Conference Proceedings – XXIX International Seminar on Urban Form
ISUF 2022: Urban Redevelopment and Revitalisation. A Multidisciplinary Perspective
6th June–11th September 2022, Łódź–Kraków

Dr. Alessandro Venerandi[(1)], Dr. Vahid Zamani[(2)], Prof. Sergio Porta[(1)]

[(1)] Urban Design Studies Unit (UDSU), Department of Architecture, Faculty of Engineering, University of Strathclyde, Glasgow (UK)
[(2)] Faculty of Architecture & Urban Planning, Art University of Isfahan, Isfahan (Iran)

10.34658/9788367934039.133

On the relationship between urban form and amenities: A new perspective from Qom (Iran)

Abstract Amenities are fundamental for urban life as they promote socio-economic interactions and enhance city dynamics. Previous studies investigated the relationship between metrics of street network centrality and urban amenities. However, they hardly focused on further aspects of the built environment. A further drawback is that relationships were mainly assessed through linear models even though more complex and non-linear relationships plausibly exist. In this work, we, first, comprehensively describe the urban form of our case study, the city of Qom (Iran), through a set of 55 morphometrics computed at the plot level; second, we investigate the relationship between these metrics and density of amenities, through a set of machine learning techniques that handle non-linear behaviours. The best model explains up to 45% of the variance of the density measure, with coverage ratio, plot size, floor area ratio, street canyon width, and betweenness centrality being the top five explanatory factors. While the findings of this work do not have universal value, the methodology can be replicated to explore the same research question in different contexts. It can also be used as an evidence-based tool to inform design choices in urban redevelopment affecting the location of amenities in cities.

Keywords Urban morphology, Urban morphometrics, Machine learning, Amenities, Qom (Iran).

Introduction

Urban amenities including commerce and services are fundamental for our cities as they not only provide goods and services to residents, but also contribute to a set of tangible and intangible aspects, such as safety, street liveliness, community identity, prosperity and liveability of the urban environment. In this respect, several theories were proposed in the recent past. Jacobs (1961) and Gehl (1987) argue that the presence of amenities in streets positively contribute to neighbourhood attractiveness by promoting socio-economic interactions between urbanites and, more broadly, enhancing city dynamics. Jacobs (1961) also suggests that transparent surfaces, including shop windows, close to the streets, are of paramount importance for ensuring ‘eyes on the street’, hence informal control on the urban space against petty crimes and anti-social behaviours.

More recently, with the advent of geographic information systems and diffusion of spatial data, several studies delved deeper into this topic by quantitatively exploring the relationship between morphological features of cities and density of amenities. Hillier et al. (1993) suggest that the preferential location of shops and services is directly associated with the so-called ‘natural movement’ of pedestrians in the public space, which, in turn, is shown to be linearly correlated to higher levels of ‘spatial integration’, a measure of network centrality that simultaneously assesses proximity and interconnectivity of street segments in a specific area. Porta et al. (2009), Produit et al. (2010) and Wang et al. (2011) rely on similar techniques of network analysis to assess the relationship between density of amenities in Bologna (Italy), Barcelona (Spain) and Baton Rouge (LA, US) respectively and several metrics of street network centrality.

Results in the former show significant correlations (Pearson's $r > 0.7$) especially between betweenness centrality, a measure of through-movement in urban space that is computed at the scale of the entire city, and density of commerce and services measured through a kernel density estimation function at a resolution of 300 meters. Similar results are found in Barcelona, where a composite measure of centrality, including betweenness, closeness and straightness, reflecting simultaneously being ‘intermediary, straight and critical’ in the street network, is found to be positively associated (Pearson's $r > 0.6$) with retail activities.

In Baton Rouge, closeness centrality shows the strongest correlation with land use density, straightness the next and betweenness the last. A more nuanced analysis investigates the relationship between the kernel densities of different shop categories and street network centrality in Changchun (China) (Wang et al. 2014). Specialty stores show the strongest correlation with closeness centrality; supermarkets and department stores with betweenness centrality; consumer product stores (wholesale stores) with straightness centrality.

The outcomes of these works are undoubtedly a step forward in the study of city dynamics and, more specifically, of the relationship between the preferential location of amenities and street centrality. However, as important as it is, the latter is only one aspect of urban form and, while statistically significant correlations were indeed found, a large part of potential correlates are still unknown because they have never been explored. Furthermore, the statistical techniques used for carrying out these studies are linear (correlation and regression). It might well be, though, that relationships do not behave in a linear fashion, given the complexity of cities and the intricacy of underlying and inter-related phenomena. Araldi et al. (2020) investigated the relationship between a more comprehensive set of metrics of urban form and retail distribution in a non-linear fashion. However, the focus of this study was mainly on the comparison between different modelling solutions rather than on understanding how the tested metrics of urban form behaved in relation to retail distribution, in the case study under examination.

In this paper, we use an urban morphometrics approach to generate a comprehensive, multidimensional numerical description of urban form, to then explore its relation with amenities (commerce and services) *momepy* (Fleischmann 2019), a recently developed open-source package for morphometric analysis, is used to comprehensively describe urban form through a set of 55 metrics and analyse their relationship with density of amenities in the city of Qom (Iran). More specifically, all metrics are computed at the plot level and measure a wide variety of spatial attributes, from street canyon width to plot size, from betweenness to straightness centrality. Furthermore, to avoid the assumption of linearity, we use a composite machine learning approach to handle non-linearity and model possible interactions among metrics. Outcomes show that a selection of 38 morphometrics can explain up to 45% of the variance of the density of amenities, with 16 morphometrics having an impact of roughly 75% on the magnitude of the model output. In line with the studies mentioned above, global betweenness and straightness are important explanatory factors. However, plot coverage ratio, plot size, floor area ratio and street canyon width comparatively show stronger impacts.

The work presented in this paper presents a novel, more comprehensive and robust approach to the study of the intricate relationship between urban form and amenities which includes not only metrics of street network centrality, but also tens of others that capture features of the urban fabric and relies on non-linear statistics. It furthermore paves the way to further analyses of this kind in order to ascertain whether similar relationships hold in different contexts. Finally, the identification of the physical characters of the urban environment most strongly associated with amenities can potentially inform urban design and planning strategies aimed at the redevelopment and revitalisation of urban form at different scales (from the neighbourhood to the entire city).

The remainder of this paper is structured as follows. First, we describe the urban morphometrics method and the machine learning techniques utilised to model spatial relations: we compute, on one hand, a comprehensive set of morphometrics, on the other, a metric of density of amenities and, finally, we analyse the relationship between the two in a non-linear fashion. Second, we illustrate the application of this methodology to the city of Qom: we present the case study, the datasets used, the outputs of the composite machine learning approach and the interpretations of such outputs. Finally, we discuss possible implementations of the methodology in an urban redevelopment and revitalisation perspective and conclude with final remarks.

Methodology

This section presents the general methodological framework for analysing the relationship between features of urban form and amenities in any context. It relies on two main steps: (i.) computation of a comprehensive set of morphometrics to describe the urban form of the city as well as a metric of density of amenities; (ii.) modelling the relationship between the two.

Measuring urban form and amenities

As mentioned in the introduction, several studies analyse the relationship between street network and amenities in cities. However, the former is only one aspect of the much larger, nuanced and complex spatial entity we call ‘urban form’. Since reducing the descriptive dimensions of the phenomenon results in potential selection biases, urban morphometric research has recently explored a more comprehensive approach (Fleischmann, Romice, et al. 2020) aimed at operationalizing *all* metrics of urban form (i.e. ‘morphometrics’) that are a) available in literature and b) viable and compliant with the research principles and framework.

The point is to avoid upfront by-theory discrimination of morphometrics. Related to this, an open-source Python package named *momepy* (Fleischmann 2019) has been recently developed to boost the continuous accumulation of shared knowledge about morphometrics. The package makes it possible to measures different aspects of the three main morphological components of cities, i.e., plot, building and street (Moudon 1997), via six main categories of information extracted from the urban design literature: dimension, shape, spatial distribution, intensity, connectivity and diversity (Fleischmann et al. 2021). In dimension, *momepy* includes, among others, building footprint, plot size and length of perimeter walls, while in connectivity proportion of 3-way intersections, local meshedness and all three Multiple Centrality Analysis (MCA) metrics of street centrality, i.e. betweenness, closeness and straightness (Porta et al. 2006). Metrics can be computed for local geographies to capture spill-over effects typical of spatial phenomena (Legendre 1993) and behaviours across different scales.

Momepy usually relies on two input datasets: street network (cleaned of transportation-related geometry, e.g. dual carriageways), and buildings (with heights). However, it can be adapted to sub-optimal quality or availability of data. Venerandi et al. (2021) showed, for example, that it is possible to carry out an accurate morphometric analysis at city level relying on building footprints only. The spatial unit for which all morphometrics are computed is the morphological cell, which is obtained via a Voronoi tessellation-based partitioning of space from building footprints (Fleischmann, Feliciotti, et al. 2020). This system was put in place to overcome theoretical and geometrical inconsistencies in the definition of plots (Kropf 2018). However, it is at the discretion of the researcher whether to use these automatically generated spatial units, proper cadastral parcels or other official boundaries.

Density of amenities measures the amount of commerce and services per areal unit. This can be computed in several ways depending on the granularity of the input data. For example, if amenities are made available as georeferenced points (like in OpenStreetMap)[1], one can compute the ratio between the number of points in a specific unit and the area of such a unit. If one has information on the area of the plot and the total floor area of amenities pertaining to it, one can potentially compute a measure of commercial floor area ratio by dividing the latter by the former.

A machine learning ensemble method

Having obtained a comprehensive set of morphometrics to describe the built environment, the next step requires investigating the relationship between them and the density of amenities. This is a 3-step process inspired by previous research about urban form and house prices in the French Riviera (Venerandi et al. 2019), consisting in: (i.) identifying the most relevant morphometrics to model density of amenities via sequential forward selection; (ii.) using gradient boosting to model the relationship between the selected morphometrics and density of amenities and (iii.) interpreting the behaviours of each morphometric via an additive feature attribution method.

Sequential Forward Selection

While the computation of tens of morphometrics provides a comprehensive description of the urban form under examination, it can also potentially generate issues at the modelling stage. Using a too large number of explanatory variables can increase computational time and result in data redundancy, which, in turn, can produce noise and overfitting. To contravene this, our methodology relies on the use of the Sequential Forward Selection (SFS) procedure (Raschka 2018), a feature selection technique that adds one variable at the time, based on a regressor performance, until an optimal number of variables, within a specified interval, is reached. In the context of this work, the target is a measure of density of amenities and the candidate variables are the morphometrics computed through the *momepy* package.

For statistical consistency, the regressor should be the same as the one used in the next step, i.e. gradient boosting, which will be detailed in the next section. Finally, to obtain a statistically robust selection of morphometrics, train and test sets and *k*-fold cross-validation must be implemented. The former consists in splitting the dataset in a subset used for training the model (usually 80% of the data) and in one for testing it (usually 20% of the data). The latter consists in repeating this process for 10 different train and test partitions.

Gradient Boosting

Once the most relevant morphometrics are selected, the next step consists in modelling the relationship between such morphometrics and density of amenities in the case study under examination. To do so, we propose the use of gradient boosting (Friedman 2001), a non-parametric machine learning technique relying on multiple decision trees for prediction. The interesting aspect of this technique compared to more standard ones (e.g. random forest) is that, rather than averaging the performances of single base models, it progressively fits new decision trees to reduce the error made at previous steps, leading to better predictions.

Since the aim is to model a continuous variable (i.e. density of amenities), the reduction of these errors relies on a gradient descent that optimises a cost function based on a least squares regression pointing to the negative gradient direction. As in SFS, *k*-fold cross-validation and train and test sets must be used to ensure the robustness of model and results.

[1] Wiki, *Key: amenity*, online https://wiki.openstreetmap.org/wiki/Key:amenity (access: 13.11.2023).

To avoid overfitting, the maximum depth of decision trees is set to a third of the number of morphometrics used in the model, i.e. a widely accepted value in the machine learning community (Franklin 2005). Finally, to quantify the explanatory power of the model, an overall adjusted R squared value can be computed by averaging the adjusted R squared values associated with each test set.

Interpretative tool

Since the aim of this work is to understand the relative importance of features of urban form in relation to density of amenities, a further machine learning technique, i.e. SHapley Additive exPlanations (SHAP) (Lundberg, Lee 2017) is used to interpret the results of the gradient boosting algorithm. SHAP is part of a set of statistical techniques (i.e. additive feature attribution) that explains a specific prediction from a model through the selection of interpretable features sampling and then fits a linear model in the local area around this prediction. By collecting these feature importance values at the local level for the entire study area, SHAP can quantify the general behaviours (negative or positive) that the selected morphometrics have in relation to density of amenities.

Application

In this section, we describe how we applied the general methodology presented above to the specific case of the city of Qom. We start by presenting the case study and the two datasets used to carry out the analysis. Second, we illustrate the morphometrics used to comprehensively describe Qom's urban form and the measure of density of amenities extracted from the two input datasets. We then illustrate the application of the machine learning ensemble method to investigate the relationship between morphometrics and density of amenities.

Case study

Qom is a historical city if Iran located approximately 140 km south of Tehran (the capital of the country) (Figure 1). Several phases of urban development took place throughout its long history, creating an overall rich and stratified urban form. The very first city core, called Shahrestan and built during the Sasanian dynasty (224 to 651), was apparently located south-east of the current one, which was established later and developed in the Islamic era (Saeidnia, 1986). From its creation until 1925 (end of the Qajar dynasty), the city of Qom had a slow growth. However, during the Pahlavi dynasty (1925 to1979) and especially after the Islamic Revolution (1979), Qom changed considerably both in terms of geographical and population sizes and urban structure, through the addition, for instance, of several radial axes (e.g., Azar, Chaharmardan, Shah Ebrahim).

After the Islamic Revolution, Qom witnessed a constant influx of migrants from the less advantaged north-western regions of Iran (Markazi, Hamedan, Zanjan) and east Azerbaijan, with most of them settling in an area located north-west of the Qomrood river, a seasonal river cutting the city in south-west north-east direction, increasing city size further mainly through the construction of informal neighbourhoods.

Before the modern age, Qom has been progressively built through the addition of largely self-sufficient neighbourhoods units consisting of several interrelated elements: a central square, a mosque, a bazaar (i.e., commercial structure with several small shops organized on a tight network of covered streets), a water storage system, a local workshop and mainly residential ordinary urban fabric (Tavassoli 2016). This neighbourhood unit was thus a fundamental and constitutive component of Qom both morphologically and, more importantly, socially, providing sense of identity and belonging. Indeed, the historical core of the city consisted of several of these interrelated neighbourhood units.

However, in the Pahlavi dynasty and subsequent period, large infrastructural works, mainly consisting of large Haussmannian urban arteries cutting through the historical fabric, separated and fragmented these recognizable units. More recently, in the last fifty years, this traditional model has been largely substituted by more repetitive and monotonous urban patterns, characterised by multi-storey, dense, almost exclusively residential buildings and gridded street layouts: these largely rely on pre-existing neighbourhoods (including their traditional commercial fabrics), newly planned commercial strips and shopping malls for the provision of goods and services. The layered complexity of the city of Qom constitutes a particularly interesting case study for investigating the relationship between urban form and density of amenities and understand which specific morphometrics – and to what extent – are more strongly associated with the latter.

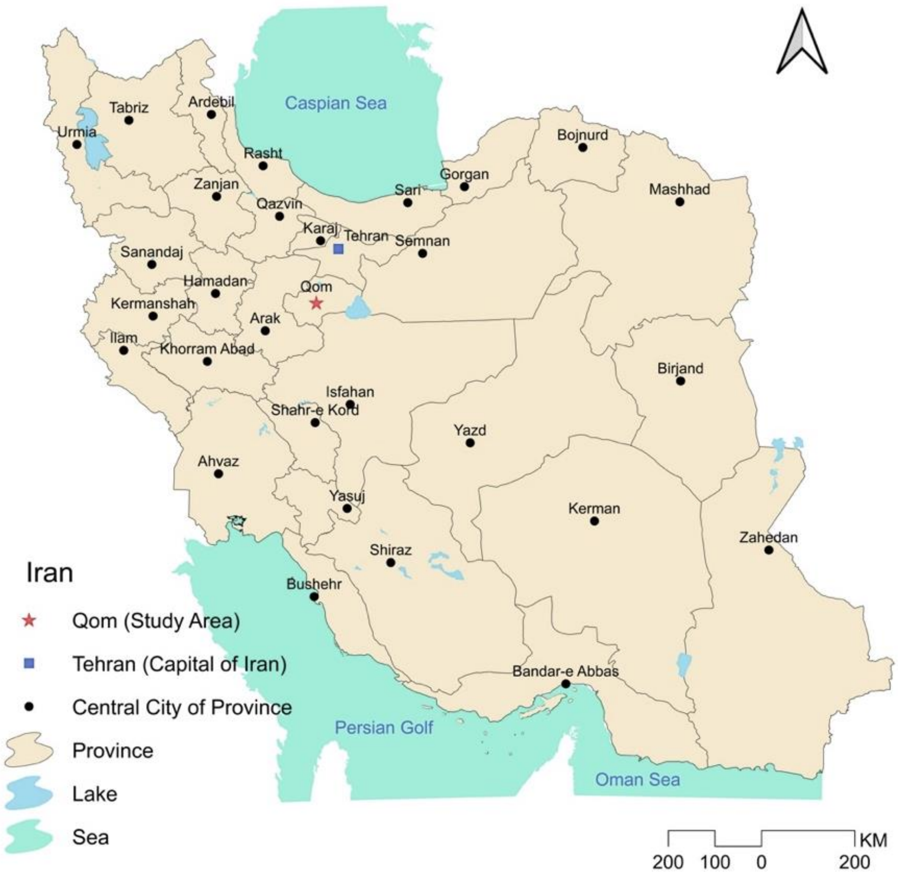

Figure 1. Geographical context of Qom
Source: author's own work.

Datasets

The analysis presented in this paper is built on two vector datasets: a) plots (or cadastral parcels), drawn as polygons, and b) street segments, i.e. the ensemble of lines connecting street intersections, making up the street network of Qom. Both datasets were provided by the municipality of Qom, with the former dating back to 2020 and the latter to 2019. Both covered the entire municipality, however, due to uneven coverage in the peripheral parts of the city (i.e., presence of one of the two datasets but not of both), only plots and streets within the area circumscribed by the main city ring road (i.e. Imam Ali highway) were considered in the analysis.

Apart from polygons, the dataset of plots also contains morphological information on plot layout, e.g. number of floors of the building pertaining to the plot, main land use, total floor area occupied by amenities. Both datasets have been cleaned of topological and geometrical issues, e.g. duplicated/invalid geometries and false intersections (as generated, for example, by erroneously separated street segments where no intersection is actually present). The resulting datasets consist of 35,827 street segments and 227,165 plots.

Computing morphometrics and density of amenities

Since the building footprint layer was not available in Qom, only 55 morphometrics were computed out of the 74 that a typical urban morphometric session would otherwise work with: these are the maximum number of metrics computable in *momepy* from the available input datasets. Note that the dataset of plots also contained information on the areas and heights of the buildings pertaining to each plot. Hence, we were able to add three morphometrics (i.e. building area, height and volume) which would have been otherwise impossible to compute without the building footprint layer. The full list of such morphometrics is provided in Table 1, together with labels, physical element of reference, scale of computation (context) and broad categorisation. For explanations on the morphometrics formulas, we refer the reader to (Fleischmann et al. 2021).

Table 1. The 55 morphometrics used to comprehensively describe the urban form of Qom, together with labels, physical element of reference, scale of computation (context) and broad categorisation

Morphometric	*Label*	*Element*	*Context*	*Category*
area	sdbAre	building	building	dimension
height	sdbHei	building	building	dimension
volume	sdbVol	building	building	dimension
longest axis length	sdcLAL	plot	plot	dimension
area	sdcAre	plot	plot	dimension
circular compactness	sscCCo	plot	plot	shape
equivalent rectangular index	sscERI	plot	plot	shape
solar orientation	stcOri	plot	plot	distribution
street alignment	stcSAl	plot	plot	distribution
coverage area ratio	sicCAR	plot	plot	intensity
floor area ratio	sicFAR	plot	plot	intensity
length	sdsLen	street segment	street segment	dimension
width	sdsSPW	street profile	street segment	dimension
height	sdsSPH	street profile	street segment	dimension
height to width ratio	sdsSPR	street profile	street segment	shape
openness	sdsSPO	street profile	street segment	distribution
width deviation	sdsSWD	street profile	street segment	diversity
height deviation	sdsSPH	street profile	street segment	diversity
linearity	sssLin	street segment	street segment	shape
area covered	sdsAre	street segment	street segment	dimension
plots per meter	sisBpM	street segment	street segment	intensity
area covered	sddAre	street node	street node	dimension
weighted neighbours	mtcWNe	plot	neighbouring plots (queen)	distribution
area covered	mdcAre	neighbouring plots	neighbouring plots (queen)	dimension
reached plots	misRea	neighbouring segments	neighbouring segments	intensity
reached area	mdsAre	neighbouring segments	neighbouring segments	dimension
degree	mtdDeg	street node	neighbouring nodes	distribution
mean distance to neighbouring nodes	mtdMDi	street node	neighbouring nodes	dimension
reached plots	midRea	neighbouring nodes	neighbouring nodes	intensity
reached area	midAre	neighbouring nodes	neighbouring nodes	dimension
gross floor area ratio	licGDe	neighbouring plots	plot queen neighbours 3	intensity
weighted reached blocks	ltcWRB	neighbouring plots	plot queen neighbours 3	intensity
area	ldkAre	block	block	dimension
perimeter	ldkPer	block	block	dimension
circular compactness	lskCCo	block	block	shape
equivalent rectangular index	lskERI	block	block	shape
compactness-weighted axis	lskCWA	block	block	shape
solar orientation	ltkOri	block	block	distribution
weighted neighbours	ltkWNB	block	block	distribution

weighted plots	likWBB	block	block	intensity
betweenness centrality	gloBET	street network	streets within network	connectivity
straightness centrality	gloSTR	street network	streets within network	connectivity
local meshedness	lcdMes	street network	nodes 5 steps	connectivity
mean segment length	ldsMSL	street network	segment 3 steps	dimension
cul-de-sac length	ldsCDL	street network	nodes 3 steps	dimension
reached plots	ldsRea	street network	segment 3 steps	dimension
node density	lddNDe	street network	nodes 5 steps	intensity
reached plots	lddRea	street network	nodes 3 steps	dimension
proportion of cul-de-sacs	linPDE	street network	nodes 5 steps	connectivity
proportion of 3-way intersections	linP3W	street network	nodes 5 steps	connectivity
proportion of 4-way intersections	linP4W	street network	nodes 5 steps	connectivity
weighted node density	linWID	street network	nodes 5 steps	intensity
compactness-weighted axis	lddARe	street network	nodes 3 steps	dimension
local closeness centrality	lcnClo	street network	nodes 5 steps	connectivity
square clustering	xcnSCI	street network	nodes within network	connectivity

Source: author's own work on the basis of Fleischmann et al. 2021.

On the other side, density of amenities in each plot was computed by dividing the total floor area dedicated to commerce and services by the area of the plot (both data are present in the official dataset provided by the municipality of Qom). This metric can be considered the 'commercial' counterpart of floor area ratio. Due to length constraints, presenting maps for all the morphometrics computed in this study is not feasible. In Figure 1, we thus present three examples, i.e. the plot's longest axis length, the floor area ratio and the betweenness centrality and the metric of density of amenities (i.e. commercial floor area ratio). Visual inspection highlights that higher commercial densities tend to be located on main and local thoroughfares, such as Imam Khomeini Street and Shahid Motahhari Street (Zangaraki area), but also to spread locally in specific neighbourhoods, such as in the area around Qom's Old Bazaar.

Selecting the best morphometrics

Having computed the 55 morphometrics and density of amenities, the next step requires identifying which of the former are best to model the latter. The SFS technique is thus implemented on the 227,165 plots for which the metrics were computed by using a gradient boosting regressor and splitting the dataset in train (80% of the observations) and test (20% of observations) sets, in a 10-fold cross validated regime. SFS is set to search for the best combination of metrics in a range between 3 and 40. Since the variable to be modelled is continuous (i.e. density of amenities), negative mean squared error is utilised as scoring system to evaluate the performance at each iteration. SFS selects a total of 38 morphometrics: plot area (sdcAre), plot longest axis length (sdcLAL), plot circular compactness (sccCCo), plot equivalent rectangular index (sscERI), coverage area ratio (sicCAR), floor area ratio (sicFAR), plot's weighted neighbours (mtcWNe), area covered by neighbouring plots (mdcAre), gross floor area ratio (licGDe), weighted reached blocks (ltcWRB), plot's street alignment (stcSAl), building height (sdbHei), block area (ldkAre), block's circular compactness (lskCCo), block's compactness-weighted axis (lskCWA), block orientation (ltkOri), block's weighted plots (likWBB), street length (sdsLen), street profile average width (sdsSPW), street profile average height (sdsSPH), street profile average openness (sdsSPO), street profile average height deviation (sdsSHD), area covered by street segment (sdsAre), plots per meter of street (sisBpM), reached area by neighbouring streets (mdsAre), plots reached by neighbouring streets at three topological steps (ldsRea), node degree (mtdDeg), local meshedness (lcdMes), proportion of 4-way intersections (linP4W), proportion of cul-de-sac (linPDE), local closeness (lcnClo), square clustering (xcnSCl), mean distance to neighbouring nodes (mtdMDi), plots reached by neighbouring nodes at three topological steps (lddRea), sum of plot areas around node (sddAre), reached plots by neighbouring nodes (midRea), betweenness centrality (gloBET), straightness centrality (gloSTR).

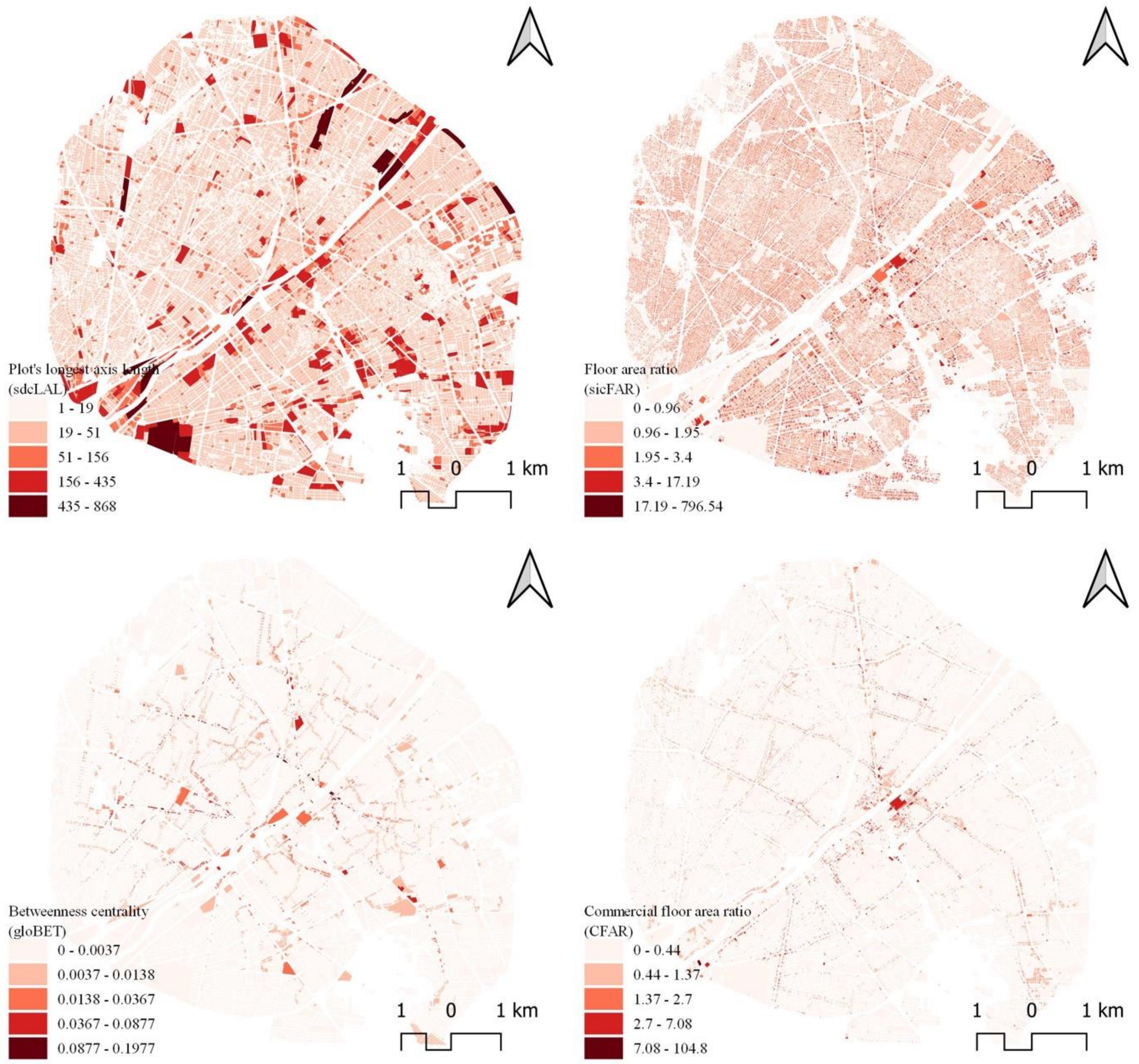

Figure 2. Examples of three morphometrics (plot's longest axis length, floor area ratio, betweenness centrality) and density of amenities (commercial floor area ratio). Gray corresponds to smaller values. Red corresponds to bigger values
Source: author's own work.

Modelling selected morphometrics and density of amenities

The 38 morphometrics selected via the SFS technique are then used to model the commercial floor area ratio via the gradient boosting algorithm. As in the previous step, in order to obtain statistically robust results, 10-fold cross validation is used on consecutive splits of the dataset in train and test sets (80% and 20% of observations respectively). The maximum number of metrics to consider at each split and the maximum depth of decision trees are set to a third of the total number of metrics used in the model (in this case, 13), a value commonly accepted in the machine learning community (Trevor et al. 2009). The total number of decision trees is set to 128 as trial-and-error tests showed this number to be optimal both in terms of performance and to avoid overfit. Finally, the quality of each split in the decision trees is evaluated through mean squared error with improvement score by Friedman. The average adjusted R squared values for the 10 train and test sets are 0.90 and 0.38 (+/– 0.07), respectively, meaning that the best train model can explain 90% of the variance of commercial floor area ratio in Qom, while the best test model can explain 45% of the latter.

Interpreting the model through SHAP

To investigate further this model and the behaviours of the 38 morphometrics across the case study, SHAP is implemented on the outputs obtained at the previous step, by firstly applying the Tree Explainer function on the fitted model and, secondly, by computing SHAP values for each metric. Figure 3 summarises the absolute impacts (measured as percentages) that the 38 morphometrics have on commercial floor area ratio in Qom. The first 16 have alone a 75% impact on the model output magnitude. Figure 4 shows the impacts (positive or negative) of such metrics on commercial floor area ratio across the case study. Vertical dispersions of points mean high concentrations of plots with similar SHAP values. Conversely, slim dispersions correspond to low concentrations of plots with similar SHAP values. Since gradient boosting is a non-linear modelling technique, metrics do not have a uniform behaviour across the case study. However, general trends can still be traced by looking at the dominant colour patterns in Figure 4.

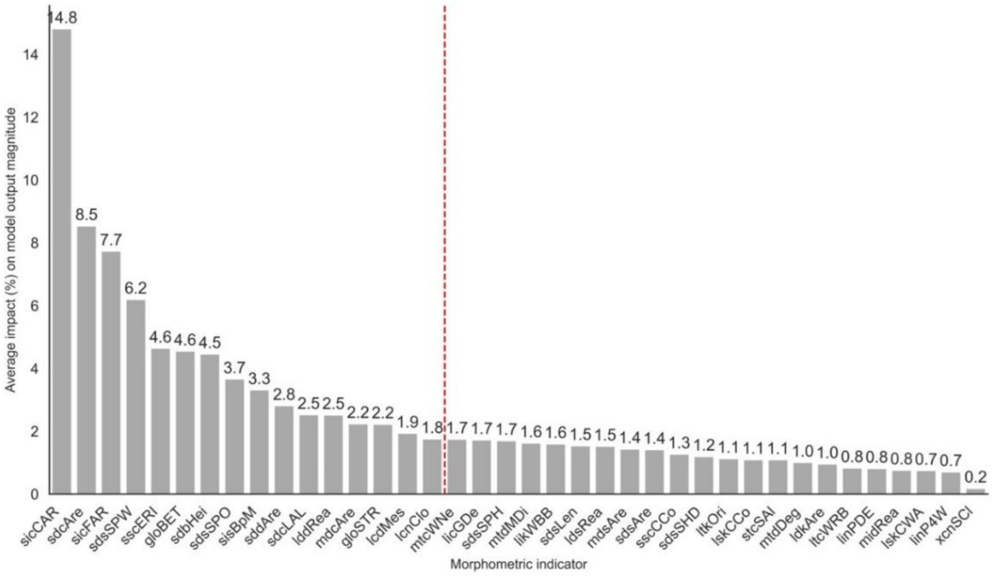

Figure 3. Relative impacts measured as percentages of the 38 morphometric metrics on commercial floor area ratio in Qom. Metrics to the left of the dashed red line have alone a 75% impact on the model output magnitude
Source: author's own work.

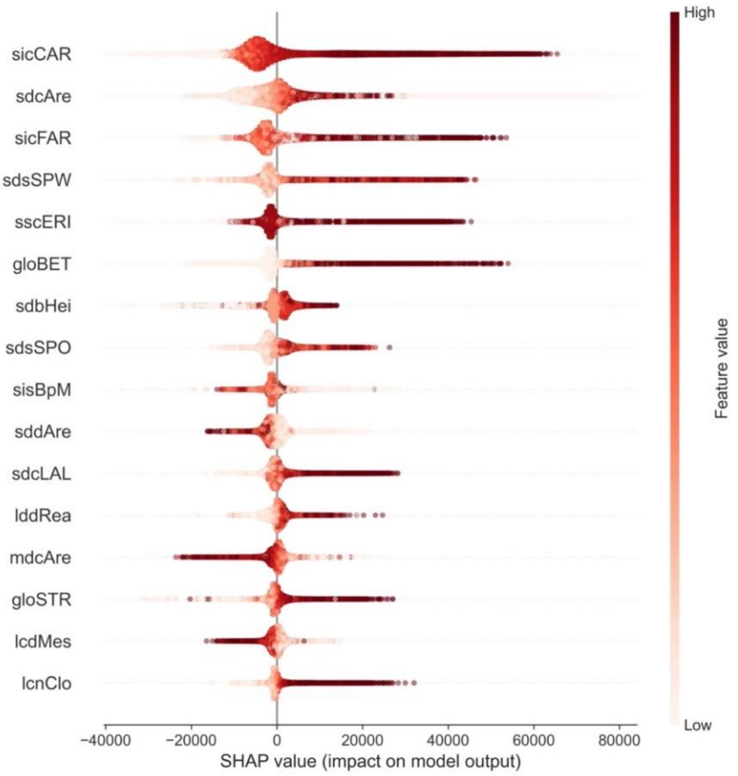

Figure 4. Impacts of the top 16 morphometrics on density (commercial floor area ratio) of amenities in Qom. Values on the x-axis represent the impact (positive or negative) of each morphometric.
The darker the colour, the greater the value of the morphometric
Source: author's own work.

It would be impossible to analyse the relationship between each of the 38 morphometrics and commercial floor area ratio in the space of a single article. For this reason, we focus our interpretation on the top 16 morphometrics which, as previously illustrated, have a 75% impact on the overall magnitude of the model output, thus accounting for a large part of the model's explanatory power. The interpretation is organised around four main topics: *built--up density*, *shape of plots*, *urban grain*, and *street network centrality and configuration*.

Built-up density

The outcomes of the gradient boosting model show that metrics of built-up densities, in particular coverage area ratio (sicCAR), floor area ratio (sicFAR) and building height (sdbHei), are among the strongest predictors of density of amenities (commercial floor area ratio) in Qom, whereby greater values of the three metrics correspond to more square meters of commercial floor area per areal unit. This, to a certain extent, is to be expected since denser urban areas tend to host more commerce and services (de Nadai et al. 2016). However, this relationship should not be taken for granted since many new neighbourhoods in Qom tend to be characterised by coarser and less dense commercial tissues such as supermarkets and shopping malls lying on large plots, often surrounded by parking lots.

Shape of plots

Morphometrics referring to the shape of plots, plot area (sdcAre), equivalent rectangular index (sscERI), and plot's longest axis length (sdcLAL), are also important predictors of commercial floor area ratio. More specifically, smaller plot areas and plots that tend to have more rectangular and elongated shapes are associated with more commercial floor area per areal unit. This finding seems again to recall traditional processes of city building, particularly that of medieval European towns where denser, longer and more elongated plots with their short edges abutting on main commercial streets were resulting emergent elements of the Conzenian adaptive cycle of change, i.e. the 'burgage cycle' (Conzen 1960).

Urban grain

Among the top 16 morphometrics, three of them, i.e. area covered by the plots surrounding a street intersection (sddAre), area covered by neighbouring plots (mdcAre), number of plots reached from street intersections at three topological steps (lddRea), measure the grain of the urban fabric. More specifically, smaller values of the former two and greater values of the latter are associated with more commercial floor area per areal unit, suggesting a relationship between finer grained urban fabrics, characterised by small plots close to one another and higher densities of amenities.

Street network centrality and configuration

Greater values of all the metrics of street network centrality considered in the model, i.e. betweenness centrality (gloBET), straightness centrality (gloSTR), and local closeness (lcnClo), are associated with more commercial floor area per areal unit, meaning that locally well-connected street layouts with a clear hierarchy, rectilinear axes, and more potential though-passage are related to higher densities of amenities. This is in line with previous research focusing on European cities, e.g. Barcelona (Produit et al. 2010) and Bologna (Porta et al. 2009), hinting to the existence of morphogenetic processes that go beyond specific geocultural contexts. In particular, betweenness centrality is the strongest predictor among the three metrics of street network centrality and the sixth in terms of overall importance.

In terms of street network configuration, in Qom we observe lower levels of local meshedness (lcdMes), which express less gridded – i.e. more 'dendritic' (Marshall 2004) – street layouts, in association with more commercial floor area per areal unit. This is in contrast with patterns observed in western cities, where more gridded and therefore better interconnected configurations tend to be related to more commerce and services (Lunecke, Mora 2018). One interpretation of this phenomenon relates to the particular evolution of urban form in Islamic cultural regions. Differently from western cities, dendritic structures typically emerge in inner areas that are *bounded by main streets* – or 'sanctuary areas' (Mehaffy et al 2010), where prevailing cultural values stressing seclusion and privacy of the extended family realm appear to have enhanced processes of progressive street closures and cul-de-sac creation (Remali, Porta 2017).

This is reflected in the low meshedness values detected on plots abutting on the mains streets, since meshedness is computed in a 5-steps context (see Table 1) that spans well over the plots lying deeper in the sanctuary areas. For example, we show the (relatively low) local meshedness and (relatively high) density of amenities in the area around Enqelab (Chaharmardan) Street, in the historical core of Qom (Figure 5).

Indeed, in Qom, not differently than in any western historical area, most amenities concentrate on main thoroughfares, which demonstrates that, as far as amenities preferential location is concerned, what really counts is the higher density and centrality that comes with interconnectedness at the *main street* network level, while little impact can be attributed to the local level of the sanctuary area.

Street canyon width (sdsSPW), openness (sdsSPO) and plots per meter of street (sisBpM) are three further relevant metrics, whereby larger street section, more permeable street edges and less plots per meter of street are associated with more commercial floor area per areal unit. This finding is in line with what we found for the metrics of street network centrality: in particular, higher densities of amenities are predominantly present in symbolically representative urban spaces, i.e. Qom's radial axes, main and secondary throughfares which not only have greater values of betweenness, straightness and closeness but also tend to have larger street sections, more permeable street frontages due to the presence of several small spaces dedicated to green areas or parking lots and comparatively lower density of plots than in more residential streets of Qom.

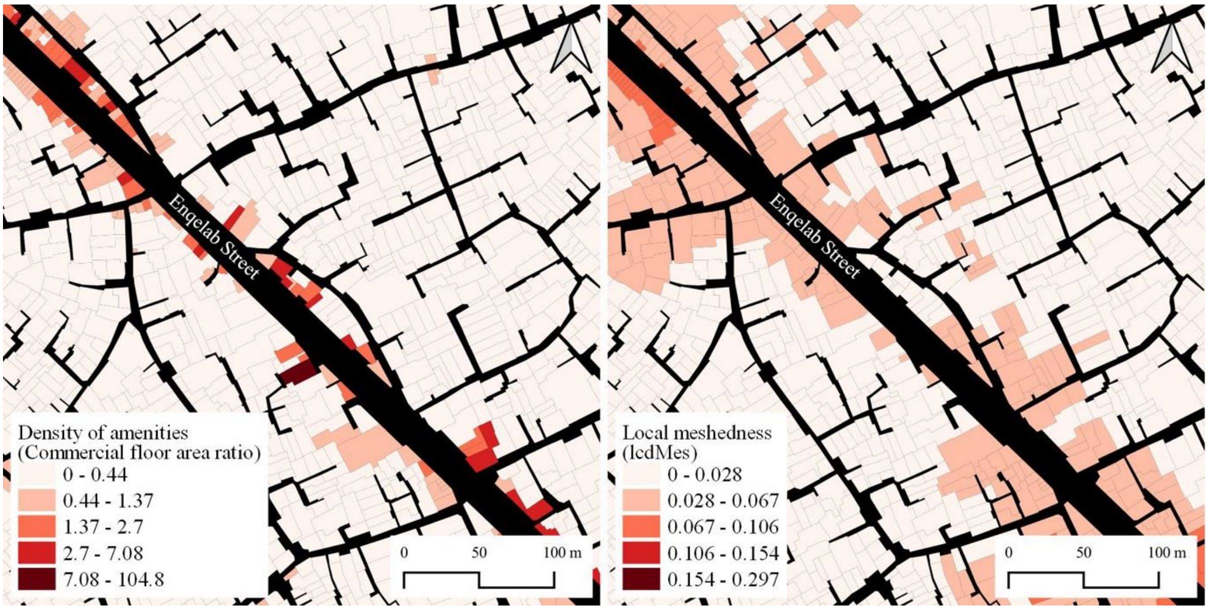

Figure 5. Amenities in Qom (left) tend to concentrate in plots located on main throughfares, which have relatively small values of local meshedness (right) due to dendritic street layouts within sanctuary areas
Source: author's own work.

Discussion

With this analysis, we were able to characterise the typical features of urban form associated with higher densities of amenities in Qom, by considering not only metrics of street network connectivity as in previous studies (Hillier et al. 1993; Porta et al. 2009; Produit et al. 2010; Wang et al. 2011, 2014), but a comprehensive set of descriptors of the urban fabric (morphometrics). In particular, we show that commerce and services in Qom tend to be predominantly related to areas characterized by a mix of high built-up density, fine grained elements and representative urban spaces, such as main and secondary throughfares. These spaces are also characterized by higher street centrality. From an urban design, redevelopment and revitalisation perspective, the machine learning model built in this work can be useful for scenario-based masterplanning. More specifically, different urban revitalisation plans can be tested to model where higher concentrations of amenities will likely be located based on the plot characteristics proposed by-design. The outputs of these predictions are not prescriptive but mapping these outcomes at plot level can be used to inform and guide the design process and also engage/involve local stakeholders and communities through visual material.

The replicability of the methodology proposed in this work makes it possible to i) analyse the relationship between features of urban form and density of amenities virtually anywhere, to understand whether the patterns found in Qom hold or different relationships emerge; ii) in policies of urban redevelopment, the gradient boosting model trained on the ensemble of plots of an entire city can be used for scenario testing to predict the commercial floor area ratio in the project area. The spatial distribution/concentration of urban amenities is not only associated with the morphology of cities, but also with many other factors, including socioeconomic levels, perceived safety at street level, existing regulations. Depending on data availability, future work may thus integrate these further aspects in the model to obtain a more complete picture of what drives higher concentrations of commerce and services in specific city parts.

Conclusions

Urban amenities are of fundamental importance for providing goods and services to city dwellers, but also, contribute to the vibrancy of streets and neighbourhoods. This, in turn, has been shown to have positive impacts on a wide array of socioeconomic and well-being dynamics, affecting for example the prosperity and safety of urban communities. However, as of today, a comprehensive and systematic analysis of the relationship between urban form and density of amenities is missing since most existing studies are either purely qualitative or focus on a reduced set of metrics of urban form, typically street network, and only use linear techniques of spatial analysis. In this work, we proposed a quantitative methodology to i) describe the urban form under examination, through a comprehensive set of 55 metrics of urban form ('morphometrics') at the plot level measuring not only the street network, but also a large array of other urban form elements; ii) model the relationship between these morphometrics and density of amenities (as measured by commercial floor area ratio) in a non-linear fashion, through the use of sequential forward selection, gradient boosting and SHAP.

This methodology has been applied to the Iranian city of Qom. Results showed that more square meters of commercial floor area per plot are associated with urban structures featuring a higher plot coverage and built-up density, a finer grained urban fabric characterised by comparatively smaller and more elongated plots, and better-connected (both globally and locally), predominantly rectilinear main and secondary throughfares. While these findings hold for Qom, the method's replicability paves the way for further analyses exploring the extent to which similar patterns would emerge across different geo-cultural contexts or change on case-by-case basis.

Finally, from an urban redevelopment and revitalisation perspective, a model trained to learn the patterns of urban form and density of amenities of an entire city can help with scenario testing, to predict density of amenities at plot level that different design schemes would *facilitate*. In this approach, in fact, design decisions set the 'environmental' conditions that give a 'selective advantage' to certain desired land-use configurations to emerge and consolidate in time: in fact, an underlining evolutionary approach to urban design, as opposed to the conventional rational-comprehensive of the western modern tradition (Porta et al. 2018; Porta, Romice 2014; Romice et al. 2022).

Acknowledgements

This work was supported by the Axel and Margaret Ax:son Johnson Foundation as part of "The Urban Form Resilience Project' at the University of Strathclyde.

References

Araldi A., Venerandi A., Fusco G., (2020), *Count Regression and Machine Learning Approach for Zero-Inflated Over-Dispersed Count Data. Application to Micro-Retail Distribution and Urban Form*, Computational Science and Its Applications – ICCSA 2020, Springer: Cham, Switzerland, pp. 550–565.

Conzen M., (1960), *Alnwick, Northumberland: A Study in Town-plan Analysis*, Transactions and Papers, London: George Philip.

de Nadai M., Staiano J., Larcher R. et al., (2016), *The death and life of great Italian cities: a mobile phone data perspective*, Proceedings of the 25th international conference on world wide web, pp. 413–423.

Fleischmann M., (2019), *momepy: Urban Morphology Measuring Toolkit*, 'Journal of Open Source Software', vol. 43, no. 4, p. 1807.

Fleischmann M., Romice O., Porta S., (2020), *Measuring urban form: Overcoming terminological inconsistencies for a quantitative and comprehensive morphologic analysis of cities*, 'Environment and Planning B: Urban Analytics and City Science', vol. 48, no. 8.

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

Fleischmann M., Feliciotti A., Romice O. et al., (2021), *Methodological foundation of a numerical taxonomy of urban form*, 'Environment and Planning B: Urban Analytics and City Science', vol. 49, no. 4.

Franklin J., (2005), *The elements of statistical learning: data mining, inference and prediction*, 'The Mathematical Intelligencer', vol. 27, no. 2, pp. 83–85.

Friedman J.H., (2001), *Greedy function approximation: A gradient boosting machine*, 'The Annals of Statistics', vol. 29, no. 5.

Gehl J., (1987), *Life between Buildings: Using Public Space*, New York: Van Nostrand Reinhold.

Hillier B., Penn A., Hanson J. et al., (1993), *Natural Movement: Or, Configuration and Attraction in Urban Pedestrian Movement*, 'Environment and Planning B: Planning and Design', vol. 20, no. 1, pp. 29–66.

Jacobs J., (1961), *The Life and Death of Great American Cities*, Random House.

Kropf K., (2018), *Plots, property and behaviour*, 'Urban Morphology', vol. 22, no. 1.

Legendre P., (1993), *Spatial Autocorrelation: Trouble or New Paradigm?* 'Ecology' vol. 74, no. 6.

Lundberg S., Lee S.I., (2017), *A Unified Approach to Interpreting Model Predictions* [in:] I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (eds.) *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, Neural Information Processing Systems Foundation, Inc. (NeurIPS).

Lunecke M.G.H., Mora R., (2018), *The layered city: Pedestrian networks in downtown Santiago and their impact on urban vitality*, 'Journal of Urban Design', vol. 23, no. 3, pp. 336–353.

Marshall S., (2004), *Streets and Patterns*, London: Routledge.

Moudon A.V., (1997), *Urban morphology as an emerging interdisciplinary field*, 'Urban Morphology', vol. 1, no. 1, pp. 3-10.

Porta S., Romice O., (2014), *Plot-based urbanism: towards time-consciousness in place-making* [in:] C. Mäckler, W. Sonne (eds.) *Dortmunder Vorträge Zur Stadtbaukunst [Dortmunder Lectures on Civic Art]: New Civic Art*, Sulgen, pp. 82–111.

Porta S., Crucitti P., Latora V., (2006), *The network analysis of urban streets: a primal approach*, 'Environment and Planning B: Planning and Design', vol. 33, no. 5, pp. 705–725.

Porta S., Latora V., Wang F. et al., (2009), *Street centrality and densities of retail and services in Bologna, Italy*, 'Environment and Planning B: Planning and Design', vol. 36, no. 3. 450–465.

Porta S., Romice O., Feliciotti A. et al., (2018), *Big box, short life: little box, long life: the democracy of resilience: plot-based urbanism, evolution and informal participation*, London: Create Streets.

Produit T., Lachance N., Strano E. et al., (2010), *A Network Based Kernel Density Estimator Applied to Barcelona Economic Activities* [in:] D. Taniar (ed.) *Computational Science and Its Applications–ICCSA 2010*, Springer Berlin Heidelberg, pp. 32–45.

Raschka S., (2018), *MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack*, 'Journal of Open Source Software', vol. 3, no. 24.

Romice O., Rudlin D., al Waer H. et al., (2022), *Urban design as a specialised, evidence-based, coordinated educational and professional endeavour*, 'Proceedings of the ICE-Urban Design and Planning', vol. 175, no. 4, pp. 179–198.

Saeidnia A., (1986), *The Origin of the City of Qom* [in:] *The Cities of Iran*, Tehran: Jahad Daneshgahi.

Tavassoli M., (2016), *Iranian Cities in Islamic Period, the Middle Ages* [in:] *Urban Structure in Hot Arid Environments*, Springer, pp. 19–58.

Trevor H., Robert T., Jerome F., (2009), *The elements of statistical learning: data mining, inference, and prediction*, Spinger.

Venerandi A., Fusco G., Tettamanzi A. et al., (2019), *A Machine Learning Approach to Study the Relationship between Features of the Urban Environment and Street Value*, 'Urban Science', vol. 3, no. 3.

Venerandi A., Romice O., Chepelianskaia O. et al., (2021), *Urban morphometrics and the intangible uniqueness of tangible heritage. An evidence-based generative design experiment in historical Kochi*, 'Heritage', vol. 4, no. 4, pp. 4399–4420.

Wang F., Antipova A., Porta S., (2011), *Street centrality and land use intensity in Baton Rouge, Louisiana*, 'Journal of Transport Geography', vol. 19, no. 2, pp. 285–293.

Wang F., Chen C., Xiu C. et al., (2014), *Location analysis of retail stores in Changchun, China: A street centrality perspective*, 'Cities', vol. 41, pp. 54–63.