A graph-based approach for the structural analysis of road and building layouts

Mathieu Domingo, Rémy Thibaud & Christophe Claramunt

Taylor & Francis
Taylor & Francis Group

A graph-based approach for the structural analysis of road and building layouts

Mathieu Domingo [iD][a], Rémy Thibaud [iD][b] and Christophe Claramunt [iD][b]

[a]Department of Computer Science, University of Pau and Pays de l'Adour, Pau, France; [b]Naval Academy Research Institute, Brest Cedex, France

ABSTRACT

A better understanding of the relationship between the structure and functions of urban and suburban spaces is one of the avenues of research still open for geographical information science. The research presented in this paper develops several graph-based metrics whose objective is to characterize some local and global structural properties that reflect the way the overall building layout can be cross-related to the one of the road layout. Such structural properties are modeled as an aggregation of parcels, buildings, and road networks. We introduce several computational measures (*Ratio Minimum Distance, Minimum Ratio Minimum Distance, and Metric Compactness*) that respectively evaluate the capability for a given road to be connected with the whole road network. These measures reveal emerging sub-network structures and point out differences between less-connective and more-connective parts of the network. Based on these local and global properties derived from the topological and graph-based representation, and on building density metrics, this paper proposes an analysis of road and building layouts at different levels of granularity. The metrics developed are applied to a case study in which the derived properties reveal coherent as well as incoherent neighborhoods that illustrate the potential of the approach and the way buildings and roads can be relatively connected in a given urban environment. Overall, and by integrating the parcels and buildings layouts, this approach complements other previous and related works that mainly retain the configurational structure of the urban network as well as morphological studies whose focus is generally limited to the analysis of the building layout.

1. Introduction

The representation and analysis of the spatial structure of geographic spaces are relatively old but still open research issues for many environmental and urban sciences. Among many subjects to explore, a better comprehension of the relationship between the structure and function of urban and suburban spaces is one of the directions that offer many research challenges for geographical information science on the one hand, and novel methodological opportunities for environmental and geographical sciences on the other hand.

The spatial distribution of land ownership, buildings and roads is a key factor in the emergence of patterns in space and time (Hawbaker et al. 2005; Oliveira 2016). The way that a given urban space might be either planned or self-organized is closely related to the underlying properties of the geographical space, but overall the forms that appear are also the result of many cultural and social interactions. In particular, a close analysis of the coherence between street accessibility and building density is particularly relevant when evaluating the appropriateness of a given urban layout when considered from urban design and planning perspectives. This is a crucial issue to understand the interactions between urban and suburban spaces and human development, the spatial arrangements and morphology that emerge, and the role and organization of the road network that constitutes the backbone of the whole urban system. Indeed, urban spaces have been long shaped by human beings, by developing activities, building, routes, and so forth thus generating forms, spatial arrangements, and clusters. Cities have been long analyzed from a structural point of view. For instance, Borrego et al. (2006) categorized three types of city structures (i.e. compact, disperse, corridor) to study the impact of land use patterns on urban air quality, and vice versa. A similar approach has been developed by Steiniger et al. (2008), but the focus was mainly oriented to the analysis of the built environment and emerging forms. Despite their interest, none of these approaches take into account the specific role and properties of the underlying road network in the building layout.

However, when considering the spatial and structural patterns that emerge from a given urban space, the distribution of the built environment and its relationships (i.e. accessibility) with the underlying street network are key factors to take into account.

This is a key component when the objective is to provide a better understanding of the overall urban layout, and this has been hardly considered, to the best of our knowledge, in previous urban research and studies (Ewing, Pendall, and Chen 2003; Wells and Yang 2008) Building layouts can be characterized by a measure of compactness that combines population and building characteristics (i.e. size, density, distribution, clustering) (Tsai 2005; Berghauser-Pont and Haupt 2007; Ye and Van Nes 2014; Raimbault 2018). Roads subdivide space and the built environment, while creating displacement opportunities, and generate different forms such as regular or irregular subdivisions, thus reflecting different social, cultural, and economic patterns (Marshall 2004). For instance, Hawbaker et al. (2005) highlight the relationship between road density (i.e. km/km^2) and housing density (i.e. units/km^2). Similarly, the space matrix method quantifies building types at the urban block level as defined by the underlying road network (Berghauser-Pont and Haupt 2010). However, although most of these works propose some valuable classifications of some urban structural properties and forms, the intimate relationships between the structural properties of the road network and the buildings and parcels distribution are still not taken into account.

The development of structural and computational approaches for the representation of urban spaces has been largely favored by the emergence of graph-based networks analysis and space syntax since the early eighties (Hillier and Hanson 1984; Batty and Howes 2001; Herold, Couclelis, and Clarke 2005; Gastner and Newman 2006; Batty 2013; Barabási 2016). A given urban environment is modeled as a graph from a primal approach where edges represent streets, and nodes represent street intersections, or as a dual graph where edges represent connection between street intersections and streets are represented as nodes (Porta, Crucitti, and Latora 2006a, 2006b; Marshall 2016). The main interest of the primal approach relies in its intuitive component and the fact that it can be easily matched with geographical data models. However, this intersection-centered approach generates many road segments, thereby leading to a loss of continuity when representing the underlying urban network. Conversely, the axial line approach keeps the notion of road continuity although road intersections are not directly materialized. Moreover, as noted by Porta, Crucitti, and Latora (2006b), longer roads are likely to appear relatively central, being not always the case because the notion of distance is not directly taken into account. Overall, these space syntax approaches support derivation of many graph-based metrics, either local or global, that favor the analysis of centrality, cluster, and many indices that can be related to several urban properties and functional characteristics (Penn 2003; Batty and Rana 2004; Jiang and Claramunt 2004; Crucitti, Latora, and Porta 2006; Turner 2007; Ye, Li., and Liu 2018). Despite the wide success and rapid development of space syntax, especially when correlating functional properties with the underlying space structure, the network-based view of an urban environment can be considered as an over simplification of the city (Ratti 2004).

When considering urban and suburban contexts, the structural spatial properties cannot be limited to some basic network properties, and additional metrics should be considered when modeling the network configuration, as well as additional structural elements such as parcels and buildings that cannot be separated (Buhl et al. 2006). In particular, the influence of the spatial configuration of parcels and buildings are notably important in less dense environments such as suburban areas where the underlying road network is less important than in urban areas. In such environments, hamlets and villages often appear as the conjunction of some long-term and self-organized process in which buildings, parcels, and the road network constitute the basic abstraction primitives to consider. Overall, built environments encompass complex spatial organizations that should be studied according to the structural properties that emerge from these basic primitives, at different levels of scale (Chowell et al. 2003; Latora and Marchiori 2003).

The objective of the research presented in this paper is to characterize some local and global properties that reflect the way the overall building layout can be cross-related to the one of the road layout. This paper significantly extends a preliminary graph-based model introduced in a previous work (Domingo, Thibaud, and Claramunt 2013) by the introduction of several graph-based metrics whose objectives are to evaluate the compactness of the road network and built environment, as well as their interconnection. The modeling approach is applied at different levels of granularity to potentially identify several local and global clustering properties. With the application of these metrics, a distinction is made between homogeneous and heterogeneous connected networks, that is, parts of the network in which roads are relatively well connected, and the ones with sub-parts of the network that are loosely connected. A preliminary application of these metrics to an illustrative case study shows a series of correlations between the buildings density and the compactness of the road network, and also identifies a series of road segment outliers with respect to their relation with the building layout.

The remainder of the paper is organized as follows. The next section presents the main principles of our modeling approach and introduces a series of graph-based metrics that evaluate the local and global properties of the underlying road and building
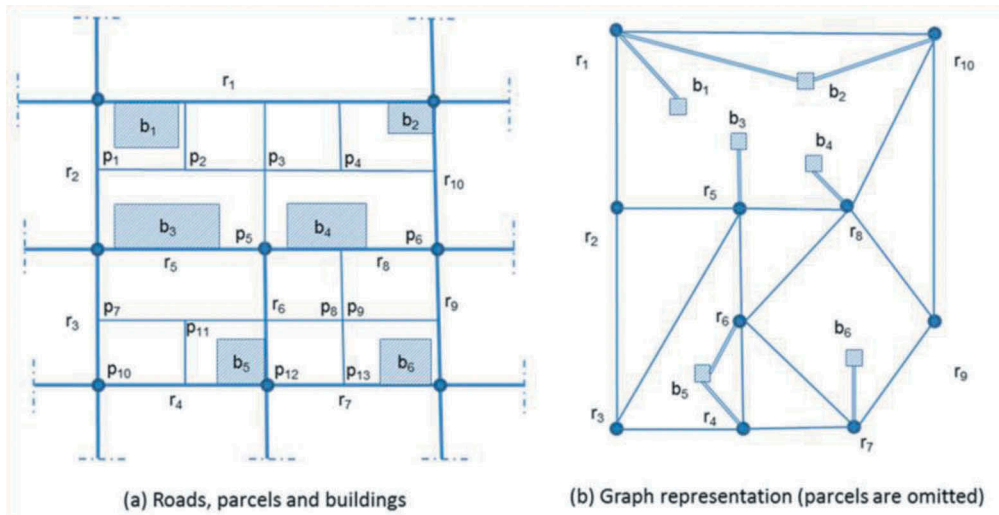
Figure 1. Graph-based modeling approach. (a) Schematic land configuration example (a) with $B= \{b_1, b_2, \ldots b_6\}$ $P= \{p_1, p_2, \ldots p_{13}\}$ and $R= \{r_1, r_2, \ldots r_{13}\}$; (b) Graph representation example (adapted from Domingo, Thibaud, and Claramunt 2013).

networks. These metrics make a difference between compact and non-compact neighborhoods as well as between different degrees of road and building connections and densities in the urban network. Section 3 introduces a classification that characterizes different road and building layouts. Section 4 develops an application of these metrics to an experimental case study, while finally Section 5 summarizes the contribution and outlines further works.

2. Graph-based approach

2.1. A graph model

Let us introduce the main principles and properties of the graph-based modeling approach introduced in a previous work (Domingo, Thibaud, and Claramunt 2013). An urban and suburban space is formally modeled as a space S defined in $\mathbb{R}^2$, and is composed by three categories of regions, namely buildings, parcels, and roads. Let us denote B the set of buildings, P the set of parcels, and R the set of roads. A road is here defined as a route segment (including turns) limited by junctions. A series of assumptions and constraints are modeled between these regions, they are as follows:

- Each building b of B is included in a parcel p of P or in the union of several parcels $p_1, p_2, \ldots p_n$ of P.
- The sets P and R form a partition of S ($P \cap R = \varnothing$ and $P \cup R = S$).
- The union of all roads of R forms a connected region.
- Two given buildings, respectively two parcels, respectively two roads, or respectively one parcel and one road are either disconnected or externally connected as defined by the RCC8 algebra (Cohn et al. 1997).

The relations and constraints above provide the foundations of the graph-based modeling approach (as shown in Figure 1). First, all regions, parcels, and roads constitute the elementary nodes of the graph. Second, edges between these nodes are derived from the application of the relations identified above. Two buildings (respectively two parcels, respectively two roads) in the graph are connected by an edge when they are externally connected. Similarly, a parcel is connected by an edge to a road when they are externally connected. A building is connected by an edge to a parcel when they intersect, and a building is connected to a road through a parcel when that building intersects that parcel externally connected to that road. More formally, let E_B denote the set of edges that connect buildings, E_P the set of edges that connect parcels, E_R the set of edges that connect route segments, E_{BP} the set of edges that connect buildings to parcels, E_{PR} the set of edges that connect parcels to roads, and E_{BR} the set of edges that connect buildings to roads. A peculiarity of this modeling approach is that it allows to derive different graphs and overall a complete graph that combines roads, parcels, and buildings with all edges derived from all relations mentioned above, such graph denoted $G_{BPR} = (V_{BPR}, E_{BPR})$ is given as $V_{BPR} = B \cup P \cup R$ and $E_{BPR} = E_B \cup E_P \cup E_R \cup E_{BP} \cup E_{PR} \cup E_{BR}$.

2.2. Metric compactness operator

2.2.1. Definitions

Many metrics introduced so far by graph-based approaches allow to evaluate some centrality, clustering, and other structural properties of a network (Sabidussi 1966; Hillier and Hanson 1984; Jiang and Claramunt 2004; Watts and Strogatz 1998; Porta, Crucitti, and Latora 2006a, 2006b). For instance, the

average path length of a given node of the network is given by the average of the shortest distance between this node and all other nodes of the network, while the *betweeness centrality* of a node is defined as the number of the shortest paths that go through that node. These two metrics evaluate how distant is the rest of the network from a given node, and how often a given node lies in shortest paths, respectively.

Under similar principles, we would like to evaluate how a given road is connected to all other roads of a given urban layout. Another objective is to evaluate the connectivity of each road at different levels of abstraction by considering the whole to sub-parts of the road network. The main idea behind this approach is to evaluate how clustered is the network at different levels of granularity. While centrality metrics evaluate the role of the different roads with respect to the whole network, our objective is to evaluate how clustered are the roads in the network, at different levels of granularity. This will allow us to make a difference between locally-to-globally well connected roads and loosely connected roads in the network (note that parcels are not considered at this preliminary stage). The connectivity of a given road is here quantitatively based on an evaluation of the distance between this road and part of the urban road network. To take into account different levels of abstraction and then to evaluate some local to global measures of connectivity of a given road, these metrics are evaluated on some given parts of the whole urban road network. Let us successively introduce these measures as well as their respective semantics:

(1) *How distant is a given road to a given part of an urban network?*

A first metric derives the *minimum distance* $D_{\min}^X(r_i)$ that allows a given road $r_i \in R$ with $1 \leq i \leq n$, to reach X percent of the roads of the road network R with $|R| = n$. It is given as follows:

$$D_{\min}^X(r_i) = \text{Min}\left\{ d \Big| \left(\frac{\left|\left\{r_j | d \geq d_{r_i r_j}\right\}\right|}{|R|} \geq \frac{X}{100} \right) \right\}$$

where d is a distance, Min{} is a function that returns the minimum value of a set, and $d_{r_i r_j}$ denotes the distance (center to center) between the roads r_i and r_j. The peculiarity of this metrics is that it is relatively flexible as the choice of the number of roads to reach is given by a parameter X: for instance and for $X = 50$, and for a given road r_i, $D_{\min}^X(r_i)$ given by the minimal distance that allows to reach 50% of the roads from the road r_i. The interest of that parameter X is that it gives the opportunity to analyze the structural role played by the given road in the network at different levels of granularity, that is, locally for low X value and globally for high X value.

(2) *What is the minimum distance that allows all roads to reach a given part of a road network?*

To follow-up, the previous metric is generalized to derive its minimum value when applied to all roads of a given network. The *Minimum Ratio Minimum Distance* ($D_{\min}^X$) of a road network R is defined as the smallest value of $D_{\min}^X(r_i)$ with $1 \leq i \leq n$ with $|R| = n$. It is given as follows:

$$D_{\min}^X = \text{Min}\left(\left\{D_{\min}^X(r_i)\right\}\right)$$

The semantics of the *Minimum Ratio Minimum Distance* is that it gives the minimum distance that allows any road to reach X part of the roads of the whole network. For a given road network, the shorter the distance, the more compact the road network will be. Another evaluation of the road network compactness is also derived by a quantitative value as follows.

(3) *How many roads a given road reaches for a Minimum Ratio Minimum Distance $D_{\min}^X$?*

For a given road r_i, the *Metric Compactness $MC(r_i)$* denotes the number of roads reached at a *Minimum Ratio Minimum Distance $D_{\min}^X$* of r_i. It is given as follows:

$$MC(r_i) = \left| \left\{r_j | d_{r_i r_j} \leq D_{\min}^X\right\} \right|$$

where $d_{r_i r_j}$ denotes the distance between the roads r_i and r_j with r_i, $r_j \in R$. Overall, this parameter evaluates the degree of connectivity of a given road, but with respect to only a X part of the roads of the whole road network.

2.2.2. Structural analysis

One peculiarity of the measure of *Metric Compactness* is its flexibility given by the value X defined by the user, and then its capability for taking into account different levels of granularity. The smaller X should emerge local patterns, whereas the higher X values should highlight global patterns. Figure 2 illustrates an application of the *Minimum Ratio Minimum Distance $D_{\min}^X$* to differentiate "more-connective" and "less-connective" structures as defined by Marshall (2004), with "less-connective" structures corresponding to sub-networks loosely connected with the rest of the network. Given the threshold value X, the distribution of $D_{\min}^X$ along the whole road network can highlight connected versus disconnected parts of the network, and the overall homogeneity of this road network. When a given network is relatively homogeneous, $\frac{\Delta D_{min}^X}{\Delta X}$ tends to be a constant value, while this is not the case for heterogeneous networks. The example exhibited by Figure 2
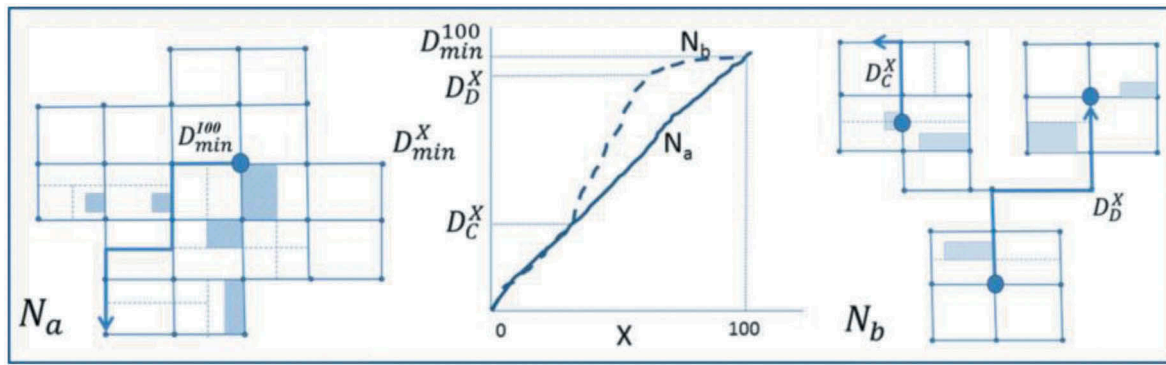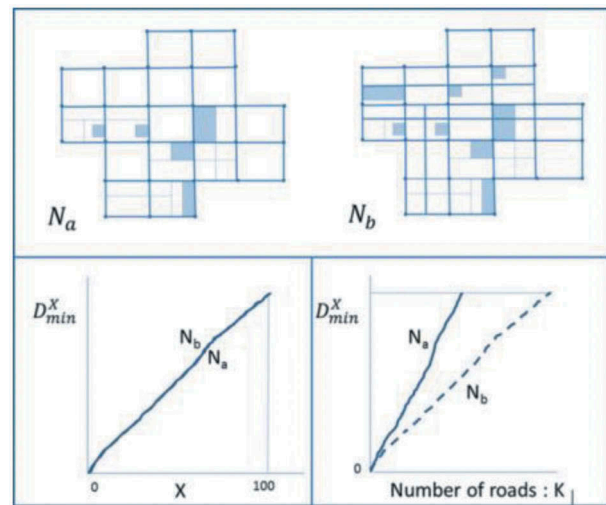
Figure 2. Influence of the road network structure on the *Minimum Ratio Minimum Distance* for a well-connected structure (N_a) and less-connected structure (N_b). Step distances and number of roads are equivalents for both structures.

reveals two changes of slope denoted as D_D^X and D_C^X, which are defined as follows:

- D_C^X denotes the distance that characterizes the size of a local cluster, that is, the maximum distance from the center to the boundary of a given cluster. Values lower than D_C^X denote short connections within a given local cluster, whereas values higher than D_C^X denote distances that materialize a connection toward the exterior of the cluster.
- D_D^X denotes the distance that connects a given cluster to another cluster (i.e. respectively to the center of one cluster to the center of another cluster).

Overall, when considering two given clusters, the value $(D_D^X - D_C^X)$ gives the distance between them. The *Global Minimum Distance* $D_{\min}^{100}$ (i.e. for $X = 100$) denotes the minimum distance that allows each road to reach all other roads. As illustrated in Figure 2 where distances and number of roads are similar, the *Global Minimum Distance* is independent of the road network structure.

The *Minimum Ratio Minimum Distance* evaluates the overall structure of a road network given a part X of connected roads that reflects the extent of the neighborhood considered. As illustrated in Figure 3, for the same spatial extension the difference of road density does not appear on the relationship between D_{min}^X and X. Conversely, when considering the number of roads K (i.e. for a specific $D_{\min}^X$), the number of roads reached is lower for a lower density structure. In other words, the relation between X and the measure of *Minimum Ratio Minimum Distance* can exhibit a clustering road-network structure but not some density variations of the road network. In fact, the notion of clustering is a parameter commonly applied by graph theory and that evaluates the degree to which nodes in a graph tend to cluster together (Watts and Strogatz 1998; Jiang and Claramunt

Figure 3. Influence of roads density on the *Minimum Ratio Minimum Distance*. Na: low density road network, Nb: high density road network. Distances are equivalents for both structures.

2004; Radicchi et al. 2004). However, current approaches as often generally applied in the scope of space syntax theories are generally limited to topological relationships without taking additional metrics. In a related work, Porta, Crucitti, and Latora (2006b) combined a measure of closeness centrality as applied to road intersections and as introduced by Wasserman and Faust (1994) with some distance metrics to extract local clusters. Our approach is different from this work in several respects as we first consider road segments as nodes, and as the connectivity of each road is evaluated at different levels of granularity, by considering different parts of the urban road network. This gives a much more qualitative approach rather than the ones relying on purely distance metrics.

2.3. Building road density

The spatial distribution of a building's layout can be characterized by a density metric, typically a number

of buildings per km^2 (Hawbaker et al. 2005; Gonzalez-Abraham et al. 2007). While valuable when considered as overall density values, such metrics are not directly interrelated to the underlying network structure. To combine the respective roles played by the road network, parcels, and buildings, let us introduce a measure of *Building_Road Density* where the buildings density is evaluated on top of the road network, and based on the number of buildings connected to each road. The *Building_Road Density* is derived from the graph-based spatial modeling approach previously introduced in Section 2.1, where buildings and roads are connected in the graph through their respective connections with parcels. The graph model, where parcels are key elements, gives the connections (E_{BR}) from the buildings to the roads through the parcels as follows:

$$E_{BR} = \{(b_i, r_j) \in B \times R | \exists p_k \in P, \text{Asso}BP(b_i, p_k) = \text{true} \land \text{Asso}PR(p_k, r_j) = \text{true}\}$$

with $\text{Asso}BP(b_i, p_k)$ the association where a building b_i is associated to a parcel p_k if and only if the building b_i intersects the parcel p_k, and $\text{Asso}PR(p_k, \ r_j)$ the association where parcel p_k is associated to a road r_j if and only if the parcel p_k is externally connected to the road r_j.

Each building is affected by a weight $W(b_i)$ with $W(b_i) = 1$ for a building connected with only one road. For building connected with several roads (i.e. associated to a parcel externally connected with several roads) the weight $W(b_i)$ is partly and equally distributed to each connected roads:

$$W(b_i) = \frac{1}{|\{r_k \in R | (b_i, r_k) \in E_{BR}\}|}$$

As such *Building_Road Density* is given as follows:

$$BRD(r_j) = \frac{\sum_{b_i \in B \,|\,(b_i, r_j) \in E_{BR}} W(b_i)}{Length(r_j)}$$

where *Length(r$_j$)* denotes the length of the road r_j.

A combination of *Metric Compactness* and *Building Road Density* should support the analysis of spatial road network structures, at two levels of granularity as follows:

- at a global scale, the capability to evaluate the homogeneity versus heterogeneity of a given network and,
- at a local scale, the level of the compactness of the road versus building layout associated to this road.

3. Buildings to road layout classification

One peculiarity of the measures of *Metric Compactness* and *Building Road Density* is that they clearly take into account the underlying role of the urban network. To illustrate further the concept, we

introduce a classification based on three different levels for *Building Road Densities* and two different levels for *Metric Compactness*. The thresholds defined to generate these categories were applied using a k-means approach following (MacQueen 1967) low to high scores of *Metric Compactness* and *Building Road Density*. *Building Road Density* is classified as follows:

- The set B_0 corresponds to roads without connected buildings (i.e. no building):

$$B_0 = \{r_i \in R | \exists 0 b_j \in B, (b_j, r_i) \in E_{BR}\}$$

- The set B_1 corresponds to roads connected with one to several buildings:

$$B_1 = \{r_i \in R | \exists b_j \in B, (b_j, r_i) \in E_{BR}\}$$

We use the k-means method on Building Road Density to define two subsets B− and B+, where $B_1 = \{B-, B+\}$. These two subsets are defined as follows:

$$B- = \{r_i \in R, BRD(r_i) < Min(B+)\}$$

$$B+ = \{r_i \in R, BRD(r_i) > Max(B-)\}$$

From the *Metric Compactness*, we define two configuration types based on a k-means method. These configurations are defined as follows:

- The set *MC-* denotes roads with lower *Metric Compactness* values:

$$MC- = \{r_i \in R, MC(r_i) < Min(MC+)\}$$

- The set *MC+* denotes roads with higher *Metric Compactness* values:

$$MC+ = \{r_i \in R, MC(r_i) > Max(MC-)\}$$

From these configuration types, let us propose a classification of roads and buildings layout with six distinct classes illustrated in Figure 4. Four of these classes correspond to spatial patterns where one can observe coherence between the road network and buildings respective densities:

- The two classes B+MC+ and B−MC+ denote relatively dense spaces with high compact road network and buildings.
- The two classes B_0MC − and B−MC − denote relatively open spaces with low compact road network associated with few buildings or not.

Conversely, two of the classes correspond to peculiar spatial configuration where the road network structure and the building layout are inconsistent:
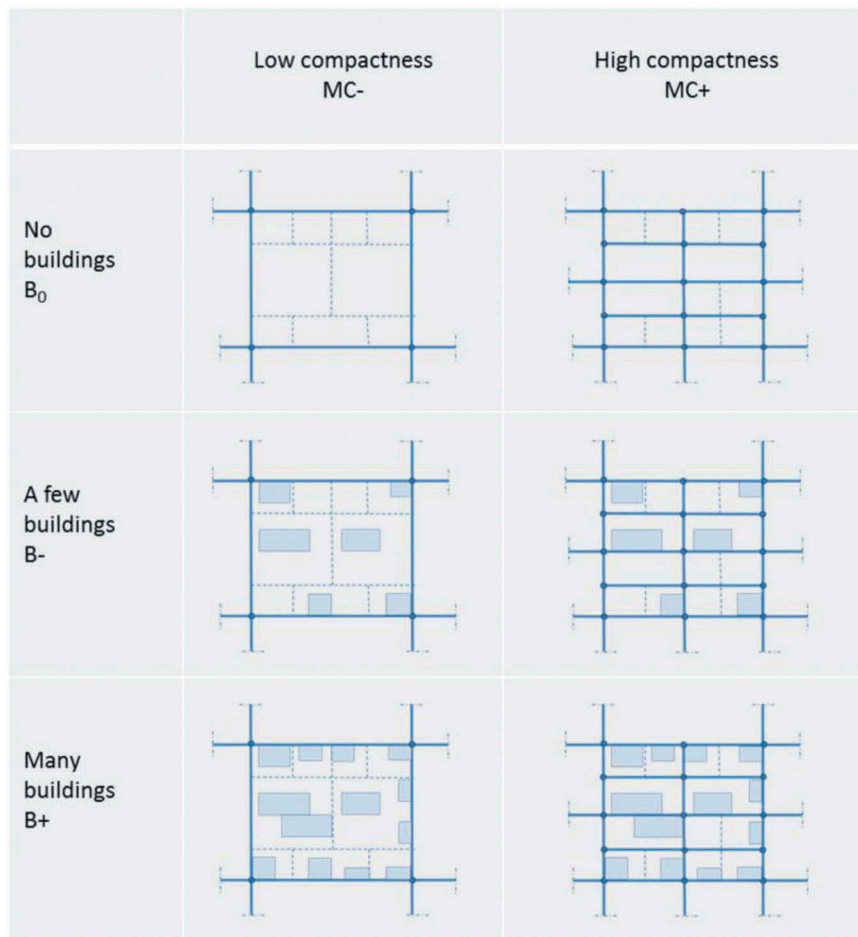
Figure 4. Spatial structures based on buildings vs. roads densities.

- The class B_0MC+ denote high compact roads network but with a few buildings.
- The class $B+MC-$ denote large number of buildings but with a non-compact road network.

4. Experimental

To evaluate our approach, we selected a region that includes relatively dense and relatively open areas. The geographical database used as an experimental setup is derived from the city of Guisseny, a village in North West France (Figure 5). Bounded to the North by the English Channel, the Guisseny background map includes 1026 buildings (houses or others), 2598 cadastral parcels, and 244 road segments connected to 213 road junctions. Dominated by rural and agricultural activity, this region has undergone urbanization over the last 30 years related to the attractiveness of the coast. Thus, while the center of the village and the furthest parts of the coast retained their appearance, the area changed with more or less coherence with the construction of a large number of buildings (main or secondary dwellings occupied by a population attracted by the place) and the modification of the road network.

4.1. Metric compactness

4.1.1. Local and global influences

As introduced in a previous section, the *Metric Compactness* ($MC(r_i)$) evaluates the connectivity of a road r_i when compared to the rest of the road network, the connectivity of a road been evaluated for X percent of the network. Therefore, for all roads, the *Metric Compactness* is computed for $1 \leq X \leq 100$ to evaluate local compactness (i.e. low X values) to global compactness (i.e. high X values). The examples exhibited by the maps presented in Figure 6 illustrate different *Metric Compactness* for different and representative X values. Different levels of granularity are illustrated by $X = 10$, $X = 25$, $X = 45$ and $X = 100$ to potentially reflect different patterns. In this figure, for each difference in X, variation in *Metric Compactness* is illustrated by change in color for four classes (<25%, <50%, <75%, <100%) related to comparison between *Metric Compactness* and the number of roads to be reached. For example, for $X = 10$ the number of roads that can be reached is 24, and the red color materializes roads with $MC(r_i) \geq 18$ (i.e. 75% from 24).

By applying this analysis at different levels of granularity, and from local to global levels, the *Metric*
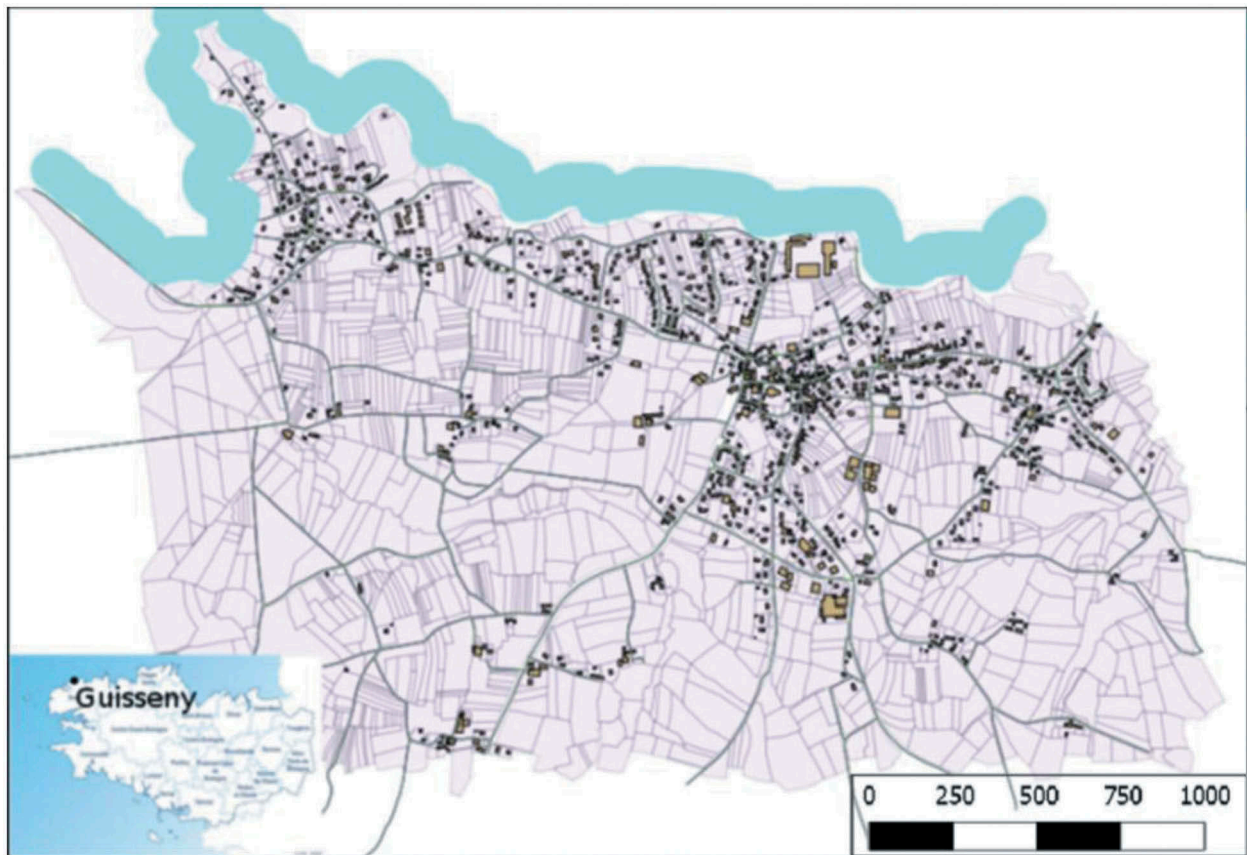
Figure 5. Guisseny North map with buildings, parcels, and road network.

Compactness underlines the historical village as the most connected place and materializes its progressive extension or at least interactions toward different directions. At the local level (i.e. $X = 10$ and 25), the number of most connected roads (in red and yellow) is respectively 26 and 52, while at a more global level (i.e. $X = 45$) the number is 94.

This analysis also reveals an overall homogeneity of this network without any emergence of additional clusters apart from the historical center. This trend is confirmed by the relative linearity of the *Minimum Ratio Minimum Distance* $D_{\min}^X$ curve exhibited in Figure 7. In fact, as no significant change of slope can be observed, there is not any emergence of additional clusters apart from the historical center (a break of slope for $X = 95$ is considered as a side-effect). Nevertheless, the outcome is different if one takes into account the roads that can reach at least 25% (i.e. green roads) of the numbers of roads still for $X = 10$. One can observe the emergence of a few areas scattered along the costal domain, this being confirmed for $X = 25$ as emerges a better connectivity of the coastal domain compared to the rest of the area. Overall, one can observe that the *Metric Compactness* reveals different magnitudes of roads connectivity when even observed at the local level as well as a sort of diffusion of connectivity values along the network. Overall, it also appears that most of the roads are relatively well connected to the rest of the network.

This case study example also shows that the threshold X has a direct impact on the number of connected areas that appears. While the *Metric Compactness* makes a difference between compact and non-compact roads, different threshold X values highlight different connectivity patterns and clustering effects. The main idea behind this approach is to reflect significant as well as valuable trends in the roads layout at either local (i.e. $X = 10, 25$) or global levels (i.e. $X = 45$).

4.1.2. Road network structure influence

Overall, the *Metric Compactness* and *Minimum Ratio Minimum Distance* favor the identification of differences in connectivity in the road network. The figures that emerge reflect that the village of Guisseny is spatially distributed around its historical center. To evaluate the effect of the underlying spatial structure on the measure of *Metric Compactness*, we modify the road network by deleting three roads in the central part of the studied area (Figure 8). This should modify the connectivity of the graph between the South-Western and Center East-parts of the road network. Figure 9 shows the differences of $D_{\min}^X$ and X values for original and altered datasets. The results that
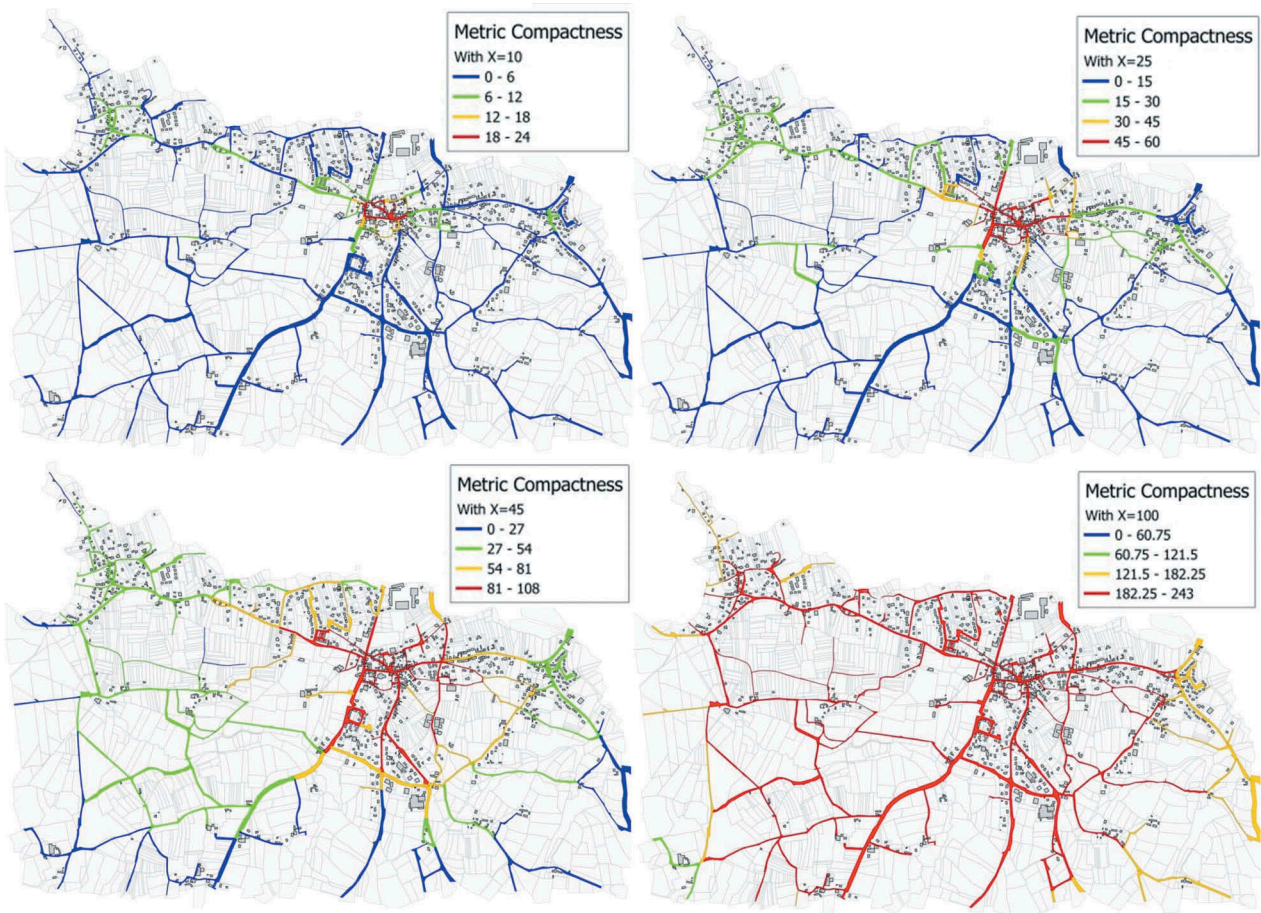
Figure 6. Metric compactness with X = 10, 25, 45, 100. For each road, color corresponds in four classes (<25%, <50%, <75%, <100%) related to comparison between the value in Metric Compactness and the number of roads to be reach.
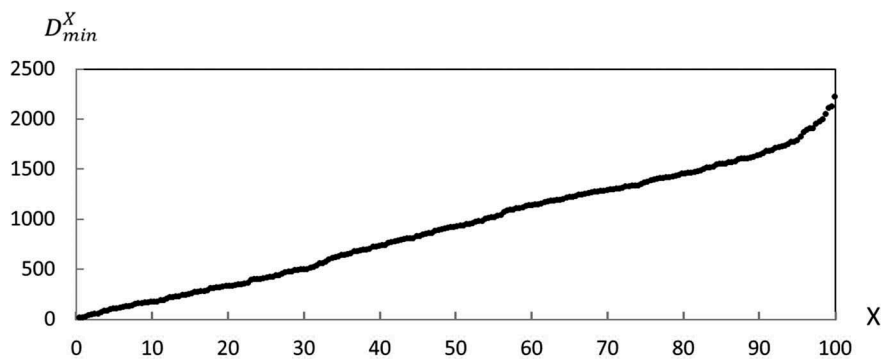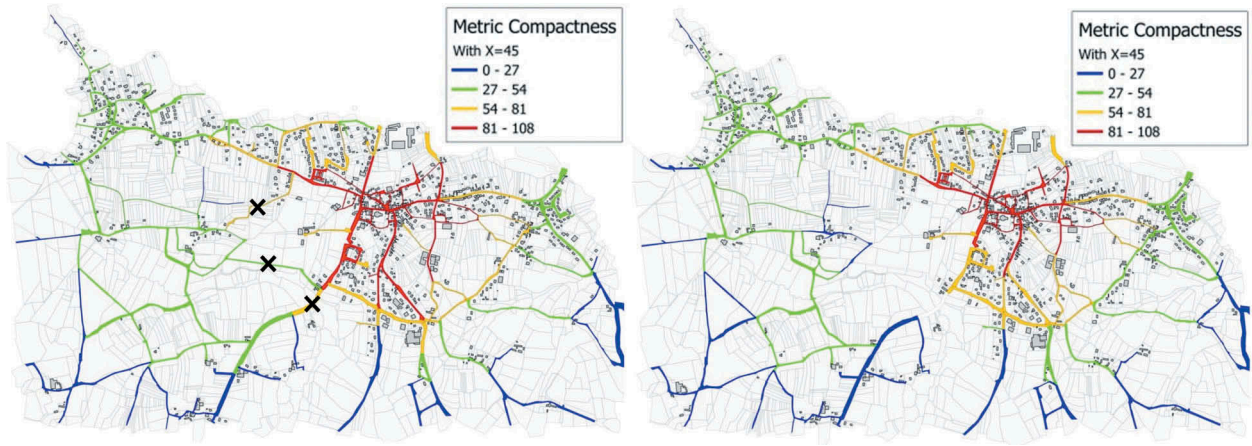

Figure 7. Variation of $D_{\min}^X$ in function of X.

emerge shows consistent $D_{\min}^X$ values for X <50, but divergent D_{min}^X values for X >50 with a change of slope for the altered dataset for $D_{\min}^X$ = 950 m.

According to the definition of the *Minimum Ratio Minimum Distance*, one can observe that for lower values of X (i.e. X <50) the parts of the road network explored for a given road are the ones at a distance less than 950 m from this road. Indeed, and as illustrated by Figure 8 for X =45, we observed similar *Metric Compactness* values for both the original and altered datasets. It can be concluded that the modification of the road network does not influence the

level of connectivity of the roads at the local scale for a threshold distance of less than 950 m. On the other hand, it appears that this threshold close to 1 km exhibits a sort of boundary limit of the urban structures that give form to different parts of the village of Guisseny.

For higher values of the threshold (i.e. X >50), the relationship between D_{min}^X and X diverges (Figure 9). Regarding the altered dataset, the local cluster that appears with that threshold of 950 m is now loosely connected with the rest of the network. At the global level, there is
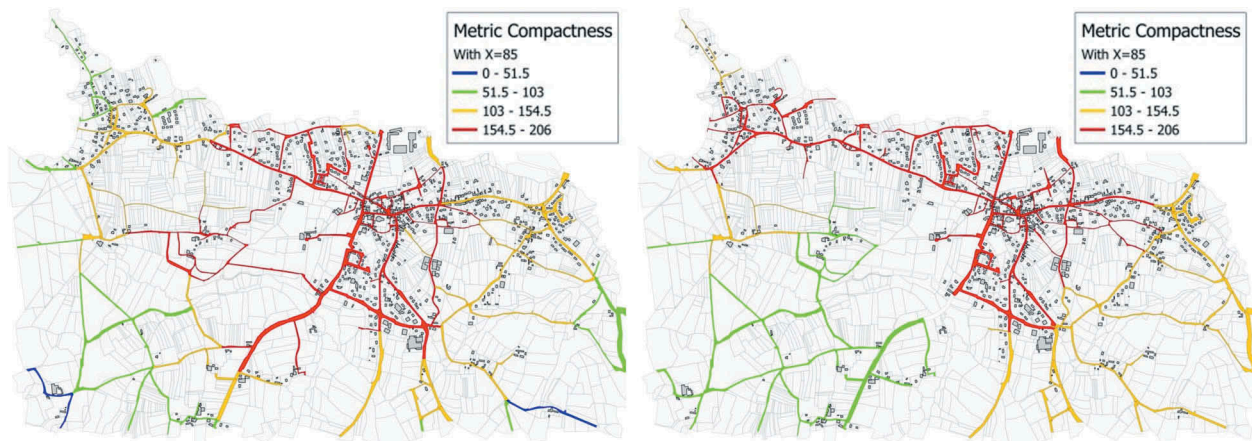
Figure 8. Metric compactness on original and altered datasets with $X = 45$. The three roads removed are denoted by a black cross.
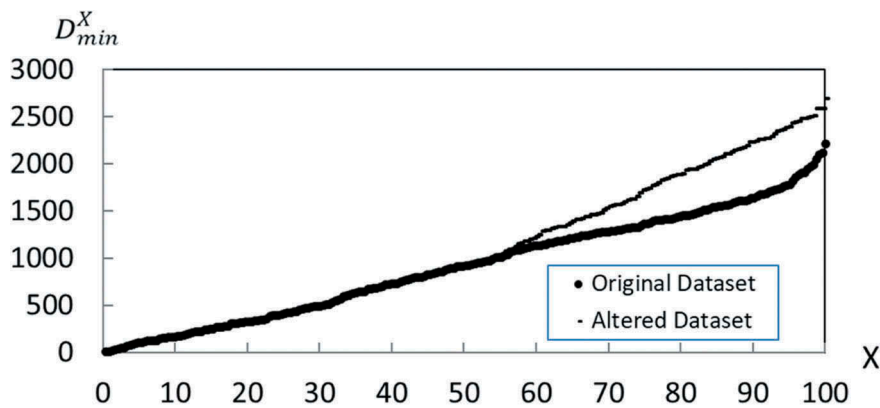
Figure 9. Variation of D_{min}^X in function of X on the original and altered road networks.

a significant change in the overall road network connectivity as illustrated in Figure 10 with a threshold value of $X = 85$: the altered road network exhibits a much lower connectivity when compared to the initial urban network. In particular, we observe that the removal of three roads decreases the connectivity of the South-Western parts of the road network thus creating a less-connective part in the network, but while this increases the connectivity of the North-Western parts of the same road network. For the altered dataset, this highlights the importance of the Northern unique pathway between the East and West sides of the road network. Conversely, and when considering the original dataset, the connectivity at the global level is favored by the related grid structure of the urban network as illustrated by Figure 10 left.

Figure 10. Metric Compactness on original datasets (left) and altered datasets (right) with $X = 85$.

4.2. Classification of the roads and buildings layouts

The next step of our modeling approach relies in the cross-analysis of the urban network structure and its relationship with the underlying built environment density. For that purpose, we applied the classification introduced in Section 3 and derived from the local metrics of *Building Road Density* and *Metric Compactness* applied with a threshold value of $X = 10$. Figure 11 shows the result of the analysis with the 6 classes previously identified. Several trends can be identified. First, for the main central roads, the results confirm that the coherence one might expect between the respective connectivity of the road networks and the buildings density:

- High *Metric Compactness* values (MC+) are mainly coherent with the ones of the *Building Road Density* (B+). The combination of these two parameters clearly identifies an area with high concentration in buildings and connected roads that appear to be the village center.
- Low *Metric Compactness* values (MC−) are mainly related with low *Building Road Density* (B-) or no-building roads (B0). These values reveal a relative open area corresponding to suburban and rural areas in our application, in particular in the more recent buildings area along the coastal domain.

Nevertheless, the classification reveals also some relative inconsistencies corresponding in one part to roads highly connected (MC+) but associated with a low number of buildings (B−) (in yellow) and in other part to roads lowly connected (MC−) but associated with a high number of buildings (B+) (in orange). For the former, roads are located around the village center, and one can expect this situation to evolve with new houses around the village. However, for the latter, roads are located relatively far away from the village and related with recent suburban arrangements. In other words, in these places, the roads connectivity does not appear sufficient enough to secure appropriate accessibility to the buildings. This can be then considered as an urban planning drawback.

Overall, these preliminary results show the potentiality of the combination of *Building Road Density* and *Metric Compactness* to extract regularities and anomalies in the spatial structure of a built environment. While the method depends on the choice of threshold values (i.e. X) to distinguish local to global structures, the experiment shows that a few iterations are sufficient to identify a relevant threshold that outlines structural patterns in the road and building layouts.

5. Conclusions

The representation and analysis of the spatial structure of a geographic space are still open research issues for many

Figure 11. Roads classification.

environmental and urban sciences. Such developments might be extremely valuable for scientific studies oriented to the exploration of the relationship between the overall structure and function of urban and suburban spaces. A valuable distribution of the buildings and roads layout should reflect the best compromise between environmental constraints and human activities. This is one of the main reasons that motivate the search for the development of a structural and computational approach that takes into account not only the properties of the road networks, but also the ones of the distribution of the parcels and buildings in an urban environment.

The research presented in this paper is based on the modeling of different levels of topological relationships applied on different elementary entities (i.e. buildings, parcels, roads) that together form the core of a given urban space. The objective of this approach is to favor a structural and spatial analysis that will reveal several global and local patterns that emerge from the spatial layout of a given urban system, and this at different levels of scale. First, to take into account the overall road network structure, we introduced several metrics (*Ratio Minimum Distance*, *Minimum Ratio Minimum Distance*, and *Metric Compactness*) that evaluate the capability for a given road to be connected with the whole road and different sub-parts of the network. In particular, the Metric Compactness reveals the emerging sub-network structure. Furthermore, the *Metric Compactness* allows to make a difference between homogeneous and heterogeneous connected networks. Related to the concept of less-connective and more-connective spatial structure introduced by Marshall (2004), homogeneous networks denote the ones in which connectivity values are relatively equivalent in the whole network, while heterogeneous networks are likely to reveal some sub-parts of the network loosely connected while others more connected. Second, we introduced a computational and structural evaluation of the building layout that takes into account the relationship with the underlying roads network. We introduce a measure of *Building Road Density* that is evaluated on top of the road network, thanks to an integration of the respective graph-based connection with parcels and buildings. Combined with the *Metric compactness*, this metric characterizes the correlation between the building density and the compactness of the road network. Such classification makes also the difference between consistent and inconsistent configurations of the urban layout. The former denotes either road segments whose high compactness matches high buildings density or road segments whose low compactness matches low buildings density of buildings. The latter denotes either road segments whose high compactness is not reflected by the buildings density or road segments whose low compactness is also not reflected by the buildings density.

Overall, such classification favors the detection of regularities and inconsistences in a given urban layout, this being of potential interest for urban planning and studies. While applied so far to a representative small demonstrative case study, the approach shows promising potential to reveal different patterns in an urban layout, especially when combining the respective cross-configurations of the building, parcel, and road network layouts. The whole might provide a direction that complements current space syntax approaches that mainly retain the configurational structure of the urban network as well as morphological studies whose focus is generally limited to the analysis of the building layout. Indeed, and to favor and develop a better understanding of the relationship between the underlying urban structure of a given urban network and emerging human activities, additional socio-economical values should be integrated as additional variables to take into account by our modeling approach. This is left to further work and experimental studies still to perform. Our on-going work is currently applied to the cross-comparison of the evolution of different urban layouts as well as the exploration of additional metrics.

Notes on contributors

Mathieu Domingo is a PhD in computer science, currently a research assistant and full stack developer at the University of Pau and Pays de l'Adour in France. His research interests are in the areas of spatio-temporal models and theories, environmental, and urban GIS.

Rémy Thibaud is a PhD in geophysics, currently an associate professor in computer science and GIS of the French Naval Academy Research Institute. His research interests are in the areas of geographical information systems, spatio-temporal modeling and database, and marine geophysics.

Christophe Claramunt is a professor in Computer Science and deputy director of the French Naval Academy Research Institute. His research interests include spatio-temporal models and theories, semantic and cognitive-based GIS, Web and wireless GIS systems, and maritime, environmental, and urban GIS.

ORCID

Mathieu Domingo 🔾 http://orcid.org/0000-0001-9649-6474
Rémy Thibaud 🔾 http://orcid.org/0000-0001-8599-2384
Christophe Claramunt 🔾 http://orcid.org/0000-0002-5586-1997

References

Barabási, A. L. 2016. *Network Science*. Cambridge, UK: Cambridge University Press.
Batty, M. 2013. *The New Science of Cities*. Cambridge, MA, USA: MIT Press.

Batty, M., and D. Howes. 2001. "Predicting Temporal Patterns in Urban Development from Remote Imagery." In *Remote Sensing and Urban Analysis*, edited by J. P. Donnay, M. Barnsley, and P. Longley, 185–204. London: Taylor and Francis Group. doi:10.4324/9780203306062_chapter_10.

Batty, M., and S. Rana. 2004. "The Automatic Definition and Generation of Axial Lines and Axial Maps." *Environment and Planning B – Planning and Design* 31 (4): 615–640. doi:10.1068/b2985.

Berghauser-Pont, M., and P. Haupt. 2007. "The Relation between Urban Form and Density." *Urban Morphology* 11 (1): 62–65.

Berghauser-Pont, M., and P. Haupt. 2010. *Spacematrix: Space, Density and Urban Form*. Rotterdam, Netherlands: NAI Publishers.

Borrego, C., H. Martins, O. Tchepel, L. Salmin, A. Monteiro, and A. I. Miranda. 2006. "How Urban Structure Can Affect City Sustainability from an Air Quality Perspective." *Environmental Modelling and Software* 21 (4): 461–467. doi:10.1016/j.envsoft.2004.07.009.

Buhl, J., J. Gautrais, N. Reeves, R. V. Sole, S. Valverde, P. Kuntz, and G. Theraulaz. 2006. "Topological Patterns in Street Networks of Self-Organized Urban Settlements." *The European Physical Journal B* 49 (4): 513–522. doi:10.1140/epjb/e2006-00085-1.

Chowell, G., J. M. Hyman, S. Eubank, and C. Castillo-Chavez. 2003. "Scaling Laws for the Movement of People between Locations in a Large City." *Physical Review E* 68 (6). doi:10.1103/physreve.68.066102.

Cohn, A. G., B. Bennett, J. Gooday, and N. M. Gotts. 1997. "Qualitative Spatial Representation and Reasoning with the Region Connection Calculus." *GeoInformatica* 1: 275–316. doi:10.1023/A:1009712514511.

Crucitti, P., V. Latora, and S. Porta. 2006. "Centrality Measures in Spatial Networks of Urban Streets." *Physical Review E* 73 (3): 036125. doi:10.1103/physreve.73.036125.

Domingo, M., R. Thibaud, and C. Claramunt. 2013. "A Graph-Based Model for the Representation of Land Spaces." Paper presented at *the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Orlando, USA, November 5–8,516–519. doi:10.1145/2525314.2525317.

Ewing, R., R. Pendall, and D. Chen. 2003. "Measuring Sprawl and Its Transportation Impacts." *Transportation Research Record* 1831 (1): 175–183. doi:10.3141/1831-20.

Gastner, M. T., and M. E. Newman. 2006. "The Spatial Structure of Networks." *The European Physical Journal B–Condensed Matter and Complex Systems* 49 (2): 247–252. doi:10.1140/epjb/e2006-00046-8.

Gonzalez-Abraham, C. E., V. C. Radeloff, R. B. Hammer, T. J. Hawbaker, S. I. Stewart, and M. K. Clayton. 2007. "Building Patterns and Landscape Fragmentation in Northern Wisconsin, USA." *Landscape Ecology* 22 (2): 217–230. doi:10.1007/s10980-006-9016-z.

Hawbaker, T. J., V. C. Radeloff, R. B. Hammer, and M. K. Clayton. 2005. "Road Density and Landscape Pattern in Relation to Housing Density, Land Ownership, Land Cover, and Soils." *Landscape Ecology* 20 (5): 609–625. doi:10.1007/s10980-004-5647-0.

Herold, M., H. Couclelis, and K. C. Clarke. 2005. "The Role of Spatial Metrics in the Analysis and Modeling of Urban Land Use Change." *Computers, Environment*

and Urban Systems 29 (4): 369–399. doi:10.1016/j.compenvurbsys.2003.12.001.

Hillier, B., and J. Hanson. 1984. *The Social Logic of Space*. Cambridge, UK: Cambridge University Press.

Jiang, B., and C. Claramunt. 2004. "Topological Analysis of Urban Street Networks." *Environment and Planning B* 31 (1): 151–162. doi:10.1068/b306.

Latora, V., and M. Marchiori. 2003. "Economic Small-World Behavior in Weighted Networks." *The European Physical Journal B*32: 249–263. doi:10.1140/epjb/e2003-00095-5.

MacQueen, J. 1967. "Some Methods for Classification and Analysis of Multivariate Observations." Paper Presented at the Fifth Berkeley Symposium onMathematical Statistics and Probability, Berkeley, USA, 1(14), 281–297.

Marshall, S. 2004. *Streets and Patterns*. London: Routledge (Taylor & Francis Group).

Marshall, S. 2016. "Line Structure Representation for Road Network Analysis." *Journal of Transport and Land Use* 9 (1): 29–64. doi:10.5198/jtlu.2015.744.s240.

Oliveira, V. 2016. *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*. Berlin, Heidelberg: Spinger-Verlag.

Penn, A. 2003. "Space Syntax and Spatial Cognition or Why the Axial Line?" *Environment and Behavior* 35 (1): 30–65. doi:10.1177/0013916502238864.

Porta, S., P. Crucitti, and V. Latora. 2006a. "The Network Analysis of Urban Streets: A Dual Approach." *Physica A: Statistical Mechanics and Its Applications* 369 (2): 853–866. doi:10.1016/j.physa.2005.12.063.

Porta, S., P. Crucitti, and V. Latora. 2006b. "The Network Analysis of Urban Streets: A Primal Approach." *Environment and Planning B: Planning and Design* 33 (5): 705–725. doi:10.1068/b32045.

Radicchi, F., C. Castellano, F. Cecconi, V. Loreto, and D. Parisi. 2004. "Defining and Identifying Communities in Networks." *Paper Presented at the National Academy of Sciences of the United States of America* 101 (9): 2658–2663. doi:10.1073/pnas.0400054101.

Raimbault, J. 2018. "Calibration of a Density-Based Model of Urban Morphogenesis." *PloS one* 13 (9): 1–23. doi:10.1371/journal.pone.0203516.

Ratti, C. 2004. "Space Syntax: Some Inconsistencies." *Environment and Planning B: Planning and Design* 31 (4): 487–499. doi:10.1068/b3019.

Sabidussi, G. 1966. "The Centrality Index of a Graph." *Psychometrika* 31 (4): 581–603. doi:10.1007/bf02289527.

Steiniger, S., T. Lange, D. Burghardt, and R. Weibel. 2008. "An Approach for the Classification of Urban Building Structures Based on Discriminant Analysis Techniques." *Transactions in GIS* 12 (1): 31–59. doi:10.1111/j.1467-9671.2008.01085.x.

Tsai, Y. H. 2005. "Quantifying Urban Form: Compactness versus 'Sprawl'." *Urban Studies* 42 (1): 141–161. doi:10.1080/0042098042000309748.

Turner, A. 2007. "From Axial to Road-Centre Lines: A New Representation for Space Syntax and a New Model of Route Choice for Transport Network Analysis." *Environment and Planning B: Planning and Design* 34 (3): 539–555. doi:10.1068/b32067.

Wasserman, S., and K. Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.

Watts, D. J., and S. H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393 (6684): 440–442. doi:10.1038/30918.

Wells, N. M., and Y. Yang. 2008. "Neighborhood Design and Walking: A Quasi-Experimental Longitudinal Study." *American Journal of Preventive Medicine* 34 (4): 313–319. doi:10.1016/j.amepre.2008.01.019.

Ye, Y., D. Li., and X. Liu. 2018. "How Block Density and Typology Affect Urban Vitality: An Exploratory Analysis in Shenzhen, China." *Urban Geography* 39 (4): 631–652. doi:10.1080/02723638.2017.1381536.

Ye, Y., and A. Van Nes. 2014. "Quantitative Tools in Urban Morphology: Combining Space Syntax, Spacematrix and Mixed-Use Index in a GIS Framework." *Urban Morphology* 18 (2): 97–118.