Describing the Residential Valorisation of Urban Space at the Street Level. The French Riviera as Example

Alessandro Venerandi[1(✉)] and Giovanni Fusco[2]

[1] Université Côte-Azur, ESPACE, Nice, France
alessandro.venerandi@univ-cotedazur.fr
[2] Université Côte-Azur, CNRS, ESPACE, Nice, France
giovanni.fusco@univ-cotedazur.fr

Abstract. There is a growing concern regarding the use of relatively coarse units for the aggregation of various spatial information. Researchers thus suggest that the street segment might be better suited than areal units for carrying out such a task. Furthermore, the street segment has recently become one of the most prominent spatial units, for example, to study street network centrality, retail density, and urban form. In this paper, we thus propose to use the street segment as unit of analysis for calculating the residential valorisation of urban space. To be more specific, we define a protocol that characterises street segments through a measure of central tendency and one of dispersion of prices. Moreover, through Bayesian clustering, it classifies street segments according to the most probable combination house type-valuation to provide a picture of local submarkets. We apply this methodology to the housing transactions exchanged in the French Riviera, in the period 2008–2017, and observe that outputs seem to align with local specificities of the housing market of that region. We suggest that the proposed protocol can be useful as an explorative tool to question and interpret the housing market, in any metropolitan region, at a fine level of spatial granularity.

Keywords: House prices · Street network · Summary statistics · Bayesian clustering · French Riviera

1 Introduction

The scale at which socioeconomic indicators are computed is of paramount importance, especially if they then guide urban interventions. In France, social mix policies have been harshly criticized because they targeted spatial units that were too coarse (i.e., entire municipalities) and seemed to produce segregation at a smaller level of spatial granularity [1]. In studying the relationship between poverty and accessibility in Atlanta (GA, US), Carpenter and Peponis [2] highlighted how even census tracts and city blocks could hid differences, which became apparent only when considering street segments[1].

[1] A street segment is the line connecting two intersections in a street network.

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12252, pp. 505–520, 2020.
https://doi.org/10.1007/978-3-030-58811-3_37

In the last 60 years, the street segment has gained recognition as one of the most relevant spatial units of analysis because is the physical component at the basis of the functioning of urban systems [3–6]. Thus, several indicators were computed at such level of spatial granularity. In Urban Design, for example, researchers defined and computed several street-based metrics of network centrality (e.g., betweenness, closeness) [5]. Similarly, in the field of Geography, academics quantified several aspects of urban form and commerce at the street level and summarised this information through Bayesian clustering [7, 8]. However, to the best of the authors' knowledge, street-based indicators of the housing market have never been defined. Aggregations of such information are usually made at areal level or through surface interpolation [9]. However, such aggregations never concerned street segments.

In this paper, we propose a novel protocol to aggregate information on house prices for street segments, to study the relative residential valorisation of urban space, street by street. To be more specific, such protocol, firstly, renders house prices comparable across different years and house types; secondly, it aggregates such information for streets segments by means of a measures of central tendency and one of dispersion; thirdly, it uses Bayesian clustering to categorise streets based on their most probable price range-house type. To test the protocol, we implemented it on the housing transactions exchanged in the French Riviera, in the period 2008–2017, and observed that patterns of value and dispersion at the street level and clusters of price range and house type seem to correspond to current real estate and socioeconomic dynamics. For example, the outputs of the proposed protocol well identify historically well-off sets of streets (e.g., Promenade des Anglais and Carré d'Or in Nice and the eastern part of Cannes) as well as more popular ones (e.g., Vieux-Nice, Grasse and Vallauris).

We argue that the protocol presented in this paper might be useful to realtors and public officials to obtain an alternative reading of the housing market, which, in turn, might inspire innovations in sales strategies and policy making. Furthermore, it might be helpful to researchers who want to analyse the relationship between any kind of street-based phenomenon and the housing market. Related to this, our protocol would also avoid statistical issues associated with the aggregation of data computed for different spatial units. Several studies explored the relationship between housing transactions and street network centrality, for example Chiaradia et al. [10] and Narvaez et al. [11], but their focus was still on transactions seen as points in urban space, overlooking data aggregation issues at the level of street segments.

The reminder of this paper is structured as follows: we firstly present related works, pointing out the relevance of the street segment as spatial unit of analysis; secondly, we illustrate the methodology to characterise the housing market at the street level; thirdly, we present the data used to test the methodology and its pre-processing; fourthly, we lay out the results; finally, we conclude with final remarks and possible future work.

2 Related Works

This section firstly provides a brief excursus on the relevance of the street segment as spatial unit of analysis in various fields. Secondly, it presents previous works that focused on the computation of street-based measures of several urban phenomena. Thirdly, it illustrates known aggregation techniques of house prices, at different spatial resolutions.

2.1 The Street Segment: A Brief Retrospective

The street segment has for long been considered one of the most relevant spatial units by several researchers. Conzen [3], one of the founders of the Anglo-German school of Urban Morphology, theorised a technique of town plan analysis focused on three main components: the streets and their organisation into a street system (i.e., street network); the plots and their arrangement into street blocks; buildings as block plans. In his view, these three elements were keys to understand the form and dynamics of urban settlements. The Italian school of Urban Morphology developed a similar theory and much of its research on the form and evolution of settlements focused on the "fascia di pertinenza" ("pertinent strip"), that is the street with all the plots that abut on it [12]. Several influential works that focused on the role of streets in cities were also developed in US context, in the second half of the XXth century. Jacobs [4], renown urban theorist and activist, considered streets as the most important spatial entities in cities as their form and configuration not only affected human interactions, but also the success of neighbourhoods, in terms of liveability and economic prosperity. Lynch [13] based much of his imageability theory on the visual perception that pedestrians have at the street level. More recently, an entire new branch of quantitative research focused on the analysis of streets, and the network they create, from a graph theory perspective [5, 6, 14].

2.2 Examples of Street-Based Measures

Several works in different domains focused on the definition and implementation of street-based measures. Porta et al. [5], for example, formalized a set of centrality measures to study the configuration of street networks. Such measures are based on the assumption that a street could be central, with respect to all the other streets in a network, in different ways: being central as being characterized by through-movement (i.e., betweenness centrality) or by high interconnectedness in a small space (i.e., closeness centrality). These very same measures, however with different names (i.e., choice for betweenness and integration for closeness), are also at the basis of the Space Syntax methodology [6].

Researchers in the field of Urban Geography defined several street-based measures to quantify aspects of urban form and commerce [7, 8]. The former are, for example, the fragmentation of street fronts and canyon effect [7]. Measures of commerce at the street level were, for example, the richness of retail categories within 300 m of a store location or the ratio between the number of stores belonging to a specific category and

the total number of stores within the same radius [8]. Here again, retail becomes a street-dependent phenomenon, both for its description and understanding.

2.3 Spatial Aggregation of House Prices

Classic techniques for the aggregation of house prices focus on areal units (e.g., census areas). In France, for example, Meilleurs Agents (i.e., a private company that offers real estate services) and Notaires de France (i.e., the French association of notaries) provide the average[2] and median price[3] per square meter, respectively, for official administrative boundaries called IRIS. In England, the London Datastore provides information on the average and median house prices[4] for census areas of various sizes, the finest of which is at the intra-neighbourhood scale, i.e. Lower Layer Super Output Area (LSOA).

A more advanced technique of aggregation of house prices is the Kernel Density Estimation (KDE), which interpolates values, usually through a Gaussian function, at a chosen resolution. McMillen [9], for example, used such a technique to pre-process 25,633 housing transactions of single-family homes in Chicago (IL, US), before testing different modelling solutions.

In this section, we illustrated that there exist different types of aggregation methods of house prices, however, none of them focused on the street segment. In the next section, we present a novel protocol to carry out this task and provide not only information on a measure of central tendency, but also on dispersion of values and on the belonging to a given residential submarket.

3 Method

To characterise the housing market at the street level, our methodology requires three subsequent steps:

- data pre-processing;
- aggregation of house prices at the street level through a measure of central tendency and one of dispersion;
- application of Bayesian clustering to classify each street based on its most probable price range and house type.

We provide more details on each of these steps in the following subsections.

3.1 Data Pre-processing

The study of the relative valorisation of subspaces, being them districts, census tracts, city-blocks, or streets always needs aggregation of house prices. However, such data

[2] https://www.meilleursagents.com/prix-immobilier/provence-alpes-cote-d-azur/.

[3] https://www.immobilier.notaires.fr/fr/prix-immobilier?periodeReferences=1ettypeLocalisation=COMMUNEetcodeInsee=06088.

[4] https://data.london.gov.uk/dataset/average-house-prices.

for different house types and years are not directly comparable for two main reasons. Firstly, each calendar year has specific economic trends, namely inflation and housing market cycles (i.e., recession or upturn), which affect the prices of properties on a yearly basis. Secondly, different house types have different markets and this unevenly affects the valuations. For example, small flats tend to be sold more frequently than larger homes because they usually have higher demand and involve short term investments. On the contrary, larger properties tend to be sold less frequently than smaller homes because they typically have lower demand and they are associated with long term investments. The average price per square meter tends also to be structurally higher for smaller flats due to very technical reasons (i.e., even the smallest flat needs sanitary and cooking equipment, which proportionally weigh more on the average price per surface unit compared to larger properties). The very notion of an average price per square meter can thus be challenged when applied to such diverse sections of the housing market. To address these issues and have prices comparable across several years and house types our method requires three subsequent steps:

- segmenting the data by year of transaction;
- segmenting the data by house type;
- computing ventiles and deciles of house prices for each subset year of transaction-house type.

At the end of this process, the valuations are comparable as they have been segmented by typology and year and are expressed in relative terms, that is they are no longer nominal values but classes of prices (i.e., ventiles and deciles). We will explain the use of such classes of prices in the next paragraphs.

3.2 Computation of Median and Dispersion

To obtain information on the central measure and dispersion of prices at the street level, the next step of our protocol requires to, firstly, associate each data point to the street segment to which they belong and, secondly, aggregate such information at the street level, through the computation of a measure of central tendency and one of dispersion. The first of such measures is the median of the ventiles, which provides information on the mid-point of the local distribution of the relative housing valuation, for each street. The second measure is the difference between the 9th decile (D9) and the 1st decile (D1) of the distribution of the ventiles. This provides information on the range of valuations in each street. Our method requires to compute these statistics for the ensemble of each street segment and its immediate neighbouring segments, directly connected to it. This provides a smoothing of local statistics, which takes into account the morphological impact of closely connected street segments and the "hybrid" situation of properties laying on street corners. Since our protocol requires the computation of the D9-D1 statistic, street segments with too few transactions should not be considered. We thus propose to consider only those – and relative neighbouring streets – with at least 10 transactions. This also recognises the limits, in terms of statistical knowledge, of the relative valorisation of street segments when using transaction data: below a certain threshold, our knowledge is too limited and, as we will see in the Conclusions, issues of uncertainty might arise.

3.3 Identifying Classes of Housing Market

To identify different classes of housing markets at the street level, the last step of our methodology requires the application of a well-known technique for classification purposes (i.e., Bayesian Networks) to the double description (i.e., house type and value) associated to each street segment. More specifically, we describe each of these spatial elements with a total of 17 variables: 10 correspond to the percentages of properties falling in each decile of the distribution of house prices; 7 represent the percentages of properties for each house type.

Several statistical and machine learning methods can cluster instances into classes. However, we argue that Bayesian Networks are better suited, in this specific case, for two main reasons. They can identify subsets of records presenting a precise pattern on a limited number of features whereas other methods usually look for homogeneous behaviour of records across all features [15]. A further advantage is that the clustering uses probabilities defined on a finite probability space. We can thus identify elements which belong to a given cluster with low uncertainty (for example, when the probability of assignment is higher than 0.9) and more "hybrid" ones, which have a higher probability of belonging to other clusters.

Operationally, Bayesian Networks work as follows. A star-shaped network is built between all the measured variables and a newly implemented non observable node, which plays the role of a cluster variable. Such variable is linked to each individual variable through oriented arcs, in what is called a naive Bayes classifier [16]. In such a network, each variable becomes independent of the value of every other knowing the cluster, hence the name "naive". The conditional probability tables linking the cluster variable to each observed variable are determined through an Expectation-Maximisation (EM) algorithm [17]. These probabilistic parameters are recursively used to assign to each given record its most probable cluster. Two different schemes can guide the EM algorithm. The easiest one consists in predetermining the number of clusters required, as in k-means clustering. The EM algorithm is here guided by likelihood maximisation of the clustering model, given the data. By repeating the clustering process for different number of clusters, the algorithm obtains a series of optimal solutions. A more computationally intensive scheme consists in applying a random walk among all possible solutions, considering different number of clusters. The EM algorithm is, in this case, guided by a Minimal Description Length score [18], which combines clustering likelihood to a penalisation factor for the number of clusters used in the model (to avoid overfitting), and its result is a unique optimal solution. Both schemes were applied within our research.

4 Application to the French Riviera

The French Riviera is a coastal metropolitan area in South-Eastern France, stretching over 60 km from the French-Italian border to the Esterel mountains. It includes several towns and cities, namely, from East to West, Menton, Monaco, Nice, Antibes and Cannes. After being the birthplace of aristocratic tourism in the XIX century, and of mass tourism since the 1930s, this metropolitan area of more than 1 million people has

presently one of the tensest real estate markets in France, after the one of Paris, the capital city. The Principality of Monaco is one of the most expensive markets in the whole of Europe, but it will be excluded from our analysis, given its fiscal and regulatory specificity. The municipalities of the French Riviera are required by law to favour residential mix through their social housing projects, but they are all far from attaining the 25% threshold imposed by the ALUR Law for 2025 [19]. Lack of generational mix is also becoming a concern on the French Riviera. Because of the high demand for small secondary homes, the coastal cities seem increasingly unable to cater for families with children. These are thus leaving the most central areas and moving to suburban and exurban areas in the hinterland which, in turn, have difficulties in coping with the sudden increase of demand for day-care and schooling. Gentrification of city-centres and further development of the tourism economy through Airbnb add to this trend. If issues of social and generational mix were to be addressed (as the public policies state), then it is fundamental to understand the spatial logic of the private housing sector in this area.

We thus applied the methodology presented above to the housing transactions exchanged in the French Riviera, between 2008 and 2017. To carry out our study, we accessed two official data sources: PERVAL, which contains information on housing transactions, and BD TOPO, which contains the street network of the study area, in vector format. We provide more details on these next.

4.1 Datasets

PERVAL

The PERVAL is a proprietary dataset containing recorded housing transactions in France[5]. It is curated and kept updated by the association Notaires de France, the French association of notaries. PERVAL is a very rich geo-referenced dataset containing tens of features for each transaction. The information provided goes from the habitable surface to the energetic label, from the number of rooms to the presence of a basement, from the condition of properties to the number of toilets. For the purpose of this study, we obtain PERVAL data for the French Riviera and its close hinterland, for the period 2008–2017. This resulted in a total of 150,116 transactions, 124,236 of which are flats, while 25,880 are single houses. The average price (taxes included) of a flat is €205,075, that of a single house is €711,747.

BDTOPO

The DB TOPO is an official dataset, provided in shapefile format, that contains vector representations of multiple elements of the landscape and the built environment, such as, streets, natural features, and buildings[6]. It is issued and kept constantly updated by the Institute National de l'Information Geographique et Forestiere (the French national institute of Geography and Forestry). For the purpose of our study, we obtain the BD TOPO dataset for the area under study, which contains 98,297 junction-to-junction street segments, with an average length of 131 m.

[5] https://www.perval.fr/.

[6] http://professionnels.ign.fr/bdtopo.

4.2 Pre-processing

A first analysis of the PERVAL dataset revealed some issues, for example, spatial inaccuracies were present for some housing transactions. Moreover, it did not include housing typologies, a requirement of our methodology. A pre-processing has thus been necessary before proceeding further. Such procedure mainly included: performing a new geocoding based on the addresses provided with PERVAL; cleaning and repairing the dataset; and assigning housing typologies to each transaction. We provide more details on each step next.

To perform the new geocoding, we matched the addresses of each housing transaction contained in PERVAL with the ones contained in the Base Address Nationale (BAN), the official dataset of French addresses. At the end of this process, each transaction had a set of geographic coordinates derived from BAN, which was then used to re-geocode the entire dataset. Although BAN is accurate, the spatial precision was particularly low for 25,078 transactions (e.g., some transactions were geo-coded at the municipal level). We thus filtered them out.

Having carried out this step, we performed some further cleaning and repairing. The PERVAL dataset included transactions of properties that were not the subject of this study, that is retirement homes and properties under special contractual agreements. The values of such properties could be quite different from the ones sold at normal market prices and could thus bias the analysis. For this reason, we discarded them. At the end of this process, the total number of transactions was 113,781, 95,556 of these were flats, while 18,225 were single houses.

As we mentioned above, the PERVAL dataset did not include information on the house type of each transaction, a datum required by our protocol. However, it contained information on the broad category of each property (i.e., single house or flat) and the number of rooms. We thus assigned a house type to each transaction based on this information and created the following categories: T1 for studios and 1-bedroom flats, T2 for 2-bedroom flats, T3 for 3-bedroom flats, T4 for 4-bedroom flats, T5 for flats with 5 bedrooms or more, small houses (PM, petites maisons in French) for single houses with up to 4 rooms, and big houses (GM, grandes maisons in French) for single houses with 5 or more rooms. For 1,770 transactions, typology had to be estimated from dwelling surface since the number of rooms was not available in the original database.

5 Results

5.1 Median and Dispersion of House Prices in the French Riviera

Having assigned a dwelling type to each transaction, we segmented the PERVAL dataset by year and house type. We thus obtained 60 separate datasets, one for each combination year of transaction - house type. Following the proposed protocol, we computed ventiles of house prices for each subset. To then aggregate the information at the street level, we firstly assigned to each transaction the ID of the closest street segment and, secondly, we computed median and dispersion of ventiles for units

comprising each street segment and its directly linked segments, whenever a minimum of 10 transactions were found in such spatial unit.

We present the median of ventiles of prices for the whole study area and a zoom on the city of Nice in Fig. 1. The information provided by these maps reveals the overall valorization-devalorization of the street segments within the study area. Furthermore, we present the dispersion of ventiles of prices for the whole study area and a zoom on the city of Nice in Fig. 2. This latter set of maps visualize a less explored aspect of housing prices, that is the capacity of streets to support a wide variety of values (streets that tend to red) or to specialize in a specific market segment (streets that tend to blue).

A visual inspection of these maps allows several considerations on the relationship between the outputs and the current socioeconomic as well as real estate market conditions in the French Riviera. By taking Nice as example, we note that the historically well-off neighborhood of Cimiez, characterized by villas and detached blocks of flats, has many streets in the top three classes of ventiles (i.e., 14–16, 16–18, 18–20). We observe a similar pattern in the Carre d'Or, a very central neighborhood built in the XIX century, with elegant blocks of 5–6 floors. For what concerns devalued streets (i.e., streets in classes 1–2, 2–4, 4–6), the output seems to correctly identified those of the historic core of Nice (i.e., Vieux-Nice), which is a heritage district but lacks, at times, of contemporary living standards (e.g., more generous habitable surfaces, natural light). The very same reason probably holds true also for the devalued streets of some ancient settlements outside of Nice, such as Grasse, Vence, and Vallauris.

For what regards the dispersion of ventiles of prices (Fig. 2), the outputs of our method show that the traditional centers of the coastal cities, characterized by short street segments, are more capable of supporting ranges of house prices (i.e., classes 14–16, 16–18, 18–19) than the peripheral and suburban areas, which, conversely, are characterized by longer street segments. This holds true for Nice, Menton, Antibes and even Cannes, with the exception of the luxurious Boulevard de La Croisette and nearby streets, which tend to have a more homogeneously highly priced housing market. By allowing a greater mix of housing valuations, the streets of the coastal city-centers might thus favor more socioeconomic mix. However, there are also sets of central streets that tend to more stably retain high or low valuations and thus probably attract the same population segments. The former are, for example, the above mentioned Boulevard de La Croisette and nearby streets in Cannes, the latter can be found around the main train station of Nice. For what concerns more peripheral/suburban streets, we observe homogeneously high valorization in the capes (Cap Ferrat, Cap-d'Ail, Cap Martin, Cap d'Antibes), coastal areas of outstanding scenic beauty and homogenous devalorization in the streets near the two above mentioned social housing estates (i.e., Ariane and Les Moulins) and at the bottom of valleys, such as Route de Canta Galet and Boulevard de la Madeleine in Nice, which tend to have unfavorable micro-climatic conditions (i.e., lower amount of day-light, higher levels of pollution).

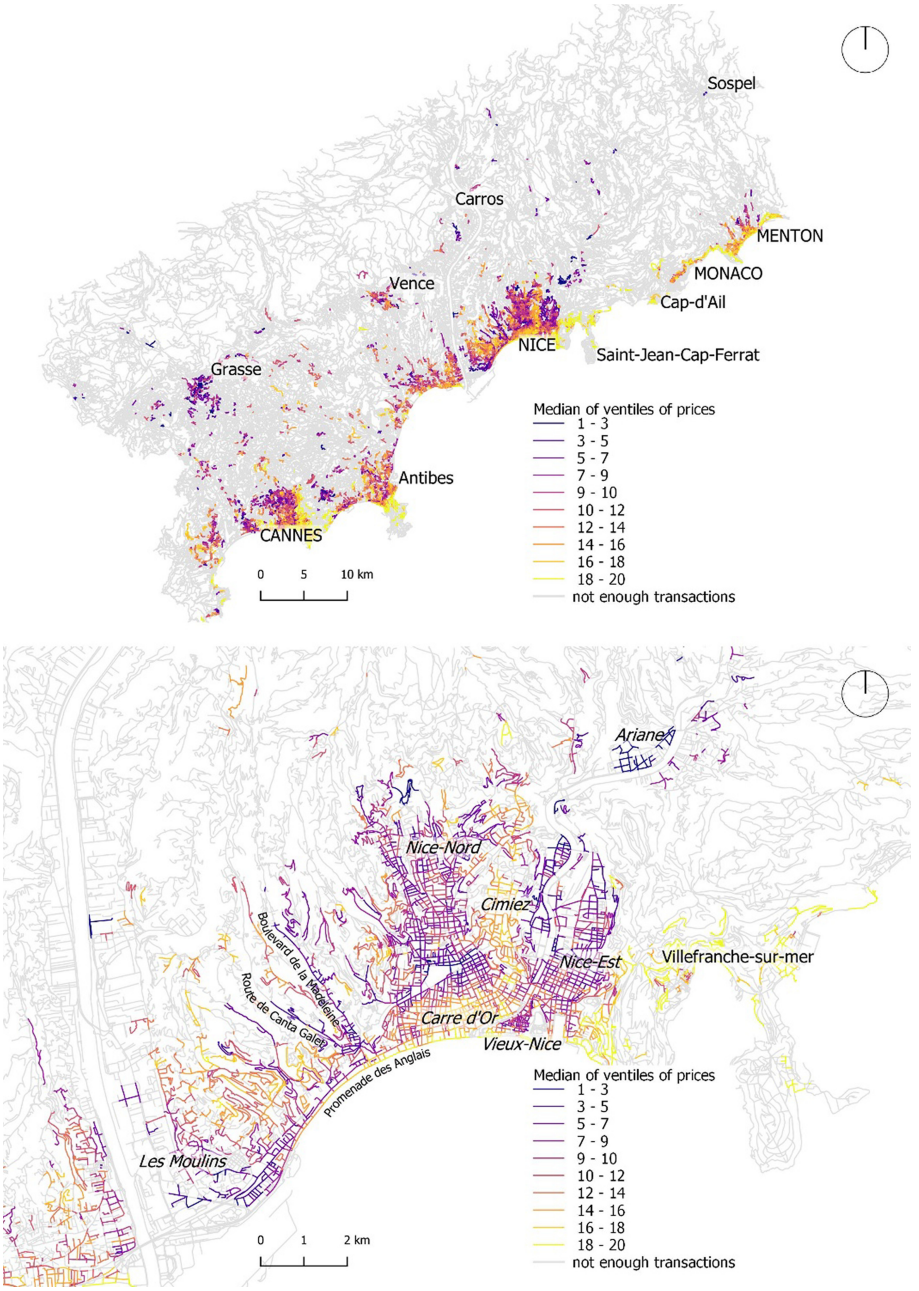

Fig. 1. Median of ventiles of house prices at the street level, for the whole study area and the city of Nice.

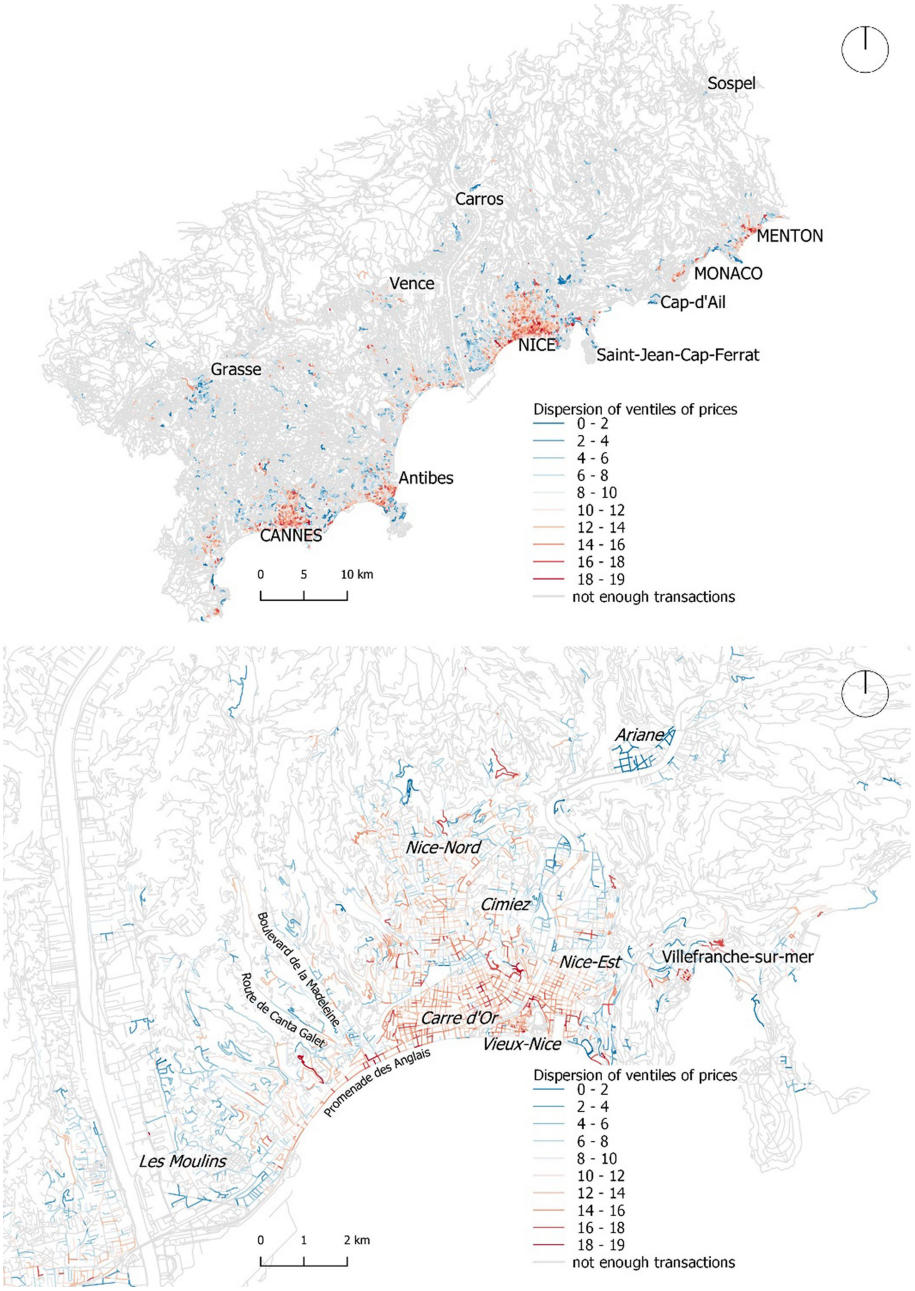

Fig. 2. Dispersion of ventiles of house prices at the street level, for the whole study area and the city of Nice. (Color figure online)

5.2 Classes of Housing Markets in the French Riviera

The analyses conducted so far allow a fine description of the spatial logic of housing valorization and devalorization on the French Riviera, which is potentially very relevant in terms of social mix. However, generational mix (or lack of) cannot be understood only in terms of valorization, but has to addressed also from the perspective of house types. To have a description of the housing market at the street level that integrates both value and house types, we applied Bayesian Networks to the previously computed 10 variables of deciles of prices and the 7 variables of house types (i.e., T1, T2, T3, T4, T5, PM, and GM). However, before performing the computation we weighted each of the house type variable 1.5 to balance the amount of information provided by the 10 variables associated with the deciles of prices. The random walk search for the best clustering, output by the Bayesian Network, converged to a solution with 10 classes. The final output had a contingency table fit score of 27% (i.e., the amount of information contained in the 17 variables which was captured by the cluster variable). We interpreted the probabilistic profiles of the 10 clusters and defined the following labels:

- Cluster 1 (C1): small to medium flats with value above average;
- Cluster 2 (C2): very small flats with mixed values;
- Cluster 3 (C3): medium to big flats with medium value;
- Cluster 4 (C4): big flats and houses with very high value;
- Cluster 5 (C5): small flats with value above average;
- Cluster 6 (C6): flats of mixed sizes with low value;
- Cluster 7 (C7): flats and houses with mixed values;
- Cluster 8 (C8): flats of mixed sizes with very high value;
- Cluster 9 (C9): medium to big flats with very low value;
- Cluster 10 (C10): houses with low value.

In Fig. 3, we map the clusters output by the Bayesian Networks on the street segments of the whole study area and the city of Nice. By visually inspecting these maps, we observe that the output of the Bayesian Networks seems to add a layer of information to the results for median and dispersion previously showed. More specifically, we observe that the groups of streets uniformly characterized by above average valuations in Fig. 1 (i.e., the eastern part of Cannes, the city center of Nice, and the area around Menton) are actually more nuanced, if housing typologies are accounted for. The eastern part of Cannes is characterized by very expensive flats, however of mixed sizes (i.e., C8). In the city center of Nice and in Menton, flats are moderately expansive, but also small to medium-sized (i.e., C1 and C5). There are also differences in groups of streets evenly characterized by low and median values in Fig. 1. For example, in Nice-Nord and Nice-Est, the housing market is relatively devalorized. However, there are notable changes within these submarkets, namely the presence of sets of streets with medium to big flats (i.e., C9), very small ones (i.e., C2), and flats of mixed sizes (i.e., C6). Interestingly, higher dispersion of prices, which tends to be more present in central areas (see Fig. 2), seems to affect streets characterized by very diverse housing typologies and not only small flats.

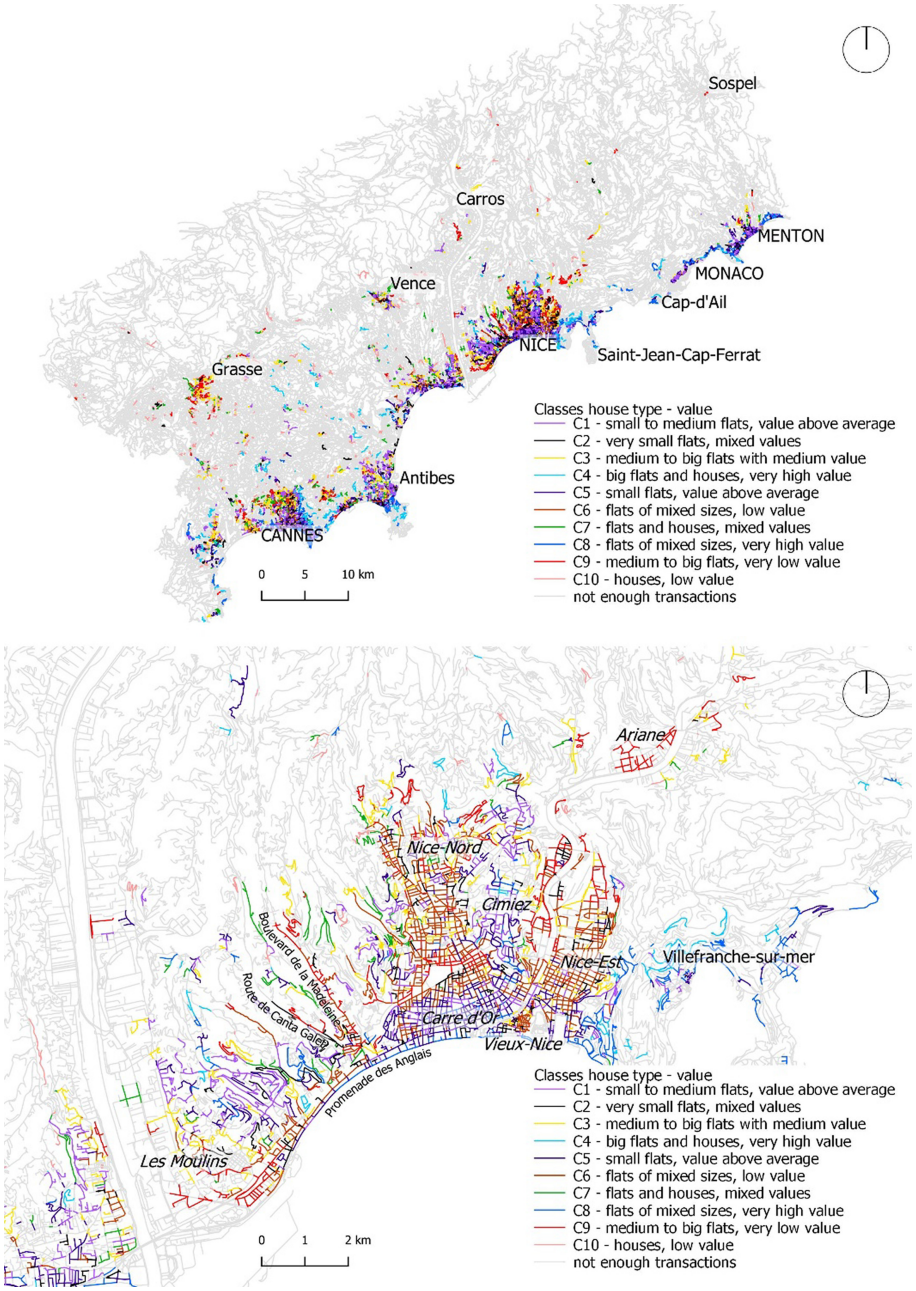

Fig. 3. Classes value - house type, for the whole study area and the city of Nice.

We also remark the peculiarity of clusters C2 and C7 as they both offer a wide range of valuations, with C7 also offering a mix of house types. Due to these characteristics, we suggest that streets belonging to C7 and, to a lesser extent C2, are likely to be more prone to generational and socioeconomic mix, an aspect which is considered fundamental for promoting a more equal society and thriving neighborhoods by several urbanists and researchers. See, for example, the works by Matan and Newman [20] and Andersson [21].

6 Conclusions

In the last 60 years, the street segment has gained recognition as one of the most relevant spatial units of analysis in different fields (e.g., Urban Morphology, Urban Design). Several street-based measures have been defined at this level of spatial granularity. However, as of today, a street-based measure of house prices has yet to be defined. In this paper, we proposed a protocol to carry out precisely this task by aggregating information on housing transactions at the street level in three ways: through a measure of central tendency (i.e., median of prices), through a measure of dispersion of prices, and through classes value-house type. We implemented such protocol on the housing transactions occurred in the French Riviera, in the period 2008–2017, and found patterns that seemed to correspond to current socioeconomic and real estate market specificities. For example, the patterns of high and low median valuations seem to match the urban landscapes of better-off and less-advantaged neighbourhoods, respectively. The dispersion of prices seem, instead, to be especially concentrated in areas more intensely urban, regardless of the house types. Finally, the classes value-house types seem to add further knowledge to the characterisation provided by the median and dispersion of prices. For example, although both the eastern part of Cannes and the centre of Nice are characterised by medium to high median values, they can be seen as different local markets. The former tends to have flats of different sizes, while the latter only small to medium sized ones. We suggest that the protocol proposed in this paper might be useful to a wide range of actors, from realtors, who might want to explore alternative readings of the housing market, to public officials, who might desire to use its outputs in policy making. Finally, academics might want to analyse the relationship between the outputs of our protocol and existing or novel street-based measures.

7 Future Work

We consider the outputs presented in this paper only the first phase of a wider research endeavour on urban housing markets. In fact, we envision three main research directions. Firstly, our protocol was able to identify streets with greater dispersion of prices. We thus hypothesised that such dispersion was associated with more socioeconomic mix. Future work might test: whether this pattern is confirmed through quantitative analysis, for example, through the analysis of the relationship between dispersion of prices and socioeconomic indicators extracted from census data. Secondly, since our

methodology ignores street segments with a limited number of housing transactions, we envision to use uncertainty-based methodologies to infer the housing market characteristics of streets with partial information. Thirdly, the interpretations of the outcomes presented in this paper were based on expert knowledge. Future work should ascertain whether such interpretations hold from a quantitative standpoint, for example, through the implementation of linear or non-linear models that explore the relationship between the proposed street-based measures of housing values and morphological, positional and configurational characteristics of the built environment. Comparison with other approaches like Geographically Weighted Regression [22] could also be carried out, as long as the street segment is kept as the spatial unit of analysis.

Acknowledgements. This research was funded by the French Government, through the National Research Agency, under the Investissements d'Avenir IDEX UCA JEDI, with reference number ANR-15-IDEX-01. The authors would like to thank Pr. Andrea Tettamanzi, at I3S Laboratory in Université Côte d'Azur, and Dr. Denis Overal, director of the R&D at Kinaxia, for their support and insightful suggestions.

References

1. Jaillet, M.C., Perrin, E., et Ménard, F.: Diversité sociale, ségrégation urbaine, mixité. Research report Plan Urbanisme Construction Architecture (PUCA) (2008)
2. Carpenter, A., Peponis, J.: Poverty and connectivity. J. Space Syntax **1**(1), 108–120 (2010)
3. Conzen, M.R.G.: Alnwick, Northumberland: a study in town-plan analysis. Trans. Papers (Inst. Brit. Geogr.) **27**, iii–122 (1960)
4. Jacobs, J.: The Death and Life of Great American Cities. Random House, New York (1961)
5. Porta, S., Crucitti, P., Latora, V.: The network analysis of urban streets: a primal approach. Environ. Plan. B Plan. Des. **33**(5), 705–725 (2006)
6. Hillier, B.: Space is the Machine: a Configurational Theory of Architecture. Space Syntax, London (2007)
7. Araldi, A., Fusco, G.: Decomposing and recomposing urban fabric: The city from the pedestrian point of view. In: Gervasi, O., et al. (eds.) ICCSA 2017. LNCS, vol. 10407, pp. 365–376. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62401-3_27
8. Araldi, A., Fusco, G.: Retail fabric assessment: describing retail patterns within urban space. Cities **85**, 51–62 (2018)
9. McMillen, D.P.: Changes in the distribution of house prices over time: structural characteristics, neighborhood, or coefficients? J. Urban Econ. **64**(3), 573–589 (2008)
10. Chiaradia, A., Hillier, B., Barnes, Y., Schwander, C.: Residential property value patterns in London: space syntax spatial analysis. In: Proceedings to the 7th International Space Syntax Symposium, Stockholm, Sweden (2009)
11. Narvaez, L., Penn, A., Griffiths, S.: Space syntax economics: decoding accessibility using property value and housing price in Cardiff, Wales. In: Proceedings of the Eighth International Space Syntax Symposium, Santiago, Chile (2012)
12. Caniggia, G., Maffei, G.L.: Architectural composition and building typology: interpreting basic building. Alinea Editrice, Florence, Italy (2001)
13. Lynch, K.: The Image of the City. MIT press, Cambridge (1960)

14. Batty, M.: Cities as Complex Systems: Scaling, Interaction, Networks, Dynamics and Urban Morphologies. Research report Centre for Advanced Spatial Analysis (CASA), no. 131 (2009)
15. Fusco, G., Perez, J.: Bayesian Network clustering and self-organizing maps under the test of Indian districts. A comparison. Cybergeo Eur. J. Geogr. (2020). https://doi.org/10.4000/cybergeo.31909. http://journals.openedition.org/cybergeo/31909
16. Duda, R., Hart, P.: Pattern Classification and Scene Analysis. Wiley, New York (1973)
17. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc. Ser. B **39**(1), 1–38 (1977)
18. Rissanen, J.: Information and Complexity in Statistical Modeling. Springer, New York (2007)
19. Levasseur, S.: La loi SRU et les quotas de logements sociaux 15 ans après, quel bilan? Research report OFCE (Sciences Po) no. 54 (2015)
20. Matan, A., Newman, P.: People Cities: The life and Legacy of Jan Gehl. Island Press, Washington (2016)
21. Andersson, C.: Public space and the new urban agenda. J. Public Space **1**(1), 5–10 (2016)
22. Manganelli, B., Pontrandolfi, P., Azzato, A., Murgante, B.: Using geographically weighted regression for housing market segmentation. Int. J. Bus. Intell. Data Min. **9**(2), 161–177 (2014)