ISSN 2029-6991 / eISSN 2029-7009

VILNIUS TECH
Vilnius Gediminas Technical University

GEODESY and CARTOGRAPHY

2025
Volume 51
Issue 3
Pages 141–145
https://doi.org/10.3846/gac.2025.21261

UDC 528.44

CLUSTER ANALYSIS OF FACTORS INFLUENCING THE VALUATION OF REAL ESTATE OBJECTS

Serhii NESTERENKO[1], Yuliia RADZINSKA[1], Maryna PILICHEVA[1], Alena PALAMAR[2]✉

[1]*Department of Land Administration and Geoinformation Systems, O. M. Beketov National University of Urban Economy in Kharkiv, Marshala Bazhanova Str. 17, 61002 Kharkiv, Ukraine*
[2]*Department of Geodesy, Kryvyi Rih National University, Vitalyi Matusevich Str. 11, 50027 Kryvyi Rih, Ukraine*

Article History:
- received 03 April 2024
- accepted 28 May 2025

Abstract. The current real estate market analysis reveals challenges in valuation methods, procedure adequacy, and evolving technological approaches. Uncertainty arises from using localised methods for valuing individual real estate objects. A significant concern is the reliability and completeness of valuation data. Researchers emphasise market-driven aspects as trends in real estate valuation. Features for valuation are identified through quantitative characteristics, uncovering components and their nature. The research analyses foreign and domestic practices for real estate object valuation. Challenges include understanding methodological and informational support through mathematical methods and addressing factors affecting real estate object valuation. The need for cluster analysis to identify factors affecting real estate object valuation is recognised. To implement cluster analysis of factors affecting real estate valuation, a method is proposed involving the development of classification features, optimal typological grouping, and clustering implementation technology. Six groups of factors were chosen: spatial formation, urban planning provision, environmental impact, investment indicators, infrastructure provision, and limiting characteristics. An agglomerative process calculated the distance matrix between clusters of factors. The MacQueen *k*-means clustering method determined final clusters, confirming the validity of the proposed factor groups. The clustering of factors affecting real estate valuation was based on obtained distance data. The result identifies a high level of factors influencing real estate object valuation. Nine coincidences justify this in their clustering with four units of factors influencing real estate object valuation.

Keywords: real estate, cluster analysis, valuation, spatial formation, urban planning, indicators.

✉Corresponding author. E-mail: ***pikilnyak@gmail.com***

1. Introduction

At the current stage of real estate market development, there are many issues related to valuation methods (Petrakovska & Lizunova, 2009; Petrakovska, 2010; Drapikovskyi et al., 2015; Smyth et al., 1993), adequacy of the chosen valuation procedure, and changes in technological approaches to real estate valuation. Uncertainty problems persist with local methods of valuing individual real estate objects (Mamonov et al., 2020).

As pointed out by domestic researchers, significant challenges include the reliability and completeness of valuation data (Petrakovska & Lizunova, 2009; Petrakovska, 2010; Drapikovskyi et al., 2015; Hubar et al., 2017; Hubar, 2012; Mingming et al., 2019; Pieri et al., 1996; Nesterenko et al., 2019).

Some researchers highlight the influence of market trends on real estate valuation (Mingming et al., 2019; Pieri et al., 1996; Nesterenko et al., 2020; Molodchenko, 2010). It is known that identifying peculiarities in real estate valuation is performed based on quantitative indicators and the identification of components and their nature.

However, considering the views of some practitioners, there is a problem in identifying and gaining a deeper understanding of the methodological and informational support for valuation work through the application of a mathematical approach and addressing the analysis of factors affecting the valuation of real estate objects (Hubar et al., 2017; Hubar, 2015; Smyth et al., 1993). Methodological regulation of real estate valuation is equally essential. As stated in the study (Hubar, 2012), methodological principles for real estate valuation consider the requirements imposed on valuation entities involved in the valuation process of real estate objects of various ownership forms and emphasise compliance with national standards and principles of such valuation. An essential element to be

considered in real estate valuation is the requirements of methodological and regulatory documents that directly influence the valuation process and consider a range of subjective factors related to a specific real estate object. These factors may include spatial, urban planning, environmental characteristics, investment indicators, indicators of infrastructure provision for real estate objects, and the consideration of restrictive characteristics of real estate use.

Some indicators used in the practice of real estate valuation are obtained through the accumulation of statistical data obtained from relevant urban planning authorities, land resources agencies, environmental services, pre-assessment or valuation of similar real estate objects, utilisation of infrastructure indicators, and indicators of specialised enterprises responsible for the maintenance of real estate objects.

In cases where reliable data is unavailable for the valuation process, expert analysis methods, which rely on experts' opinions to assess the real estate object, are used. In this regard, using the latter method requires the selection of specialised experts who are specialists in data selection. Data sampling is necessary to ensure the results' reliability, and experts' opinions must be harmonised.

As noted in the study (Meteshkin et al., 2020; Mamonov et al., 2020), various factors can have a decisive impact on the valuation of real estate objects. Therefore, there is a need to conduct a cluster analysis of these factors and identify sets of factors affecting the valuation of the real estate object (Hubar, 2015; Nesterenko & Vyatkin, 2017).

For the valuation of real estate objects, the number of factors for analysis can reach several tens, hundreds, or even more. Moreover, the number of features characterising a specific real estate object can also reach dozens. Hence, visual (random) analysis of the data matrix with a significant number of indicators and features of the real estate object is practically inefficient. Thus, there is a need for aggregation, concentration of input factors, and analysis of the structural characteristics of real estate objects. Addressing these tasks can be achieved through modern multifactor classification methods.

Tools and methods of multifactor classification enable grouping the factors that influence the valuation of real estate objects, taking into account significant structural and typological features and the distribution pattern of factors within a given system of classification features. This type of classification aims to group the factors affecting the real estate valuation based on their similarities, ensuring that each group is as distinct as possible in terms of content.

Thus, multifactor classification methods for the factors affecting the valuation of real estate objects are used to allocate the total factors into homogeneous groups. It should also be noted that various random related attributes characterise each factor affecting the real estate valuation.

2. Material and methods

The clustering method (Figure 1), which consists of four stages to implement cluster analysis of the factors affecting the valuation of real estate objects, was proposed.

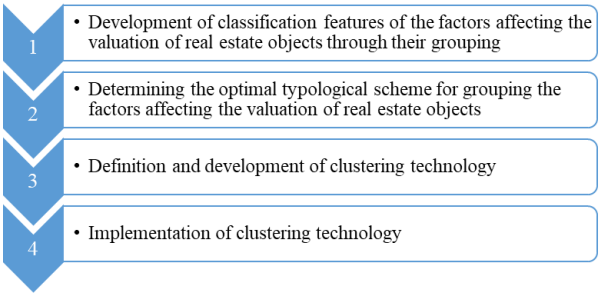

Figure 1. Clustering method for factors affecting the real estate valuation

In the first stage, classification features of factors affecting the valuation of real estate objects were determined by grouping them according to six characteristics (Meteshkin et al., 2020):

- Spatial formation and use of real estate objects;
- Urban planning provisions for the use of real estate objects;
- Environmental impact from the use of real estate objects;
- Investment indicators determining the attractiveness of the real estate object;
- Infrastructure provision for the use of the real estate object;
- Restrictive characteristics ensuring the safety of the real estate object's use.

Each group of factors affecting the valuation of real estate objects was assigned a separate cluster (p). The number of clusters (n) in each group was determined by its constituent elements. To merge factors into one cluster, similarity and the closest proximity of indicators of its individual components were determined, reducing the number of clusters to one.

In the next stage of implementing the clustering method, typological scenarios of grouping factors affecting the valuation of real estate objects were determined. The agglomerative process of calculating the Δ matrix was applied as the selected grouping method, with distances between data clusters for affecting factors being calculated (Figure 2).

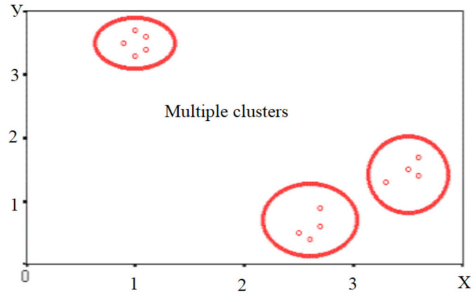

Figure 2. Scatter plot of clusters X and Y, relative units

After calculating the distances between influence clusters in real estate valuation, the centre of one of the clusters $x_l \in X$ was selected, and a hypothetical cluster S was formed with factors at lower levels and a greater distance from it than its radius $\tau > 0$. The distances to individual affecting factors on real estate valuation within the cluster were constrained by models:

$$\text{if } d_{ix_1} \le \tau, \text{ then } x_i \in S; \quad (1)$$

$$\text{if } d_{ix_1} > \tau, \text{ then } x_i \notin S, \quad (2)$$

where d_{ix_1} – is the distance from the factor x_i affecting the valuation of real estate objects to the cluster's centre x_1.

In the third stage, the identification and development of a technology for clustering factors affecting the valuation of real estate objects were performed. The MacQueen *k*-means clustering method was applied as the clustering method, which allowed determining the number of final clusters and confirming the reliability of the proposed groups of analysis factors (Figure 3).

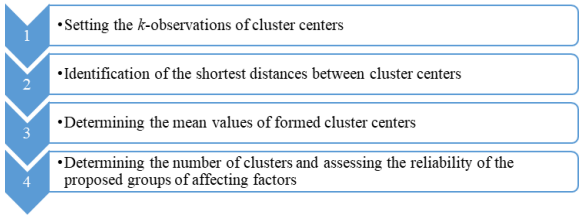

Figure 3. The MacQueen *k*-means method

To determine the *k* observations of cluster centres, the elements that form the factors affecting the valuation of real estate objects for each group were initially identified and assigned typological characteristics, subsequently allowing them to be identified as separate clusters (Hubar, 2015).

A necessary condition for conducting cluster analysis is determining the intervals of factors affecting real estate valuation by specifying the ranges of cluster influence (Table 1).

Table 1. Ranges of intervals for cluster influence factors on real estate valuation in relative units

Ranges of intervals	Cluster impact levels
0.01–2	Very low
2.01–4	Low
4.01–6	Moderate
6.01–8	High
8.01–10	Influential

During cluster analysis using the *k*-means method by MacQueen, the cluster centres of factors affecting real estate valuation were determined. These calculations allowed for the elimination of distorted results of cluster analysis and changes in the accuracy of cluster centres. According to the specified approach, the set of factors affecting real estate valuation (X) was represented by indicators of their performance in a cluster (n), and they corresponded to the following characteristics (m):

$$X = \{X_1, X_2, ..., X_m\}. \quad (3)$$

According to the *k*-means method of MacQueen, the set of factors affecting the valuation of real estate objects (X) was described by a set of observation vectors of cluster analysis (Equation (4)).

$$X_j, \quad j = \overline{1,m}. \quad (4)$$

The cluster analysis sample included 5 ranges of influence based on typological features in accordance with the functioning of factors affecting the valuation of real estate objects. As a result, 5 clusters were obtained (k = 5).

The matrix of cluster analysis of factors affecting the valuation of real estate objects was described by the Equation (5):

$$X = [X_1, X_2, ..., X_m] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{im} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nm} \end{bmatrix}, \quad (5)$$

where x_{ij} – is the j-th parameter of the i-th object; n – performance indicators in the cluster of factors affecting the valuation of real estate objects.

3. Results

Based on the constructed matrix of cluster analysis of factors affecting the real estate valuation using the *k*-means method by MacQueen, assigned reference parameter values were determined for each primary cluster. These parameters were specified as the average values for each interval of factor impact on real estate valuation. Each reference value corresponds to the centre of a cluster. Grouping of the closest objects based on typological features of the affecting factors was performed around the reference values. The results of the obtained clustering parameters for the factors affecting the real estate valuation are presented in Table 2.

Table 2. Results of clustering parameters for the factors affecting the real estate valuation

Clusters of influence factors	Initial parameters	Average range values	Range weight	Average high cluster values	Average maximum cluster values
Very low	1.999	0.999	1	1.001	1.501
Low	3.999	2.999	2	3.001	3.501
Moderate	5.999	4.999	3	5.001	5.501
High	7.999	6.999	4	7.001	7.501
Significant	10	8.999	5	9.001	9.501

The clustering of factors affecting the real estate valuation was performed based on the obtained data of distances between centroids and clustering objects using matrix calculation data (5). In this process, reference values were added to the clustering results of factors affecting the real estate valuation. Taking into account the reference parameters and determined clustering centroids, the matrix for calculating the factors affecting the valuation of real estate objects is represented as Equation (6):

$$X = \left[X_1, X_2, ..., X_m\right] = \begin{bmatrix} x_{11} & x_{12} & ... & x_{1j} & ... & x_{1m} \\ x_{21} & x_{22} & ... & x_{2j} & ... & x_{2m} \\ ... & ... & ... & ... & ... & ... \\ x_{i1} & x_{i2} & ... & x_{ij} & ... & x_{im} \\ ... & ... & ... & ... & ... & ... \\ x_{n1} & x_{n2} & ... & x_{nj} & ... & x_{nm} \\ ... & ... & ... & ... & ... & ... \\ x_{n+5,1} & x_{n+5,2} & ... & x_{n+5,j} & ... & x_{n+5,m} \end{bmatrix}. \quad (6)$$

The distances between individual clusters, affected by a real estate factor, were determined based on the Euclidean metric according to Equation (7):

$$d_{ij} = \frac{1}{m}\sqrt{\sum_{k=1}^{m}\left(z_{ik} - z_{jk}\right)^2}, \quad (7)$$

where d_{ij} – is the distance between the i-th and j-th clusters of factors affecting the assessment of real estate objects; m – is the number of factor operating ranges; z_{ik} – is the value of the i-th cluster for the k-th typological feature; z_{jk} – is the value of the j-th cluster for the k-th typological feature.

Clustering was performed until a complete distribution of potential variations in the factor affecting the assessment of real estate objects was obtained for each group within a cluster.

The results of the cluster analysis of factors affecting the valuation of real estate objects are presented in Table 3.

Table 3. Results of cluster analysis of factors affecting the valuation of real estate objects

	Projected clustering parameters for affecting factor					
	Very low	Low	Moderate	High	Influential	Total
Actual data	0	0	1	4	2	2
Forecasted data	1	2	2	5	6	5
Number of matches	1	4	6	9	1	4

4. Discussion

The obtained clusters have been classified based on actual and forecast data. In this context, actual data is represented by clustering six typological types of factors into five groups. In contrast, forecast data is determined by calculating benchmark elements of clusters using the MacQueen k-means clustering method, considering the distance of affecting factors from the cluster centres.

As a result of the conducted cluster analysis, a high level of affecting factors on the valuation of real estate objects has been identified. Nine coincidences justify this in their clustering with four units of affecting factors on the valuation of real estate objects.

5. Conclusions

As a result of the research, an analysis of existing foreign and domestic practices for the valuation of real estate objects was conducted. A problem was identified in determining and thoroughly understanding the methodological and informational support of valuation work by applying mathematical methods and addressing issues related to analysing factors affecting the valuation of real estate objects. The necessity of conducting their cluster analysis and identifying sets of affecting factors on the valuation of real estate objects has been recognised.

The clustering method to implement the cluster analysis of factors affecting the valuation of real estate objects has been proposed. This method involves the development of classification features of affecting factors on the valuation of real estate objects by grouping them, determining the optimal typological grouping scheme of affecting factors and defining and developing the clustering technology and its implementation.

For the determination of classification features, six groups of factors were chosen: spatial formation and utilisation of real estate objects, urban planning provision for the use of real estate objects, environmental impact from the use of real estate objects, investment indicators determining the attractiveness of real estate objects, infrastructure provision for the use of real estate objects, and limiting characteristics for ensuring the safety of the use of real estate objects. An agglomerative process for calculating the distance matrix between clusters of data factors was applied as the chosen grouping method.

The MacQueen k-means clustering method was applied as the clustering method, allowing for the determination of the number of final clusters and confirming the validity of the proposed analysis factor groups. The clustering of affecting factors on the real estate appraisal was performed based on the obtained distance data between the centres and objects of clustering using the calculated matrix data.

As a result of the conducted cluster analysis, a high level of affecting factors on the valuation of real estate objects has been identified. Nine coincidences justify this in their clustering with four units of affecting factors on the valuation of real estate objects.

References

Drapikovskyi, O. I., Ivanova, I. B., & Krumelis, Yu. V. (2015). *Otsinka nerukhomosti: navchalnyy posibnyk* [Real estate appraisal: Educational manual]. TOV "Sik Hrup Ukraina".

Hubar, Yu. (2012). Rozrobka pidkhodiv i metodiv kadastrovoyi otsinky nerukhomosti naselenykh punktiv [Development of approaches and methods for cadastral valuation of real estate in settlements]. *Modern Achievements of Geodesic Science and Industry*, *2*(24), 146–150.

Hubar, Yu. (2015). Clustering implementation of the methodology for the construction of economic and mathematical regression model of valuation real estate. *Geodesy, Cartography and Aerial Photography*, *82*, 110–135. https://doi.org/10.23939/istcgcap2015.02.110

Hubar, Yu., Hubar, A., & Tkachenko, S. (2017, December 14–17). The estimation of the typical level of transactions prices of real estate objects by research methods (on the example of the town of Gorodok of the Lviv region). In *GeoTerrace-2017* (pp. 197–200), Lviv.

Mamonov, K., Nesterenko, S., Frolov, V., & Troyan, V. (2020). The research of the impact of environmental factors on the use of underground real estate. In *International Conference of Young Professionals "GeoTerrace-2020"* (pp. 1–5). European Association of Geoscientists & Engineers. https://doi.org/10.3997/2214-4609.20205739

Meteshkin, K., Shipulin V., Nesterenko S., & Kobzan, S. (2020). Features of underground real estate use: Infrastructural and regional aspects. *Geodesy and Cartography*, *69*(1), 53–64. https://doi.org/10.24425/gac.2020.131081

Mingming, W., Nesterenko, S., & Shterndok, E. (2019). Modelling of the parcel pattern impact. *IOP Conference Series: Earth and Environmental Science*, *272*(3), Article 032085. https://doi.org/10.1088/1755-1315/272/3/032085

Molodchenko, T. H. (2010). Analiz dzherel formuvannya dokhodnoyi chastyny mistsevykh byudzhetiv u rozvynenykh krayinakh [Analysis of the sources of formation of the revenue part of local budgets in developed countries]. *Communal Economy of Cities*, *94*, 330–334.

Nesterenko, S. V., Shchepak, V. V., Kariuk, A. M., & Mishchenko, R. A. (2019). System of designing livestock small-volumetric cooperative buildings. *IOP Conference Series: Materials Science and Engineering*, *708*, Article 012015. https://doi.org/10.1088/1757-899X/708/1/012015

Nesterenko, S., & Vyatkin, K. (2017). The study of land management and geographic information support of municipal building in Ukraine. *Technology Audit and Production Reserves*, *1*(4(33), 24–28. https://doi.org/10.15587/2312-8372.2017.94381

Nesterenko, S., Mishchenko, R., Shchepak, V., & Shariy, G. (2020). Public cadastral maps as a basis for a construction of the building general layout. In V. Onyshchenko, G. Mammadova, S. Sivitska, & A. Gasimov (Eds.), *Lecture notes in civil engineering: Vol. 73. Proceedings of the 2nd international conference on building innovations* (pp. 161–170). Springer. https://doi.org/10.1007/978-3-030-42939-3_18

Petrakovska, O., & Lizunova, A. (2009, February 11–13). *Urban planning in private property conditions in Ukraine* [Paper presentation]. International Academic Group on Planning, Law and Property Rights, Third Conference, Aalborg, Denmark.

Petrakovska, O. (2010, April 11–16). *Restrictions of urban land use in Ukraine* [Paper presentation]. FIG Congress 2010, Sydney, Australia. https://www.fig.net/resources/proceedings/fig_proceedings/fig2010/papers/ts06e/ts06e_petrakovska_4491.pdf

Pieri, C., Dumanski, J., Hamblin, A., & Young, A. (1996). *Land quality indicators* (World Bank discussion papers WDP 315). The World Bank. https://doi.org/10.1596/0-8213-3511-1

Smyth, A. J., Dumanski, J., & Spendjian, G. (1993). *FESLM: An international framework for evaluating sustainable land management: A discussion paper.* Land and Water Development Division, Food and Agriculture Organization of the United Nations.