints to calculate the gross floor area (GFA). The accessible GFA is then calculated by adding up all GFA within the set radius which in the next step is used as the nominator A in the simple fraction of density A/B using what is called the cumulative opportunity measure[6] (Bhat 2001) to calculate accessible FSI. For accessible GSI the area of the building footprints is used as nominator. To make sure only GFA respective building footprints are included that really can be reached via the network, the amounts are loaded on address points which then are used to do the calculations (see figure 1). Secondly, the denominator B is arrived at by calculating the area reached through the network within the set radius. The area is defined by the convex hull connecting all end points of the street segments that are reached (figure 1)[7]. The distance thresholds used are 500m, 1000m, up to 5km in steps of 500m.

Distance, as described above, is measured using non-motorized network models. Non-motorized models include all streets and paths that are accessible for people walking or cycling, including those that are shared with vehicles. All streets where walking or cycling is forbidden, such as motorways, highways, or high-speed tunnels, are not included in the model[8]. We processed the original Road-Centre-line[9] maps with two objectives: first to create line segment maps which we could use in analysis and second, to create comparable representations of the street network in all cities, both in the types of roads included as in the level of detail[10].

The software used for editing and analysis are FME, ArcGIS, Mapinfo and the plugin software PST[11].

Statistical analysis

To be able to describe and compare the cities based on multi-dimensional density results, we performed two main statistical analyses using R software (Gil and Serra, 2017): Principal Component Analysis (PCA) as a means of reducing the ten scales to the most representative scales; and k-medoids clustering to classify the addresses based on their individual density profile, obtaining a density typology. The PCA is necessary to ensure that the variables used for clustering have the lowest possible correlation, which is desirable because co-linearity is known to bias clustering results.

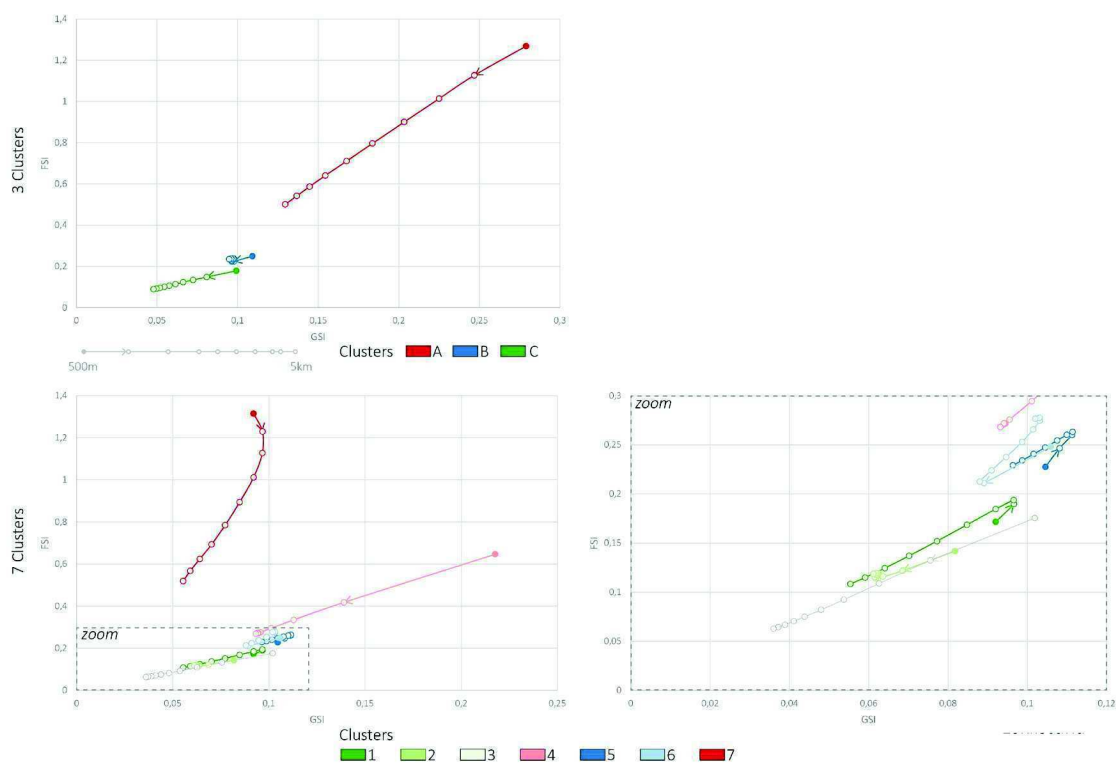For the statistical analysis of the results we

Figure 2.
Scatter plots "Spacematrix" with on the y-axis accessible FSI and on the x-axis accessible GSI showing the
results of the clustering analysis (mean values for every scale and cluster).

only consider the address points contained in the study area and ignore all others to avoid using values affected by edge effect. As a first step, FSI and GSI are converted to z-scores to be able to compare and combine them irrespective of the value ranges. Next we run PCA on the density results for all cities following the method proposed by Serra (2013a; 2013b) and developed further by Berghauser Pont et al. (2017). The amount of principal components is chosen using a scree plot where the aim is to find the least amount of components that still explain most of the variance of the original variables. These components are then input to the next step of unsupervised classification with clustering. We use the k-medoids clustering algorithm (also known as Partitioning Around Medoids, PAM) similar to k-means. The k indicates that the algorithm can identify any number of clusters which must be defined a priori. A range of k values is tested and a silhouette analysis is used to study the separation distance between the resulting clusters where values near +1 indicate that the sample is far away from the neighboring

clusters while a value of 0 indicates that the clusters are highly similar.

The clustering analysis is performed on the data of the three cities combined using the results of the PCA[12]. Once the cluster solutions are calculated we create a series of boxplots of the density profile of every cluster, plotting along the y-axis the cluster's FSI values and along the x-axis the GSI values. To make these boxplots more readable, only the mean values for each cluster and radius are presented.

Results: multi-scalar density types

The first step of the statistical analysis (PCA) reduced the 10 radii of analysis into three rotated components that meet the two scree plot test conditions and explain 98% of the total variance of the original variables. The scales these principal components represent are the neighbourhood (i.e. radius 500m), the larger district (i.e. radii 1km-2,5km) and the city scale (i.e. radii 3km-5km).

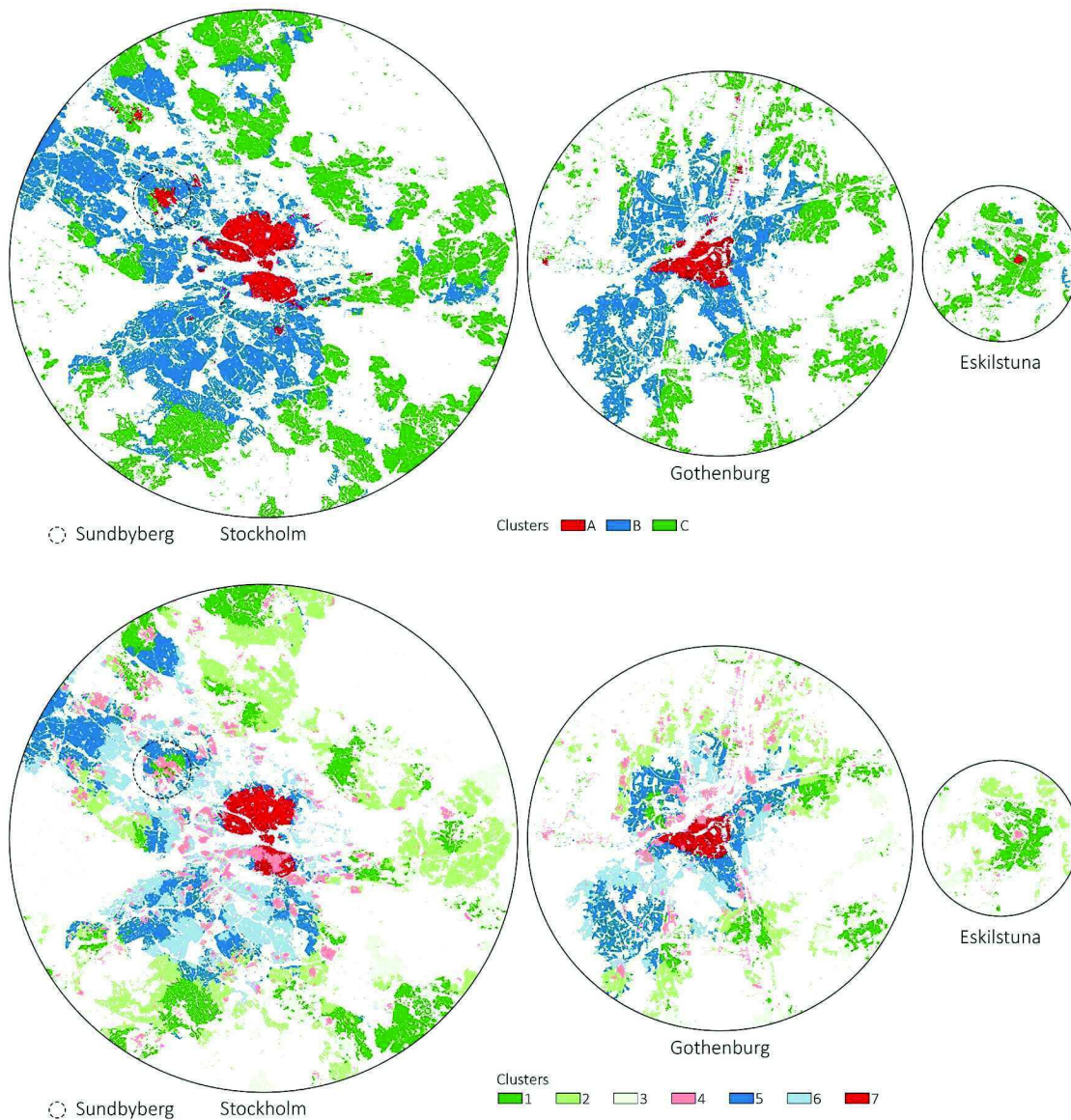These principal components are then used to

Figure 3.
Spatial distribution of three and seven clusters in Stockholm, Gothenburg and Eskilstuna.

identify multi-scalar density types in the three cities. The amount of clusters (k) was chosen by the identification of an elbow in the silhouette plot where both three and seven clusters were identified as relevant. We tested both and will discuss them in turn below.

Three clusters

The three clusters show meaningful differences, both in the scatter plot (Figure 2) as in the spatial distribution (see maps in figure 3). Cluster A captures the inner city and could be referred to as the first ring of density, cluster B captures the suburban ring surrounding the inner city (i.e. second ring of density) and cluster C is found in the urban areas at the edges of the cities (i.e. third ring of density). Their profiles as shown in the boxplots in Figure 2 show that cluster A starts with a high density (both FSI and GSI) which drops gradually when the radius of analysis increases. Cluster B has a consistently medium density while cluster C starts with the lowest density that drops gradually to even lower levels when increasing the distance from 500m to 5km.

The three rings of density can be clearly identified in Gothenburg and Stockholm,

where the rings follow a concentric structure with the highest density in the city cores. Further, Stockholm has some sub-centers in the periphery which are not found so clearly in Gothenburg. One of them is Sundbyberg municipality which today is part of the Stockholm urban area. The municipality prefers to call itself a city, which, however, has no legal significance. But as we can see in the density analysis, the area is of similar type as the core of Stockholm and spatially we can indeed speak of Sundbyberg as city because being part of the first ring of density.

In Eskilstuna, the smallest city in our analysis, the second ring of density (cluster B) is absent. Eskilstuna has a dense center, though very small, surrounded by low dense areas that decrease gradually. In other words, Eskilstuna jumps from the first to the third ring of density. This is probably related to the size of the city where one reaches the countryside.

Seven clusters

When using the seven clusters instead, we see that we get a similar spatial division in three so called rings of density where the first ring (cluster A) is sub-divided in two separate clusters: the "urban core" (cluster 7) with a high FSI dropping gradually without decreasing the level of GSI so much and the "sub-center" (cluster 4) with high GSI values, gradually decreasing when the scale of analysis increases (see figure 2 and 3). Further, some sub-centers (cluster 4) emerge in the other rings of density (cluster B and C) and now actually differentiating Stockholm's city center and Sundbyberg where the first is part of cluster 7 and the latter part of cluster 4.

Within the second ring of density (cluster B), we can identify two sub-clusters (clusters 5 and 6). Both have a relative high density (FSI and GSI) and do not change so much when increasing the distance, but there is also an interesting difference in their density profiles. The density in cluster 5 increases first (until radius 1km) to start decreasing at the higher radii. Cluster 6 does exactly the reverse and decreases first (until radius 1km) to increase in density at the higher radii. In other words, cluster 5 represents a denser locality within the

second ring of density while cluster 6 represents the periphery reaching higher densities only at a larger distance.

The earlier discussed cluster C (i.e. third ring of density) is divided into three sub-clusters where cluster 2 has the most stable density throughout all scales, representing homogenous villa areas. Cluster 1 and 3 start with similar density values where cluster 1 first increases slightly in density (until radius 1,5km), to then slowly start decreasing and cluster 3 steadily decreases to the lowest density of all clusters at the highest radius (i.e. 5km). Cluster 1 is thus the most "urban" of the third ring of density while cluster 3 represents the peri-urban areas.

Conclusion and discussion

We have shown that working with types is a fruitful way to compare cities. They summarise the complexity of the two density variables measured at various scales and reveal the individual identity of these types. Further, it highlights some differences in the way the three cities are structured at these different scales. Stockholm and Gothenburg show great similarity although Stockholm has more sub-centers in the second and third ring. It is worth mentioning that in Gothenburg a cluster of the third ring of density (cluster 1) is located very close to the urban core (cluster 7) which captures the fragmented city structure very well. The most prominent difference when comparing the three cities, is between the two larger cities and Eskilstuna. The densest cluster ("city core", cluster 7) is not present and the center of Eskilstuna ends up in the cluster "sub-center" (cluster 4). This discrete identification of the density profiles including cross-scale dynamics makes the presented typology an important complement to more traditional density analysis.

A comparison of the local density types following Berghauser Pont and Haupt (2010) with the here proposed multi-scalar density types shows that the types to some extend overlap. This is not surprising because the local density, measured at radius 500m, is also included in the multi-scalar typology discussed in this paper. Some information is
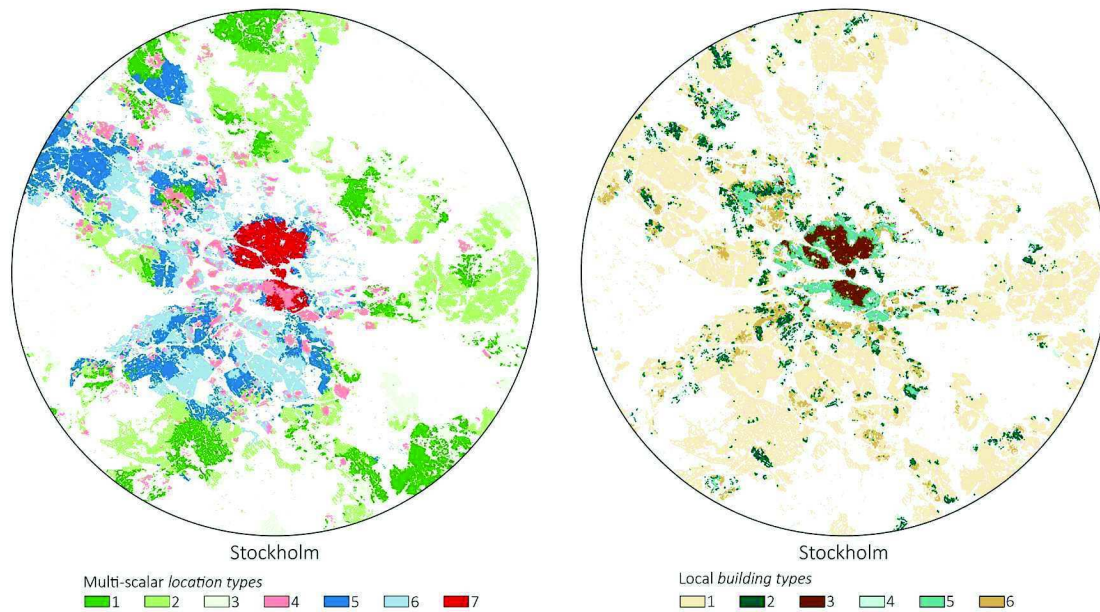
Figure 4.
Comparison of the spatial distribution of multi-scalar location types and local building types in Stockholm
(clusters building types: 1 sprawl, 2 suburban, 3 city centre, 4 compact low, 5 compact city, 6 modernistic).

lost using the multi-scalar approach as can be see in figure 4, but the multi-scalar typology also captures more subtle variations, which appeared to be uniform in the local scale. For example the local density type 1 ("sprawl") is now greatly differentiated, depending on its performance throughout scales, or in other words, because of the local types of another kind in the surroundings. We can say that the local density analysis describes characteristics such as building types and the multi-scalar approach presented in this paper describes location types, including what building types are found next door so to speak.

Besides the clustering in types, it is worth mentioning how in two of the three cities, the three basic density rings were found and that, underneath the visible morphologic variability, their spatial structures are characterized by only three basic spatio-functional scopes (i.e. the three density rings). Further, the main aspect of note is the absence of one of these scopes in Eskilstuna which might be something typical for this size of cities. We should study a larger set of cities to be able to draw more irrevocable conclusion.

A lot of interesting new questions were discussed while working at this paper and studying the results of which we want to mention two.

Firstly, when seeing the interesting grouping of cities, we are eager to add more cities. One trajectory is to add more Swedish cities, to see whether we can speak of typical patterns for different sizes of cities. Another trajectory is to add cities from other European cities and other continents to study differences and similarities between them.

Secondly, we can add the other core variables of spatial urban form that will make the trilogy of the main elements of urban form, that is the street, the plot and the building, complete. A Path typology using network centrality measures was recently developed by Berghauser Pont et al. (2017) and the patterns of plots is under development (Bobkova et al. 2017). In relation to this, it would be interesting to develop a typology based on the pairing of these density types, paths and plot patterns using cross tabulation. The same density clusters can hold very different paths and, depending on this grouping, might perform very differently in terms of social and economic outcome. Correlating these ´combined types´ with pedestrian and vehicle flows and economic activities is a natural and extremely interesting next step.

Notes

1   http://www.scb.se/hitta-statistik/statistik-efter-amne/befolkning/befolkningens-sammansattning/befolkningsstatistik/pong/tabell-och-diagram/kvartals--och-halvarsstatistik--kommun-lan-och-riket/kvartal-4-2016// (accessed 16-05-2017).

2 Urban morphological zones (UMZ) are defined by Corine land cover classes considered to contribute to the urban tissue and function. A UMZ can be defined as "a set of urban areas laying less than 200m apart" (source: http://www.eea.europa.eu/data-and-maps/data/urban-morphological-zones-2006 (download date 13-7-2016)

3 The convex hull of a set of points, lines or polygons is a closed region (i.e. polygon) which includes all the points, lines or polygons on its interior.

4   Source: Lantmäteriet (https://www.lantmateriet.se/)

5 Buildings with no or incorrect heights (too high, too low, zero or negative values) were corrected using Google Street View or similar online services. In cases where it was impossible to find the building height using above-mentioned methods, we used a buffer around each building separately and considered the average height value of the surrounding buildings as the height of the building(s) in question. The buffer distance chosen varied according to the average density of buildings in that area.

6 This is also known as the contour measure or isochrone measure and in the software used for this paper (PST), it is called Attraction Reach.

7 At the smallest radius (500m), the reached area can be smaller than the reached built up areas due to the layout of the streets (long streets without crossings). The values for GSI > 1 are taken out of the further analysis as they would disturb the statistical analysis.

8 All streets or paths are represented with one line irrespectively of the number of lanes or type, meaning that parallel lines representing a street and a pedestrian or a cycle path running on the side, are reduced to one line. The reason is that these parallel lines are nor physically or perceptually separated, and thus are accessible and recognized from pedestrians as one "line of movement" in the street network. If there are obstacles or great distance between parallel streets and paths, then the multiple lines remain. The aim is to make a skeletal network that better represents the total space which is accessible for pedestrians to move, irrespectively of the typical separations or distinctions of streets and paths.

9 The Road-Centre-line maps used to make the non-motorized network for Stockholm and Eskilstuna originated from the NVDB (Nationell Vägdatabas) and were downloaded from Trafikverket, the official road authority for Sweden (https://lastkajen.trafikverket.se, date of download 15-5-2016, last update 8-11-2015). For Gothenburg, the respective Road-Centre-line map originated from Open Street Maps (openstreetmap.org, http://download.geofabrik.de, date of download 29-4-2016), because the NVDB did not provide enough detail for the non-motorized network, as in the other cities.

10 We followed the same editing and generalizing procedure for all maps aiming to remove errors and to increase comparability between networks. This process included removing duplicate and isolated lines, snapping and generalizing. The snapping threshold used was 2m (end points closer than 2m were snapped together). The generalizing threshold used was 1m (successive line segments with angular deviation less than 1m were merged into one).

11 PST is a plugin tool developed at KTH and Chalmers, available via www.smog.chalmers.se/pst

12 Due to the size of the data set, we used a faster version of the PAM algorithm, called CLARA (Kaufman and Rousseeuw, 1990), which runs on samples of the entire data set.

References

Barbano, G. and Egusquiza, A. (2015). ´Interconnection between scales for friendly and affordable sustainable urban districts

retrofitting´. 6th International Building Physics Conference, IBPC 2015

Barthelemy, M. (2015). ´From paths to blocks: New measures for street patterns´. Environment and Planning B: Planning and Design. doi: 10.1177/0265813515599982

Berghauser Pont, M., Stavroulaki, G., Gil, J., Marcus, L., Serra, M., Hausleitner, B., Olsson, J., Abshirini, E., Dhanani, A. (2017). ´Quantitative comparison of cities: Distribution of street and building types based on density and centrality measures ´. Proceedings XI space syntax conference, Lisbon.

Berghauser Pont, M. and Haupt, P. (2010), Spacematrix. Space, density and urban form (Rotterdam: NAi Publishers).

Bhat, C.; Handy, S.; Kockelman, K.; Hahmassani, H.; Chen, Q.L.; Weston, L. (2000). Urban accessibility index: literature review. Centre for Transportation Research. Research Project. University of Texas at Austin. Project 7-4938.

Bobkova, J., L. Marcus, M. Berghauser Pont, (2017). "Spatial Capacity: Introducing multivariable measures of urban form that influences urban diversity", The 11th Space Syntax Symposium (SSS11), July 2017 in Lisbon.

Caniggia, G. and Maffei, G. L. (2001). ´Architectural Composition and Building Typology: Interpreting Basic Building´. In: Marzot, N., Kropf, K., Scheer, B. and Fraser, S.J., ed., 1st ed. Englsih. Firenze: Alinea Editrice.

Conzen, M.R.G. (1960). Alnwick, Northumberland: A Study in Town-Plan Analysis. London: Philip (in series: The Institute of British Geographers, publication 27).

Gil, J. and Serra, M. (2017). ´Street-network-typologies´. https://doi.org/10.6084/m9.figshare.5044462.v1 (figshare).

Jansson, Å. and Colding, J. (2007). ´Tradeoffs between Environmental Goals and Urban Development: The Case of Nitrogen Load from the Stockholm County to the Baltic Sea´, Ambio, Vol. 36, No. 8.

Lozano, E. (2007). 'Density in Communities, or the Most Important Factor in Building Urbanity', in: M. Larice and E. Macdonald (eds.), The Urban Design Reader (Oxon: Routledge), 312-327

Marcus. L. and Berghauser Pont, M. (2017). ´Cities as accessible densities and diversities – adding attraction variables to configurational analysis´. Proceedings XI space syntax conference, Lisbon.

Martin, L. and March L. (1972). Urban Space and Structures. Cambridge: Cambridge University Press.

Panerai, P., Castex, J., and Depaule, J. C. (1977). Formes urbaines, de l'Iˆlot a` la barre. Paris: Dunod.

Panerai, P., Demorgon, M., &nd Depaule, J. C. (1999). Analyse urbaine. Marseille: Edition Parentheses.

Serra, M., Gil, J. and Pinho, P. (2016). ´Towards an understanding of morphogenesis in metropolitan street-networks´. Environment and Planning B: Planning and Design. doi: 10.1177/0265813516684136

Serra, M. (2013a). Anatomy of an Emerging Metropolitan Territory - Towards an Integrated Analytical Framework for Metropolitan Morphology. PhD thesis, Faculdade de Engenharia da Universidade do Porto.

Serra, M. (2013b). ´Tackling the structure of very large spatial systems: space syntax and the analysis of metropolitan form´. The Journal of Space Syntax, 4(2), pp. 179-196.

UN Habitat (2015), https://unhabitat.org/a-new-strategy-of-sustainable-neighbourhood-planning-five-principles/

UN (2014), http://www.un.org/sustainabledevelopment/sustainable-development-goals/

Whitehand, J. W. R. (2012). ´Issues in urban morphology´, Urban Morphology, 16(1), 55-65.