AESTIMUM 79, Dicembre 2021: 75-102

Veronica Alampi Sottini[1], Elena Barbierato[1,*], Irene Capecchi[1], Tommaso Borghini[2], Claudio Saragosa[2]

[1] *Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Italy*
[2] *Department of Architecture (DIDA), University of Florence, Italy*

E-mail: *veronica.alampisottini@unifi.it, elena.barbierato@unifi.it, irene.capecchi@unifi.it, tommaso.borghini@unifi.it, claudio.saragosa@unifi.it*

Keywords: *Urban visual quality, Urban indicators, Geographically Weighted Regression, Random Forest, Google Street View, Flickr, Urban planning*
Parole chiave: *Qualità visiva urbana, Indicatori urbani, Regressione Geograficamente Ponderata, Random Forest, Google Street View, Flick,; Pianificazione urbana*
JEL codes: *O21, R14*

*Corresponding author

Assessing the perception of urban visual quality: an approach integrating big data and geostatistical techniques

Human well-being is affected by the design quality of the city in which they live and walk. This depends primarily on specific physical characteristics and how they are aggregated together. Many studies have highlighted the great potential of photographic data shared on the Flickr platform for analyzing environmental perceptions in landscape and urban planning. Other researchers have used panoramic images from the Google Street View (GSV) web service to extract data on urban quality. However, at the urban level, there are no studies correlating quality perceptions detected by social media platforms with spatial geographic characteristics through geostatistical models. This work proposes the analysis of urban quality in different areas of the Livorno city through a methodological approach based on Geographical Random Forest regression. The result offers important insights into the physical characteristics of a street environment that contribute to the more abstract qualities of urban design.

1. Introduction

It is well known that people live well in environments that they recognize and perceive as pleasant, comfortable, and safe. Human well-being is influenced by the physical characteristics of the surrounding urban space and how they aggregate with each other (Alexander et al., 1977; Lynch, 1960). European cities are generally built in different periods with distinctive architectural styles. The visual quality of the urban space for each era and for each zone of the city is influenced by different variables. In the various zones of the city, the visual quality of urban spaces can be explained by geographical and morphological macro-elements, such as coastlines, waterways or hills. These characteristics influence the visual quality of urban spaces in limited areas and not of the whole urbanized area.

According to Radovic (2003) the physical structure of the city implies "a complex set of built elements, space and environment, units and assemblages, which united and connected in an integrated urban system, create the atmosphere and environment for the complex processing of urban life". Therefore, visual percep-

DOI: 10.36253/aestim-12093

tion, understood as the subjective presentation of objective reality, has always been a complex and highly sensitive issue in the architectural and urban design process. Resources and visual effects play a dominant role in the identification of cultural, socio-economic, identity and communal values of the built environment, as the value and meaning of the built space is manifested predominantly through the subjective view of that space (Perovic & Folic, 2012).

Many studies have focused on researching visual perception using photographic data shared on social platforms (Alampi Sottini et al., 2018; Dunkel, 2015; Quercia et al., 2014; Zhou et al., 2015), others have used indicators to obtain information on urban visual quality using panoramic images from the Google Street View (GSW) web service (Yin and Wang, 2016), but, at the urban level, there are no studies that correlate perceptions of visual quality detected by social media platforms with spatial geographic characteristics through geostatistical models.

Therefore, this paper proposes a geostatistical approach using Geographical Random Forest regression on the Tuscan city of Livorno. this has been analysed city because allows us to assess the visual quality of urban space in very diverse geographical areas. In fact, despite its relatively small size, the city of Livorno consists of a rather heterogeneous mosaic of neighbourhoods with peculiar characteristics due to different construction periods. For this reason, it is an appropriate study area to test a first version of the model to assess the visual quality of urban spaces.

The proposed methodological approach consists of 3 macro-phases: the first one aims at obtaining the indices that compose the urban visual quality perceived by users using photos shared on Flickr; the second one involves calculating the indicators that constitute the urban visual quality using data from both Google Street View, LiDAR data and geographic data; the last one consists in applying two geostatistical models a global random forest and a geographic random forest to differentiate the results for each neighbourhood of the city.

The objective of the study is to test the proposed methodological approach by understanding its strengths and weaknesses and to understand what methodological aspects are needed for the spatial component in the regression models used. The final goal is to provide useful information not only to researchers but also to public and private sectors to develop projects, standards and guidelines to improve the visual quality of urban design in cities.

2. Literature review

Many scholars over time have focused on understanding what visual elements citizens positively perceive as they walk and experience their cities, in order to obtain the indices that make up an image of spaces in which there was a constant observer-environment relationship (Lynch, 1960). The visual perception of architecture and urban planning of cities throughout history has been studied by multiple theorists (Arnheim, 1977; Cullen, 1959; Ittelson, 1960; Lynch, 1960; Rossi, 1966; Spreiregen, 1965; Stea, 1978; Winters, 1999), who not only indicated the im-

portance of the human-environment relationship, but also the importance of creating images in the memory of the users themselves.

According to Ewing and Handy (2009; 2006), the visual quality of an urban design depends on physical characteristics such as the sidewalk width, street width, traffic volume, tree canopy, building height, and number of people. Gavrilidis et al. (2016) selected six urban landscape components for visual assessments and used a five-level Likert scale to visually evaluate each of the six landscape components. Talavera-Garcia and Soria-Lara (2015) developed an alternative walking index called the quality of pedestrian level of service (Q-PLOS), based on the visual quality of urban design for pedestrians, and its relationship with walking needs. However, one major limitation with developing such urban design indicators is that they require unfeasible or inefficient large-scale field observations in terms of time and cost. Recently, Yin and Wang (2016) explored the potential of big data and big data analytics with respect to the current approaches to measuring streetscape features. By applying machine learning algorithms on Google street view (GSV) imagery, the authors objectively generated three measures for visual enclosures.

However, there remains a methodological limit in the literature in regard to identifying an efficient model for relating the perception of well-being deriving from the visual quality of urban space to its physical and architectural characteristics (as measured with the above-mentioned dimensions and indicators).

Lately, with the evolution of Internet, several approaches have been developed for using the so-called big data (Jin et al., 2010) made available by the "social media" platforms.

The correlation between perceived urban landscape quality and the density of photo data shared on the Flickr platform has been demonstrated by numerous studies. Dunkel (2015) highlighted the usefulness of information associated with photos shared on for analyzing environmental perceptions in landscape and urban planning. Zhou et al. (2015) automated the detection of places of interest in multiple cities based on the spatial and temporal characteristics of Flickr images.

Some authors used questionnaires as a complementary survey to the density of data shared on Flickr. Quercia et al. (2014) correlated the emotions stated by the questionnaires with the emotional perception of some paths in London. Similarly, Alampi Sottini et al. (2018) in a research in Livorno used a questionnaire administered via virtual reality headset based on differential semantics techniques and found a correlation between perceived emotion and snapshot point density of shared photographs. Recent studies have used panoramic images taken from Google Street View (GSV) web service to extract urban quality data. In a pioneering work, Naik et al. (2014) correlated the safety perception in the city with indices deriving from Gabor-like filters calculated on a sample of over 1 million Google Streetview. Doersch et al. (2012) used linear Support-Vector Machines to automatically extract representative architectural elements from Google Street View images in Paris and Prague. Steinmetz et al. (2019) used GSV to detect 40, mainly binary, micro-scale features, which influence the walkability of a city, and used them to create a tool to measure the visual quality of urban space. Zhou et

al. (2019) combined street view data with deep learning technologies and mapped residents' perception of walkability in Shen Zhen, China.

Geographic information shared on the internet by different sources (social media platforms, geographic internet services such as google street view, etc.) is an important resource for studying the city's characteristics, which explain the visual perception of urban quality. However, the plethora of existing studies leaves some important gaps. First of all, at the urban level, there are no studies that correlate the visual perception of quality detected by social media platforms with the geographical spatial characteristics through geostatistical models. These studies, widespread in natural and rural environments (for a review see Oteros-Rozas, 2018), are difficult to carry out in cities as a single model cannot explain the importance of the features that define the quality of spaces throughout the urban surface.

While in natural and rural contexts, the statistical model Random Forest (RF) has been used in multiple researches due to its excellent performance (Genuer et al., 2008; Gromping, 2009; Probst et al., 2019; Strobl et al., 2008), in the urban context there are very few studies that manage to apply it, because a single model cannot explain the importance of the characteristics that define the quality of spaces throughout the urban area.

The most interesting application has been adopted in the field of geography with Geographic Weighted Regression (GWR), which is an adaptation of linear regression to be more suitable for spatial data (Fotheringham et al., 2003),

and later with Geographic Random Forest (GRF), applying the same principle to RF (Georganos et al., 2019). GRF is a relatively recent concept used in the environmental field, but it has great potential as it allows selecting the most meaningful indicators for different areas of the city.

For these reasons, we proposed a methodological approach that combines two versions of the random forest (RF) model: traditional for predictive purposes and geographical for exploration purposes.

3. Materials and methods

3.1 Study area

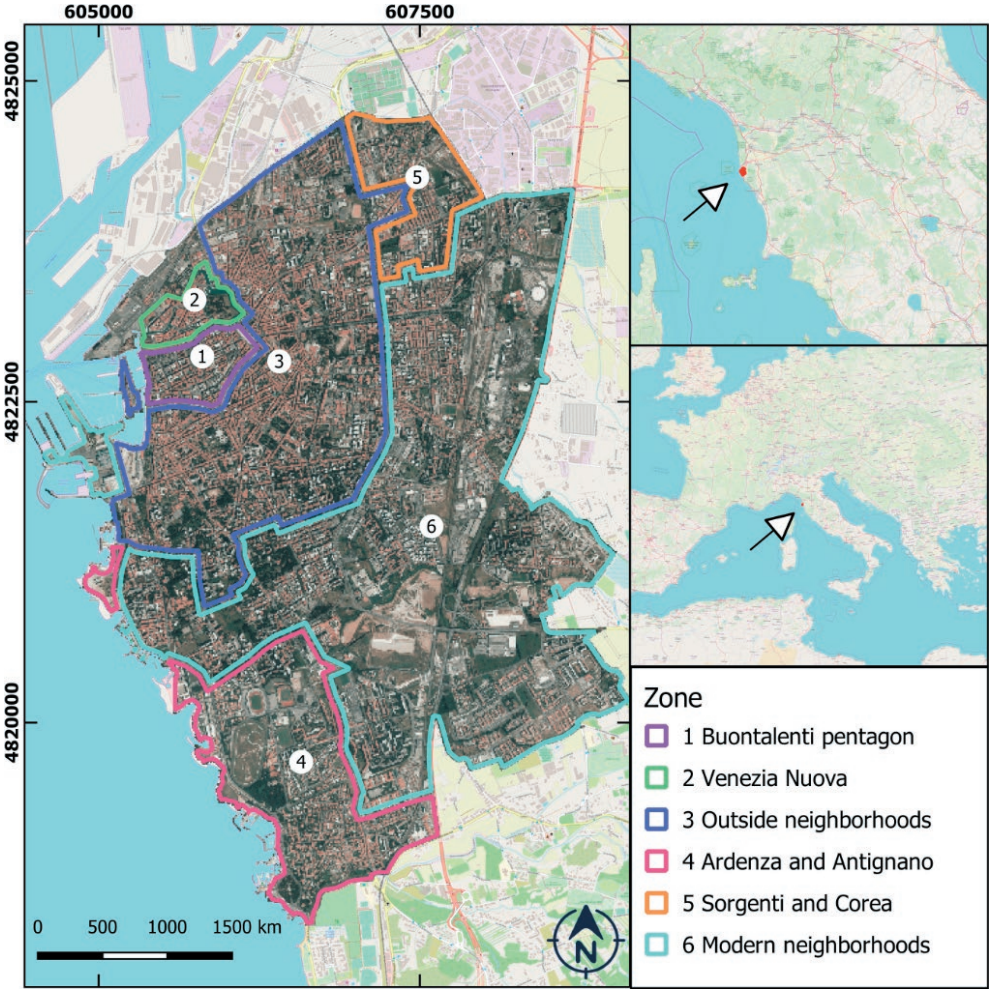
Livorno is a town in the region of Tuscany, in the central Italy. Until the second half of the sixteenth century Livorno was only a small village around a cove. In the 16th century the Medici family contributed in a decisive way to the development of Livorno and its port system with the intention of making it the main seaport of Tuscany. Bernardo Buontalenti was therefore commissioned to design a new fortified city around the original nucleus of Livorno, with an imposing system of moats and bastions that gave the city a pentagonal shape (Figure 1 number 1). A following expansion took place at the beginning of the XVII century with the realization of a new quarter called Venezia Nuova, because of the presence of many canals (Figure 1 number 2). The growth of the city outside the pentagonal walls began in 1700, with the duke Pietro Leopoldo (Figure 1 number 3). In the

nineteenth century Livorno became a destination for seaside tourism and an elegant waterfront was created that transformed the ancient villages of Ardenza and Antignano into neighbourhoods (Figure 1 number 4). In the Fascist period Livorno became a city and an industrial port and this made necessary the construction of new residential quarters for workers called Sorgenti and Corea (Figure 1 number 5). In the middle of 1900 there was the last consistent urban expansion of the city with the creation of modern suburbs (Figure 1 number 6).

We believe that Livorno is an appropriate study area for our methodology for the following reasons.

1) It is relatively small and this has allowed us to reduce computer processing time in the application of Deep Learning based models.

Figure 1. Study area.



2) It presents a heterogeneous mosaic of neighbourhoods built in different eras, thus with peculiar characteristics.
3) Presents macro-spatial characteristics that are found in many other cities, such as waterfront. The perimeter of the study area only includes the urbanized territory of the municipality, as the topic of this study focuses on urban quality. Thus, industrial areas to the north of the city have been excluded, as well as some neighbouring areas that are too far from the city center, and those bordering rural areas.

3.2 Methods

The proposed methodology is synthetically divided into the following four macro phases. Each of them is characterized by specific micro phases detailed in subsequent sections.

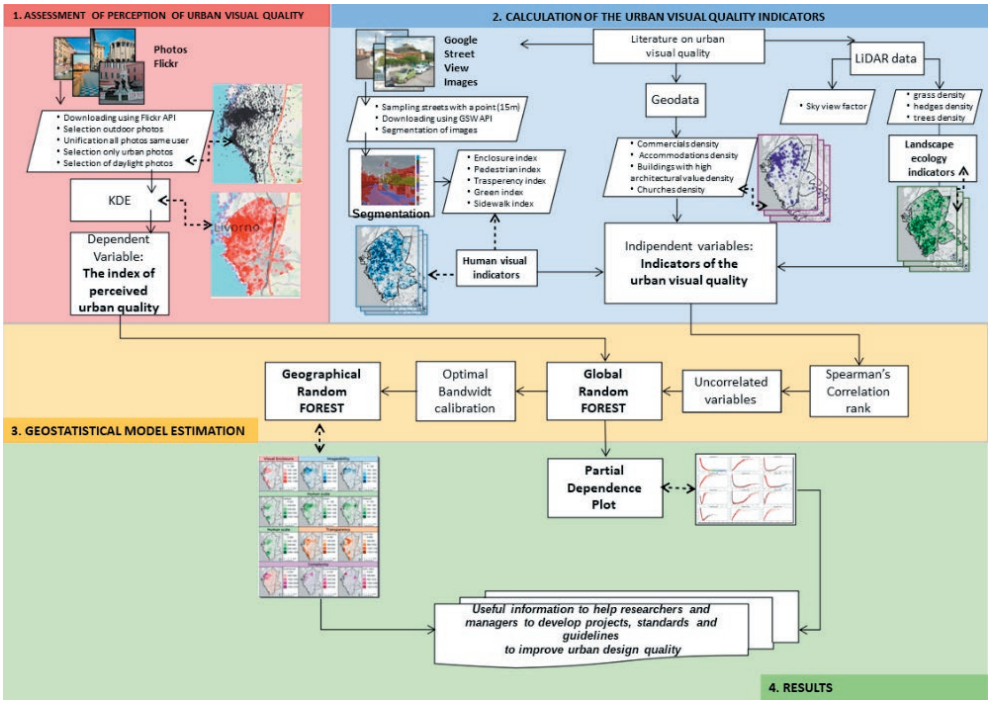
1. Assessment of the perception of urban visual quality: the visual quality of urban spaces was obtained by calculating of the density map of the photo shooting points shared on Flickr (variable depending on the geostatistical model).
2. Calculation of the urban visual quality indicators: using different methodologies to obtain indices that built urban visual quality (dependent variables of the geostatistical model).
3. Geostatistical model estimation: a global Random Forest (RF) model for predictive purposes and a local Geographic Random Forest (GRF) model for exploration purposes, to understand which are the most significant indices of urban visual quality for the whole city and which are the most important dimensions that characterize the visual perception of the different urban neighbourhoods. Figure 2 shows a flowchart of the work.

3.2.1 Assessment of the perception of urban visual quality

Previous research showed that the information contained in the social platform of Flickr can be used to assess the perceived quality of urban spaces. The download of the geographical coordinates of the photo shotting points shared on Flickr is characterized by the following steps.

1) Download through an R language program based on Flickr API the metadata of the shooting points of photos shared from 2005 to 2018. The downloaded data were: photo code, owner code, geographic coordinates, date, time, title, TAG description, geo_context parameter.
2) Selection of photos taken outdoors by setting the "geo_context parameter" in the API to "outdoors".
3) Unification of all photos taken by the same user on the same date and with the exact coordinates.
4) Selection of images not related to urban quality by filtering tags with specific keywords.
5) Selection of photos taken with daylight, using the Flickr API to know the date

Figure 2. Flowchart of the work.



and time of the shot. The time of sunset was calculated using the R {suncalc} library.

The final geo-database of points was transformed into a density surface, using a kernel density estimation analysis (KDE) (Chen and Shaw, 2016).

To calculate the probability that a photo could be taken in a given location, we performed a kernel density analysis for the data, using the geographical locations of the photos. The kernel density was used to estimate the intensity of the points, by creating a smooth surface using a bivariate probability density function. The kernel estimator is defined as:

$$f(x) = \frac{1}{nh}\sum_{i=1}^{n} K\frac{x-x^i}{h} \tag{1}$$

Here, n is the total number of points, h is the bandwidth that determines the amount of smoothing, K is the kernel function, x is the location of the estimation, and xi is a known point location.

The kernel function K can have different forms. The triangular function we used in the analysis is given below:

$$K = \begin{cases} \frac{b-|x-x^i|}{b^2} x\epsilon[x^i - b, x^i + b] \\ 0otherwise \end{cases} \tag{2}$$

The case under study used a triangular Kernel with a bandwidth b = 50 m. The kernel was then overlapped with a 100-meter side hexagonal grid, chosen because of its topological and geometric properties (Feick and Robertson, 2015; Patil, et al., 2000). The size of the grid was identified by applying the Abstract Method (Stamps III, 2001) according to the number of blocks in the study area.

$$r = \sqrt{\frac{S}{0.866 \cdot NB}} \tag{3}$$

where r is the radius of the grid, S is the total area of the area, NB is the number of blocks.

3.2.2 Calculation of the urban visual quality indicators

In the vast amount of literature on urban design quality and structural properties, we referred to the classification of the indicators proposed by Ewing and Handy (2009). According to their study, visual quality indicators were divided into the following five conceptual categories.

(a) Visual enclosure. Outdoor spaces are defined and shaped by vertical elements, which interrupt viewers' lines of sight. Christopher Alexander et al. (1977, p. 106) stated that «an outdoor space is positive when it has a distinct and definite shape, as definite as the shape of a room, and when its shape is as important as the shapes of the buildings which surround it». According to this definition, visual enclosure indicators are: sky view factor and enclosure index.

(b) Imageability. Kevin Lynch (1960, p. 9) defines imageability as the quality of a physical environment that evokes a strong image in an observer: "it is that shape, color, or arrangement which facilitates the making of vividly identified, powerfully structured, highly useful mental images of the environment".

(c) Human scale. Following Ewing and Handy (2009, p. 77), «Human scale refers to a size, texture, and articulation of physical elements that match the size and proportions of humans and, equally important, correspond to the speed at which humans walk. Building details, pavement texture, street trees, and street furniture are all physical elements contributing to human scale».

(d) Transparency. "Transparency refers to the degree to which people can see or perceive what lies beyond the edge of a street and, more specifically, the degree to which people can see or perceive human activity beyond the edge of a street." (id., p. 78).

(e) Complexity. "Complexity results from varying building shapes, sizes, materials, colours, architecture and ornamentation" (id., p. 79) but also "The presence and activity of people add greatly to the complexity of a scene" (id., p. 80).

According to the above and also using additional indicator classifications found in the bibliography (Ewing et al. 2006; Gavrilidis et al. 2016; Talavera-Garcia and Soria-Lara, 2015; Yin and Wang, 2016), we selected the following quality indicators in our study (Bernetti et al., 2020):

(a) Visual enclosure: sky view factor; enclosure index.
(b) Imageability: pedestrian index; distance from churches; pavement index.
(c) Human scale: grass density, hedges density, trees density, green index, sidewalk index.
(d) Transparency: transparency index, distance from coast-line.
(e) Complexity: distance from commercials; distance from accommodation land use; distance from buildings with high architectural value.

Most of these indicators selected are based on data available on the Internet. The remaining were based on geodatabases made available by public administrations. The methodologies used to calculate our indicators are the following:

– Deep learning segmentation was applied to the enclosure index, pedestrian index, cyclist index, road crowdedness index, building crowdedness index, and transparency index.
– Landscape ecology indicators for the grass density, hedges density and trees density.
– Kernel Density Estimation for urban points of interest (POIs) defined as commercials, accommodations, buildings with high architectural value, and churches.
– Sky View Factor was applied to LIDAR data.
– GIS application for coastline distance.

3.2.2.1 Human Visual Indicators

To efficiently detect the streetscape of the city of Livorno through images taken from the GSV platform, we sampled all the streets of Livorno with a point every 15 meters. Adopting a 60° field of view (FOV) (similar to human, Yin and Wang, 2016) to cover a 360° panoramic view of the surrounding environment, we downloaded 6 images with azimuth = 0, 60, 120, 180, 240 and 300 for each sampled point. Sampling was carried out using procedures based on GRASS and R.

A segmentation procedure was applied to the images obtained. Segmentation of an image in digital image processing is the process of partitioning an image into meaningful regions. It is used to obtain a more compact representation, to extract objects or as a tool for image analysis and allows you to partition digital images into sets of pixels. To segment the GSV images we employed a pre-trained network of MATLAB software based on the Deeplabv3 + network trained using a collection of images containing street level views obtained while driving (Brostow et al. 2009). To facilitate training, CamVid's 32 original classes have been grouped into 11 classes as follows: "Sky", "Building", "Pole", "Road", "Sidewalk", "Tree", "Symbol Sign", "Fence", "Car", "Pedestrian" and "Cyclist".

Based on the theories and recent literature mentioned above, we proposed a method to evaluate four indicators related to the visual quality of urban design, and to calculate the indices for each GSV photo. The indices calculated were as follows (Bernetti et al., 2020).

1. Enclosure index (Encl), defined as the degree to which streets and other public spaces are visually defined by buildings, walls, trees, and other elements.

Usually, people positively perceive an urban space when they are able to recognize a distinct shape defined by vertical elements that interrupt their view (Alexander et al., 1977). It was calculated using the following equation:

$$Encl = \frac{\sum_1^6 Bn + \sum_1^6 Tr + \sum_1^6 Pol}{\sum_1^6 Rd + \sum_1^6 Pv} \quad (4)$$

where *Bn* is the number of building pixels; *Tn* is the number of tree pixels; Rd is the number of road pixels; *Pv* number of pavement pixels, and *Polis* is the number of road signs pixels.

2. Pedestrian and cyclist, index (Ped), defined as the degree to which people can see or perceive human activity: The perception of other people makes the environment more comfortable and safe.

$$Ped = \frac{\sum_1^6 Bc + \sum_1^6 Pd}{\sum_1^6 Car} \quad (5)$$

where *Bc* is the number of bicyclist pixels; *Pd* is the number of pedestrian pixels; and *Car* is the number of car pixels.

3. Transparency index (Trp), defined as the degree to which people can see what lies beyond the edge of a street or other public space. The wide view ensures that people can observe objects that are far away, allowing observers not only to recognize the world around them but also to feel safer moving around in the environment.

$$Trp = \frac{\sum_1^6 Sky}{\sum_1^6 TotPix} \quad (6)$$

Where *Sky* is the number of sky pixels and *TotPix* is the total pixel number in the image.

4. Green index (Green), defined as the extent to which the visibility of street vegetation can influence pedestrian psychological feelings. The therapeutic effects of natural environments is well-known and extensively reported in the literature. Several cities have been equipped with healing gardens and green roads for restoration from stress (Pouya et al., 2015).

$$Green = \frac{\sum_1^6 Veg}{\sum_1^6 TotPix} \quad (7)$$

where*Veg* number of vegetation pixels.

5. Sidewalk index (Sidewalk), defined as the extent to which the visibility of pavement and fences influences pedestrian psychological feelings:

$$Sidewalk = \frac{\sum_1^6 Pav + \sum_1^6 Fenc}{\sum_1^6 Road} \quad (8)$$

where *Pav* is the number of pavement pixels and *Fenc* is the number of fences pixels.

3.2.2.2 Landscape ecology indicators

The indicators used in our analysis were the percent density of grassland, hedgerows, urban forest in each hexagon of the grid. The indicators were calculated using multispectral and lidar remote sensing data according to the methodology defined by Barbierato et al. (2019).

Urban vegetation cover was identified through normalized difference vegetation index (NDVI) analysis, taking into account only healthy vegetation extracted based on NDVI values greater than or equal to 0.2 (Rodgers III et al., 2009).

The result was presented as a Boolean map with a resolution of 1 m (similar to LiDAR data), in which the value of 0 indicated the absence of vegetation, and a value of 1 indicated its presence. The results were spatialized on the hexagonal grid. As urban green areas are characterized by various types of vegetation with different ecologic and perceptive functions, we distinguished these types according to their height values.

To obtain the height of the vegetation, we made an overlay operation between the NDVI binary map and a normalized digital surface model generated from LIDAR data. The result of this operation was a raster map divided into three height classes. The first class (from 0 to 0.40 m) represented grass, the second (from 0.40 to 3 m) was classified as hedges, and the third (greater than 3 m) was classified as trees.

The indicators we used were the percentages of green landscapes of class i (P_i) with $i = \{\text{grass, hedges, trees}\}$. The former allowed us to understand the percentage of plant cover of each grid hexagon. The operation is as follows:

$$P_i = \frac{\sum NDVI_{j,i}}{H} \tag{9}$$

In the above, $NDVI_{j,i}$ is the j-th pixel in the NDVI raster map classified on class i, and H is the total hexagon area.

3.2.2.3 Kernel density estimation (KDE) of the urban points of interest (POIs)

We have identified commercial, housing, churches and buildings of high architectural value as urban Points of Interest (POI). The location of the POIs was taken from the OpenStreetMap (OSM) database.

We calculated the territorial density of the POIs using a KDE procedure. KDE has been widely used in POI data analysis (Li et al. 2013). Lian et al. (2014) demonstrated that through KDE, it is possible to identify areas of influence of POIs, as related to areas of activity of the users. The POIs derived from OSM were implemented in density maps by applying a triangular KDE with a bandwidth of 500 m. The results were spatialized on the hexagonal sampling grid.

In according to Chan et al. (2021), the equation is the following:

$$K(q, p_i) = \max(1 - \frac{1}{b} * dist(q, p_i), 0) \tag{10}$$

where b and $dist(q,p_i)$ denote the bandwidth of the triangular kernel function and Euclidean distance, respectively.

3.2.2.4 The sky-view factor

Sky-View Factor (SVF) indicates the portion of the sky that is visible from an observation point. The higher the SVF, the greater the heat loss to the atmosphere. Sky-View Factor (SVF) is a widely used parameter to describe urban climatology at high resolution scales, several studies have employed SVF as a parameter of urban design quality (Lindberg and Grimmond, 2010; Nasrollahi and Shokri, 2016; Yang et al. 2007). We calculated SVF through Quantum GIS software with the urban multi-scale environmental predictor (UMEP) plug-in (Lindberg et al. 2017) by superimposing reconstructing through Lidar data the 3D model of the city of Livorno.

3.2.2.4 GIS application to calculate coastline distance

The coastline distance is an index was calculated with Qgis software using the public administration geodatabase data. The coastline, with its promenade represents a significant geographical macro-characteristic for the visual perception of the local quality of the urban space.

3.2.3 Geostatistical model estimation

The regression models used the Flickr photo density indicator as the dependent variable, and the urban design quality indicators as independent variables. For this purpose, we combined two versions of the random forest (RF) model: traditional and geographical.

The RF regression model has some advantages over traditional statistical methods: it allows to treat complex relationships between predictors that can arise with large amounts of data, and is able to process nonlinear relationships between predictive variables. The main limitation is the difficulty in the direct interpretation of the results, since the explicit ensemble model is represented by hundreds (sometimes thousands) regression trees. To overcome this difficulty, Friedman (2001) proposed the use of partial dependence plots, allowing for visualization of a suitable RF model through its mapping from feature space to prediction space. Welling et al. (2016) proposed a new methodology called “forest floor” using feature contributions (FC), a method to decompose trees by splitting features and then performing projections. The advantages of the forest floor approach over the partial dependence plots are that the interactions are not masked as averaging. As a result, interactions that are not visualized in a given projection can be located. Forest floor was implemented in the {foresFloor} library in the statistical programming language R (Team, 2015).

Even though RF is a well-functioning and generalizable algorithm, the vast majority of its implementations are not spatialized. Georganos et al. (2021) proposed a geographical implementation of the RF, called the geographical RF (GRF), as a disaggregation of RF in the geographical space in the form of local sub-models.

The equation for a typical GRF model is:

$$y_i(u) = f\left(\beta_{0,i}(u_i, v_i)x_{i,0}, \beta_{1,i}(u_i, v_i)x_{i,1}, \cdots, \beta_{m,i}(u_i, v_i)x_{m,}, e\right) \quad (11)$$

In the above, $\beta_{m,i}(u_i,v_i)$ is the non-linear prediction of an RF model calibrated on location i and u_i, v_i are the coordinates, e is an error term and $x_i \ldots x_n$ are the independent variables (indices). The area in which the sub-model operates is called the neighborhood (or kernel), and the maximum distance between a data point and its kernel is called the bandwidth (Brunsdon, 1998). The bandwidths for geographical models can be user-specified, or can be determined via automated procedures (e.g., cross-validation), provided an objective function exists (Akaike, 1973; Fotheringham, 2018; Hurvich, 1998). With the data set organized on a regular hexagonal tessellation, we set an adaptive kernel bandwidth to include the N hexagons closest to the observation/calibration hexagon. We estimated the bandwidth using the Akaike information criterion.

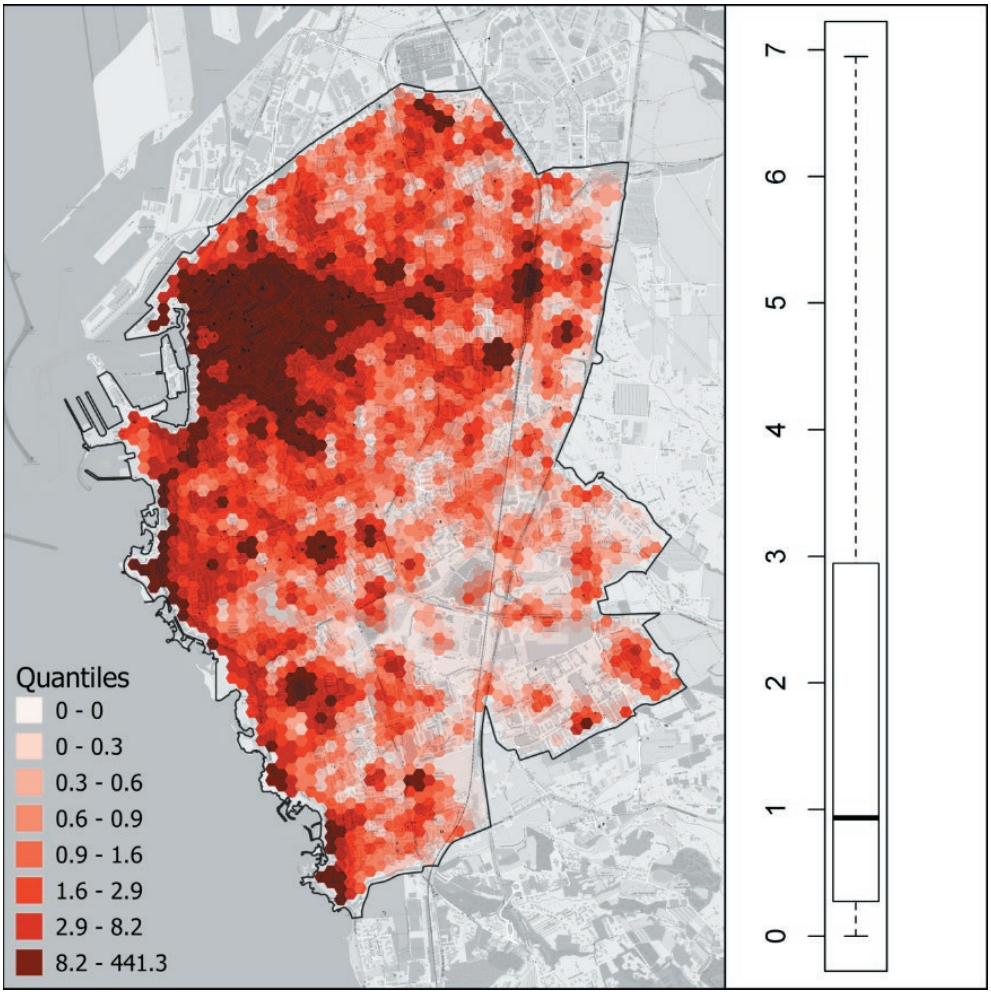
To avoid overfitting problems, we used two distinct approaches. We set a large number of trees (500), with 5 variables each and maximum number of terminal node trees in the forest equal to 4 (Cutler et al. 2012; Fotheringham and Park, 2018; Scornet, 2017). We then extracted three sub-datasets from the original dataset: a training set (70% of total observations), used for learning the RF model; a validation set (15% of total observations) for verification of model parameters and a test set (15% of total observations) evaluation of model performance. Unlike RF, we used the GRF as a purely exploratory (rather than predictive) tool. GRF is a local decomposition of the RF and, therefore, the results can be mapped using the entire data set without training/test divisions (for better visualization), to study the local importance of the individual indicators and the performance distributions of the local models. The global and geographical RF models were calculated using the R library {randomForest} and {SpatialML} packages, respectively. The procedures are available as supplementary material.

4. Results

4.1 Assessment of the perception of urban visual quality

The raw database contained approximately 23,063 photo localizations taken in the period from 2005–2017, and the final filtered database contained 11,008 observations. Figure 3 shows a dependent variable map based on the KDE Flickr photo index.

Figure 3. Map of the dependent variable.



The index recorded a maximum value of 26.07 photo/m^2 with an average value of 3.33 photo/m^2, a median value of 0.93 photo/m^2, and first and third quartiles of 0.27 and 2.94 photo/m^2, respectively. Thus, the data had a very asymmetrical frequency distribution.

4.2 Calculation of the urban visual quality indicators

Using the Googleway library, we downloaded 17,196 geo-tagged images, relating to 2,866 sampling points acquired in 2018. Figure 4 shows an example of the segmentation process.

Figure 4. Example of segmentation process.

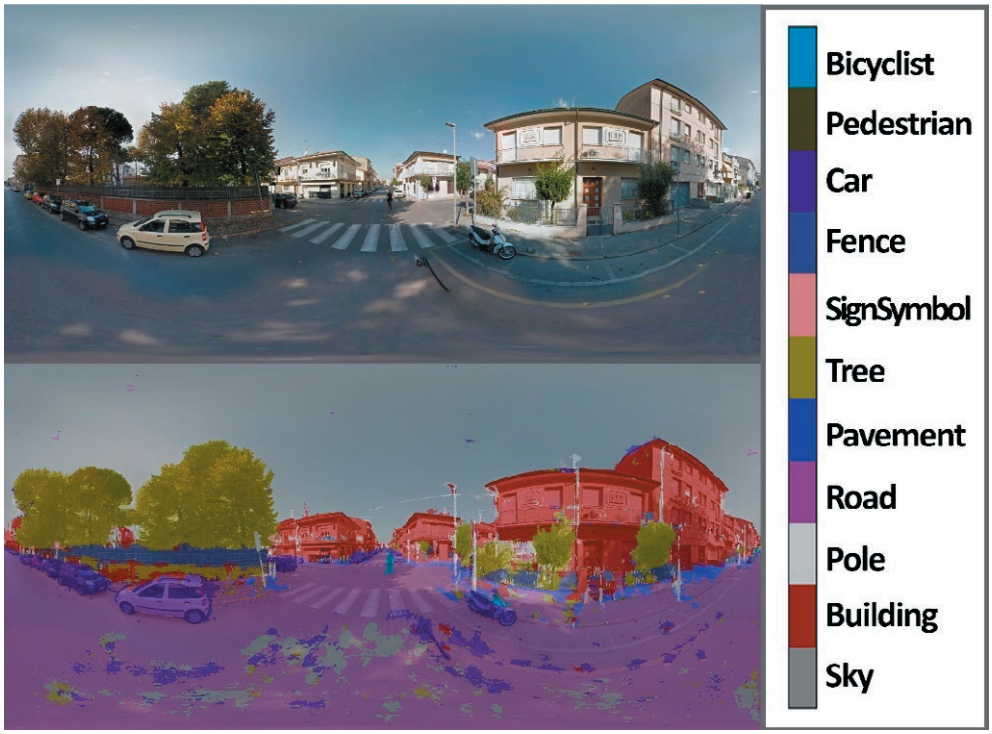


Figure 5. Maps of the indices calculated through the segmentation of the GSV images.

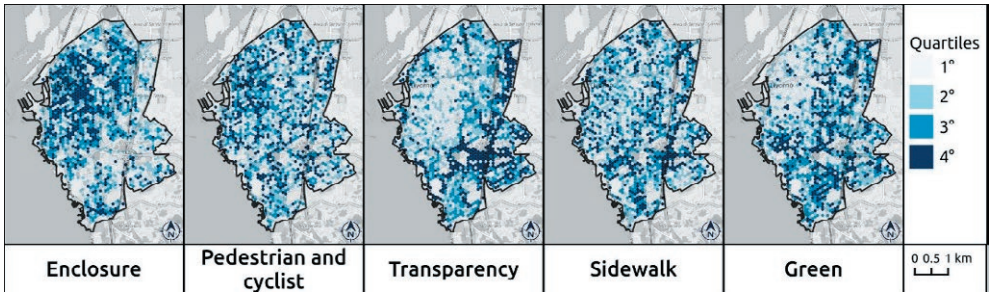


Figure 5 shows the maps of the indices calculated via the segmentation of the GSV images.

Figure 6 shows the maps of the three landscape indices. Figure 7 shows the density indices of the buildings intended to influence the perceived quality of the urban environment. Lastly, Figure 8 shows the two geographical indices linked to the "visual enclosure" and "transparent" dimensions.

Figure 6. Maps of landscape ecology indices.

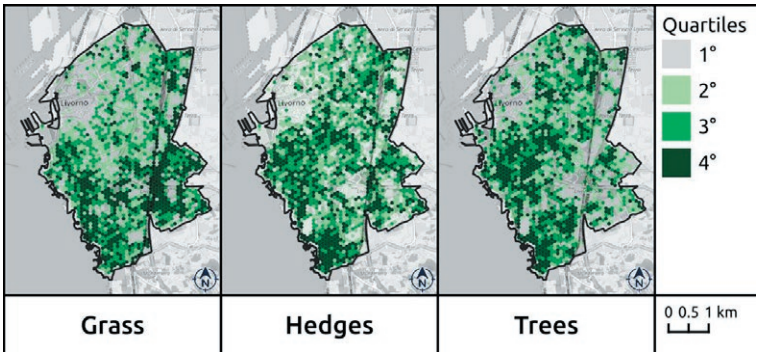


Figure 7. Maps of density indices.

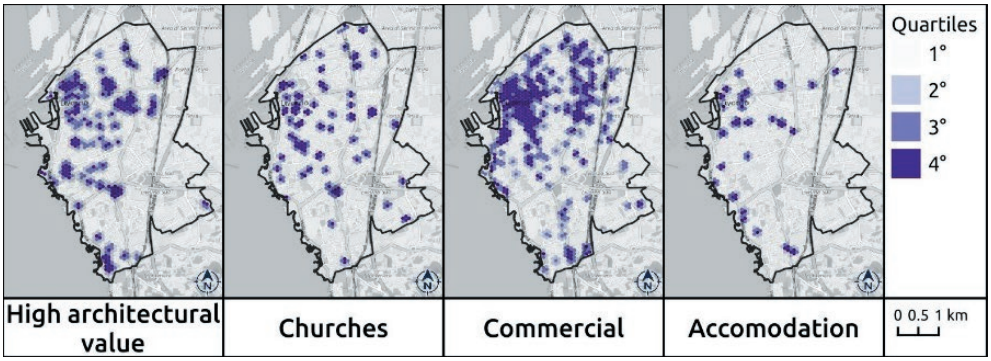
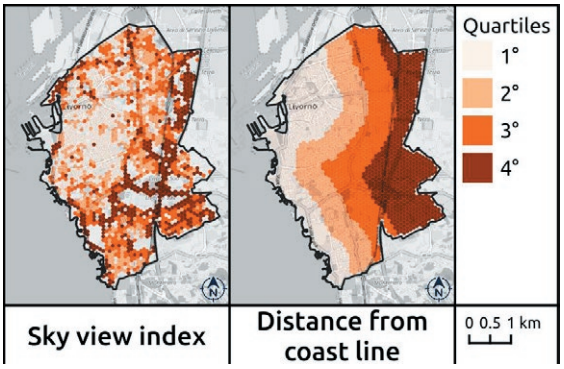


Figure 8. Maps of sky view factor (SVF) and of distance from coastline.



4.3 Random forest (RF) models

The first step in the RF algorithm was to test for multicollinearity between variables using Spearman's degree of correlation. We kept all variables that

showed Spearman's correlation below 0.7, a limit used in the literature in applying random forest models using GIS (Chen et al., 2018; Kamusoko & Gamba, 2015). Variance Inflation Factor (VIF, Table 1) analysis also shows that there is no multicollinearity among the independent variables, as all VIFs are less than 10 (Kim, 2019). Table 1 shows the VIF values for each index.

The final set of 12 variables is the following:

a) Visual enclosure: enclosure index.
b) Imageability: pedestrian index; churches density.
c) Human scale: hedges density, trees density, green index, sidewalk index.
e) Transparency: transparency index, distance from coastline.
e) Complexity: commercial land use density; accommodation land use density; density of buildings with high architectural value.

The global (non-geographical) RF model had a McFadden pseudo R-square value of 0.911 for the training set, 0.56 for the testing set, and 0.566 for the validation set (Table 2). These results can be considered satisfactory on a general level. The most important predictors, in decreasing order of the percent increase in mean squared error (% increase in MSE) and node purity, are shown in Table 2. The variable that contributes most to explaining the perception of urban visual quality in the global model is the distance from the coastline, followed by the commercial land use density and the hedges density.

Figure 9 illustrates the partial dependency plots for each predictor. The partial dependency diagram shows the marginal effect that a predictor has on the expected result of a model. A partial dependency graph can show whether the relationship between the dependent variable and the predictor is linear, monotonous or more complex. Most of the graphs show non-linear relationships of the dependent variable in relation to the predicted perception of urban quality.

Table 1. Variance Inflation Factor.

Index	VIF
Pedestrian	7.877074
Trees density	6.611743
Hedges density	5.297603
Green index	4.918989
Enclosure index	4.048282
Sky view factor	3.995972
Sidewalk index	3.779627
Commercial land use density	3.765017
Churches density	3.586181
Coast distance	2.728047
Density of building with high architectural value	2.482175
Accommodation land use density	1.697885

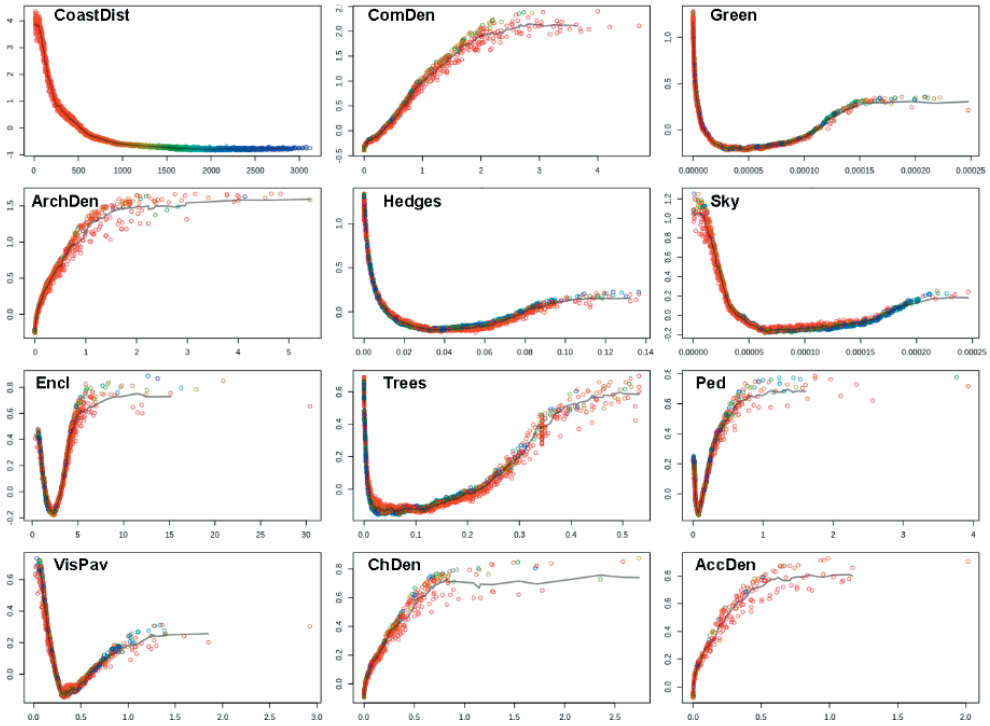
Table 2. Results of the non-geographical random forest (RF) regression model.

Predictor	% IncMSE	Node Purity
Coast distance	103.04	30,090.52
Commercial land use density	50.82	12,289.57
Hedges density	27.81	7,133.00
Density of building with high architectural value	26.75	5,904.94
Trees density	20.90	5,216.73
Green index	13.19	4,918.20
Transparency index	17.56	3,770.87
Sidewalk index	6.57	3,171.20
Enclosure index	10.59	3,117.37
Pedestrian index	8.16	2,752.35
Churches density	4.30	1,319.50
Accommodation land use density	6.44	1,236.92
Residual Sum of Squares	7551.707	
Pseudo R-squared Training set	0.91	
Pseudo R-squared Validation set	0.56	
Pseudo R-squared Testing set	0.57	
Moran index Global model	0.53	

For example, the graph shows that the effect of coastline on quality perception is geographically localized. In fact, the curve decreases rapidly to zero less than 1 km from the sea. Similarly, an increase in the density of churches and buildings with high architectural value is associated with an increase in urban quality perception. Interestingly, some variables seem to have partial dependency plots with two distinct segments. For example, the green index, edge density, sidewalk index, and transparency index have a decreasing trend followed by an increasing trend, and this can be explained by the fact that the presence of these two variables has positive effects on the perception of the visual quality of a city, but only beyond a certain threshold value. Another interesting result is obtained in the case of the enclosure index. The partial dependence graph shows that the perception of urban spaces is correlated to two cases: null or very low enclosure index values (typically in the promenade) and enclosure index values equal to about 5-7 (typical of the central areas of the city). Beyond this value, the observations (represented by small colored circles) are very scattered and, therefore, the model does not give reliable indications.

We tested the spatial distribution of the global model residues and found a Moran index of 0.53, indicating the presence of spatial autocorrelation. This justified the use of a local GRF model to identify the variations in the importance of the predictors in space.

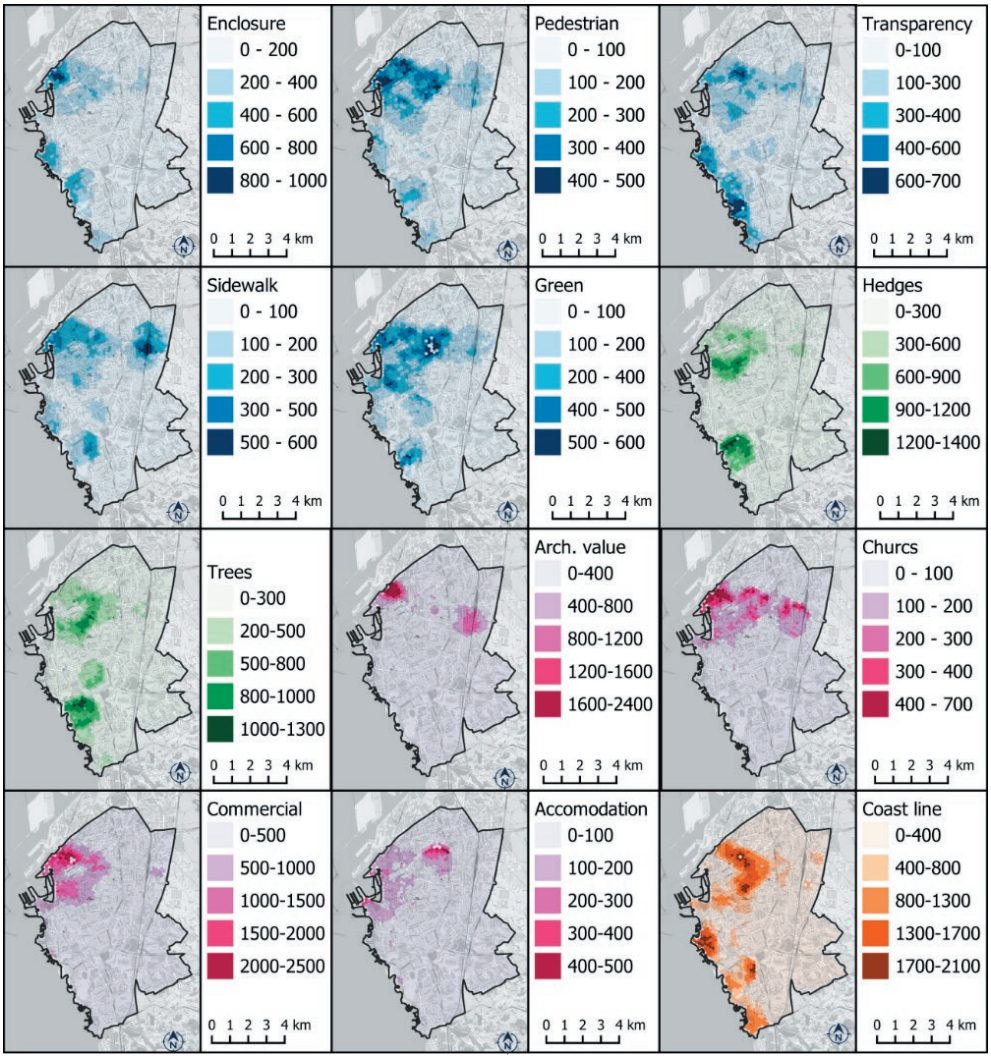
Figure 9. Plots for the random forest (RF) variables. Panel titles designate which variable is being plot along the x-axis: (CoastDist) distance from coastline, (ComDen) Commercial land use density, (Green) Green index, (ArchDen) Density of buildings with high architectural value, (Hedges) Hedges density, (Sky) Transparency index, (Encl) Enclosure index, (Trees) Trees density, (Ped) Pedestrian index, (VisPav) Sidewalk index, (ChDen) Churches density, (AccDen) Accommodation density. Panel titles also include the R2 (leave-one-out goodness of fit) of the average feature contribution line (denoted in black). The color gradient is applied in all panels along the distance from the coastline, passing through red-yellow-green-blue with increasing distance.



The results of the bandwidth optimization suggested an optimal bandwidth of 95 cells (i.e., for each of the 2,517 cells, a local RF model should be calibrated using data from the nearest 95 cells). The use of the GRF model reduced the spatial autocorrelation of the residues, with a Moran index of 0.17, and a pseudo R-square value equal to 93%. Figure 10 and 11 and Table 3 show respectively the geographical variation, boxplots and the statistics of the purity index of the indicators. The geographic model is therefore overall consistent with the global model. The results of the geographic model confirm the importance of the explanatory variables in terms of purity index for all positions. The R square improves from 0.91 to 0.93.

Notably, there is a strong degree of spatial interaction between each predictor, whereas the importance of each predictor varies consistently through space. At the level of individual geographic location, we obtain very diverse results relative to the ranking of the importance of the variables. The importance of the predic-

Figure 10. Geographical variation of predictors of the geographic RF (GRF) model.



tors is generally highest in areas with a greater density of photos, but with important variations among the predictors, especially in relation to the north-south gradient of the study area. Interestingly, both transparency variables seem to be more important near the coastline. Moreover, the two imageability indices appear to be strongly more predictive over the historical center of the city. Similarly, the rest of the predictive variables formulate unique spatial patterns.

More specifically, comparing Figure 2 with Figure 10, we can see how the different neighborhoods of Livorno show importance of the dimensions of urban quality strongly differentiated. The Buontalenti neighborhood has high val-

Figure 11. Boxplot of purity index of predictors in geographical RF (GRF) regression model.

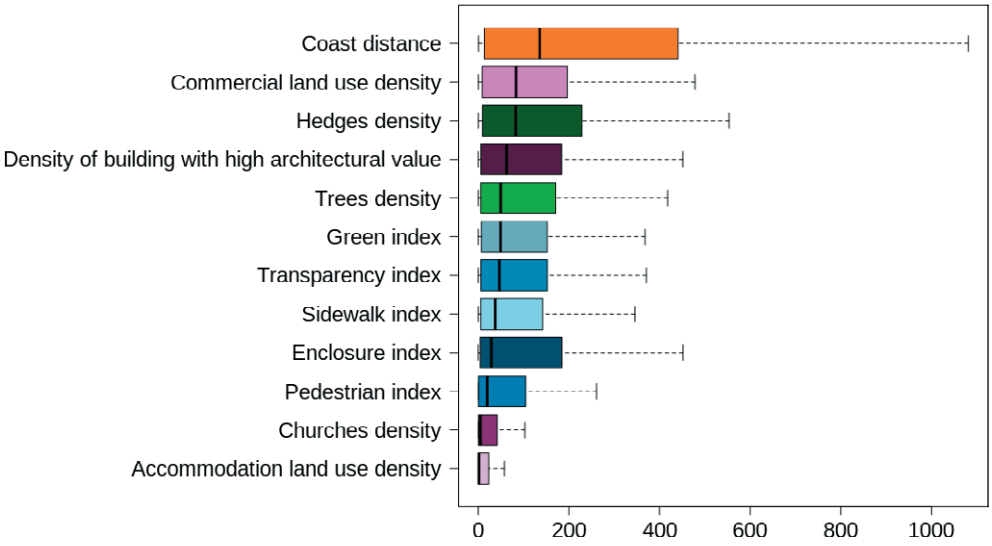


Table 3. Results of the geographical RF (GRF) regression model.

Predictors	Purity index			
	Min	Max	Mean	StD
Coast distance	0.79	2,232.39	315.11	423.52
Commercial land use density	0.00	2,604.24	192.20	361.75
Hedges density	0.52	1,455.72	170.86	244.55
Density of buildings with high architectural value	0.46	1,309.07	164.82	217.50
Trees density	0.21	732.39	113.42	133.36
Green index	0.00	2,291.88	112.93	253.59
Transparency index	0.34	668.38	108.33	131.80
Sidewalk index	0.33	1,024.51	106.14	139.69
Enclosure index	0.26	567.36	89.84	103.71
Pedestrian index	0.28	595.58	86.61	102.20
Churches density	0.00	605.35	48.26	89.82
Accommodation land use density	0.00	639.11	31.33	66.19
Statistical parameters of Geographical Random Forest				
Residual Sum of Squares (Predicted)	5895.932			
Pseudo R-squared %	93.058			
Moran index	0.17			

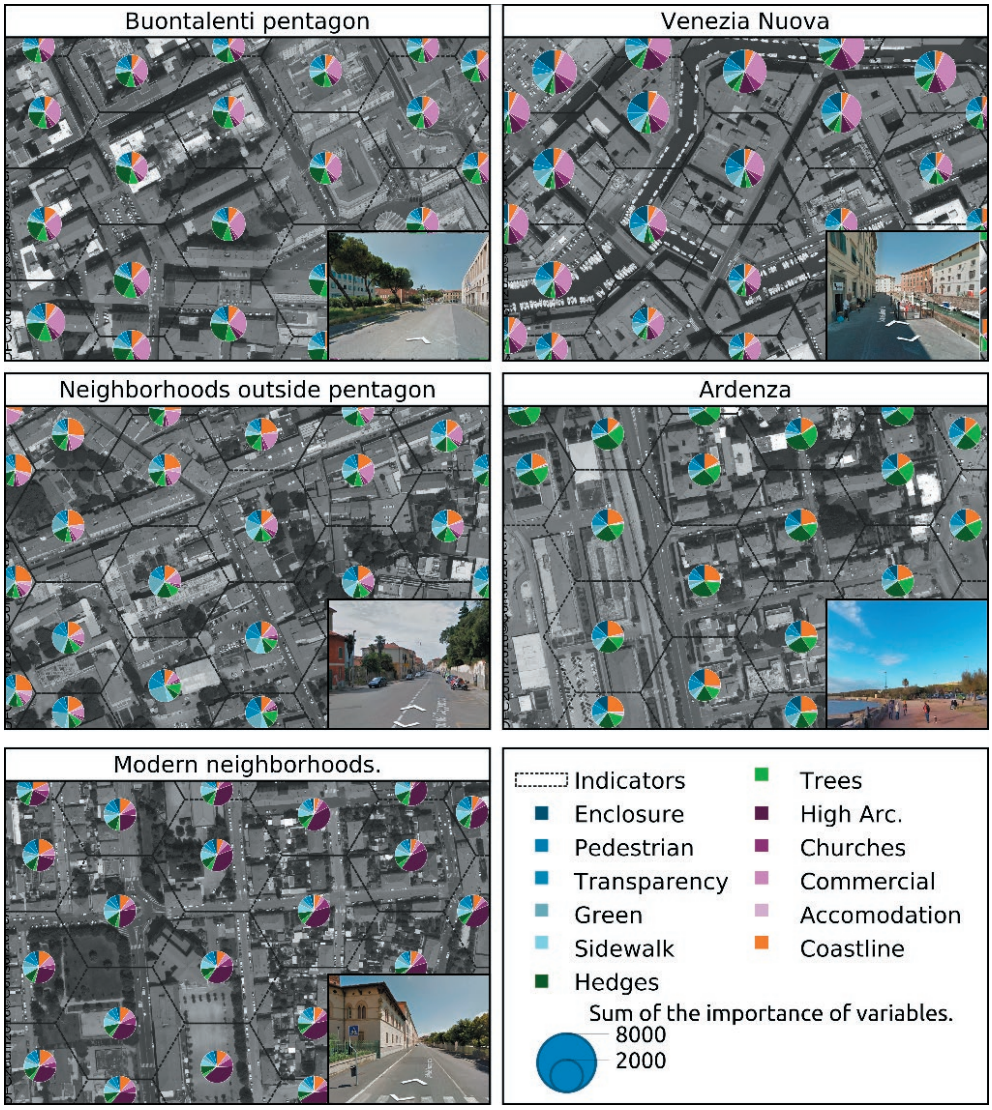
ues of the purity index for all indices related to the human scale. The Venezia Nuova neighborhood is characterized by a multidimensional urban quality with high importance of the Visual Enclosure, Imageability and Complexity indices. In the neighborhoods outside the pentagonal city the urban quality is corelated to the indices of transparency near the promenade, similarly to the neighboring neighborhood of Ardenza, which is also characterized by the importance of the indices of human scale density of hedges and density of trees. The modern neighborhoods have only a few indices with significant importance values: sidewalk, density of churches and density of architectural value. Finally, the Sargenti and Korea neighborhoods show low urban quality with no indices with appreciable importance.

5. Discussions

We combined two different RF models: a global model for predictive purposes and a local model for exploration purposes, to select the most significant indicators for different areas of the city of Livorno. The study here offers important information on the geometric, physical and environmental variables of an urban environment that contribute to more abstract qualities. The significant characteristics had the expected relationships with the other findings present in the literature, although many others proved to be statistically not significant despite initial expectations. of city design.

For instance, Figure 12 shows some case studies referring to the different periods of historical expansion of the city. The zone of the pentagon of Buontalenti (see also Figure 1) is characterized by a greater correlation between perceived urban quality and the presence of urban greenery (both trees and hedges). The perception of the quality of the spaces in the Old Venice (Figure 1 and 12) is instead related to the commercial density index and the enclosure index. Therefore, in this part of the city it may be appropriate to implement urban projects for the redevelopment and maintenance of the facades and to improve the visual quality of the shops. The neighbourhoods outside the pentagon of Buontalenti (Figure 1 and 12) do not reveal urban quality elements such as to influence the perception of spaces. On the other hand, the neighbourhoods of modern expansion (Figure 1 and 12), despite having an urban layout similar to the previous one, appear to have a higher perceived quality due to the widespread presence of modern buildings but of architectural value. Finally, as regards the Ardenza zone (Figure 1 and 12), the analysis confirms the correlation between the perceived quality of public spaces and the characteristic indicators of the seaside promenade: distance from the coastline and the presence of urban greenery (trees and hedges). The results of this study can be used in different of ways. Municipal planners and administrators can achieve a more detailed and complete understanding of an urban environment. The map can provide good support for the definition of urban planning policies.

Figure 12. Some case studies of indicators correlated with urban quality.



6. Conclusions and future developments

The proposed approach is able to provide useful information to identify and evaluate the geometric, physical and environmental characteristics of public spaces that most determine the perceived urban visual quality at the planning scale.

We believe that our study has helped to demonstrate that the perceived quality of the city is influenced by many physical, geometric and environmental vari-

ables in a complex way. From a general point of view, data shared by social media combined with data deriving from Google's Internet services and remote sensing data can provide a useful tool to improve our understanding of the relationship between humans and the urban environment.

Another advantage is the ability of using publicly available data with a much lower cost if compared to a traditional analogous survey through questionnaires. A survey through questionnaires can evaluate a greater number of subjective variables, for example linked to the sentiment that arouses perception. However, this study clearly has some limitations.

The method has been tested on a small city and this allowed to have acceptable processing times, but for larger cities there are two possible solutions that will be implemented in the future development of the research. The first is to tile the study area; since the methodology is based on a local geographical model, this should not affect the estimation of the model parameters. A second solution would be to use a reduced sample of images downloaded from GSV. The pros and cons of the two hypotheses will have to be evaluated. In the future both methods will be applied in Florence to verify their advantages and limitations.

Furthermore, other limitations are related to the characteristics of social media data. Social media are not only about young people, but it seems that only this social group is actively involved, older people may be mostly recipients of content and not its creators. Moreover, the Flickr platform does not allow to obtain social and personal information about the individual user in order to segment visual preferences by age groups and other social variables.

An additional weakness is that the presence of autocorrelation was tested only by Moran's index calculation. Further investigations will be needed to test for the presence of spatial heterogeneity through measures of similarity between associations of values (covariance, correlation or difference) and associations in space (contiguity), including Lagrange multiplier tests (LM-lag and LM-error). The results of these tests could allow us to compare the results obtained through GRF models with other spatial regression models, such as the Spatial Lag Model (SLM) or the Spatial Error Model (SEM).

After all, the results of the methodology can help to implement more efficient sample designs, stratifying the city in relation to the importance of the indicators used as independent variables, thus reducing the sample size and therefore the costs. The synergistic implementation of the two methodologies will constitute a first future development of this research to give useful information to help researchers and managers to develop projects, standards and guidelines to improve urban visual quality in the cities.

References

Adams, M. A., Ryan, S., Kerr, J., Sallis, J. F., Patrick, K., Frank, L. D., & Norman, G. J. (2009). Validation of the Neighborhood Environment Walkability Scale (NEWS) items using geographic information systems. *Journal of Physical Activity and Health*, 6(s1), S113–S123.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In Parzen, E., Tanabe K., & Kitagawa G. (Eds). *Selected Papers of Hirotugu Akaike*. Springer Series in Statistics (Perspectives in Statistics). New York, NY, Springer.

Alampi Sottini, V., Barbierato, E., Bernetti, I., Capecchi, I., Cipollaro, M., Sacchelli, S., & Saragosa, C. (2018). Urban landscape assessment: a perceptual approach combining virtual reality and crowdsourced photo geodata. *Aestimum*, (73), 147–171.

Alexander, C., Ishikawa, S., & Silverstein, M. (1977). *A Pattern Language*. Oxford University Press.

Arnheim, R. (1969). *Visual Thinking*. Berkeley, University of California Press.

Asihara, Y. (1983). *The Aesthetic Townscape*. Cambridge, MA, MIT Press.

Barbierato, E., Bernetti, I., Capecchi, I., & Saragosa, C. (2019). Remote sensing and urban metrics: an automatic classification of spatial configurations to support urban policies. In *Earth Observation Advancements in a Changing World,* Torino, Italy, p. 187. Italian Society of Remote Sensing.

Bernetti, I., Alampi Sottini, V., Bambi, L., Barbierato, E., Borghini, T., Capecchi, I., & Saragosa, C. (2020). Urban niche assessment: an approach integrating social media analysis, spatial urban indicators and geo-statistical techniques. *Sustainability*, 12(10), 3982.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.

Brostow, G. J., Fauqueur, J., & Cipolla, R. (2009). Semantic object classes in video: a high-definition ground truth database. *Pattern Recognition Letters*, 30(2), 88–97.

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D*, 47(3), 431–443.

Cerin, E., Saelens, B. E., Sallis, J. F., & Frank, L. D. (2006). Neighborhood Environment Walkability Scale: validity and development of a short form. *Medicine & Science in Sports & Exercise*, 38(9), 1682–1691.

Chan, T. N., Ip, P. L., U, L. H., Tong, W. H., Mittal, S., Li, Y., & Cheng, R. (2021). KDV-Explorer: A near real-time kernel density visualization system for spatial analysis. *Proceedings of the VLDB Endowment*, 14(12), 2655–2658.

Chen, J., & Shaw, S. L. (2016). Representing the spatial extent of places based on Flickr photos with a representativeness-weighted Kernel Density Estimation. In Miller, J., O'Sullivan, D., & Wiegand, N. (Eds). *Geographic Information Science*. GIScience 2016. Lecture Notes in Computer Science, vol 9927. Springer, Cham.

Chen, W., Zhang, S., Li, R., & Shahabi, H. (2018). Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. *Science of the total environment*, 644, 1006–1018.

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision* (ECCV) (pp. 801–818).

Cullen, G. (1959). *Townscape*. London, Architectural Press.

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In Zhang, C., & Ma, Y. (Eds). *Ensemble Machine Learning*. Boston, MA, Springer. (pp. 157–175).

Doersch, C., Singh, S., Gupta, A., Sivic, J., & Efros, A. A. (2012). What makes Paris look like Paris?. *ACM Transactions on Graphics*, 31(4), 101.

Duncan, D. T., Aldstadt, J., Whalen, J., Melly, S. J., & Gortmaker, S. L. (2011). Validation of Walk Score® for estimating neighborhood walkability: an analysis of four US metropolitan areas. *International Journal of Environmental Research and Public Health*, 8(11), 4160–4179.

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape and Urban Planning*, 142, 173–186.

Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: urban design qualities related to walkability. *Journal of Urban Design*, 14(1), 65–84.

Ewing, R., Handy, S., Brownson, R. C., Clemente, O., & Winston, E. (2006). Identifying and measuring urban design qualities related to walkability. *Journal of Physical Activity and Health*, 3(s1), S223-S240.

Feick, R., & Robertson, C. (2015). A multi-scale approach to exploring urban places in geotagged photographs. *Computers, Environment and Urban Systems*, 53, 96–109.

Fotheringham, A. S., & Park, B. (2018). Localized spatiotemporal effects in the determinants of property prices: a case study of Seoul. *Applied Spatial Analysis and Policy*, 11(3), 581–598.

Fotheringham, A. S., Brunsdon, C. and Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships.* Chichester, John Wiley & Sons.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 29(5), 1189–1232.

Gavrilidis, A. A., Ciocănea, C. M., Niţă, M. R., Onose, D. A., & Năstase, I. I. (2016). Urban landscape quality index–planning tool for evaluating urban landscapes and improving the quality of life. *Procedia Environmental Sciences*, 32, 155–167.

Genuer, R., Poggi, J.M. and Tuleau, C. (2008). Random Forests: some methodological insights, *arXiv,* arXiv:0811.3619. Available at: https://arxiv.org/pdf/0811.3619.pdf (Accessed 1 July 2021).

Georganos, S., Grippa, T., Gadiaga, A.N., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., Wolff, E., & Kalogirou, S. (2019). Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. *Geocarto International,* 36(2), 1–16.

Georganos, S., Grippa, T., Niang Gadiaga, A., Linard, C., Lennert, M., Vanhuysse, S., Mboga, N., Wolff, E., & Kalogirou, S. (2021). Geographical random forests: a spatial extension of the random forest algorithm to address spatial heterogeneity in remote sensing and population modelling. *Geocarto International*, 36(2), 121–136.

Gibson, J. J. (1979). The Ecological Approach to Visual Perception. Boston:Houghton Mifflin.

Gromping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, 63(4), 308–319.

Hauthal, E., & Burghardt, D. (2016). Using VGI for analyzing activities and emotions of locals and tourists. In *Proceedings of the Link-VGI Workshop in Connection with the AGILE,* June 14, 2016, Helsinki, Finland.

Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2), 271–293.

Hyun, C. U., Lee, J. S., & Lee, I. (2013). Assessment of hydrogen fluoride damage to vegetation using optical remote sensing data. *International Society for Photogrammetry and Remote Sensing Spatial Inf. Sci*., XL-7/W2, 115–118.

IGM, I. (2015). Airplane equipped with an UltraCam XP aerial digital cam-era installed on a gyro-stabilized mount. In *Pattern Recognition: 7th Mexican Conference*, MCPR 2015, June 24-27, 2015, Mexico City, Mexico Proceedings (Vol. 9116, p. 145). Cham, Springer.

Ittelson, W. H. (1960). *Visual Space Perception*. New York, Springer.

Jin, X., Gallagher, A., Cao, L., Luo, J., & Han, J. (2010). The wisdom of social multimedia: using Flickr for prediction and forecast. In *Proceedings of the 18th ACM international conference on Multimedia*, October 2010, ACM (pp. 1235–1244).

Kamusoko, C., & Gamba, J. (2015). Simulating urban growth using a Random Forest-Cellular Automata (RF-CA) model. *ISPRS International Journal of Geo-Information*, 4(2), 447–470.

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiol.* 72, 558–569.

Li, L., Goodchild, M. F., & Xu, B. (2013). Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and geographic information science*, 40(2), 61–77.

Lian, D., Zhao, C., Xie, X., Sun, G., Chen, E., & Rui, Y. (2014, August). GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, August 2014, ACM (pp. 831–840).

Lindberg, F., & Grimmond, C. S. B. (2010). Continuous sky view factor maps from high resolution urban digital elevation models. *Climate Research*, 42(3), 177–183.

Lindberg, F., Grimmond, C. S. B., Gabey, A., Huang, B., Kent, C. W., Sun, T., & Chang, Y. Y. (2017). UMEP-An integrated tool for city-based climate services. In *21ST International Congress of Bio-*

meteorology, September 3-6, 2017, Durham, United Kingdom (p. 51). International Society of Biometerology.
Lynch, K. (1960). *The image of the city*. Cambridge, MIT Pres.
Moura, F., Cambra, P., & Gonçalves, A. B. (2017). Measuring walkability for distinct pedestrian groups with a participatory assessment method: A case study in Lisbon. *Landscape and Urban Planning*, 157, 282–296.
Naik, N., Philipoom, J., Raskar, R., & Hidalgo, C. (2014). Streetscore-predicting the perceived safety of one million streetscapes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 779–785).
Nasrollahi, N., & Shokri, E. (2016). Daylight illuminance in urban environments for visual comfort and energy performance. *Renewable and Sustainable Energy Reviews*, 66, 861–874.
Oteros-Rozas, E., Martín-López, B., Fagerholm, N., Bieling, C., & Plieninger, T. (2018). Using social media photos to explore the relation between cultural ecosystem services and landscape features across five European sites. *Ecological Indicators*, 94, 74–86.
Patil, G. P., Myers, W. L., Luo, Z., Johnson, G. D., & Taillie, C. (2000). Multiscale assessment of landscapes and watersheds with synoptic multivariate spatial data in environmental and ecological statistics. *Mathematical and Computer Modelling*, 32(1-2), 257–272.
Tenerelli, P., Demšar, U., & Luque, S. (2016). Crowdsourcing indicators for cultural ecosystem services: a geographically weighted approach for mountain landscapes. *Ecological Indicators*, 64, 237–248.
Peng, X., & Huang, Z. (2017). A novel popular tourist attraction discovering approach based on geo-tagged social media big data. *ISPRS International Journal of Geo-Information*, 6(7), 216.
Perovic, S., & Folic, N. K. (2012). Visual perception of public open spaces in Niksic. *Procedia-Social and Behavioral Sciences*, 68, 921–933.
Pouya, S., & Dem˙Irel, Ö. (2015). What is a healing garden?. *Akdeniz Üniversitesi Ziraat Fakültesi Dergisi*, 28(1), 5–10.
Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest.
Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(3), e1301.
Quercia, D., Schifanella, R., & Aiello, L. M. (2014). The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city. In *Proceedings of the 25th ACM conference on Hypertext and social media,* September 2014, ACM (pp. 116-125).
Radovic, R. (2003). *Forma grada: osnove, teorija i praksa*. Beograd, Orion art; Novi Sad, Stylos.
Rodgers III, J. C., Murrah, A. W., & Cooke, W. H. (2009). The impact of hurricane Katrina on the coastal vegetation of the Weeks Bay Reserve, Alabama from NDVI Data. *Estuaries and Coasts*, 32(3), 496–507.
Rossi, A. (1966). *The Architecture of the City*. Cambridge, The MIT Press.
Scornet, E. (2017). Tuning parameters in random forests. In *ESAIM: Proceedings and Surveys*, 60 (pp. 144–162).
Segal, M., & Xiao, Y. (2011). Multivariate random forests. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 80–87.
Spreiregen, P. D. (1965). *Urban design: the architecture of towns and cities*. New York, McGraw-Hill.
Stamps III, A. E. (2001). Evaluating enclosure in urban sites. *Landscape and Urban Planning*, 57(1), 25–42.
Stea, D. (1978). Environmental perception and cognition: toward a model for :mental maps. In Kaplan, R., & Kaplan, S. (Eds.). *Humanscape: environment for people*. Boston, MIT Press.
Steinmetz-Wood, M., Velauthapillai, K., O'Brien, G., & Ross, N. A. (2019). Assessing the micro-scale environment using Google Street View: the virtual systematic tool for evaluating pedestrian streetscapes (virtual-STEPS). *BMC Public Health*, 19(1), 1–11.
Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T. Zeileis, A. (2008). Conditional variable importance for random forests. *BMC bioinformatics*, 9(1), 307.
Talavera-Garcia, R., & Soria-Lara, J. A. (2015). Q-PLOS, developing an alternative walking index. A method based on urban design quality. *Cities*, 45, 7–17.

Team, R. C. (2015). *R: A language and environment for statistical computing*. Vienna, Austria, R Foundation for Statistical Computing.

Walden-Schreiner, C., Leung, Y. F., & Tateosian, L. (2018). Digital footprints: incorporating crowdsourced geographic information for protected area management. *Applied Geography*, 90, 44–54.

Welling, S. H., Refsgaard, H. H., Brockhoff, P. B., & Clemmensen, L. H. (2016). Forest floor visualizations of random forests. *arXiv*, arXiv:1605.09196. Available at https://arxiv.org/pdf/1605.09196.pdf (Accessed 1 July 2021).

Winters, N. B. (1986). *Architecture is elementary: visual thinking through architectural concepts*. Salt Lake City, Gibbs Smith Publishers.

Yang, P. P. J., Putra, S. Y., & Li, W. (2007). Viewsphere: a GIS-based 3D visibility analysis for urban design evaluation. *Environment and Planning B: Planning and Design*, 34(6), 971–992.

Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: using machine learning algorithms and Google Street View imagery. *Applied geography*, 76, 147–153.

Yoshimura, N., & Hiura, T. (2017). Demand and supply of cultural ecosystem services: use of Geo-tagged photos to map the aesthetic value of landscapes in Hokkaido. *Ecosystem Services*, 24, 68–78.

Zhou, H., He, S., Cai, Y., Wang, M., & Su, S. (2019). Social inequalities in neighborhood visual walkability: using Street View imagery and deep learning technologies to facilitate healthy city planning. *Sustainable Cities and Society*, 50, 101605.

Zhou, X., Xu, C., & Kimmons, B. (2015). Detecting tourism destinations using scalable geospatial analysis based on cloud computing platform. *Computers, Environment and Urban Systems*, 54, 144-153.

Zuniga-Teran, A. A., Orr, B. J., Gimblett, R. H., Chalfoun, N. V., Going, S. B., Guertin, D. P., & Marsh, S. E. (2016). Designing healthy communities: A walkability analysis of LEED-ND. *Frontiers of Architectural Research*, 5(4), 433–452.