# ANALYTICAL METHOD FOR CORRECTION COEFFICIENT DETERMINATION FOR APPLYING COMPARATIVE METHOD FOR REAL ESTATE VALUATION

**Valentas Gružauskas**
*Kaunas University of Technology, School of Economics and Business, Digitalization Scientific Group*
*E-mail: valentas.gruzauskas@ktu.lt*

**Andrius Kriščiūnas**
*Kaunas University of Technology, Faculty of Informatics, Department of Applied Informatics*
*E-mail: andrius.krisciunas@ktu.lt*

**Dalia Čalnerytė**
*Kaunas University of Technology, Faculty of Informatics, Department of Applied Informatics*
*E-mail: dalia.calneryte@ktu.lt*

**Valentinas Navickas**
*Kaunas University of Technology, School of Economics and Business, Academic Center of Economics, Business and Management*
*E-mail: valentinas.navickas@ktu.lt*

## Abstract

Real estate valuation uses 3 main approaches: income, cost and comparative. When applying the comparative method, correction coefficients based on similar real estate transactions are determined. In practice, the coefficients and similar real estate objects are usually determined by using qualitative approach based on the valuators' experience. The paper provides an analytical method for the determination of correction coefficient, which limits subjectivity when using the comparative method for valuation. The provided analytical approach also integrates macroeconomic indicators in the calculation process. It also addresses issues when available historical real estate transaction data is limited. A machine learning approach was applied to determine the average price of real estate in the region, with the possibility of using this information to obtain correction coefficients where historical data was unavailable. Alternative research usually focuses on final price estimation of the selected real estate object; however, the valuation standard of Tegova released in 2018 does not allow for applying analytically based approaches for individual real estate object evaluation; these approaches can be used only as a supportive tool for valuators.

## 1. Introduction

The valuation process of real estate is essential for the economy, because it affects the loan and insurance markets, tax system, accounting and infrastructure development projects. For example, Arestis et al. (2017) have indicated that overvaluation in house prices can lead to financial "bubbles" (Arestis, et al., 2017). The valuation methodology of real estate is documented in the European and world standards; however, they provide only general terms of valuation and not the precise process (TVST, 2017), (TEGOVA, 2016). The government usually requires providing an explanation of how the value of real estate was determined. In order to gain an objective value, it is important for the valuators to understand the real estate market, formulate the goal of the valuation, conduct market research and select the correct valuation method (Čiburienė & Jegelavičiūtė, 2017). For the valuation of real estate, the following three main approaches are applied: cost, income and comparative. Cost and income approaches are based on information provided by the owner or bank. When applying the income approach, income capitalization or the discounted cash flow approach is usually used, although there are alternative approaches such as Monte Carlo simulation (Hoesli, Jani & Bender, 2006). In the work of valuators, the comparative approach is usually used (Jegelavičiūtė & Rimkevičiūtė, 2017). The essence of the comparative approach is to compare the valuated real estate with analogous or similar real estate objects the prices of which are known to the valuator. During the valuation process, the valuer determines correction coefficients (if necessary) when comparing the selected real estate with analogous or similar real estate properties based on selected determinants (Jegelavičiūtė & Rimkevičiūtė, 2017). The government requires reasoning to be provided for the valuation process; therefore, individual expert opinion is not reliable and can be undermined. Thus, there are two approaches when using the comparative method – qualitative and quantitative. Qualitative analysis methodology is related to expert valuation when the authors aim to survey real estate valuers, learn their opinion regarding the significance of the criteria and single out the most important criteria as well as establish the correction coefficients on that basis. Quantitative analysis, on the other hand, uses the actual data of real estate transactions and evaluates the correction coefficients with the use of mathematical models.

Scientists encounter a problem, as in the general case the expert valuation is done by performing the corrections of the set prices and introducing certain correction coefficients. In these cases, the coefficients are partially established following mass valuation done on the basis of the data of the Centre of Registers (e.g., the selection of similar objects employing the information of value maps). Thus, the change of price between mass and expert valuation practically comes down to the expert experience of the property valuator, as the substantiation of the established change has not been regulated and, in general, can depend on the very valuator of the property. Whereas while carrying out the expert property valuation and encountering specific problems, the decision is usually made by every valuator individually. Thus it either depends on his/her expert experience or is guided by the valuation experience of a separate property or business units and the made analysis. On this basis the valuation coefficients, indicators and other relative comparative values enabling the assessment of the property possessing similar characteristics are established. The calculations of the value of the same object done by different valuators can differ significantly. As an example, this problem is especially relevant when there is lack of data on the valuated object, when property valuation is being carried out on an object located in a region where the sales of similar real estate properties take place very rarely or had not taken place at all for a long time. In this case, the selection of similar real estate objects could have a large impact on the final value of the selected valuation object. Similarly, large value errors may occur due to incorrect transaction information in the system, for example, if the value of the real estate object is indicated incorrectly in the transaction document. If such a real estate transaction was included in the valuation process, it could lead to large errors between the real and determined real estate value. The future and current value of the real estate property depends not only on historical transaction data, but also on economic, infrastructure and other information from the region. In this sense, the traditional comparative approach is not very accurate, and the end value of the real estate object depends on the competence of the valuer (Navickas, Šaudys & Jegelavičiūtė, 2017).

Thus, the goal of this publication is to provide an analytical method to determine correction coefficients when applying the comparative method. The novelty of the research is a provided analytical method of how to select similar real estate transactions from historical data. Kokit and Gnat

(2019) have analyzed the effectiveness of multiple regression model application for real estate valuation (Kokot & Gnat, 2019). They have concluded that multiple regression models are only effective in well-functioning real estate markets and when there is a sufficient amount of market attributes and numerous sets of transactions. In other cases, multiple regression models are characterized by a high bias, and thus our publication proposes a machine learning method, which could provide more accurate results even when there is a limited amount of data available. In cases, when available historical data is limited, a machine learning approach is also provided to determine the average price of a real estate property depending on its purpose (e.g. residential). In practice, valuators usually use expert opinion, which is a subjective approach and cannot provide logical grounding to the calculations process, which is essential during trials. In scientific research publications, there are methods provided for value determination, though the research usually misses two aspects. On one hand, the proposed mathematical models are usually based only on historical real estate data and do not integrate social or economic regional indicators. On the other hand, the final value for the valuers does not improve their expertise and because of the current legislation, it is recommended not to use AVM as a main tool for final price determination. The application of artificial intelligence algorithms to provide the final value of a real estate property is considered a mass valuation process. In 2018, Tegova restricted mass valuation model application for the determination of final value. AVMs cannot be used to produce a valuation report that complies with European values association independently of a valuation process based, inter alia, on inspection of the property by the valuer and the application of valuation judgment by him or her. When used, AVM is never more than a tool contributing to the valuer's estimation of value, for which he remains responsible (TEGoVA, 2018). To accomplish the goal, the following objectives must be completed:

1. The approaches of real estate valuation will be overviewed;
2. The methodological approach for similar real estate transaction determination will be provided;
3. A case example of the analytical method for the comparative approach will be conducted,
4. The machine learning approach will be applied to determine average real estate prices in a region where available data is limited.

## 2. Literature review of real estate valuation

There are mainly two approaches to real estate valuation. One approach is used for mass valuation, while the other approach is used for individual valuation. An automated valuation model, which is usually used for mass valuation, can be applied. (International Association of Assessing Officers, 2011). Mass valuation is needed to determine the tax amount, provide loans or for other aims; this method is usually based on regional price estimation. Depending on the goal, a more individual real estate valuation is also being applied in the market. The approach of individual real estate valuation can be implemented by using mainly three methods, i.e., income, comparative and cost. Guidice et al. (2017) have proposed a fuzzy logic algorithm for real estate investment estimation, which is based more on the income method (Del Giudice, De Paola & Cantisani, 2017). Zujo et al. (2014) provided a real estate valuation approach based on the cost method (Zujo, Car-Pusic & Zileska-Pancovska, 2014). Dziadosz and Meszek (2015) have analysed the deterioration of real estate, which is important to the real estate valuation process (Dziadosz & Meszek, 2015). D'Acci (2019) have analyzed the relationship between real estate value and urban area, the results of which could be used for the comparative method (D'Acci, 2019). Hromada (2016) have provided a price estimation model based on historical transaction data; however, it did not integrate macroeconomic indicators in the process (Hromada, 2016). In the case of this publication, a focus on the comparative method is provided. When valuators apply the comparative approach, they determine correction coefficients by selecting similar objects and determining criteria of adjustment. The most general criteria consist of the time, location and condition of a real estate object. To determine the actual price of real estate, the valuators must understand the market and have data on it.

The valuation process of real estate requires following legislation and standard requirements issued by the worldwide valuation standards, Association for European Values and government authorities. Some countries face problems when valuating real estate because, for the substantiation of the valuation process access to data sources is needed. In some cases, the data available in the market is limited. However, in other cases, valuers tend to use their own experience rather than data available on the market. In spite of legal regulations requiring to provide the data on both the calculation process and the data used for the determination of correction coefficients, the evaluators

do not always present them in the valuation report. (Tumelionis, 2013). In cases of court trials, this subjectivity can be difficult to solve. Thus, not only an integrated data platform, but also support tools for valuators, which would limit the subjectivity in the valuation process, are needed.

Real estate valuation requires an understanding of the market and the access to various data sources. The important aspects when determining real estate value can be seen from demand and economic cycle perspective. The factors influencing real estate demand can be listed as follows:

1. Number of residents;
2. Amount of income;
3. Opportunities of getting employed;
4. Level of unemployment;
5. Ratio of landlords and renters;
6. Number of unrented premises;
7. Level of savings;
8. Level of getting loans;
9. Sales turnovers;
10. Directions of growing of expansion of cities and infrastructure;
11. Road infrastructure;
12. Topography and landscape;
13. The burden of taxation;
14. Cultural and educational objects;
15. Safety infrastructure (fire services, emergency medical service and police).

It is also important to monitor the economic situation in the region, because the purchasing power of companies and residents may change depending on the economic situation. This information is especially important when available historical transaction data from the region is limited. To identify the economic cycle, it should be needed to consider these factors:

1. The quantity and quality of the competing objects;
2. The scale of new construction;
3. The prices of undeveloped land and its supply;
4. The costs of construction and development of territories;
5. The number of sold objects;
6. The number of landlords and renters;
7. Not let floor-space and land area
8. The opportunities to employ the real estate for other purpose; The opportunities for financing construction;
9. The regulation of construction and other laws.

The access to the indicators listed above is important for the valuator in order to determine the correct value. There is a growing number of studies which apply artificial intelligence for real estate price evaluation. De Cock (2011) has applied an advanced regression analysis of Boston housing data, which consisted of 80 indicators (De Cock, 2011). Qingyuan et al. (2018) have developed a spatial housing price data estimation model based on China (Du, et al., 2018). Chiarazzo et al. (2014) have applied a neural network and integrated environmental quality and real estate data for price estimation (Chiarazzo, et al., 2014). Li et al. (2018) have applied a deep learning algorithm on real estate data and obtained high accuracy, with error ranging from 2% to 10.35% (Li Yu, et al., 2018). These works usually focus on providing the final price of a real estate property and do not provide a support tool. However, in 2018, Tegova restricted mass valuation model application for final value determination (TEGoVA, 2018).

The problem is that it is practically impossible for a valuator to obtain all this information and analyze it in a reasonable amount of time. Thus, support tools based on artificial intelligence are essential to help the valuator determine the value of the real estate object.

## 3. Data and Methods

In this part the methodological approach to determine the correction coefficient based on time, location and condition will be presented. For setting the value of a real estate object, a chosen number) of objects (e.g., 10), which are the closest to the analyzed object according to the average evaluation of similarity according to all the three criteria groups (time, location, condition), is selected.. The average

corrected price of these objects is employed for setting the value of a specific object. The adjustment of the prices of chosen objects is calculated after the evaluation of properties from the set of historical data that satisfy the criteria of similarity to the analyzed object and the selected objects.

The similarity of the objects according to time is outlined only according to the date of the transaction, though the economic situation, which is defined by the macro indicators, can be included as well. In this example, the similarity of time $\alpha_t$ is entered by a coefficient indicating what period separates the date of the transaction of the analysed object $S_i$ and the date of the transaction of the comparable object $S_j$:

$$\alpha_t = \max\left(1 - \frac{|S_i - S_j|}{10}, 0\right) \tag{1}$$

Here the difference of the dates of the transaction $(S_i - S_j)$ is expressed in a fraction of years. If the dates of the transaction are separated by more than 10 years, the similarity coefficient is equal to zero.

The evaluation of the influence of location can be categorical when location is attributed to a real category (good, bad, middling) or certain quantitative indicators-are attributed to it – for example, the number of shops, schools, kindergartens at a certain distance from the property, the distance to the nearest park, statistics of crimes in a certain radius and other indicators are included. For the automatic analysis of the influence of location, it is necessary to know the exact coordinates of the property, so that, according to them, it is possible to evaluate the quantitative indicators. If the coordinates of the object or a small territorial unit are not singled out, all the objects present in the municipality have the same location parameters and the similarity of the objects is equal to 1. In other cases, the evaluated similarity between the municipalities i and j is entered by digital estimator $\beta_{ij}$, which is obtained by clustering the municipalities with the macro indicators and obtaining the difference coefficients (for example, by using the Rand index). The estimator $\beta_{ij}$ is transformed to interval [0;1], where 0 means that the locations of the objects are the same and 1 means that they are completely different. A similar approach of clusterization of real estate regions was conducted by (Gnat, 2019). The location similarity $\alpha_p$ is calculated according to the formula:

$$\alpha_p = 1 - \beta_{ij} \tag{2}$$

The evaluation of the condition for the subgroup of the residential destination (flats) consists of quantitative indicators (floor area, number of floors and similar) and categories (type of walls, type of heating and similar), so in the evaluation of similarity, the estimators of both types must be reflected as well. At this stage, experts should be included; here they are to define the similarity of categories. For example, types of walls "timber" and "timber with framework" are similar and "reinforced-concrete slabs" differ from the first two immensely. Let's say, $n$ indicators define the condition of the object out of which $n_c$ are categorical and the remaining ones are quantitative. The similarity between the k-category of two objects is expressed in:

$$\delta_k = \begin{cases} 1, & \text{if the category coincides} \\ 0, & \text{if the category does not coincide} \end{cases} \tag{3}$$

The similarity between $I$ and $j$ objects according to k quantitative indicator $s_k$ is evaluated according to formula:

$$s_k = 1 - \frac{|r_k^i - r_k^j|}{r_k^i} \tag{3}$$

here $r_k^i$, $r_k^j$ – are the meanings of the k quantitative indicator of i and j objects accordingly. The similarity according to condition $\alpha_c$ is calculated according to formula:

$$\alpha_c = \frac{\sum_{k=1}^{n_c} \delta_k + \sum_{k=n_c+1}^{n} s_k}{n} \tag{5}$$

The criteria can be entered into the evaluation of condition according to different weights, which ought to be established by experts. If the value of the criterion for one object is very much exceeded in respect of the other object, 0 can be attributed to the value of the indicator, thus indicating that the criteria differ significantly according to this feature. The inter-similarity between objects $\alpha$ is calculated as the average value of similarity estimators defining separate categories:

$$\alpha = \frac{\alpha_t + \alpha_p + \alpha_c}{3} \qquad (6)$$

Evaluation of the location adjustment – the detailed analysis of the influence of location is not carried out due to the lack of accurate information about the location (coordinates). The location coefficient is evaluated only according to the correspondence of the municipalities of different real properties to the selected properties Evaluation of time adjustment – time analysis according to the inter-similarity coefficient for the selected objects, is made according to one feature of the object – the date of the transaction. Evaluation of the condition adjustment – condition analysis according to the inter-similarity coefficient for the selected properties is made according to categorical and quantitative indicators. The adjustment of every component is calculated according to the formula:

$$\omega = \frac{S_P - S_c}{S_C} \cdot 100\% \qquad (7)$$

here $S_P$ –the price of the support facility for a unit of area (EUR/m²), $S_c$ – the price of the comparable property for a unit of area (EUR/m²). The support facility is a property, which is the most similar to an abstract property with a respective component of the analyzed object and two other components of the comparable one. This enables the estimation of how much the component impacts on the value of the property.

For setting the value of real estate objects, the freely accessible information of Macro indicators from the Lithuanian Department of Statistics (Statistics Lithuania, 2019) and the information about the historical transactions of real estate objects accumulated by the Centre of Registers was analyzed (Centre of Registers , 2018). After making the primary data analysis, it was established that, although the information of property address was stored by every real estate transaction, the results were very inaccurate or in frequent cases not presented at all when using free of charge services for establishing coordinates due to the different formats of information. Although the setting of the value of real estate properties by evaluating the exact property location would enable much better and more accurate results to be achieved, in this survey seeking not to expand the work volume the authors have decided to analyse the data at a municipality level; thanks to it both the macro indicators information and the historical information on the sales of real estate objects can be related

For the analysis of real estate properties, a subset of historical data on flats and lots from 2008-2018 was selected. The following macro indicator data of the period were selected for further analysis: the turnover according to the place of carrying out of the activities (nonfinancial enterprises), the number of flats the construction of which was completed, the number of new residential buildings the construction of which was completed, the number of new non-residential buildings the construction of which was completed, wages (monthly), number of employees, the coefficient of vital senility, the emigrants, The birth rate, the density of the population at the beginning of the year, the immigrants, the death-rate, the number of permanent residents at the beginning of the year, the expenses of municipalities budgets, the income to budgets of municipalities, the territory (the area of the land), direct foreign investments at the end of the period and the number of registered fires.

For cases when available historical data is limited, a machine learning algorithm, which provided average price of real estate in the regions, was implemented. This information can be used as a reference when using the comparative method. In practice, valuators usually analyse the environment related to the infrastructure or use their expertise. Thus, the application of machine learning in this case would provide more reasoning to the logical explanation.

## 4. Empirical results

### 4.1. Determination of a correction coefficient for comparative approach

In the case example, the user should first indicate the real estate property and the characteristics that the valuators wish to valuate. The selected real estate property is presented in Table 1.

Based on the historical data and Formulas 1 – 7, similar real estate transaction properties have been obtained. A sample of the analytically obtained objects is provided in Table 2.

Table 1

Parameters of the Analyzed Object

| The analyzed real estate object | | The date of the transaction | 01.01.2019. | | |
|---|---|---|---|---|---|
| Municipality | Alytus | Type of walls | Bricks | Number of rooms | 2 |
| Price in EUR/m² | 492 | Type of heating | Central heating | The floor of the flat | |
| Floor area in m² | 68.48 | Water-supply | - | Number of flats in the house | 23 |
| The date of the end of the construction | 1917 | Sewerage | Communal waste water elimination | Number of floors | 3 |

*Source*: own research.

Table 2

Sample of analytically obtained similar real estate properties

| Comparable 1 | | Date of transaction | 01.08.2018. | Similarity coefficient | 0.92 |
|---|---|---|---|---|---|
| Municipality | Alytus | Type of walls | Bricks | Number of rooms | 2 |
| Price in EUR/m² | 552 | Type of heating | Central heating | Floor of the flat | 1st |
| Floor area in m² | 50.72 | Water-supply | - | Number of flats in the house | 27 |
| Date of the end of construction | 1987 | Sewerage | Communal waste water elimination | Number of floors | 4 |
| Comparable 2 | | Date of transaction | 01.05.2018. | Similarity coefficient | 0.92 |
| Municipality | Alytus | Type of walls | Bricks | Number of rooms | 2 |
| Price in EUR/m² | 501 | Type of heating | Central heating | Floor of the flat | 2nd |
| Floor area in m² | 49.9 | Water-supply | - | Number of flats in the house | 12 |
| Date of the end of construction | 1991 | Sewerage | Communal waste water elimination | Number of floors | 4 |

*Source*: own research.

Following the selected valuation property and analytically determined similar real estate properties, Formula 7 was applied to determine the correction coefficients based on the time, place and condition. The results have been provided in Table 3. The similarity coefficient according to location was equal to 1 in all cases, because it was estimated only by the municipality and all selected objects were located in the same municipality. Evidently, the location similarity coefficient should be calculated including various parameters such as whether the object is in the suburbs or a rural area, accessibility to transportation services, etc. In this example, this was not possible due to the limited data and data protection regulations. The correction coefficients were calculated with respect to the support object. Usually, the support property was another property from the dataset of comparable real estate properties, as the properties are not only similar to the analyzed real estate, but also to each other. However, the data regarding the support objects is not provided in the article considering how

(due to the amount of the analyzed data, a different support object can be employed to calculate the correction coefficient for each component. This results in three support objects for each comparable.

**Table 3**

Summary of comparative method

| No of the comparable | Similarity coefficient according to | | | Correction coefficient according to | | | Recalculated price, EUR/m² |
|---|---|---|---|---|---|---|---|
| | time | location | condition | time | location | condition | |
| **1** | 0.9581 | 1.0000 | 0.8027 | -2.90% | 0.00% | -9.24% | 485 |
| **2** | 0.9329 | 1.0000 | 0.8176 | -5.19% | 0.00% | 10.18% | 526 |
| **3** | 0.9581 | 1.0000 | 0.7857 | 88.68% | 0.00% | 30.19% | 928 |
| **4** | 0.9496 | 1.0000 | 0.7582 | -19.54% | 0.00% | -4.93% | 398 |
| **5** | 0.9915 | 1.0000 | 0.7163 | -32.87% | 0.00% | -31.00% | 289 |
| **6** | 0.9748 | 1.0000 | 0.7260 | 12.84% | 0.00% | 16.21% | 613 |
| **7** | 0.9581 | 1.0000 | 0.7168 | 48.98% | 0.00% | 2.79% | 815 |
| **8** | 0.9581 | 1.0000 | 0.7055 | -47.75% | 0.00% | -14.42% | 244 |
| **9** | 0.9666 | 1.0000 | 0.6921 | -11.38% | 0.00% | 2.99% | 491 |
| **10** | 0.9496 | 1.0000 | 0.7068 | 11.90% | 0.00% | 4.59% | 558 |
| **Average (recommended) price in EUR/m²** | | | | | | | 491 |

*Source*: own research.

The comparable 3 was excluded from the calculation of the mean value as a correction of more than 100% is recommended. Although the obtained recommended price of the property (491 EUR/m²) is similar to the price of the analyzed property (492 EUR/m²), for this object, the data for the selection of an established number of the comparable properties from the same municipality within the last year (2018) has been sufficient. If objects which are from a considerably earlier period or of significantly differing condition from other municipalities are included in the calculations, the obtained results can be considerably divergent from reality. The results can be improved by introducing weights to the categories in the evaluation of condition similarity (following expert experience) or in the calculation of inter-similarity evaluation.

### 4.2. Machine learning approach for average price determination

In cases when historical data is not available to apply the comparative approach directly, a machine learning approach can be applied. In this case, the data aggregated by municipality is used. The applied machine learning goal is to determine the average price of real estate in the municipalities using the known data and then use the obtained model to predict the average price of real estate in regions where data was not available. In this case, municipalities were used; however, if coordinates were available, more precise classification by location could be provided.

The "*Boosted Decision Tree*" algorithm was employed for the establishment of the coefficients of adjustment of time and location in different districts, where the macro indicators of the municipalities in a given year were related to the average selling price of the property. During the examination of the period of 2008-2017, the total data sample set consisted of 600 entries. For the creation of the model,

municipalities in which more than 20 sales (393 entries) were recorded, data owned in the corresponding year were selected and a model which relates macro indicators to the average price of the object per unit of area in different districts in different years has been created.

Figure 1 presents the convergence of the machine learning algorithm on the training and test data sets. From the error, it can be seen that the model is correct and can be applied in practice.
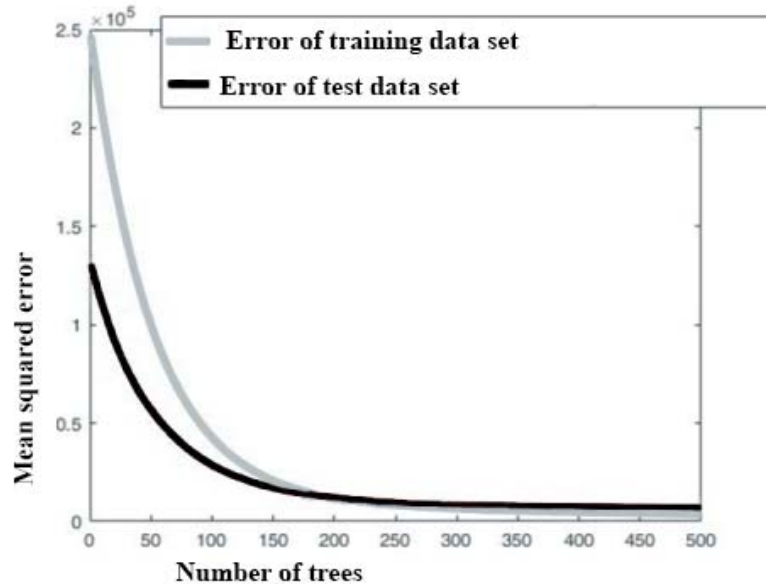


**Fig. 1.** Convergence of the machine learning algorithm. *Source*: own research.

Figure 2 illustrates the importance of macroeconomic indicators to real estate price. It can be seen that the most important indicators of price are the income and expenditure of municipalities and the density of the population. These indicators show that the algorithm has logically explained real estate price variation.
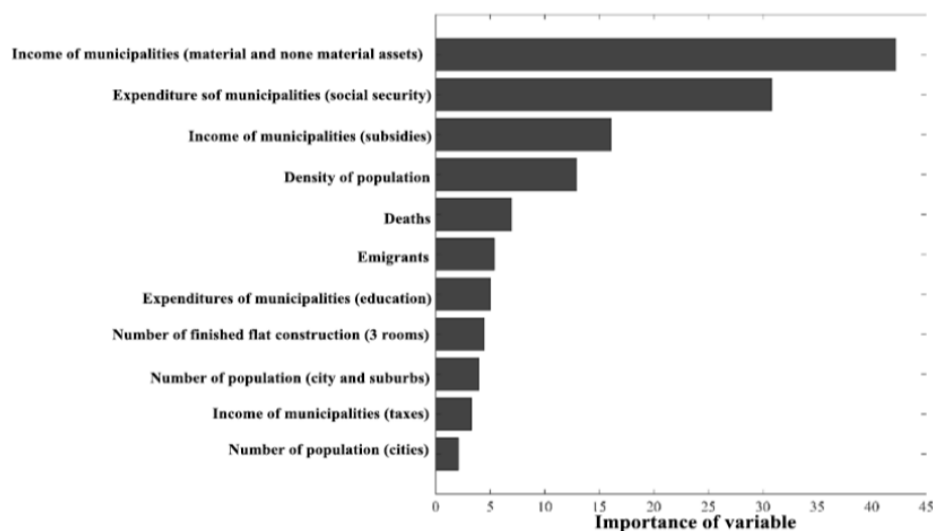


**Fig. 2.** Importance of macro indicators for real estate price. *Source*: own research.

Table 4 provides an example of municipalities by year and the average price of real estate in regions where the availability of historical data was limited.

Table 4

Example of missing data evaluation. Source: own research

| Year | Municipality | Price | Year | Municipality | Price |
|------|------------|-------|------|------------|-------|
| 2013 | Jonava D. Mun. | 276.26 | 2015 | Panevėžys C. Mun. | 625.11 |
| 2013 | Tauragė D. Mun. | 327.75 | 2015 | Širvintos D. Mun. | 309.32 |
| 2013 | Marijampolė Mun. | 282.59 | 2015 | Trakai D. Mun. | 368.78 |
| 2013 | Neringa Mun. | 1716.90 | 2015 | Zarasai D. Mun. | 280.87 |
| 2013 | Jurbarkas D. Mun. | 231.59 | 2015 | Pakruojis D. Mun. | 264.68 |
| 2013 | Utena D. Mun. | 318.73 | 2015 | Utena D. Mun. | 398.79 |
| 2013 | Druskininkai Mun. | 656.92 | 2015 | Anykščiai D. Mun. | 320.89 |

*Source*: own research.

From this chapter it can be concluded that a machine learning approach can be used to determine the average price of real estate in municipalities where available historical data was limited. Jadevicius and Huston (2015) conducted a research to analyze the effectiveness of model complexity (Jadevicius & Huston, 2015). They have concluded that increased model complexity does not necessarily yield greater forecasting accuracy, thus when implementing machine learning algorithms, variable selection is an important factor to consider. The obtained table could be used to determine correction coefficients based on the location and year. Coordinates of historical real estate data are needed to provide correction coefficients based on condition with a machine learning approach. Model testing at the municipality level has not converged, thus smaller regions for comparison should be chosen.

## 5. Discussion and conclusions

There is a growing number of artificial intelligence-based real estate valuation approaches, which mainly focus on estimating the final value of the selected real estate object (Hromada, 2016; De Cock, 2011; Du; et al., 2018; Chiarazzo; et al., 2014; Li Yu, et al., 2018). Although these approaches are suitable for improving the mass valuation model they cannot be applied to individual object valuation since the new legislation requires reasoning to be provided to the valuation process and does not allow a mass valuation model to provide the final value of a real estate object (TEGoVA, 2018). The method provided in the publication can be used as a support tool for valuators to determine correction coefficients when using the comparative method for setting real estate value. A similar method was proposed by Cheng and Hsu (2018); however, the research focused on historical transaction data and infrastructure information (Yeh & Hsu, 2018). The method did not integrate macroeconomic indicators in the determination of correction coefficient process. The research also did not consider situations when available historical data was s limited, because the research was conducted in China. This problem occurs in small markets, but not in large markets or in suburb areas where the population is declining rapidly; thus, our research provides a methodological approach to obtain correction coefficients even when available historical data is limited. Tajani et al. (2019) have proposed a market approach optimized by a weighted appraisal model; however, the model proposed by them could generate reliable outputs only if numerous comparable data were available (Tajani, et al. 2019).

In summary, it was indicated that real estate valuators usually use the comparative approach to estimate real estate prices. However, due to the limited amount of data available and undocumented valuation process, there is a tendency to use experience to determine correction coefficients. However, this provides subjectivity when evaluating real estate prices. Thus, the authors of the paper have provided an analytical method to determine correction coefficients. The limitation of the proposed model is that it was tested only on Lithuania's market, which does not validate the methodology completely. As the proposed model is based on the traditional approach, it can also be applied in other countries with the dataset of local objects. This method not only determines the final real estate price, but also provides intermediate steps which can be used as a support tool for valuators. In cases, where available historical real estate data is limited, a machine learning approach was implemented to obtain the missing average price based on specific location and time. During the application of the analytical method, the real estate data was integrated with macroeconomic indicators, which increased the accuracy of real estate price estimation. The limitation of the research is that the data

used for analysis did not have precise coordinates, thus the municipality level was used. If more detailed data were available, the results of the research could be more accurate.

**Funding**

**Conflict of interest statement**

The authors declare that there is no conflict of interest.

**6. References**

Čiburienė, J., & Jegelavičiūtė, R. (2017). Data Sharing-Way to Improve Real Estate Valuation Quality: Lithuanian Case. *Social & Economic Revue, 15*(2).

Chiarazzo, V., Caggiani, L., Marinelli, M., & Ottomanelli, M. (2014). A neural network based model for real estate price estimation considering environmental quality of property location. *Transportation Research Procedia, 3*(July), 810–817. https://doi.org/10.1016/j.trpro.2014.10.067

D'Acci, L. (2019). Quality of urban area, distance from city centre, and housing value. Case study on real estate values in Turin. *Cities, 91*(November 2018), 71–92. doi:10.1016/j.cities.2018.11.008

De Cock, D. (2011). Ames, Iowa: Alternative to the boston housing data as an end of semester regression project. *Journal of Statistics Education : An International Journal on the Teaching and Learning of Statistics, 19*(3). Advance online publication. https://doi.org/10.1080/10691898.2011.11889627

Del Giudice, V., De Paola, P., & Cantisani, G. B. (2017). Valuation of real estate investments through Fuzzy Logic. *Buildings, 7*(4), 26. Advance online publication. https://doi.org/10.3390/buildings7010026

Du, Q., Wu, C., Ye, X., Ren, F., & Lin, Y. (2018). Evaluating the effects of landscape on housing prices in urban China. *Tijdschrift voor Economische en Sociale Geografie, 109*(4), 525–541. https://doi.org/10.1111/tesg.12308

Dziadosz, A., & Meszek, W. (2015). Selected Aspects of Determining of Building Facility Deterioration for Real Estate Valuation. *Procedia Engineering, 122*(Orsdce), 266–273. doi:10.1016/j.proeng.2015.10.035

Hoesli, M., Jani, E., & Bender, A. (2006). Monte Carlo simulations for real estate valuation. *Journal of Property Investment & Finance, 24*(2), 102–122. https://doi.org/10.1108/14635780610655076

Hromada, E. (2016). Real Estate Valuation Using Data Mining Software. *Procedia Engineering, 164*(June), 284–291. https://doi.org/10.1016/j.proeng.2016.11.621

Yeh, I. C., & Hsu, T. K. (2018). Building real estate valuation models with comparative approach through case-based reasoning. *Applied Soft Computing, 65*, 260–271. https://doi.org/10.1016/j.asoc.2018.01.029

International Association of Assessing Officers. (2011). *Standard on Mass Appraisal of Real Property,* (January), 3–14.

Jegelavičiūtė, R., & Rimkevičiūtė, R. (2017). Correction coefficient of comparative method influence to real estate value. *Adjustment coeficient influence to residential real estate*, 2(2), 1–39.

Li Yu, Chen Li Jiao, Hongrun Xin, Yan Wang, K. W. (2018). Prediction on Housing Price Based on Deep Learning. *World Academy of Science. Engineering and Technology International Journal of Computer and Information Engineering, 12*(2), 10. 10.1016/j.ajpath.2011.02.029

Statistics, L. (2019). *Macroeconomic indicators of Lithuania by municipality*. Retrieved from https://www.stat.gov.lt/en

Navickas, V., Šaudys, A., & Jegelavičiūtė, R. (2017). Comparative approach application in value assessment of land areas in Lithuania. *Journal of Management, 1*, 30.

*Registry center*. (2018). Real estate historical transaction data for 2008 - 2018 period, which were obtained from R&D contract no. SV9-2070.

TEGOVA. (2016). *Europos vertinimo standartai*.

TEGoVA. (2018). Automated Valuation Models (AVMs), 1–6. Retrieved from http://www.tegova.org/data/bin/a591190c05b2c3_Geoge_Matysiak_Valuation_Report.pdf

Tumelionis, A. (2013). *Lyginamojo metodo pataisų apskaičiavimo aktualijos*, 38–54.

TVST. (2017). *Tarptautinis vertinimo standartas*, 1–107.

Zujo, V., Car-Pusic, D., & Zileska-Pancovska, V. (2014). Cost and Experience based Real Estate Estimation Model. *Procedia: Social and Behavioral Sciences, 119*, 672–681. https://doi.org/10.1016/j.sbspro.2014.03.075