XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ - BRAZIL

Identification of real estate submarkets by detecting communities in graphs

Barros P.[1], Espíndola R.P.[1]

[1]*Civil Engineering Graduate Program - COPPE, Federal University of Rio de Janeiro*
Av. Athos da Silveira Ramos, 149 - Ilha do Fundão, 21941-909, Rio de Janeiro/RJ, Brazil
pedro.barros@coc.ufrj.br, rogerio.espindola@coc.ufrj.br

Abstract. The identification of submarkets plays an important role in the context of real estate mass appraisal, to eliminate spatial autocorrelation, or the heteroscedasticity of residues in linear models, making statistical inference more reliable and allowing a better interpretation of the formation of property prices. Traditionally carried out empirically, by grouping techniques or by spatial residual modelling applied to hedonic data from properties in a specific region, studies indicate that the use of data mining models has obtained promising results to carry out this task automatically. In this work, considering a regionalized and not individual view of the real estate unit, it is proposed to discover submarkets based on the detection of communities in graphs formed by the neighborhoods of a city. From the socioeconomic and locational information of the city of Rio de Janeiro, a complex network of neighborhoods is formed, and communities are identified based on this network's modularity. The hierarchical approach to obtaining the community structure used here allows different scenarios for understanding the submarkets, an important aspect for decision makers in an area of knowledge so susceptible to uncertainty.

Keywords: Complex networks, community detection, real estate mass appraisal, submarkets.

1 Introduction

For Dantas [1], technical valuations of real estate properties are an important instrument in society and it is necessary to bring fairness and information in transactions involving real estate. Some services may be benefited from these techniques, such as: purchase, sale and lease transactions, money laundering prevention, court decisions, property taxation, investment decisions, insurance operations, separations or divisions of companies, friendly or judicial expropriations, urban planning, among others. The first appraisal engineering works known in Brazil were published in engineering technical magazines, between 1918 and 1919. The NBR 14.653:2019 [2] currently synthesizes all Brazilian standards on the subject.

Aiming to eliminate the human subjectivity in the valuations and to bring greater speed and accuracy to the process, computational models for mass property appraisals have been the subject of research for over 30 years as reviewed by Dimopoulos and Bakas [3]. Moreover, Byeonghwa and Kwon [4] state that it leads to two main approaches: regression analysis and artificial intelligence techniques. Wang and Li [5] conducted a systematic review of more than 400 articles about prediction models for real estate values between the years 2000 and 2018 and defined the current field of study for real estate appraisals as a combination of prediction techniques and geoinformation systems. One of the main strategies to perform a regression analysis in real estate is by using hedonic models.

1.1 Multiple regression analysis (MRA) - Hedonic models

As defined by Rosen [6], a hedonic model consists in splitting the item being researched into its constituent characteristics to estimate the contribution of each one. In real estate economics, it is assumed, for example, that a house can be decomposed into characteristics such as number of bedrooms, size of lot, or distance to the city center. The focus of this research is in hedonic models with different formulations for Multiple Regression Analysis (MRA). According to Bourassa *et al.* [7], Hasan [8] and Kontrimas and Verikas [9], most of the studies conducted in real estate mass evaluation were performed with them because they are appropriate for direct estimation of the relationship between price and the various characteristics of a property. Kauko [10] states that these techniques might become inefficient when dealing with aspects such as outliers, non-linearity, spatial dependence and other types of dependency such as discontinuity or fuzziness. In such cases, artificial neural networks (ANN's) can circumvent some of these problems achieving more predictive power, and therefore valuation accuracy, outperforming the traditional approaches as presented in McCluskey *et al.* [11] and Zurada *et al.* [12]. However, as most of ANN models prevent knowing the direct relationship between inputs and output, Kauko [10] notes that the price formation analysis is an unknown process with them. For this reason, McCluskey *et al.* [11] affirm that in relation to cost-effectiveness and user-friendly applicability for the valuation community, the MRA approach outperforms the 'black box' nature of an ANN technique. Nonetheless, in the context of MRA, Rodrigues [13] defends that some premises must be met to analyze the reliability of the results, such as homoscedasticity of the residuals, but the existence of spatial dependence in the residuals breaks this premise. To deal with this issue, it is possible to use methods that incorporate regional variation into the regression model or to include a variable in the model that explains regional variation, which is made by defining submarkets.

1.2 Real estate submarkets and literature review

The need of delineation of homogeneous submarkets is prevalent in most if not all applications of MRA in the field of real estate evaluations to mitigate the effects of the autocorrelation in the residuals as discussed in Miller [14]. Schnare and Struyk [15] were one of the first researchers to explore real estate submarkets and they stated that the urban housing market is a set of compartmentalized and unique submarkets with demand and supply influences likely to result in a different structure of prices in each one. Thus, space variables should consider discrete irregular areas, not continuous ones. Borst [16] noted that submarkets need not to be geographically continuous and that hedonic multiple regression models calibrated among submarkets are measurably different from one another.

However, although submarkets have been widely recognized, Zhuo *et al.* [17] state that there is no universally accepted method to identify them and Pace *et al.* [18] have suggested two ways for dealing with spatial data to obtain good results with MRA: by modelling the coefficients for the independent variables correctly or by modelling the spatial errors structure or the spatial process through applying a correction based on the error of adjacent real estate properties. Borst [16] affirms that the emphasis was placed in modelling the spatial errors structure or the spatial process for most of the models and that a spatial pattern in the model error is often an indication of misspecification of the independent variables, offering opportunity of improvement in accuracy. Several approaches to deal with spatial dependence in MRA models were presented as geostatistical kriging, Dubin [19] and Bourassa *et al.* [20]; PCA and K-means, Wu and Sharma [21]; spatial weight matrix, Pace *et al.* [18] and Ismail [22]; locally weighted regression, Clapp [23]; MRA with dummies, Goodman and Thibodeau [24] and Bourassa *et al.* [7]; trend surface analysis, Fik *et al.* [25]; decision trees, Clapp [26]; hierarchical linear model, Leishman *et al.* [27]; geographically weighted regression, McCluskey *et al.* [11]; and fuzzy algorithm, Gabrielli *et al.* [28].

Current methods of using locational and socioeconomic attributes are not enough to cancel autocorrelation of the residuals and spatial statistical methods are necessary to deal with them. Some papers describe that locational, socioeconomic and pre-defined submarkets can substantially improve model accuracy when combined to spatial statistical methods as in Dubin [19], Bourassa *et al.* [20] and Fik *et al.* [25]. Additionally, some works have achieved better results with submarkets without using spatial methods such Bourassa *et al.* [7]. So, it can be seen there is no universally accepted method to cope with the spatial autocorrelation.

2 Submarkets from a complex networks perspective

In a simplified way, a graph consists of a non-empty set V of vertices and a set A of edges, with a function that associates subsets of two elements of V with an element of A. Vertices are interpreted as the end points of an edge. Graphs can also be weighted or directed, when is assigned to each edge a value or an orientation, respectively.

A node degree is a measure of number of its connections and it is a preliminary indicator of the importance of a node. A giant component of a graph is its biggest connected subgraph and it is the focus of the most of analysis in graphs, although eventually some disconnected vertices might have research interest and are treated separately. According to Barabási [29], in complex networks vertices are generally called nodes, while the edges, connections. Also, the word network is adopted when it is used to understand a real complex system, while the term graph is related to the discussion of the mathematical representation of a network. For this work, districts of a city are considered as nodes and a similarity measure between them as the weights of the edges.

Real estate market can be described as a set of submarkets defined by distinct and interrelated market segments as presented in Rebelo [30]. The submarket concept relies on the idea of substitutability, which means that an increase in the price of one good leads to an increase in the demand for other similar one. According to Bourassa *et al.* [7], the substitutability occurs when the attributes of a pair of goods are similar and when there is equilibrium of their prices in the market. Typically, for Grigsby *et al.* [31], a real estate submarket is defined as a set of dwellings that are reasonably close substitutes for each other, but relatively far substitutes for dwellings in other submarkets. By these definitions, it was decided to set the connection between two nodes as the similarity between them. Several metrics can be used as similarity measure and the cosine similarity was chosen as a first approach for this modelling since preliminaries attempts present better results and it is a simple metric to implement and comprehend. Given two N dimension vectors $\vec{v}$ and $\vec{w}$, the cosine similarity between them is calculated as follows in eq. (1):

$$Cosine(\vec{v},\vec{w}) \;=\; \frac{\vec{v} * \vec{w}}{|\vec{v}|\,|\vec{w}|} \;=\; \frac{\sum_{i=1}^{N} v_i \times w_i}{\sqrt{\sum_{i=1}^{N} v_i^2}\,\sqrt{\sum_{i=1}^{N} w_i^2}}. \tag{1}$$

So, in this study, network communities will be treated as real estate submarkets. According to Barabási [29], a community structure of a network is uniquely encoded in its wiring diagram and communities are locally dense connected subgraphs in a network. In community detection task, the number and the size of communities are both unknown beforehand. One approach is to partition the network according to the modularities of the communities, scalar values between -1 and 1 that measure the density of links inside them as compared to links between communities and it is given, in a weighted network, by eq. (2) Newman [32]:

$$Q \;=\; \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta\left(c_i c_j\right). \tag{2}$$

Where A_{ij} represents the weight of the edge between nodes i and j, k_i is the sum of the weights of the edges attached to vertex i, c_i is the community to which vertex i is assigned, the function $\delta(u, v)$ is 1 if $u = v$ and 0 otherwise, and $m = 0.5 \times \sum_{ij} A_{ij}$.

Hence, modularity allows the assessment of a community partition and offers an optimization approach to community detection: the greater the modularity of a network partitioning, the better the detection of its communities. Yet, the number of partitions grows faster than exponentially with the network size and algorithms need to identify communities without inspecting all partitions as stated by Barabási [29]. For this work, it was chosen the Blondel *et al.* [33] algorithm to detect communities or groups of similar districts. It is a greedy algorithm based on modularity optimization of weighted graphs.

3 Dataset

The studied region is the municipality of Rio de Janeiro in Brazil which has a total area of 1,200,329 km² and a population of 6,718,903 inhabitants according to IBGE [34,35]. The municipality is subdivided into 33 administrative regions, encompassing a total of 161 districts Each district was treated as a record with 24 attributes each. Although Leishman *et al.* [27] affirm that there is some evidence that smaller areas as census tracts present better results, districts were used as the unit of information since it is more feasible to test the methodology and gather more information about them. Furthermore, geoprocessing techniques on the software QGIS were applied to fit all the data available into district's boundaries.

For Long *et al.* [36], Rebelo [30], and many other authors, the prices on the real estate market are usually described by structural, locational and socioeconomic attributes. The discussion of this work will rely only on locational and socioeconomic attributes, since the delimitation of areas and type of data available are not suitable for the requirements of structural analysis such as number of rooms, area of property, among others.

Some of the attributes used in this work were directly obtained from different sources as Brazilian census data of 2010 processed by Rio de Janeiro City Hall, Ministry of Labor and Employment; Annual List of Social Information yield by Ministry of Labor and Employment, satellite imagery sensor WorldView-2 (Digital Globe) processed by Rio de Janeiro City Hall and data from property taxation of the city. These attributes are: territorial density (pop/ha), home density (pop/dwellings), households below the poverty line (%), rented homes (%), social development index, population living in slums (%), occupied households (%), commerce jobs, service jobs, industry jobs, residential building area/total building area (%), building area/urbanized area, urbanized area (%) and plaza area (%).

Other attributes were extracted as Public Security Institute raw data edited with geoprocessing techniques using the software QGIS: passer-by robbery, vehicle theft with violence, vehicle theft and willful murder. Balneability bulletins from the State Environmental Institute were employed to create a beach quality index considering all bulletins from 2010 to 2017. Information of mass transportation companies was used to create an index by considering different weights for each mode of transport.

Some distance features of each district were also extracted. It was explored the distance to beaches with clean bathing water and the distance to centralities. Basically, a centrality is identified as an urban space which concentrates flows of transportation, services, goods, information and people, and it can be hierarchized by its relevance in a city. The centralities for the city of Rio de Janeiro were defined by other studies such as PMRJ [37] and IETS [38] and represent the magnitude and diversity of flows that occur in those areas. For example, the main center of a city, defined in the literature as CDB or Central Business District, can gather flows from every other district in a city and presents a large variety of activities in its boundaries. A metropolitan sub-center, on the other hand, also concentrates this kind of flow but it is not so significate for the entire city. With less influence, a regional sub-center will only extend its influence by attracting flows from the closer districts and it will have less activities and services when compared to the CDB. Because road network, traffic conditions and different modes of transportation could significantly affect the times of travelling around the city, these times spent were computed by Google Distance Matrix API and were employed to estimate distances more significantly than the space distance itself. So, the extracted features used are shortest distance to the city main central area, shortest distance to the metropolitan sub-centers, shortest distance to the regional centers and shortest distance to a good quality beach, all of them related to the time of travelling.

4 Results and analysis

From the dataset, records with zero values or inconsistences were filled with average values from neighboring districts. The data was standardized into the range [0,1] and none significant correlation was noted. After that, a network was created using districts as nodes and their similarities as the weights of the edges, and the communities were detected by the maximization of their modularities. After several experimental attempts, by varying thresholds for edge values, the best network partitioning is presented on Fig. 1 when the threshold was set to 0.89591, which corresponds to the third quartile of the similarities distribution, as presented in the box-plot in the same figure. As the size of a node grows as its degree increases, larger nodes represent districts with more substitutes. Seven well defined communities with 137 nodes were detected inside the giant component of the network and they were considered as submarkets. Some districts outside the giant component were excluded from this study since they are not part of the common real estate market like ecological reserves and military, university, airport or prison ones. Other ones should be considered as submarkets because they compose small components with high similarities between them or they are centralities of the city. To bring important districts with centralities aspects, such as Barra da Tijuca or the CDB itself, inside the giant component of the network would require a lower threshold of similarity, which would promote a fully-connected network since they have very low similarities with all the other districts. For this reason, they were considered as single submarkets as other districts due to their particularities, for example: slum district of Rocinha, island district of Paquetá, low constructed area as Barra de Guaratiba, mountainous area as Alto da Boa Vista, and others.

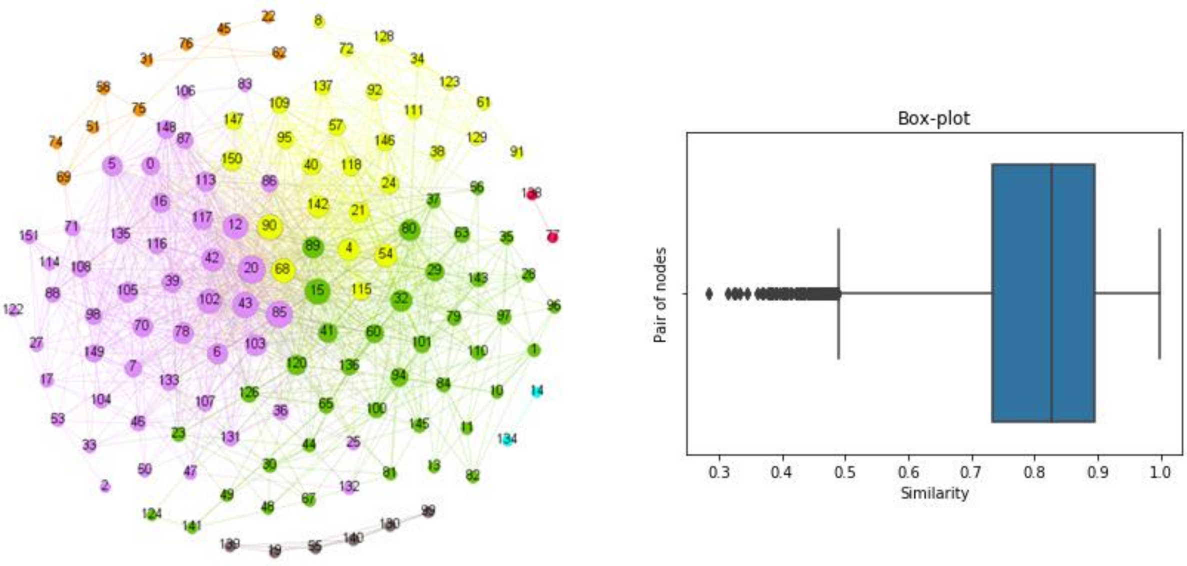

Figure 1. Complex network visualization and box-plot of the similarities

Some information to understand the profile of each community are presented in Fig. 2. Community 1 has 29 districts, such as Bangu and Realengo, with social development above the average and differs most from the others by its high values of violence related features. Despite being relatively close to the centralities of the city, it is the most distant of all communities to beaches and it also shows the minor percentage of rented homes, which is an evidence of real estate devaluation. Community 2 is only composed by the districts of Botafogo and Tijuca and, although being apart geographically, they both are regional centralities with good transport index and high job indicators. Community 3 has 48 districts such as Andaraí and Méier and it can be considered as an average community because extreme values in good or bad aspects were not detected. Community 4 has 6 districts far from the CDB, such as Camorim and Vargem Grande, which present low territorial density, low social development and constructed area, low values of violence related features, suggesting isolated districts with low demand of house owners and incipient urbanization. Community 5 has 10 districts from the most valued area of the city, such as Humaitá and Copacabana, and its average pattern reveals a very urbanized and economic active area, with good social development and well located in relation to centralities and beaches. Community 6 has 40 districts, such as Parada de Lucas and Manguinhos, and represents densely populated districts. However, it presents more commercial buildings than residential ones and more deteriorated conditions when compared to community 3, with low social development, high violence and poverty. Community 7 is composed by districts of Urca and Leme and it differs from community 5 due its more residential characteristics and absence of mass public transportation besides some normal bus lines.

To improve the real estate mass appraisal modeling in a more detailed perspective (streets, dwells), other missing variables should be collected like sea view, noise levels of the street, proximity to squares, parks, shopping centers or subway stations. However, in a smaller scale, Valente *et al.* [39] and Ismail [22] state that the number of variables needed to remove all local variation can quickly grow out of control and modelling the independent variables correctly is not an easy task because housing is a complex good. Other important issue is that districts are delineated in a way in which there is no guarantee that they represent adequate real estate submarkets at the lowest level. However, this restriction does not invalidate the methodology here proposed, as it may be applied for smaller units of information as census tracts.

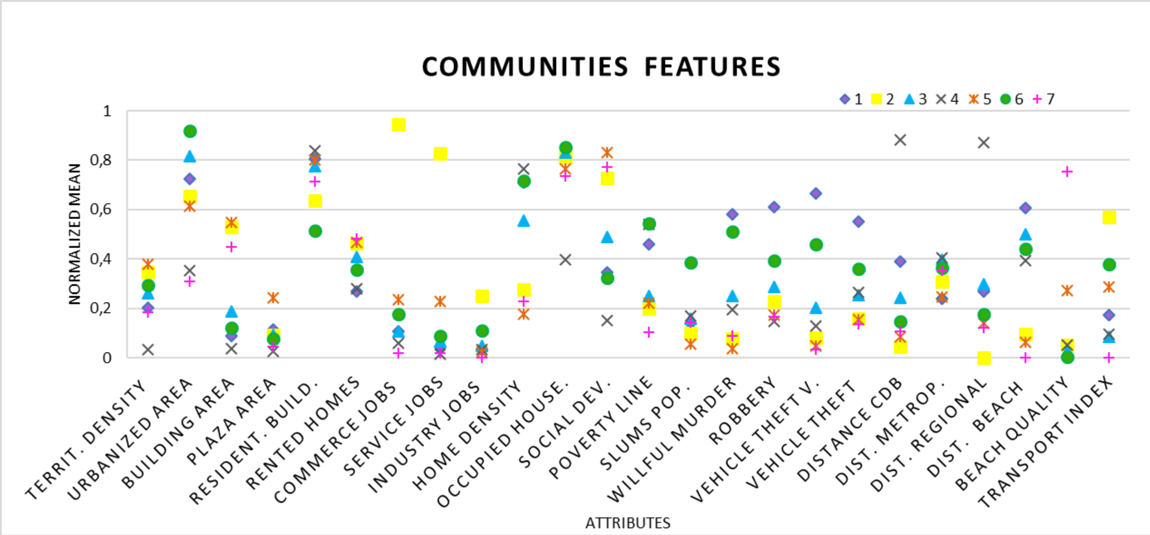

Figure 2. Normalized means of each community features by their attributes

5 Conclusions

As presented in this work, well defined submarkets of districts of a city may be identified by community detection in undirected weighted graphs using locational and socioeconomical attributes. Cosine similarity was used to evaluate weights of the edges between nodes or districts and communities were found by maximizing their modularities. The aim of this paper is to introduce a new approach to real mass appraisal modeling and to show preliminaries results. So far, this approach is suited when there is not enough available data to define submarkets by traditional techniques, allowing a researcher to use the available data for the specific region under study. Also, it promotes a better visualization of how the units of information are related. As future works, as stated by Borst [14], very little has been done for this field of research in the literature in terms of defining the optimal number of segments. So, it is also interesting to explore different metrics of similarity and criteria to set the low threshold of edges values to understand how these parameters behave in the modeling of datasets of diverse cities. This methodology might also be evaluated for smaller units of information by setting submarkets found as dummy variables and explore how it could help to reduce errors on estimating real estate prices using hedonic models.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] Dantas, R.A. *Engenharia de avaliações - uma introdução à metodologia científica*. 3ª edição. São Paulo: Pini, 2011.
[2] Associação Brasileira de Normas Técnicas. NBR 14.653: avaliação de bens. Rio de Janeiro. 2019.
[3] Dimopoulos, Thomas; Bakas. *An artificial intelligence algorithm analyzing 30 years of research in mass appraisals*. RELAND: International Journal of Real Estate & Land Planning, [S.l.], v. 2, p. 10-27, jan. 2019. ISSN 2623-4807.
[4] Byeonghwa Park; Jae Kwon Bae. *Using machine learning algorithms for housing price prediction: The case of Fairfax County*, Virginia housing data, Expert Systems with Applications, Volume 42, Issue 6, 2015, Pages 2928-2934, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2014.11.040.
[5] Wang, Daikun; Li, Victor Jing. *Mass Appraisal Models of Real Estate in the 21st Century: A Systematic Literature Review*. Sustainability, v. 11, n. 24, p. 7006, 2019.
[6] Rosen, S. (1974). "Hedonic prices and implicit markets: product differentiation in pure competition". Journal of Political Economy. 82 (1): 34–55. CiteSeerX 10.1.1.517.5639. doi:10.1086/260169. JSTOR 1830899.
[7] Bourassa, Steven C.; Cantoni, Eva; Hoesli, Martin. *Spatial dependence, housing submarkets, and house prices*. Housing Submarkets, and House Prices (June 2005), 2005.
[8] Hasan Selim, *Determinants of house prices in Turkey: Hedonic regression versus artificial neural network*, Expert Systems with Applications, Volume 36, Issue 2, Part 2, 2009, Pages 2843-2852, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2008.01.044.

[9] Vilius Kontrimas; Antanas Verikas. *The mass appraisal of the real estate by computational intelligence*, *Applied Soft Computing*, Volume 11, Issue 1, 2011, Pages 443-448, ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2009.12.003.
[10] Kauko T., (2003), *"On current neural network applications involving spatial modelling of property prices"*, Journal of Housing and the Built Environment 18, pp. 159-181.
[11] McCluskey, W.J.; McCord, M.; Davis, P.T.; Haran, M.; McIlhatton, D. *Prediction accuracy in mass appraisal: A comparison of modern approaches*. J. Prop. Res. 2013, 30, 239–265
[12] Zurada, Jozef & Levitan, Alan & Guan, Jian. (2011). *A Comparison of Regression and Artificial Intelligence Methods in a Mass Appraisal Context*. Journal of Real Estate Research. 33.
[13] Rodrigues, Sérgio da Silva. *Análise da precificação de imóveis na cidade do Rio de Janeiro utilizando Modelagem Hedônica e os efeitos da autocorrelação espacial*. Dissertação (Mestrado em Matemática Aplicada) - Escola de Matemática Aplicada, Fundação Getúlio Vargas - FGV, Rio de Janeiro, 2015.
[14] Miller, Dennis, 1974, *Review by Dennis W. Miller of Statistical Stratification of Residential Properties: Theory and Methodology*, The Application of Multiple Regression Analysis in Assessment Administration. Proceedings of a Symposium Conducted by the IAAO Research and Technical Services Department in Cooperation with John C. Lincoln Institute,130.
[15] Schnare, A. B., and R. J. Struyk. 1976. *"Segmentation in urban housing markets."* J. Urban Econ. 3 (2): 146–166.
[16] Borst. R. A. *Discovering and Applying Location Influence Patterns in the Mass Valuation of Domestic Real Property*. Doctor of Technology Thesis. Faculty of Engineering of the University of Ulster, 2007.
[17] Zhuo Chen, Z., S. H. Cho, N. Poudyal, and R. K. Roberts. 2009. *"Forecasting housing prices under different market segmentation assumptions."* Urban Stud. 46 (1): 167–187. https://doi.org/10.1177 /0042098008098641.
[18] Pace, R.K., Barry, R. and Sirmans, C.F., (1998)b. *Spatial Stat. and Real Estate*. J. of Real Estate Fin. and Eco. 17: 5-13.
[19] Dubin, R.A. *Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms*. Review of Economics and Statistics, 1988, 70:3, 466–74.
[20] Bourassa, Steven C.; Cantoni, Eva; Hoesli, Martin, 2010. "Predicting House Prices with Spatial Dependence: A Comparison of Alternative Methods" Journal of Real Estate Research, American Real Estate Society, vol. 32(2), pag 139.
[21] Wu, C., and R. Sharma. 2012. "Housing submarket classification: The role of spatial contiguity." Appl. Geogr. 32 (2): 746–756. https://doi.org/10 .1016/j.apgeog.2011.08.011.
[22] Ismail, S. (2006). Spatial autocorrelation and real estate studies: A literature review. MY J. of Real Estate, 1(1):1–13.
[23] Clapp, John M., *A Semiparametric Method for Valuing Residential Location* (August 3, 2001). Available at SSRN: https://ssrn.com/abstract=287886 or http://dx.doi.org/10.2139/ssrn.287886
[24] Goodman, A.C. and Thibodeau, T.G. *Housing Market Segmentation and Hedonic Prediction Accuracy*, Journal of Housing Economics, 2003, 12(3), 181–201.
[25] Fik, T.J., D.C. Ling, and G.F. Mulligan. *Modeling Spatial Variation in Housing Prices: A Variable Interaction Approach*. Real Estate Economics, 2003, 31:4, 623–46.
[26] Clapp, John Mantle. *"Defining Neighborhood Boundaries: The Use of Transactions Data."* (2003).
[27] Leishman, C., G. Costello, S. Rowley, and C. Watkins. 2013. *"The predictive performance of multilevel models of housing sub-markets: A comparative analysis."* Urban Stud. 50 (6): 1201–1220. https://doi.org/10 .1177/0042098012466603.
[28] Gabrielli, L.; Giuffrida, S.; Trovato, M.R. *Gaps and Overlaps of Urban Housing Sub-market: Hard Clustering and Fuzzy Clustering Approaches*. In Appraisal: FromTheoryto Practice; Stanghellini, S., Morano, P., Bottero, M., Oppio, A., Eds.; Springer: Berlin, Germany, 2017; pp. 203–219.
[29] Barabási, Albert-László. *Network science*. Cambridge University Press, 2016.
[30] Rebelo, Emília M. D. D. A. *Mercado Imobiliário e Transformações Urbanas*. (2002). 1344f. Tese (Doutorado em Engenharia Civil), Faculdade de Engenharia da Universidade do Porto.
[31] Grigsby, W., Baratz, M., Galster, G., and Maclennan, D. (1987). *"The Dynamics of Neighborhood Change and Decline*," Progress in Planning 28(1), 1-76.
[32] Newman M E J, 2004 Phys. Rev. E 70 056131
[33] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, *Fast unfolding of communities in large networks*, in Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P1000.
[34] Área da unidade territorial: Área territorial brasileira. Rio de Janeiro: IBGE, 2020.
[35] População estimada: IBGE, Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais, Estimativas da população residente com data de referência 1o de julho de 2019.
[36] Long, F., Páez, A., e Farber, S. (2007). *Spatial effects in hedonic price estimation: A case study in the city of Toronto.*
[37] Prefeitura do Rio de Janeiro, Secretaria Municipal de Urbanismo. Proposta de Metodologia para Hierarquização dos Centros e Subcentros da Cidade do Rio de Janeiro. Rio de Janeiro. 2014.
[38] Instituto de Estudos do Trabalho e Sociedade. IETS. Centralidades da Região Metropolitana do Rio de Janeiro. 2016.
[39] Valente, J., Wu, S., Gelfand, A.E. and Sirmans, C.F., (2005). *Apartment Rent Prediction Using Spatial Modelling*. Journal of Real Estate Research, 27(1): forthcoming.