

Article

From Raw Data to Meaningful Information: A Representational Approach to Cadastral Databases in Relation to Urban Planning

Francesc Valls Dalmau ^{1,*}, Pilar Garcia-Almirall ², Ernest Redondo Domínguez ¹ and David Fonseca Escudero ³

¹ Department of Architectural Representation and Visual Analysis I, Universitat Politècnica de Catalunya, 649 Diagonal Av., Barcelona 08028, Spain; E-Mail: ernesto.redondo@upc.edu

² Department of Architectural Technology I, Universitat Politècnica de Catalunya, 649 Diagonal Av., Barcelona 08028, Spain; E-Mail: pilar.garcia-almirall@upc.edu

³ Architecture Department, La Salle Campus Barcelona, Ramon Llull University, 2 Quatre Camins St., Barcelona 08022, Spain; E-Mail: fonsi@salle.url.edu

* Author to whom correspondence should be addressed; E-Mail: francesc.valls@upc.edu; Tel.: +34-934011933; Fax: +34-934016332.

External Editors: Robert Amor and Salvatore Flavio Pileggi

Received: 15 June 2014; in revised form: 19 September 2014 / Accepted: 15 October 2014 /

Published: 24 October 2014

Abstract: Digesting the data hose that cities are constantly producing is complex; data is usually structured with different criteria, which makes comparative analysis of multiple cities challenging. However, the publicly available data from the Spanish cadaster contains urban information in a documented format with common semantics for the whole territory, which makes these analyses possible. This paper uses the information about the 3D geometry of buildings, their use and their year of construction, stored in cadastral databases, to study the relation between the built environment (what the city is) and the urban plan (what the city wants to become), translating the concepts of the cadastral data into the semantics of the urban plan. Different representation techniques to better understand the city from the pedestrians' point of view and to communicate this information more effectively are also discussed.

Keywords: 3D GIS; urban planning; representation; cadaster; smart visualization; urban regeneration; semantics translation

1. Introduction

Urban morphology is constantly evolving: buildings are built, demolished, and remodeled altering the landscape of our cities through the collective action of their inhabitants. Urban Plans describe what we want our cities to be, undergoing revisions as we change the vision of the future of our cities. The analysis of the complex interaction throughout time between what the city is and what it wants to become [1–3] allows urban planners to understand both past and present state of the city to plan for a better future.

Urban planning is a complex process that involves actors from different disciplines collaborating and managing heterogeneous data from different sources [4]. The visual representation of data allows the exploration of complex datasets [5], and can improve the collaboration of the teams involved through better communication [6].

The purpose of this paper is to present a methodology to provide urban planners with analytic visualization tools to better understand the complexity of urban renewal processes [7], using publicly available cadastral data. This approach makes possible to use the methodology in multiple Spanish regions since cadastral data covers the whole Spanish territory with a uniform semantic criteria.

The methodology extends a previously developed work [8], using a multifaceted approach to the study of urban morphology in relation to planning regulations in multiple dimensions (spatial, temporal and conceptual), to visually analyze and communicate the complexity of urban development.

After a review of the literature on data representation and specifically on the representation of 3D and temporal data, followed by the introduction of the case of study, Section 2 outlines the harmonization process applied to the data due to semantic and geometric mismatches between the cadastral and planning information, Section 3 describes the methodology of the visualization techniques developed, Section 4 presents a discussion of the results obtained and their applicability, and Section 5 presents our conclusions.

1.1. Representational Approach

Cities produce a constant flow of heterogeneously structured data [9,10]; to extract and communicate meaningful information from this data it has to be converted to a medium that can be understood by human beings, which is not a trivial process [11].

From the late 1960s with the work of Bertin [12,13], a systematic approach to the study of the most adequate way to represent statistical data was developed by Cleveland [14–18], and the study of the aesthetic and conceptual properties of graphics as a vehicle to deliver meaningful information was studied by Tufte [19–21]. More recently, with the capacity of computers to store, manage and display large amounts of data, a more computer-centric approach to representing data has emerged [22–25], with the focus shifting to explaining to the machine how to map concepts to graphical elements. According to the grammar of graphics [23] definition, the representation of spatial data would limit the number of “aesthetics” because the shape and position of the graphical elements is mapped to the location and geometry of the represented elements themselves, leaving usually just color [26], and in some cases pattern, to convey the desired information.

The representation of geographic data [27–29] has specific challenges because it deals with different kinds of information [30] depending on the nature of the data, making the visualization of historical [31]

and conceptual [32] information challenging, although the results can be appreciated as an art form for their aesthetic properties [33,34].

This paper tries to address some of the difficulties in the representation of urban data, specifically regarding the interaction between built reality and urban planning, and how it affects its three-dimensional morphology, the activities that take place in the city, and its evolution in time, considering specially its impact on the pedestrians' perception of the city.

The representation of three categories of information of the building stock was explored in the case of study: spatial (the 3D volume of buildings and planning regulations), temporal (the year of construction) and conceptual (dominant use inside buildings).

The representation of urban phenomena in three dimensions is challenging, while advances have been made in the realistic representation of cities using procedural methods [35–38] or data captured from reality [39], the use of the third dimension for aesthetic purposes [40], to convey information through the abstract representation of the 3D volumes [41], or the generalization or simplification of the 3D geometry [42,43] remains less explored.

Visualizing the temporal component of spatial data is even more complex than the representation of three-dimensional data. It has been conceptually laid out [44,45] and some examples of implementation in 4D cadastral data can be found to document the lifecycle of the building stock [46].

The pedestrians' perception of the built environment, from the street level, has been described from the functional [47] and visual [48] perspectives; recently the use of geotagged images have allowed volunteers to evaluate the subjective perception of different neighborhoods [49].

Finally, some techniques have been used recently introduced to explore complex spatial 2D data, using conditioned choropleth maps to visualize multidimensional data [50] and arrays of spatially positioned maps [51] to better display dense OD matrices.

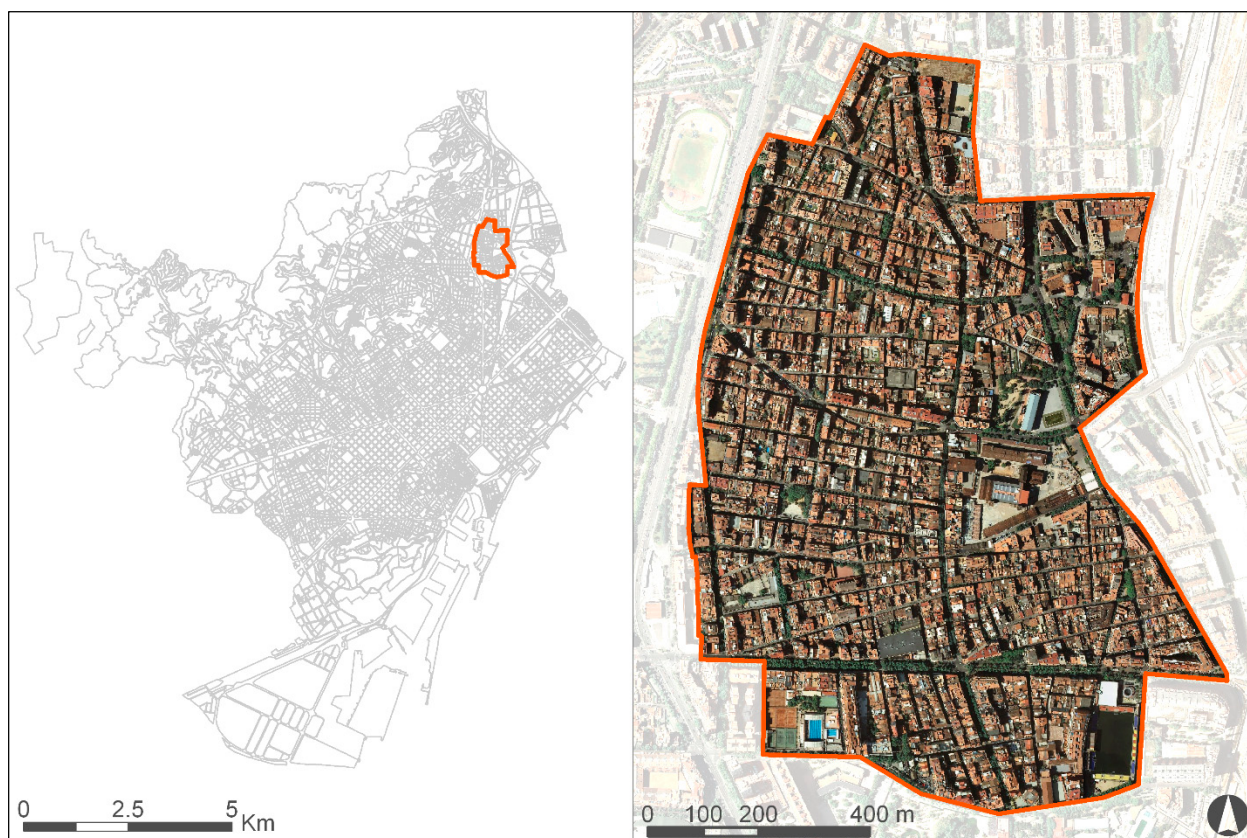
1.2. Case of Study

The case of study chosen for the development of the methodology was the old quarter of the Sant Andreu District in Barcelona. The buildings in the area of study (Figure 1) are part of the center of the former town of Sant Andreu, from which the district takes its name, and the development it underwent in the 19th century when it was incorporated to Barcelona.

This historic development resulted in a complex urban structure suitable to use as a workbench to test the methodology for the representation of 3D features, their temporal development and their perception from the pedestrians' point of view:

- Semantic correction of height information of cadastral and planning databases (Section 2.3)
- 3D representation of height information on a 2D surface, compared to traditional 2D representation (Sections 3.1 and 4.3)
- Occlusion of the built volumes from the street (Sections 3.2 and 4.5) and onto neighboring volumes (Sections 3.3 and 4.6)
- Visualization of the temporal evolution of the building stock (Section 4.8)
- Representation main of uses on the ground floor (Sections 3.4 and 4.7)

Figure 1. Area of study in the Sant Andreu District of Barcelona.



2. Materials

The Spanish cadaster provides a comprehensive inventory of all real estate properties in the Spanish territory with a common specification, and includes information about the physical configuration of every building and the uses inside them. These are some of the main aspects urban planning regulates: the mass of the buildings (their size and shape) and the allowed uses according to zoning regulations.

The comprehensive information included in cadastral data outweighed some of its limitations: (a) it contains data about the year of construction that spans more than a century (b) has building height information available; and (c) includes use information for each unit.

The objective of this study was to assess the information stored in the cadaster databases, especially in relation to the urban planning regulations, and the interpretation of the data stored in these two sources with different semantics, to be able to relate them spatially and conceptually.

2.1. Cadastral Data Import and Normalization

The Cadastral data used in this paper is publicly available in the Cadastral Electronic Site where it can be freely downloaded, excluding personal identifiable information to preserve the anonymity of real estate owners. Cartographic data can be downloaded for any municipality as a standard shapefile, but information without graphical representation is only provided in a special legacy American Standard Code for Information Interchange (ASCII) fixed width format instead of a tabular or Extensible Markup Language (XML) or CityGML [52] format.

2.1.1. Cadastral Data Import

The format in which the data was stored was a flat ASCII file with 2,856,819 lines (for the municipality of Barcelona) with 1000 characters per line. Three classes of information needed, which would correspond to three different tables in a relational database, were interleaved in the data stream: parcels (76,556 records), buildings (93,713 records) and private units (1,451,737 records). These three tables are structured hierarchically by design, where each private unit belongs to one building and each building belongs to one parcel, and each record of the lower rank table has its own unique key and the foreign key of the higher rank table it belongs to. A script had to be developed with the R programming language to parse the ASCII file and extract the raw data needed: use, floor number, year of construction and unique id.

2.1.2. Cadastral Data Normalization

The encoding of the raw data corresponding to both use and floor number had to be normalized in order to be usable for the purpose of this study.

According to the format specification, the position in the building of each unit, measured in number of floors from the ground floor is encoded as a three character alphanumeric string. In order to obtain a numeric value for the floor, a dictionary of key-value pairs was used to translate each unique key to a numeric value measuring the number of floors above street level (negative integers in the case of subterranean floors).

Two special cases (attic and over-attic) were impossible to translate with a dictionary since its interpretation depended on their context: their value is one floor (attic) or two floors (over-attic) higher than the highest of the other units in the same parcel. For these cases, a series of queries in a Structured Query Language (SQL) database had to be built to assign the correct height value.

The format specification encodes the use of each unit with three levels of detail in a tree-like structure, where each level has a more specific semantic detail than the previous one: level 1 has 10 categories, level 2 has 31 categories, and level 3 has 70 categories. Levels 2 and 3 were discarded because for the purposes of the study the detail of level 1 was adequate.

2.2. Parcel Level Aggregation

The resulting tables were not displayable for representation purposes because the cadaster does not provide an associated geometry. The most detailed geometry provided is the parcel layer, and as a consequence the data in the tables had to be summarized at the parcel level for the attributes of year of construction and use, to be able to analyze and visualize spatial patterns in the data.

2.2.1. Year of Construction

The year of construction of the units that belong to parcel is not necessarily the same for all items. Two measures were calculated to summarize this data at the parcel level: the most recent year and the weighted average.

Since we were dealing with two time series where buildings are built and regulations are modified, the most recent year of construction of all units in a parcel was necessary in order to be able to compare

which parcels were compliant at the time of a given regulation change that could potentially modify the allowed height of any given parcel.

On the other hand, a synthetic measure of building age had to be computed for each parcel. A simple arithmetic mean was not suitable because the units did not have the same built area and should contribute proportionally to their size in the average using a weighted average with the following formula (1):

$$Year_{\text{Parcel}} = \frac{\sum_{\text{Unit} \in \text{Parcel}} Year_{\text{Unit}} \cdot BuiltArea_{\text{Unit}}}{\sum_{\text{Unit} \in \text{Parcel}} BuiltArea_{\text{Unit}}} \quad (1)$$

2.2.2. Dominant Use

Each parcel can contain multiple units with different uses in each floor. The use attribute of each unit had to be summarized at the parcel level for each floor to obtain its dominant use and its degree of this dominance over the rest of the uses.

For each floor, a contingency table with the built area for each parcel and use was generated, and the dominant use (the use with a greater built area) was stored as an attribute of the parcel along with its percentage of the total built area (marginal total of all uses for each parcel).

2.2.3. Overbuilt and Underbuilt Area Measurement

It is not legally allowed to compensate overbuilt volumes with underbuilt ones inside a parcel, and accordingly aggregate calculations had to be performed independently for both situations to avoid the aggregate operations adding positive and negative numbers. This operation would be mathematically correct but it is not possible according to the regulations, which state that an unwarranted height in a part of a building is an offence independently of the same parcel having another part below the authorized height.

Formulae 2 to 6 show the aggregation operations to calculate for each parcel: the total built area (2), the maximum allowed built area in zones (3), the overbuilt area in zones (4), the underbuilt area in zones (5), and the overbuilt area in systems (6).

$$Real\ Built\ Area_{\text{Parcel}} = \sum_{\text{Subp} \in \text{Parcel}} HB_{\text{Subp}} \cdot A_{\text{Subp}} \quad (2)$$

$$Allowed\ Built\ Area_{\text{Parcel}} = \sum_{\text{FragZ} \in \text{Parcel}} HP_{\text{FragZ}} \cdot A_{\text{FragZ}} \quad (3)$$

$$Overbuilt_{\text{Parcel}} = \sum_{\text{FragZ} \in \text{Parcel}} (HB_{\text{FragZ}} - HP_{\text{FragZ}}) \cdot A_{\text{FragZ}} \quad (4)$$

$$Underbuilt_{\text{Parcel}} = \sum_{\text{FragZ} \in \text{Parcel}} |HB_{\text{FragZ}} - HP_{\text{FragZ}}| \cdot A_{\text{FragZ}} \quad (5)$$

$$Overbuilt\ in\ Systems_{\text{Parcel}} = \sum_{\text{FragS} \in \text{Parcel}} HB_{\text{FragS}} \cdot A_{\text{FragS}} \quad (6)$$

Using the following notation:

- *HB*: Real height of the building, measured in number of floors above street level.
- *HP*: Maximum height according to the plan, measured in number of floors above street level.
- *A*: Area of a polygon entity (surface of land occupied).
- *Parc*: Set of parcels inside the area of study.
- *Subp*: Set of sub-parcels.
- *FragZ*: Set of fragments from the spatial intersection between parcel and planning layers (in zones).
- *FragS*: Set of fragments from the spatial difference between parcel and planning layers (in systems).

2.3. Cadastral Cartography

The cartographic information from the cadaster is distributed in 13 shapefiles that thematically represent information of a different nature that the institution manages. Two of these shapefiles were used in this study: the built volume geometry and the parcels geometry. The built volume layer was used to relate the existing height to the maximum height allowed by the urban plan layer using an overlay operation; the parcels layer were used for representational purposes because they are the only entities that are capable to display aggregated data from non-graphical tables linked through the parcel unique identifier.

In the built volume geometry, the height of buildings is not stored as true 3D geometry [53,54] but as 2.5D cartography in a sub-parcel dataset (being sub-parcels pieces within a parcel with a distinct height from neighboring sub-parcels). Height information is stored as an alphanumeric string, which encodes several pieces of information about the sub-parcel, including the number of floors below and above street level. For example, a sub-parcel with a “-II+V” attribute has two subterranean floors and five floors above street level.

2.3.1. Height Information Correction

The height code had to be translated using a dictionary with three fields (code, height above ground, height below ground) through the interpretation of the documentation provided by the cadaster; however, since the heights had to be compared to the planning regulations, it was imperative that they were semantically comparable and not necessarily a direct representation of reality (as if it was a 3D model).

2.3.2. Semantically Adjusting Cartographic Data to Planning Regulations

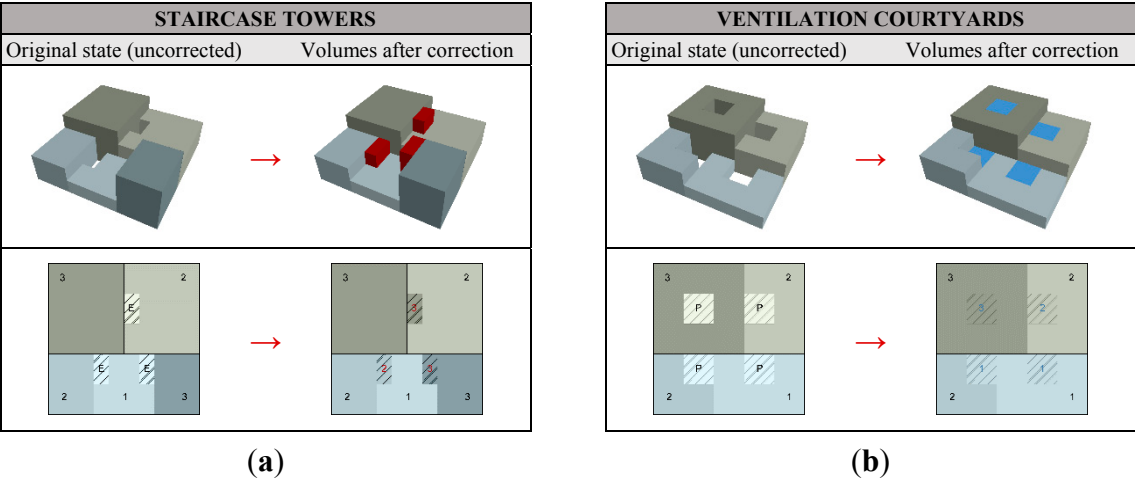
To be consistent with the nature of the planning regulations, some adjustments had to be made regarding (a) balconies; (b) ventilation courtyards; and (c) staircase towers:

- (a) The volumes of balconies over the streets were disregarded because the Metropolitan regulations do not directly regulate the maximum height of these volumes but instead the buildings they are attached to. Therefore, a balcony would not be allowed only as a consequence of the floor it was attached to not being allowed.

- (b) Ventilation courtyards are a typological device that allows placing uses like kitchens and bedrooms in the core of the building and still be able to get ventilation and natural light. Although these courtyards are geometrically voids, from the planning regulations point of view, they are a design decision of the architect of the building and should not be considered as such but instead built up to the lowest volume they serve in its own parcel.
- (c) Staircase towers are not assigned a height value but are considered built volume. Since they have to give access to the adjacent volumes, its assumed height should be at least the height of the highest volume it serves in its own parcel.

Since ventilation courtyards accounted for almost 20% of the area of all sub-parcels and staircase towers for another 2%, to get accurate results a methodology [8] had to be developed to automatically assign height values to this types of entities from their spatial context (Figure 2).

Figure 2. Correction of staircase towers (a) and ventilation courtyards (b).



The calculation of the height to be assigned involved two topological relationships: (a) adjacency to other polygons but (b) considering only polygons inside the same parcel. Since the Geographic Information System (GIS) used was non-topological, a methodology had to be implemented using Structured Query Language (SQL) from adjacency information stored in long form (origin, destination, value) instead of as a sparse matrix (Figure 3).

2.4. Planning Data

The Barcelona Metropolitan Area Urban Plan is in effect since 1976 [55] and its cartography relative to the maximum height must be interpreted [56] from the textual description according to multiple parameters: width of the street(s) the parcel faces, slope of the street(s), typology of the street, position relative to the corner of a city block, geometry this corner (arc radius, chamfer distance), zoning regulations, and the principle of the regulation granting the lesser amount of built area where there are conflicting interpretations.

The interpretation of the regulations regarding the maximum heights was performed in collaboration with the staff in the Urban Studies Bureau of the Urban Planning Department of the Barcelona City Council.

The height information of the planning layer was compared to the height information of the actual buildings layer using two Boolean spatial operations: (a) a spatial intersection, the result of which was the fragments of sub-parcel inside zones and (b) a spatial difference, the result of which was the fragments of sub-parcel inside systems (roads and parks). With the result of the overlay operations it was possible to determine the conformity to the urban plan for each of the resulting fragments (Table 1) from their building height (HB) and planned height (HP). The layers in the overlay operation had some metric imprecision which resulted in some sliver polygons that had to be cleaned up.

Figure 3. Diagram of the process to obtain the staircase towers (top) and ventilation courtyards (bottom) heights from neighboring geometry.

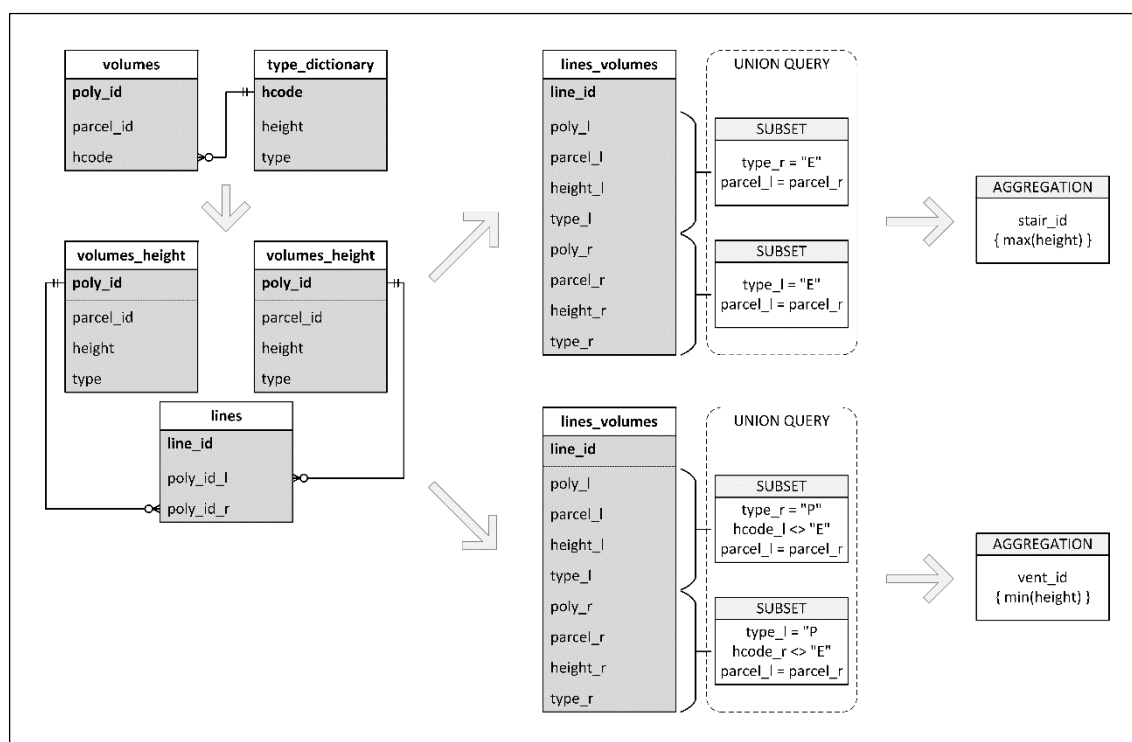


Table 1. Types of fragments from the overlay operations.

Fragment	Plan Entity	Condition	Operation	Symbology
Underbuilt	Zones	$HB - HP < 0$	Intersection	Blue hues
Conformant	Zones	$HB = HP$	Intersection	Grey
Overbuilt	Zones	$HB - HP > 0$	Intersection	Pink hues
Overbuilt	Systems	$HB > 0$	Difference	Dark green
Conformant	Systems	$HB = 0$	Difference	Light green

3. Methods

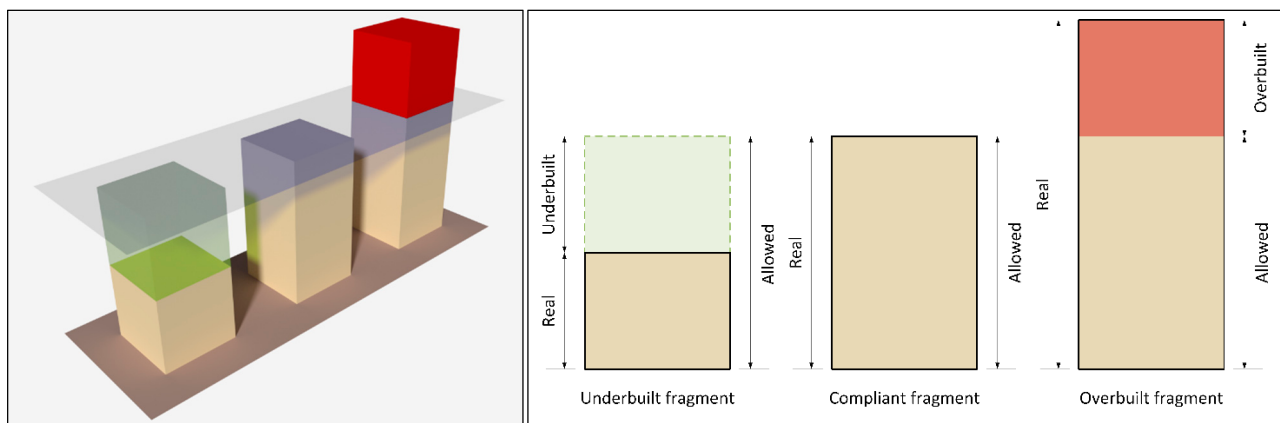
Data is processed by machines but ultimately consumed by people. Different methodologies were developed to communicate effectively the results of the study, to make them more easily understandable and to give deeper insight to the intended audience of urban planners and architects, not necessarily familiar with GIS technology but highly proficient in multiple representation techniques. Different

strategies were developed to better convey the results, using several graphic approaches to visualize data in a more natural and human-centric way.

3.1. Representation of Overlapping Information in 3D

The morphology of the buildings is not exactly what the planning regulations mandate; some parts of the city could be higher (underbuilt volumes) while some other parts are higher than what the regulations allow. It is difficult to display this information on a map because its abstract planar representation does not match the way we experience the urban environment. A criterion was developed to display the overlapping information of overbuilt and underbuilt fragments (Figure 4). The third dimension allowed the authors to display overlapping information without having to resort to 2D representation constructs such as transparency or hatching.

Figure 4. Height interpretation in the 3D model.



The result was akin to a 3D stacked bar graph, where the third dimension represented the height of the building, the bars had the shape and the position of the cartographic data it represented, and the floors overbuilt or underbuilt were displayed with a color code, adding a fourth conceptual dimension with the relationship between built and planned volumes.

3.2. Pedestrian-centric Representation of Height in 2D

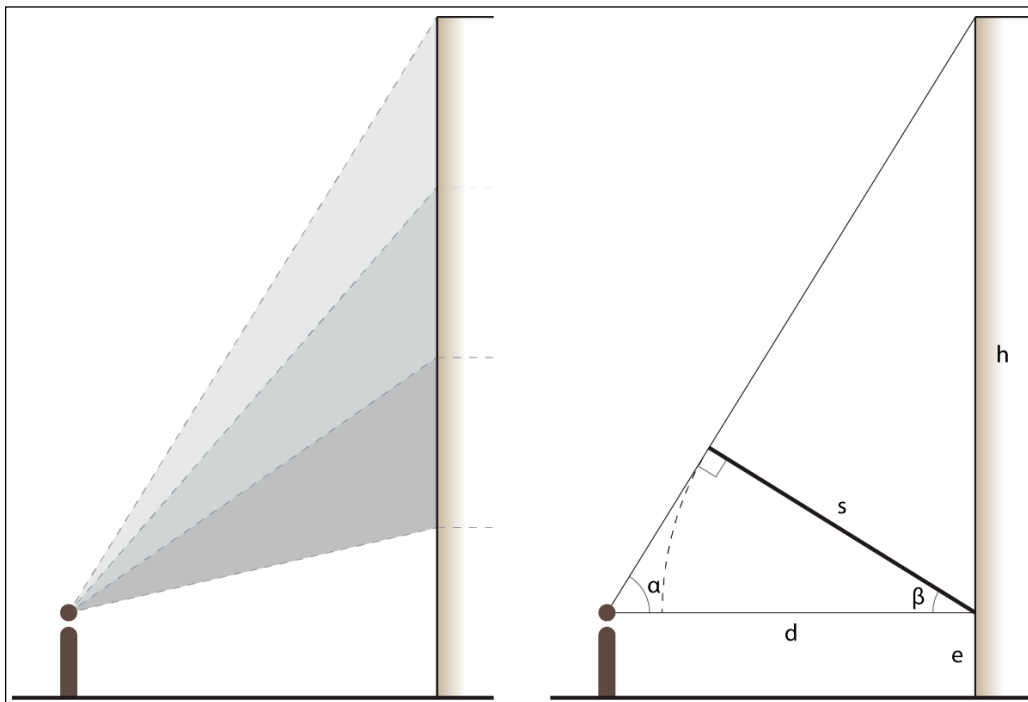
The spatial configuration from the pedestrians' point of view has been studied in the horizontal plane [57,58]. To visualize the perception pedestrians have of the height of buildings when walking along a street, a new representation technique was developed that allowed to show in a map the apparent surface that occludes the view of the sky from a pedestrian point of view.

The occlusion produced by the height of the buildings depends on the position of the highest floor relative to the eye level of the observer and it does not increase linearly with each additional floor (Figure 5, left).

The length of the segment that is the projection of the shortest plane that can occlude the highest visible point of the building (Figure 5, right) was calculated with the following formulae (7) and was implemented in Python (Code 1).

$$\begin{cases} \alpha = \tan^{-1} \frac{h-e}{d} \\ \beta = \pi/2 - \alpha \\ s = d \cdot \cos \beta \end{cases} \quad (7)$$

Figure 5. Diminishing occlusion by height (left), occlusion projection calculation (right).

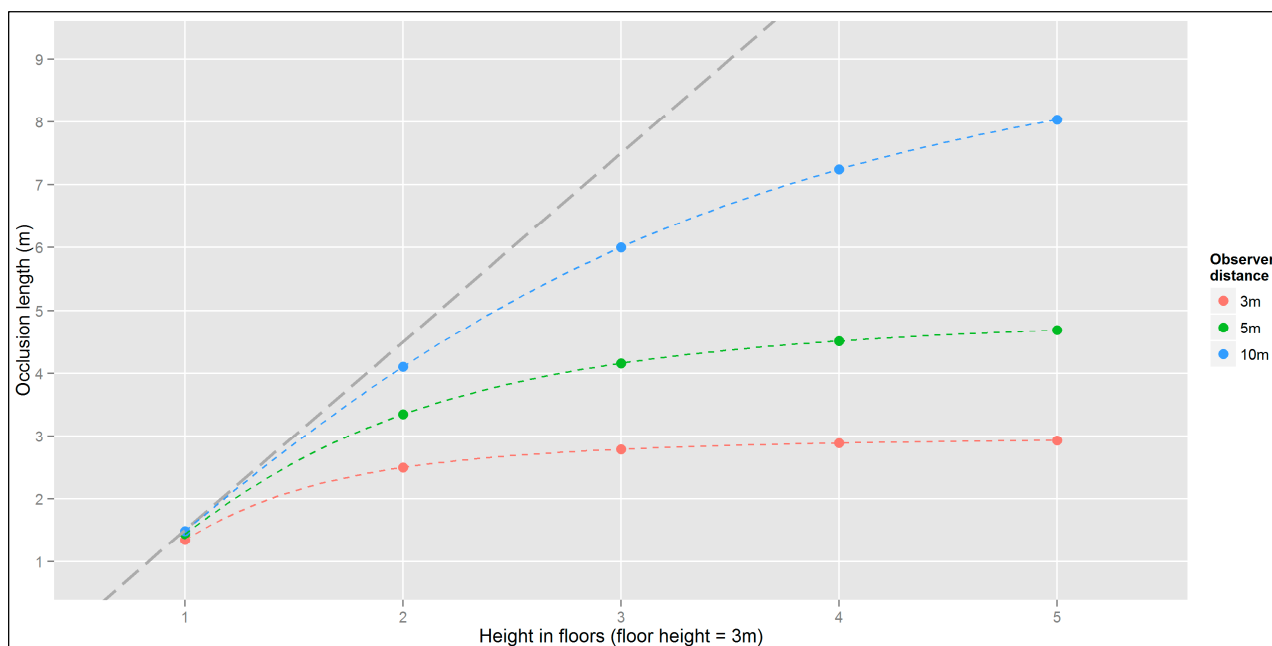


Code 1. Python function to calculate occlusion projection.

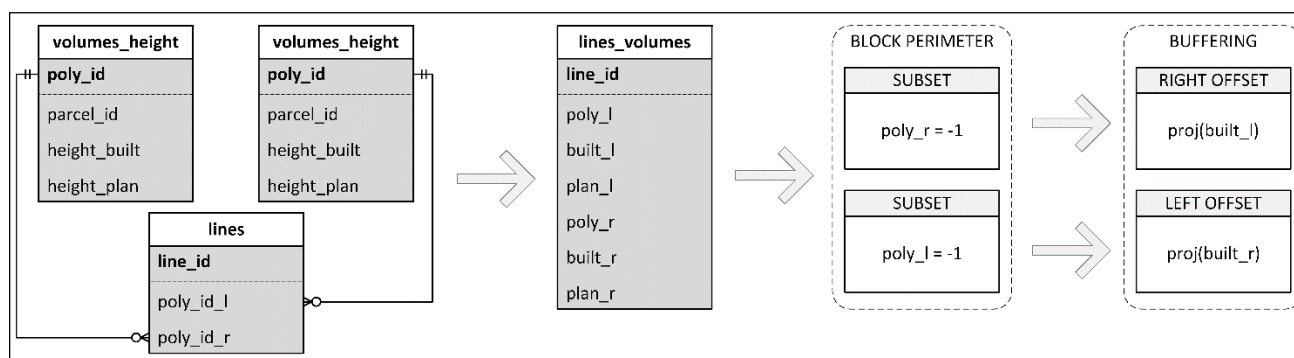
```
def segment(floors, dist, eye):
    # All numbers are assumed positive
    if floors == 0:
        return 0
    else:
        alpha = math.atan((floors * 3 - eye) / dist)
        beta = math.pi / 2 - alpha
        seg = dist * math.cos(beta)
        return seg
```

The calculation and has the following properties (Figure 6):

- For $e = 0$ is the apparent surface as seen from the pavement
- When d approaches infinity and $e = 0$, d is the projection of the height on the horizontal plane ($d = h$)
- As d increases, it has the oblique asymptote $y = h \cdot x - e$

Figure 6. Apparent occlusion for floors 1 to 5 depending on observer distance to façade.

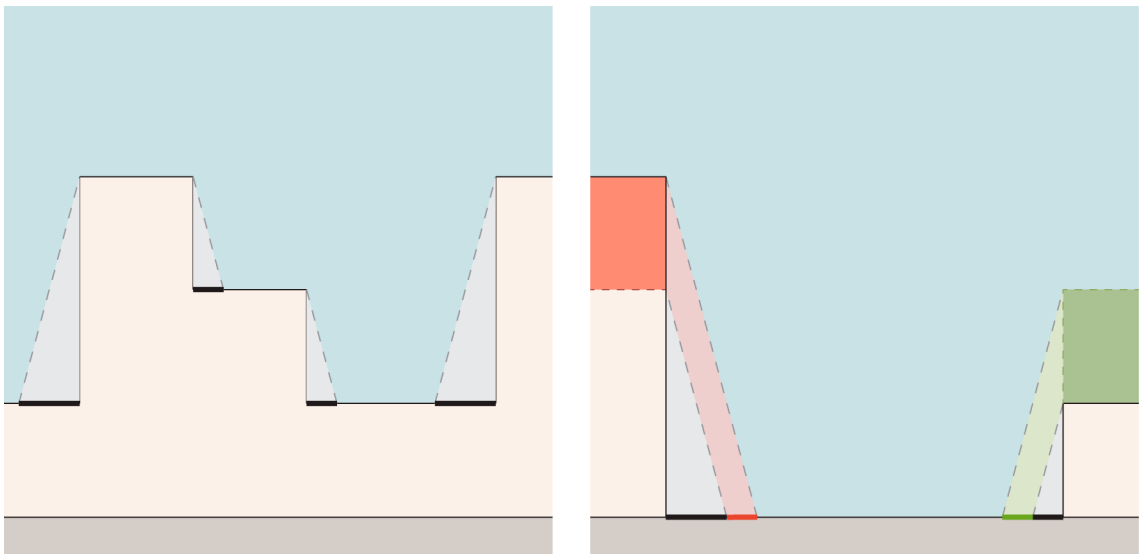
The polygon geometry of the height volumes was decomposed into lines that had the unique identifier of its right and left polygons, and the value of the apparent occlusion of each segment was assigned to the lines without neighbor (the edge of the volumes facing the street) with a series of SQL queries (Figure 7). This value was used as the offset distance in a buffer operation to create the projected entities.

Figure 7. Diagram of the method to project the apparent occlusion onto the ground plane.

3.3. View Occlusion by Neighboring Volumes

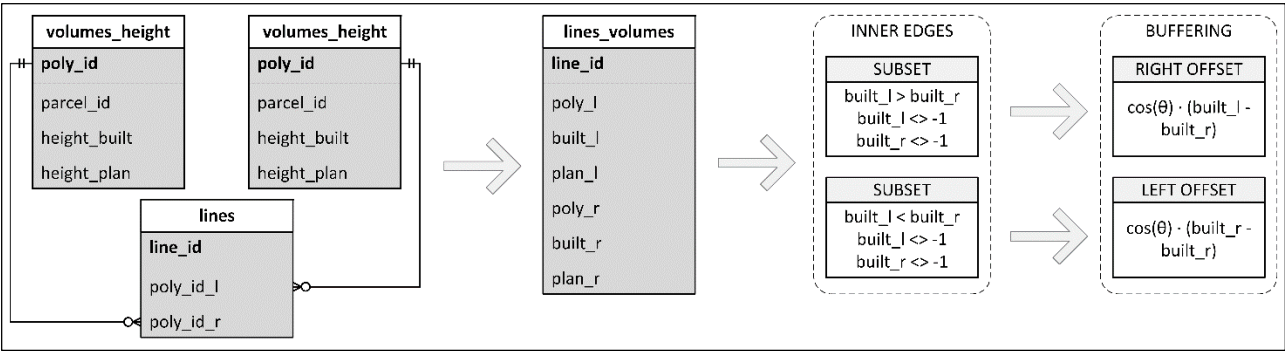
To visualize the occlusion of neighboring volumes, the height of each volume was projected onto the volumes in their immediate vicinity that had a lower height, proportional to the difference of their respective heights (Figure 8).

Figure 8. Projection of relative height onto lower volumes (left) and street (right).



As described in Section 3.2, the geometry was decomposed into lines, and with a series of SQL queries (Figure 9) each line was assigned the value of the number of heights that each neighbor rose above the roof of the lower entity. These values were buffered after being multiplied by the cosine of the angle of incidence to avoid overlapping entities.

Figure 9. Diagram of the method to project the relative height onto lower volumes.

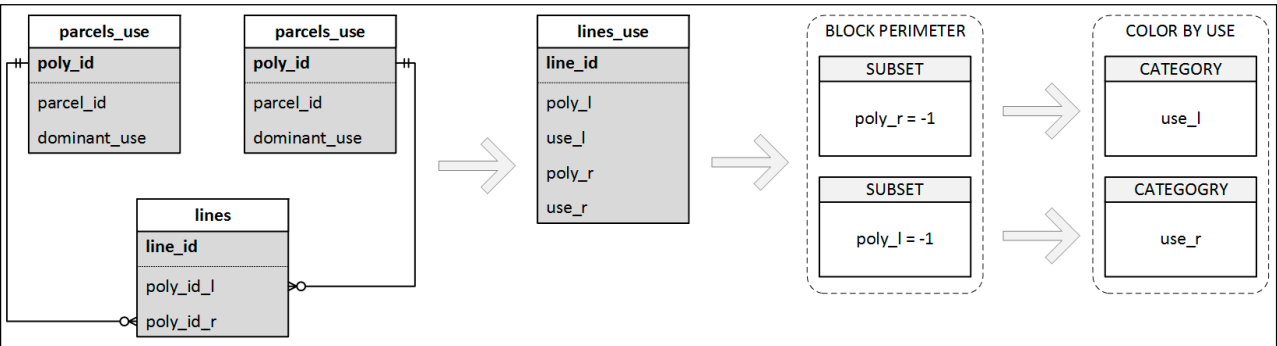


3.4. Dominant Use in Ground Floor from Pedestrian Point of View

From the pedestrian point of view, the ground floor is the part of the building it interacts more with, and defines the character of the street. To investigate the spatial distribution of uses on the ground floor, and identify the presence of commercial streets and residential streets a methodology to visualize the main use along the perimeter of the city blocks was developed.

As described in Sections 3.2 and 3.3 for the volume geometry, the parcel geometry was decomposed into lines that had two fields containing the unique identifiers of the parcels at each side. The lines that had no neighbor (the facades at the limits of the city block) were assigned the value of the dominant use of the corresponding parcel and visualized with a different color for each use (Figure 10).

Figure 10. Diagram of the process to assign the parcel dominant use to its façade(s).



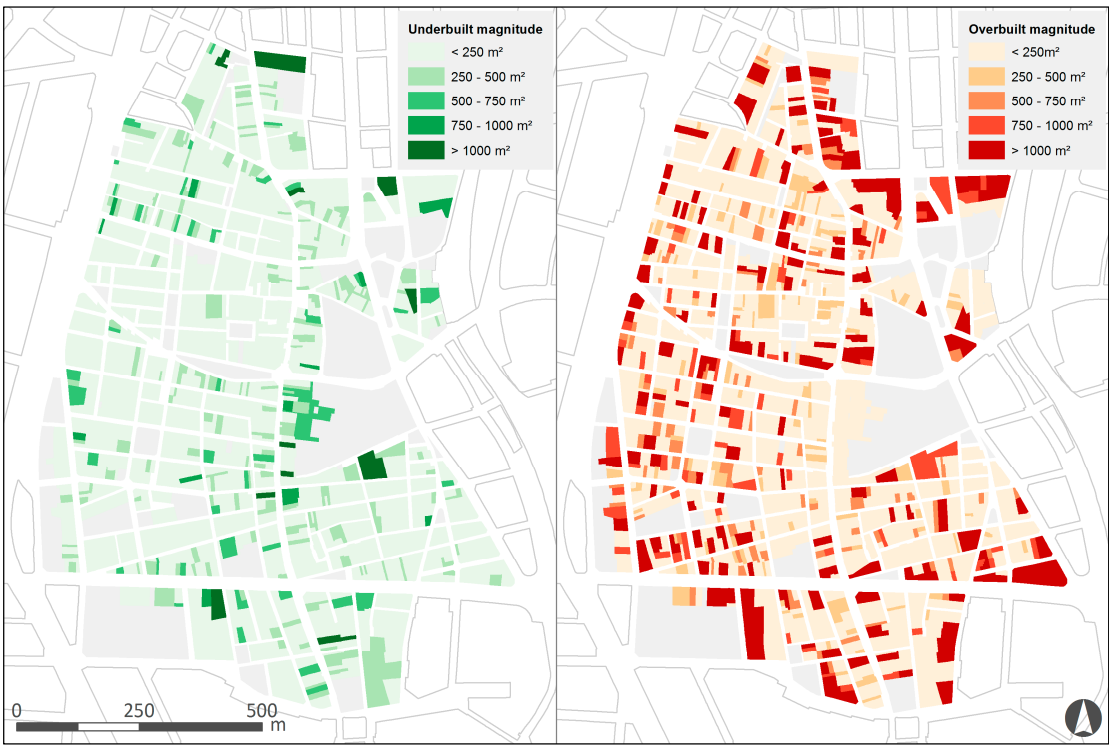
4. Results

Understanding the interactions of pedestrians when walking along a street is crucial for urban planners because it has a direct impact on their perception of the city. In this study, from Cadastral and Planning data, the authors have used different visualization techniques to represent the height of the buildings (in relation to the plan, their visibility from the street, or their visibility from other buildings), the different uses present in the built environment and the age of the edification to give new insights into an entity as complex as a city.

4.1. Representation of Aggregate Data

From the aggregation process described in Section 2.2.3, three measures of the compliance to the building regulations were obtained for every parcel: overbuilt area, underbuilt area and overbuilt area on systems (roads and parks).

Figure 11. Map of amount of underbuilt (left) and overbuilt (right) area per parcel.



To display the magnitudes of these measures in a map and investigate their spatial distribution, a choropleth map for each measure was used (Figure 11), because the position and shape of each parcel could not be used to display any information besides their own location and geometry, and the magnitudes measured had to be displayed through the use of markings (labels, hatches, symbols or color) which were the only aesthetics available.

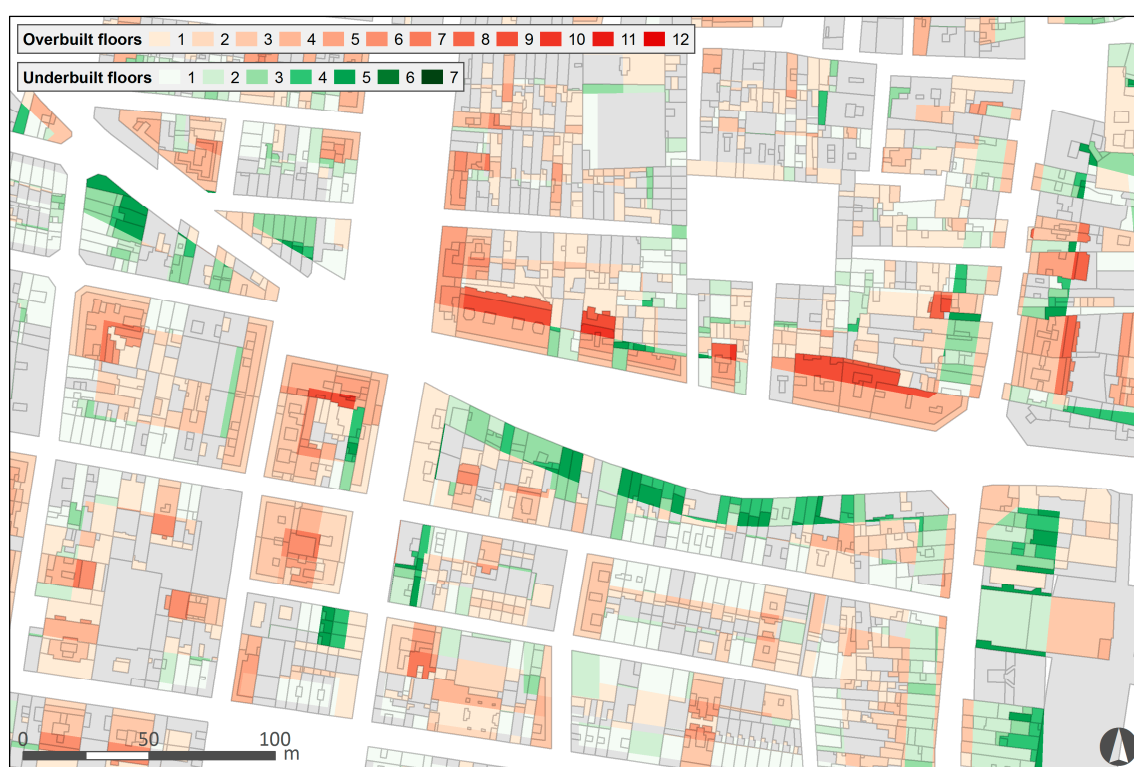
The disadvantage of this approach was that the person reading the maps had to refocus his or her attention in each of the maps individually; to assist the observer to focus on the information, the maps were displayed side by side at the same scale with the same orientation and size, and the breaks in the color categorization were the same for each map.

4.2. Representation of Fragments in 2D

The magnitude of the conformity (measured in area units) for each resulting sub-parcel fragment was represented in a map measuring number of floors beneath or exceeding the allowed height.

To visualize spatial patterns in the distribution of overbuilt and underbuilt zones in the area of study, the map displayed the number of heights above or below the planning regulation as two diverging color ramps for positive and negative numbers, with more saturated colors for higher absolute values and a neutral gray for fragments that measured exactly the same height as the urban plan specified (Figure 12).

Figure 12. Partial map of underbuilt and overbuilt fragments.



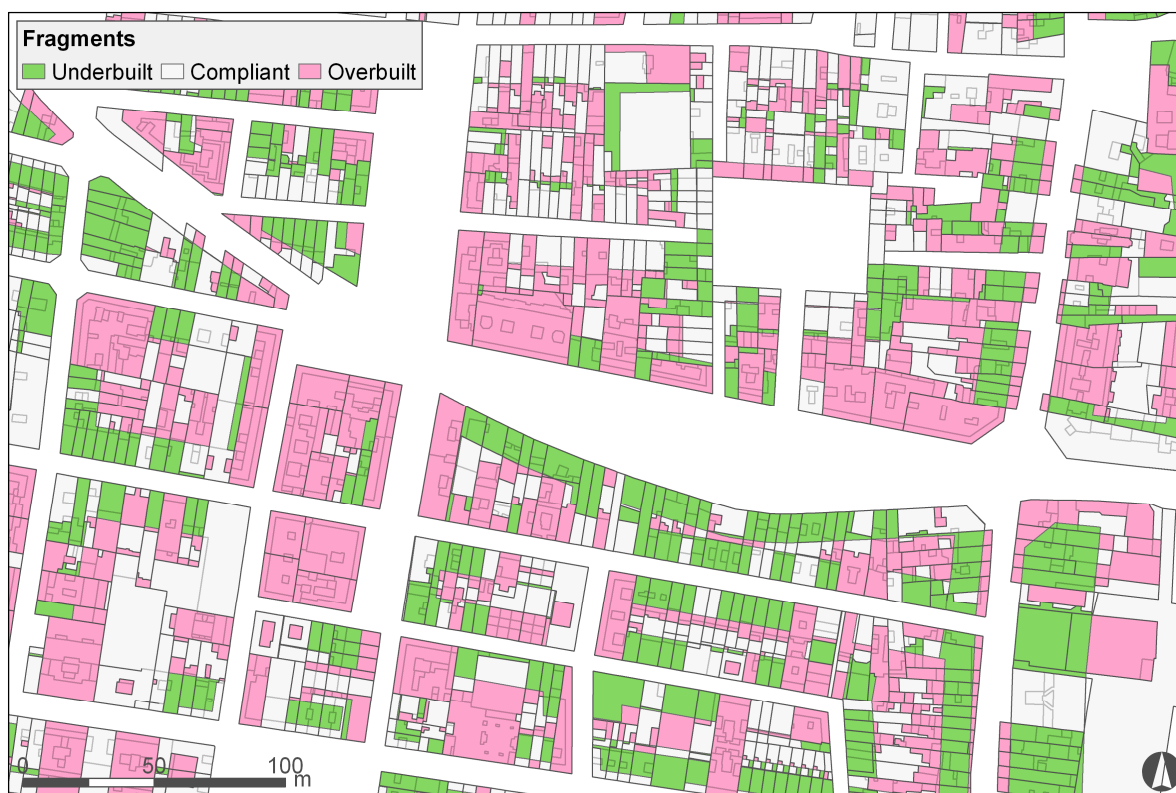
These three categories appeared visually clustered together. To check the strength of this clustering, a Moran Test to measure their spatial autocorrelation was applied to the fragments: they were divided into 3 categories: underbuilt, compliant and overbuilt (Figure 13). Three tests were performed (Table 2): (a) categorized fragments; (b) categorized fragments erasing the boundaries between equal contiguous

categories inside same parcel; and (c) categorized fragments erasing all boundaries between equal contiguous categories.

Table 2. Moran Test results.

Dissolve	Moran's I	Expected	z-score	p-value
None (a)	0.515340	−0.000060	95.380891	$<10^{-6}$
Inside Parcel (b)	0.136500	−0.000183	14.635925	$<10^{-6}$
All (c)	0.108415	−0.000142	12.782081	$<10^{-6}$

Figure 13. Partial map of fragments according to their legal status.



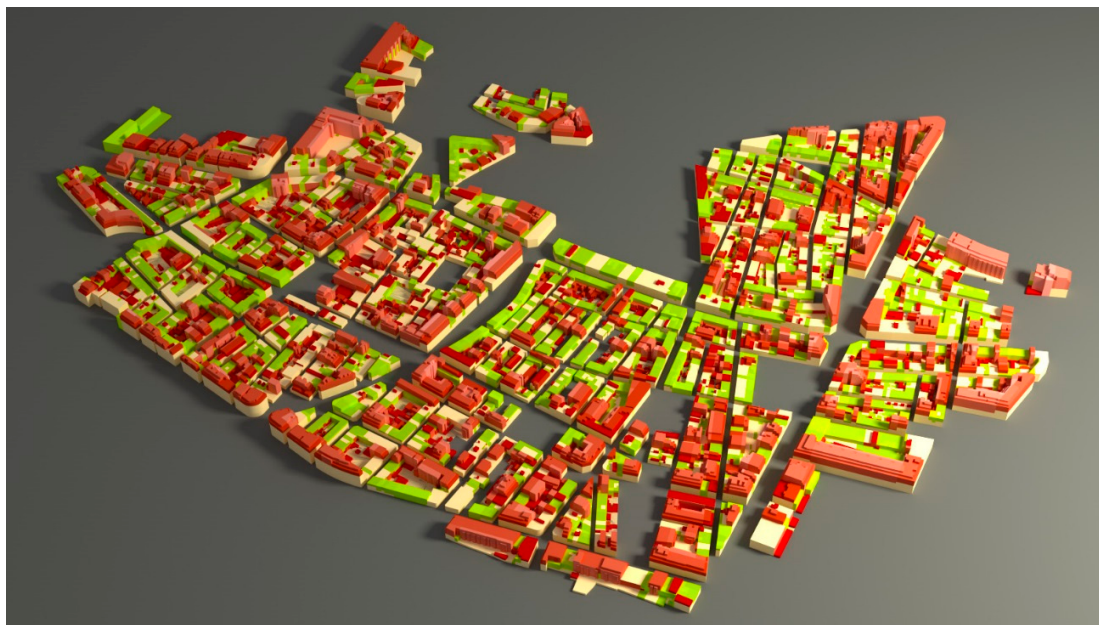
For the three cases the null hypothesis of no spatial autocorrelation is rejected with a highly statistically significant p -value. The hypothesis of the authors is that this observed clustering is not result of some kind of preferential attachment between situations of the same category, but the consequence of the renovation processes the zone has underwent, where the maximum allowed height has been raised or lowered in multiple street segments rendering the buildings in those streets higher or lower than what the plan prescribed.

4.3. Representation of Fragments in 3D

The 2D representation described in Section 4.2 was a very valuable analytical tool, but the representation of the results using 2D maps was unable to fully convey the complex volumetric information because height data had to be abstracted to be represented in plan view as color scales, hatch densities or labels. To visualize this information in a more intuitive way, a different approach had to be developed to make it easier to interpret, as described in Section 3.1.

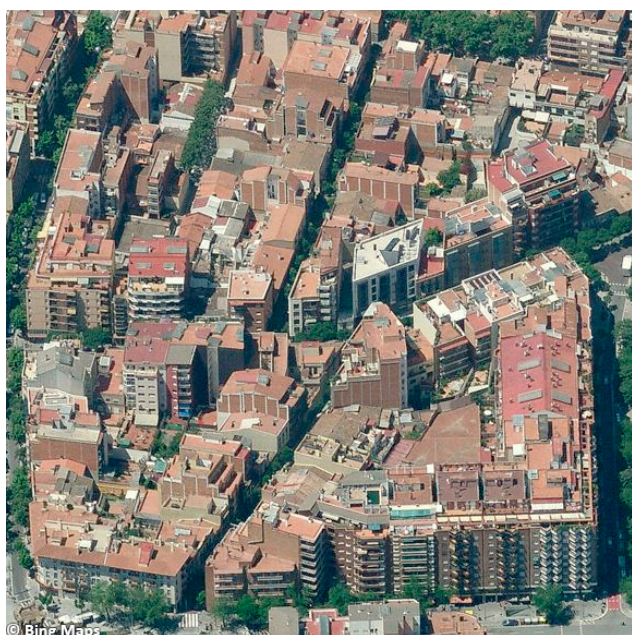
The use of 3D imagery allowed the authors to represent the volumes in a more natural and intuitive way since it matched the way we experience our cities (Figure 14), allowing the viewers to visualize and relate two concepts simultaneously (real height and planned height) much more easily than using 2D maps.

Figure 14. Overbuilt (red) and underbuilt (green) fragments in the case of study.



To visualize the accuracy of the results, an axonometric aerial photograph was compared to its corresponding 3D representation (Figure 15); in the 3D synthetic image the differences between built reality and planned city are more apparent and easier to interpret. The image on the right is not a representation of the built reality but an abstraction of the real city seen through the point of view of the urban plan.

Figure 15. (a) Aerial axonometric view (Bing Maps) and (b) 3D representation of the same area.



(a)



(b)

4.4. Map of the Projection of the Apparent Surface

To better understand the perception of pedestrians from the street, the technique described in Section 3.2 was applied to the area of study. In the resulting map (Figure 16), the length of the shaded area (dark gray) measured orthogonally from the façade plane, corresponds to the magnitude of occlusion of the corresponding built volume from the pedestrian point of view when walking along the street.

Without this representation, it is difficult to comprehend this pedestrian's perception of the public space. In the map it can be appreciated among other aspects that (a) some streets with equal width can appear wider or narrower depending on the height of the buildings along their sides; (b) some streets have a more regular skyline along their length than others; and (c) some streets have more symmetrical sides than others. This visualization should be valuable for urban planners both as an analysis and design tool.

Figure 16. Apparent occlusion projected onto the ground plane.

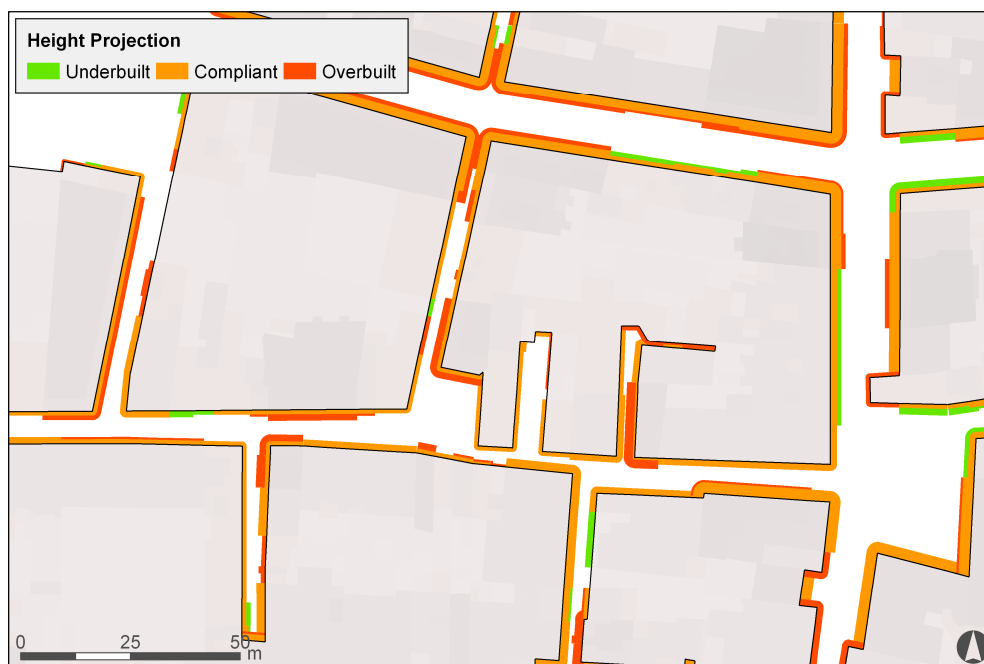


4.5. Façade Height Projection onto the Horizontal Plane

The visualization of the height of buildings from the street explained in the subsection above does not explain the relation of their geometry with the heights allowed by the urban plan: it is not possible to know whether the configuration of the cross-section of the street is a consequence of what the plan mandates or not.

To be able to get a better insight into this relation, another visualization was developed where the height was projected onto the horizontal plane at an angle, and the projection was colored with the same codification used for the 3D representation explained in Section 3.1 (Figure 4).

The resulting map (Figure 17) was capable of showing whether the skyline observed in the built environment was a consequence of what the urban plan dictates or the result of buildings being underbuilt or overbuilt, resulting in a powerful tool for urban planners.

Figure 17. Height projected onto the ground plane colored by legal status.

4.6. Occlusion by Higher Neighboring Volumes

The majority of the roofs in the area of study, as is usual in Mediterranean climates, are flat because the rainfall is rare and the high population density makes the rooftops valuable spaces of recreation and other uses.

The methodology described in Section 3.3 was used to visualize the occlusion that higher volumes have on lower ones. The result has some similarities to a shaded map because conceptually it would be the result of recording the shaded area from a light orbiting each volume at a fixed distance.

Figure 18. Occlusion of higher volumes onto neighboring lower volumes.

The resulting map (Figure 18) shows the magnitude of this occlusion in a darker color; it allows the visualization of the occlusions for each rooftop, which have an impact on the views of their surroundings, the amount of sunlight they receive, their exposure to wind and their protection from street noise.

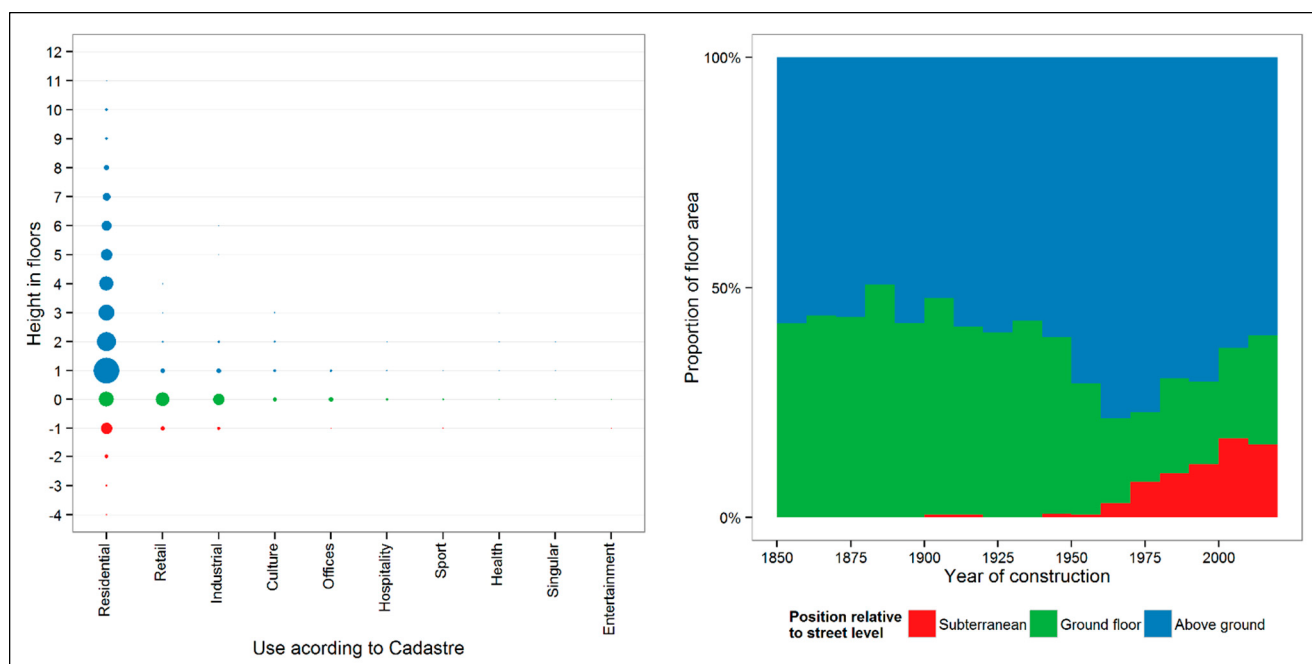
4.7. Activities on Ground Floor

Pedestrians interact primarily with the ground floor of the buildings, and different activities give streets different characters. The main activity on the ground floor was studied with the methodology described in Section 2.2.2 to be able to typify the streets according to the activities developed along them.

The area of study was markedly residential in all floors above ground (Figure 19, left), but the ground floor showed some variability in the activities developed, being the four main activities in decreasing order: residential, retail, industrial and offices.

The temporal series showed that in more recent times (from the 1970s onwards) the proportion of floor space located at the ground floor had decreased, probably because buildings had gotten taller and subterranean floors were excavated to accommodate space for parking vehicles (Figure 19, right).

Figure 19. Built area per floor and use (left) and temporal evolution relative to street level (right).



A choropleth map displaying the four main activities in each parcel was used to visualize their spatial distribution (Figure 20). Each use was assigned a color and to illustrate the degree of dominance of each use among the rest, as the percentage of the dominant use in the parcel decreased, its hue faded to white.

The resulting map clearly showed the presence of the main commercial street of the area (Gran de Sant Andreu) in the north-south axis in the center of the area of study.

To visualize the activities on the ground floor from the pedestrian point of view, the methodology described in Section 3.4 was used to display the main activity at the façade front (Figure 21). This map abandons a parcel-centric approach for a street-centric approach, and displays the different activities

where pedestrians interact with the buildings: their façade, avoiding the distortion that the size and shape of the parcels can introduce.

Figure 20. Map of parcels colored by main use at street level.

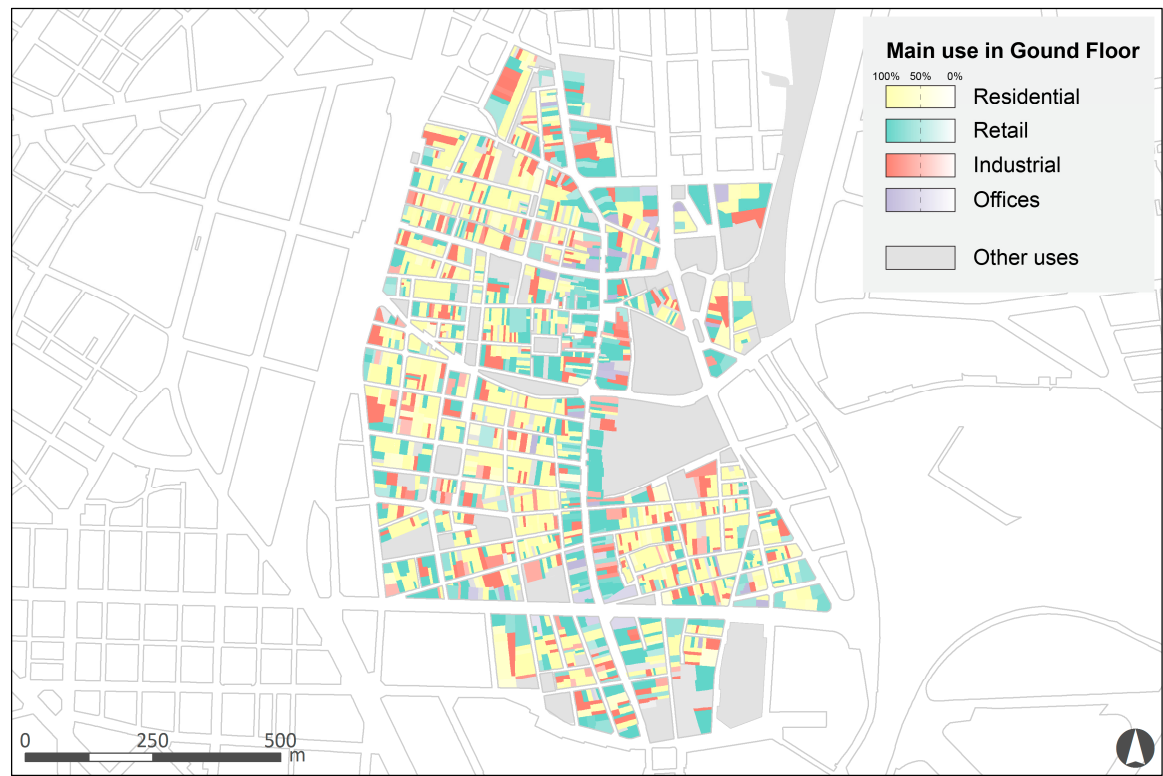
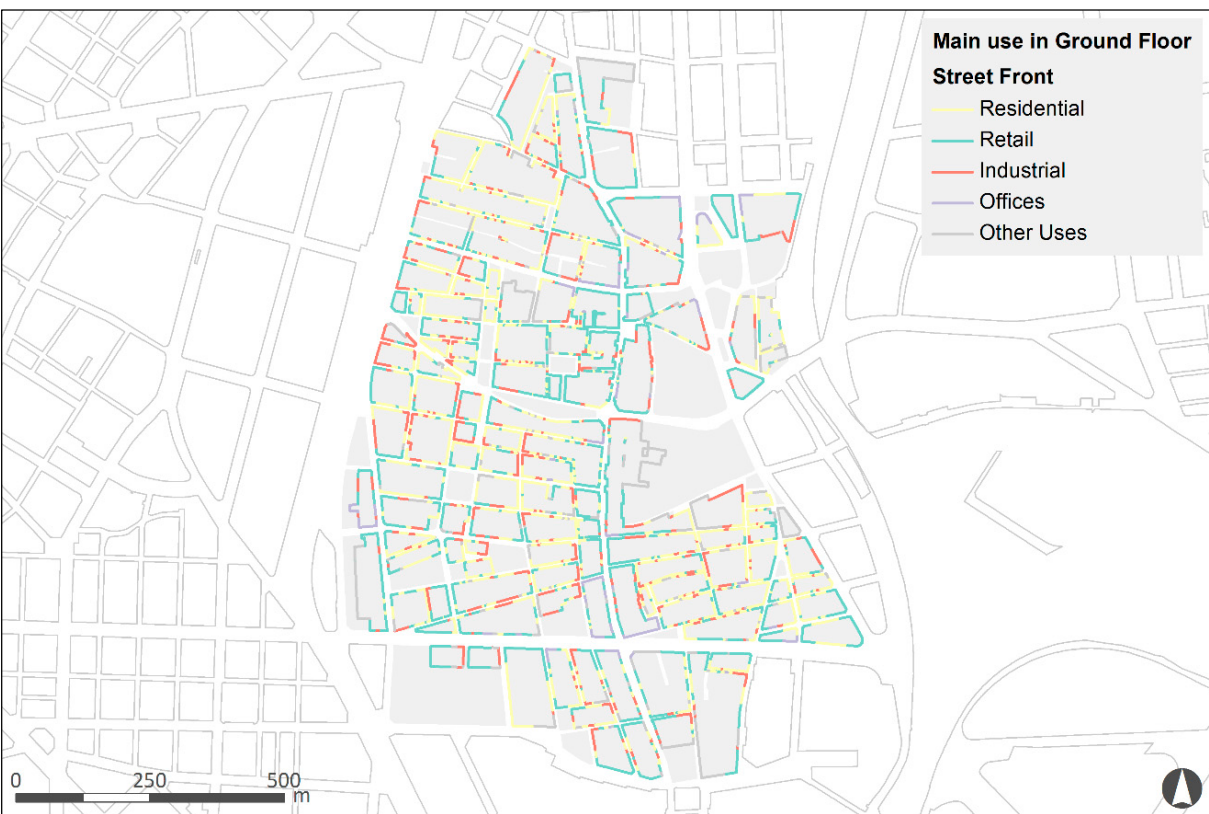


Figure 21. Map of parcel fronts colored by main use at street level.



4.8. Historic Perspective

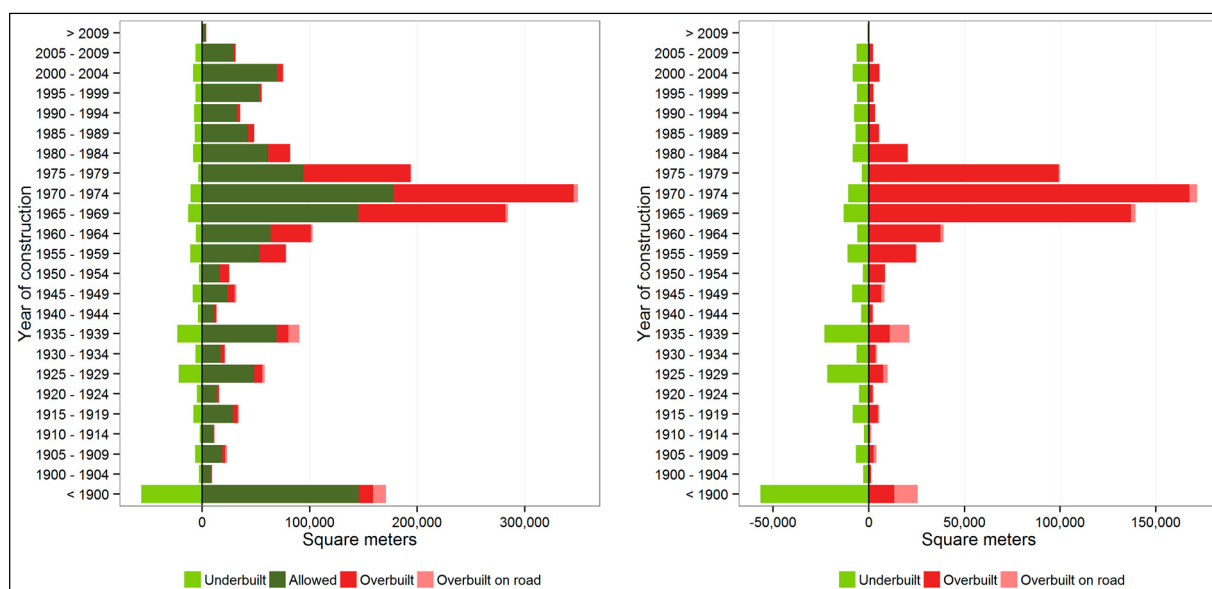
The edification age estimated per parcel as described in Section 2.2.1, was related to three other variables: planning regulations, use and vertical position inside the building. Although it was not possible to display this information on a map because of the inherent difficulties of displaying time information on a static image, the exploration of their relationship gave some insight about the historic perspective of the current situation.

4.8.1. Building Age and Planning Regulations

The calculation of the area for each situation (underbuilt, compliant and overbuilt) as described in Section 2.2.3 was aggregated in slices of five years in two graphs, conceptually similar to population pyramids (Figure 22). The first graph showed the total built area stacked with the overbuilt area on top and the underbuilt area below (Figure 22, left), using the same conceptual framework used to display the fragments height in 3D (Sections 3.1 and 4.3). The second graph (Figure 22, right) showed the same information, but excluding the allowed part to focus only on the amount of discrepancy from the plan.

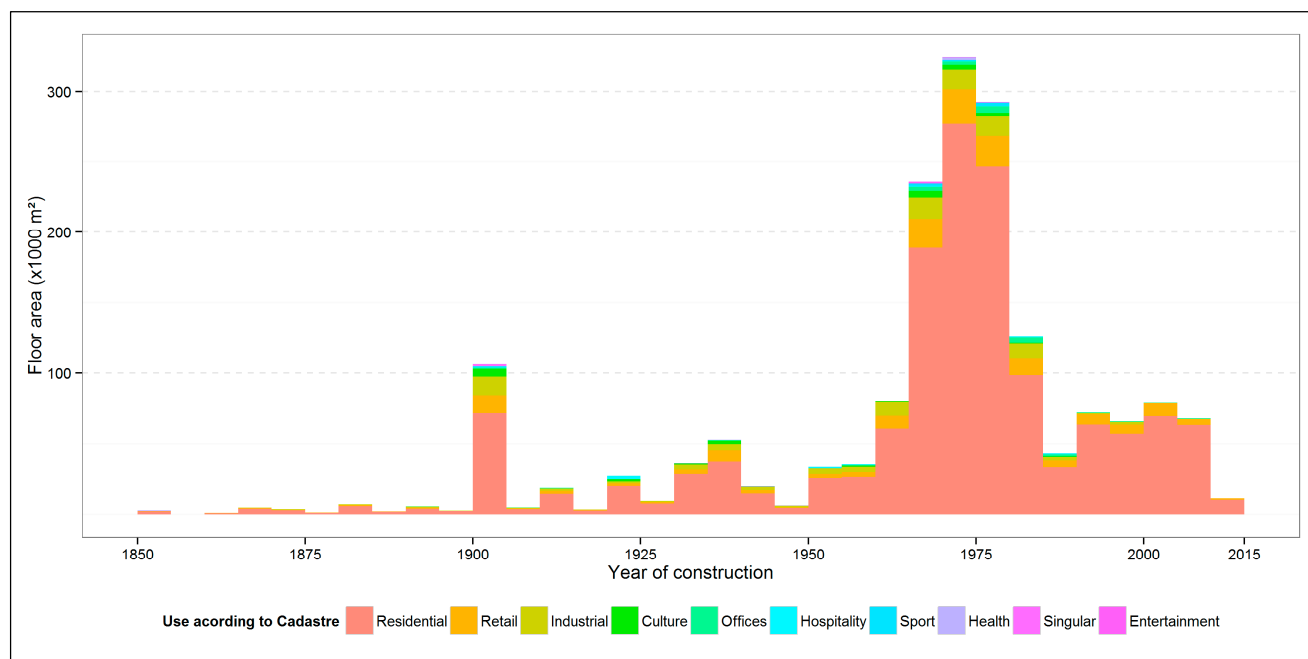
This graphs show a peak in the overbuilt area around the year the Metropolitan Plan was approved and a sharp decline afterwards. They also show a decline in the underbuilt area for buildings from the end of the Spanish Civil War to the present.

Figure 22. Distribution of legal status by year of construction.



4.8.2. Built Area per Year of Construction and Use

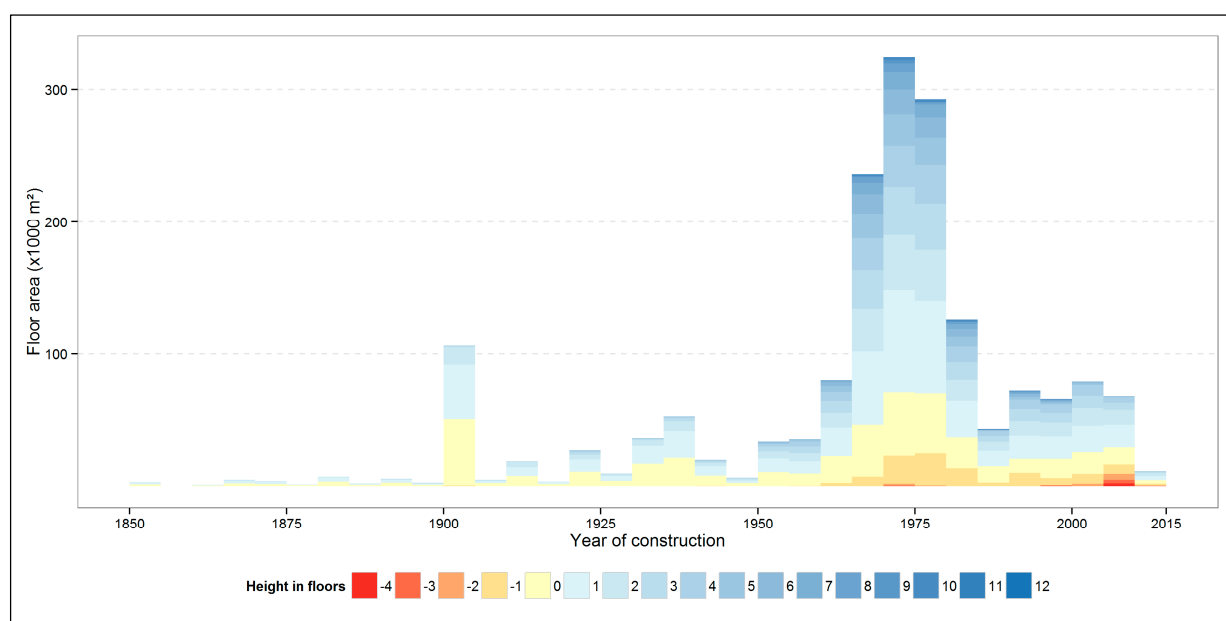
A weighted histogram of the built area for each use (Figure 23) showed the dominance of the residential use in the whole period (1850–2012) with a small but constant amount of retail and industrial uses throughout, and a consolidation of the built environment around the decade of the 1980s, with little activity afterwards.

Figure 23. Weighted histogram of built area per year of construction and use.

4.8.3. Built Area per Year of Construction and Floor

A weighted histogram of the built area for each floor (Figure 24) showed the evolution of the height of the buildings over time. The data was categorized into three categories (subterranean, ground floor and above ground) and given diverging color scales for positive and negative floor numbers.

The graph showed the highest buildings were built from the 1960s onwards, when the construction industry was able to build up to 12 floors competitively. It also shows that the majority of the ground floor area was built at the beginning of the 20th Century and around the date of approval of the Metropolitan Plan, when the construction of subterranean floors began to take off.

Figure 24. Weighted histogram of built area per year of construction and floor.

5. Conclusions

The public data from the Spanish cadaster offers a lot of useful information for urban planners. Nevertheless, its usefulness is restricted by the adoption of non-standard storage formats and the lack of semantically clear concepts in some of its sections. The presented methodological approach seeks to extract meaningful information from cadastral data using representational and semantic strategies.

The visualization techniques developed allowed better communication and understanding of the information stored in the databases and their interrelationships:

- The 3D visualization techniques allowed the discovery of patterns not obvious even for trained professionals, using the third dimension to convey information not possible in 2D, with the third dimension helping understand conceptual data.
- The 2D representation of height and use information onto the horizontal plan resulted in a very compact and synthetic visualization of the relation between the streets skyline, the planning regulations and the built environment from a pedestrian point of view.
- For data without associated cartography, multiple visualization techniques were used to extract meaningful data, in the form of charts or through aggregation operations into higher order geometric entities.
- To represent temporal information, a visualization similar to population pyramids was used for the age of buildings.

The subject described by both the urban regulations and the cadaster is the urban built environment; however, they were conceived to serve different purposes and therefore are structured differently. To be able to integrate both sources to get better insights on the evolution of the morphology of the city in renovation processes, it was necessary to translate both classes of information into a common conceptual framework.

Due to the nuances in the interpretation of building regulations and the legacy format used by the cadaster, the translation operation required a team that was able to work with both data sources. In the future, if a semantic Spatial Data Infrastructure (SDI) is in place, these databases could be automatically queried and their information extracted with less ambiguity [59], allowing a better collaboration of the professionals involved. However, the concepts in this semantic SDI need to be defined by the professionals involved in the planning process; this paper tries to assess some of the concepts that could be included in such SDI, in parallel to the work being developed by the Building Information Modeling (BIM) community to integrate building licensing [60] into their workflow.

Future Work

The model used for 3D representation assumed a flat surface without topography; to improve the accuracy and usefulness of the visualization of the results it is proposed that a Digital Elevation Model (DEM) is incorporated in the 3D model; however, to validate the geometry of the resulting more complex model, a 3D geometry checking process would have to be implemented [61–63]. In addition, a process to correct the sliver polygons resulting from artifacts in the overlay operation could be implemented [64].

To extend the methodology to regions outside the Spanish territory covered by the cadaster, the publicly available Volunteered Geographic Information (VGI) of the OpenStreetMap project [65] could be used [66], allowing the comparison of the urban fabric of different cities or neighborhoods.

Furthermore, it would be valuable to explore the usefulness of using Augmented Reality (AR) tools to visualize the information on site for planners and in public participation processes [67], and to develop tools to evaluate the environmental impact of renovation processes [68].

Acknowledgments

This research was supported by the Non-Oriented Fundamental Research Project EDU2012-37247/EDUC of the VI National Plan for Scientific Research, Development and Technological Innovation 2008-2011, Government of Spain, titled “E-learning 3.0 in the teaching of architecture. Case studies of educational research for the foreseeable future”.

The authors would like to thank Montserrat Moix Bergadà for her invaluable insights during the elaboration of this paper, and the staff in the Urban Studies Bureau of the Urban Planning Department of the Barcelona City Council for their collaboration. We also thank the anonymous reviewers for their valuable comments.

Author Contributions

Francesc Valls Dalmau, Pilar Garcia-Almirall, Ernest Redondo Domínguez and David Fonseca Escudero designed the study, developed the methodology, collected the data, performed the analysis, and wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest

References

1. Lynch, K. *The Image of the City*; The MIT Press: Cambridge, MA, USA, 1960.
2. Borja, J. *Luces y Sombras del Urbanismo de Barcelona*, 2nd ed.; Editorial UOC: Barcelona, Spain, 2010. (In Spanish)
3. Nel·lo, O. Dinàmiques urbanes, activitats emergents i polítiques públiques a la Regió Metropolitana de Barcelona. *Pap. Regió Metrop. Barc. Territ. Estratègies Planejament* **2002**, 2002, 105–114. (In Catalan)
4. Klaasen, I.T. *Knowledge-based Design: Developing Urban & Regional Design into a Science*; Delft University of Technology: Delft, The Netherlands, 2003.
5. Keim, D.A. Information visualization and visual data mining. *IEEE Trans. Vis. Comput. Graph.* **2002**, 8, 1–8.
6. Zlatanova, S.; Itard, L.; Kibria, M.S.; Van Dorst, M. A user requirements study of digital 3D models for urban renewal. *Open House Int.* **2010**, 35, 37–46.
7. Nelson, A.C.; Burby, R.J.; Feser, E.; Dawkins, C.J.; Malizia, E.E.; Quercia, R. Urban Containment and Central-City Revitalization. *J. Am. Plann. Assoc.* **2004**, 70, 411–425.
8. Garcia-Almirall, P.; Valls Dalmau, F.; Moix Bergada, M. Planned vs. Real City: 3D GIS for Analyzing the Transformation of Urban Morphology. In Proceedings of the 17th AGILE Conference on Geographic Information Science, Castellón de la Plana, Spain, 3–6 June 2014.

9. Behnisch, M.; Ultsch, A. Urban data-mining: Spatiotemporal exploration of multidimensional data. *Build. Res. Inf.* **2009**, *37*, 520–532.
10. Spielman, S.E.; Thill, J.-C. Social area analysis, data mining, and GIS. *Comput. Environ. Urban Syst.* **2008**, *32*, 110–122.
11. Wainer, H. How to display data badly. *Am. Stat.* **1984**, *38*, 137–147.
12. Bertin, J.; Barbut, M. *Sémiologie Graphique. Les diagrammes, Les réseaux, Les cartes*; Mouton: Paris, France, 1967. (In French)
13. Bertin, J. *La Graphique et le Traitement Graphique de L'information*; Flammarion: Paris, France, 1977. (In French)
14. Cleveland, W.S. Graphs in Scientific Publications. *Am. Stat.* **1984**, *38*, 261–269.
15. Cleveland, W.S. Graphical Methods for Data Presentation: Full Scale Breaks, Dot Charts, and Multibased Logging. *Am. Stat.* **1984**, *38*, 270–280.
16. Cleveland, W.S.; McGill, R. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *J. Am. Stat. Assoc.* **1984**, *79*, 531–554.
17. Cleveland, W.S.; McGill, R. Graphical Perception and Graphical Methods for Analyzing Scientific Data. *Science* **1985**, *229*, 828–833.
18. Cleveland, W.S.; McGill, R. Graphical Perception: The visual decoding of quantitative information on graphical displays of data. *J. R. Stat. Soc. Ser. Gen.* **1987**, *150*, 192–229.
19. Tufte, E.R. *Envisioning Information*; Graphics Press: Cheshire, CT, USA, 1990.
20. Tufte, E.R. *Visual Explanations: Images and Quantities, Evidence and Narrative*; Graphics Press: Cheshire, CT, USA, 1997.
21. Tufte, E.R. *The Visual Display of Quantitative Information*, 2nd ed.; Graphics Press: Cheshire, CT, USA, 2001.
22. Friendly, M.; SAS Institute. *Visualizing Categorical Data*; SAS Institute: Cary, NC, USA, 2001.
23. Wilkinson, L.; Wills, G. *The Grammar of Graphics*; Springer: New York, NY, USA, 2005.
24. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer: Dordrecht, The Netherlands, 2009.
25. Wickham, H. A layered grammar of graphics. *J. Comput. Graph. Stat.* **2010**, *19*, 3–28.
26. Harrower, M.; Brewer, C.A. ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *Cartogr. J.* **2003**, *40*, 27–37.
27. Peuquet, D.J. Representations of geographic space: Toward a conceptual synthesis. *Ann. Assoc. Am. Geogr.* **1988**, *78*, 375–394.
28. Mennis, J.L.; Peuquet, D.J.; Qian, L. A conceptual framework for incorporating cognitive principles into geographical database representation. *Int. J. Geogr. Inf. Sci.* **2000**, *14*, 501–520.
29. Goodchild, M.F.; Yuan, M.; Cova, T.J. Towards a general theory of geographic representation in GIS. *Int. J. Geogr. Inf. Sci.* **2007**, *21*, 239–260.
30. Goodchild, M.F. Geographical data modeling. *Comput. Geosci.* **1992**, *18*, 401–408.
31. Gregory, I.N. “A map is just a bad graph”: Why spatial statistics are important in historical GIS. In *Placing History: How Maps, Spatial Data and GIS are Changing Historical Scholarship*; Knowles, A.K., Ed.; ESRI Press: Redlands, CA, USA, 2008; pp. 123–149.
32. Harmon, K. *You are Here: Personal Geographies and Other Maps of the Imagination*, 1st ed.; Princeton Architectural Press: New York, NY, USA, 2003.

33. Tufte, E.R. *Beautiful Evidence*, 1st ed.; Graphics Press: Cheshire, CT, USA, 2006.
34. Harmon, K.; Clemans, G. *The Map as Art: Contemporary Artists Explore Cartography*; Princeton Architectural Press: New York, NY, USA, 2010.
35. Parish, Y.I.H.; Müller, P. Procedural Modeling of Cities. In Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '01; ACM: New York, NY, USA, 2001; pp. 301–308.
36. Wonka, P.; Wimmer, M.; Sillion, F.; Ribarsky, W. Instant Architecture. In *ACM SIGGRAPH 2003 Papers*, SIGGRAPH '03; ACM: New York, NY, USA, 2003; pp. 669–677.
37. Müller, P.; Wonka, P.; Haegler, S.; Ulmer, A.; Van Gool, L. Procedural Modeling of Buildings. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06; ACM: New York, NY, USA, 2006; pp. 614–623.
38. Müller, P.; Zeng, G.; Wonka, P.; Van Gool, L. Image-based Procedural Modeling of Facades. In *ACM SIGGRAPH 2007 Papers*, SIGGRAPH '07; ACM: New York, NY, USA, 2007.
39. Gruen, A.; Behnisch, M.; Kohler, N. Perspectives in the reality-based generation, nD modelling, and operation of buildings and building stocks. *Build. Res. Inf.* **2009**, *37*, 503–519.
40. Döllner, J.; Walther, M. Real-time expressive rendering of city models. In Proceedings of the Seventh International Conference on Information Visualization, 2003. IV 2003, London, UK, 16–18 July 2003; pp. 245–250.
41. Döllner, J.; Baumann, K.; Buchholz, H. Virtual 3D City Models as Foundation of Complex Urban Information Spaces. In *11th International Conference on Urban Planning and Spatial Development in the Information Society (REAL CORP)*; Schrenk, M., Ed.; CORP—Competence Center of Urban and Regional Planning: Vienna, Austria, 2006; pp. 107–112.
42. Glander, T.; Döllner, J. Abstract representations for interactive visualization of virtual 3D city models. *Comput. Environ. Urban Syst.* **2009**, *33*, 375–387.
43. Semmo, A.; Trapp, M.; Kyprianidis, J.E.; Döllner, J. Interactive Visualization of Generalized Virtual 3D City Models Using Level-of-Abstraction Transitions. *Comput. Graph. Forum* **2012**, *31*, 885–894.
44. Peuquet, D.J. It's About Time: A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. *Ann. Assoc. Am. Geogr.* **1994**, *84*, 441–461.
45. Peuquet, D.J. Making Space for Time: Issues in Space-Time Data Representation. *GeoInformatica* **2001**, *5*, 11–32.
46. Döner, F.; Thompson, R.; Stoter, J.; Lemmen, C.; Ploeger, H.; van Oosterom, P.; Zlatanova, S. 4D cadastres: First analysis of legal, organizational, and technical impact—With a case study on utility networks. *Land Use Policy* **2010**, *27*, 1068–1081.
47. Pushkarev, B.S. *Urban Space for Pedestrians: A Quantitative Approach*, 1st ed.; The MIT Press: Cambridge, MA, USA, 1976.
48. Hillier, B.; Penn, A.; Hanson, J.; Grajewski, T.; Xu, J. Natural movement: Or, configuration and attraction in urban pedestrian movement. *Environ. Plan. B Plan. Des.* **1993**, *20*, 29–66.
49. Saleuses, P.; Schechtner, K.; Hidalgo, C.A. The Collaborative Image of The City: Mapping the Inequality of Urban Perception. *PLoS One* **2013**, *8*, e68400.
50. Carr, D.B.; White, D.; MacEachren, A.M. Conditioned Choropleth Maps and Hypothesis Generation. *Ann. Assoc. Am. Geogr.* **2005**, *95*, 32–53.
51. Wood, J.; Dykes, J.; Slingsby, A. Visualisation of Origins, Destinations and Flows with OD Maps. *Cartogr. J.* **2010**, *47*, 117–129.

52. Gröger, G.; Plümer, L. CityGML—Interoperable semantic 3D city models. *ISPRS J. Photogramm. Remote Sens.* **2012**, *71*, 12–33.
53. Karki, S.; Thompson, R.; McDougall, K. Data validation in 3D cadastre. In *Developments in 3D Geo-Information Sciences*; Neutens, T., Maeyer, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 92–122.
54. Van Oosterom, P. Research and development in 3D cadastres. *Comput. Environ. Urban Syst.* **2013**, *40*, 1–6.
55. Corporación Metropolitana de Barcelona. *Plan General Metropolitano de Ordenación Urbana*; Comisión Provincial de Urbanismo de Barcelona: Barcelona, Spain, 1976. (In Spanish)
56. Esteban, J. *Elementos de ordenación urbana*; Edicions UPC: Barcelona, Spain, 1998. (In Spanish)
57. Turner, A. Analysing the visual dynamics of spatial morphology. *Environ. Plan. B Plan. Des.* **2003**, *30*, 657–676.
58. Hillier, B. *Space is the Machine: A Configurational Theory of Architecture*; Space Syntax: London, UK, 2007.
59. Li, W.; Goodchild, M.F.; Raskin, R. Towards geospatial semantic search: Exploiting latent semantic relations in geospatial data. *Int. J. Digit. Earth* **2012**, *7*, 17–37.
60. Benner, J.; Geiger, A.; Häfele, K.-H. Concept for building licensing based on standardized 3D geo information. In Proceedings of the 5th International 3D GeoInfo Conference, Berlin, Germany, 3–4 November 2010.
61. Kazar, B.M.; Kothuri, R.; van Oosterom, P.; Ravada, S. On Valid and Invalid Three-Dimensional Geometries. In *Advances in 3D Geoinformation Systems*; van Oosterom, P., Zlatanova, S., Penninga, F., Fendel, E.M., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 19–46.
62. Gröger, G.; Plümer, L. How to achieve consistency for 3D city models. *GeoInformatica* **2011**, *15*, 137–165.
63. Ledoux, H.; Meijers, M. Topologically consistent 3D city models obtained by extrusion. *Int. J. Geogr. Inf. Sci.* **2011**, *25*, 557–574.
64. Ledoux, H.; Arroyo Ohori, K.; Meijers, M. A triangulation-based approach to automatically repair GIS polygons. *Comput. Geosci.* **2014**, *66*, 121–131.
65. OSM Foundation Simple 3D Buildings. Available online: http://wiki.openstreetmap.org/wiki/Simple_3D_Buildings (accessed on 18 September 2014).
66. Over, M.; Schilling, A.; Neubauer, S.; Zipf, A. Generating web-based 3D City Models from OpenStreetMap: The current situation in Germany. *Comput. Environ. Urban Syst.* **2010**, *34*, 496–507.
67. Koeman, L. An Exploratory Study into the Public and Situated Visualisation of Local Data in Urban Communities. In *Proceedings of the 2014 Companion Publication on Designing Interactive Systems*, DIS Companion '14; ACM: New York, NY, USA, 2014; pp. 173–176.
68. Power, A. Does demolition or refurbishment of old and inefficient homes help to increase our environmental, social and economic viability? *Energy Policy* **2008**, *36*, 4487–4501.