

Marek Banaś¹, Józef Czaja², Janusz Dąbrowski³

The Application of Regression Analysis for Estimating the Market Value of Commercial Real Estate

Abstract: The transaction price of a land property with commercial buildings depends on both its quantitative and qualitative attributes. Quantitative attributes include surface areas of plots of land and usable floor spaces of premises and buildings with various intended purposes, as well as values of rents. Qualitative attributes are represented by the global attributes of these properties. In the analysis of the land property market with commercial buildings, all pairs that relate a transaction price to individual attributes are considered. The market value prediction is based on multiple regression analysis for a two-dimensional random variable, represented by the price and the predetermined attribute. The final market value of the property being valued is calculated as the weighted average of the market values predicted for each attribute. This research paper presents the procedure for determining the market value of land with commercial buildings, which falls within the method of statistical analysis of the market. The derived formulas and substantively justified algorithms may be the basis for market analysis and estimation of the market value of such land. This procedure has been thoroughly verified using two practical numerical examples.

Keywords: economic valuation, market analysis, commercial real estate

Received: 19 March 2021; accepted: 8 October 2021

© 2022 Authors. This is an open access publication, which can be used, distributed and reproduced in any medium according to the Creative Commons CC-BY 4.0 License.

¹ The Bronisław Markiewicz State University of Technology and Economics in Jarosław, Institute of Technical Engineering, Poland, email: marek.banas@pwste.edu.pl,

 <https://orcid.org/0000-0002-5426-0452>

² The Bronisław Markiewicz State University of Technology and Economics in Jarosław, Institute of Technical Engineering, Poland, email: jozef.czaja@pwste.edu.pl

³ The Bronisław Markiewicz State University of Technology and Economics in Jarosław, Institute of Technical Engineering, Poland, email: dabrowski.janusz.andrzej@gmail.com,

 <https://orcid.org/0000-0001-8440-7016>

1. Formulating a Research Problem

Land properties developed with building structures intended for commercial purposes, services or production, will be referred to further in this paper as land with commercial buildings. The usable floor space of such properties may include: surface areas of residential, office and commercial premises, surface areas of workshop premises and buildings, surface areas of production or warehouse buildings and facilities, as well as surface areas of ramps and storage yards. An important element of this real estate is also the surface area of their land. The main carrier of the market value of such real estate is the usable floor space of premises and commercial buildings.

All legal regulations and the methods of valuating built-up properties recommended therein and discussed Czaja and Parzych [1], always refer to unit transaction prices, which result from dividing the transaction price for the entire property by the usable floor space of the premises or the building of homogeneous purpose.

Such a unit transaction price, representing a single type of usable floor space, forms the basis for market analysis and property valuation.

It is not possible to determine one representative usable floor space for land developed with commercial buildings of various intended purposes, because each type of specific surface generates a different value of the unit market price, and thus, a different market value.

Real estate transactions involving the sale of land built-up with commercial buildings always present the total transaction price, which includes the price of all types of premises and buildings, as well as the price of a plot of land. Separation of the total transaction price into the prices of individual commercial premises and buildings, and the price of the land, is possible using a generalized market analysis based on the estimation of parametric models.

Research studies on the relationship between transaction prices of commercial real estate and their physical, location, rent and economic attributes, has been the subject of numerous scientific papers including those written by such authors as d'Amato [2] and d'Amato and Amoruso [3].

Fehribach et al. [4] studied the volatility of industrial property prices depending on their 11 physical, location, rent and economic attributes, in order to establish price indices for them. Saderion et al. [5] demonstrated the use of hedonic models to estimate the value of commercial real estate, which are based on the characteristic attributes of these properties and net operating profit. In the final part of this research paper, its authors proposed that commercial real estate analysts should exhibit great inquisitiveness as far as the selection and estimation of unit prices and capitalization rates for various types of real estate are concerned. Parzych and Czaja [6] provided the principles for parameter estimation for various elements of commercial real estate, using the weighted least squares method (WLS). The estimated parameters are unit market values of surface areas of premises having various intended purposes as well as coefficients for distinguished attributes that create the market value.

The developed algorithms were verified on numerous practical examples of commercial property valuation. Epley [7] presented the principles of using the Automatic Valuation Model (AVM) to estimate the market value of a property. The author formulated the criteria which should be met for the Automatic Valuation Model to yield correct results of the analysis of market trends.

Market analysis and valuation of land with commercial buildings, based on total transaction prices, can also be performed using multiple two-dimensional regression models, for which it is necessary to define appropriate quantitative and qualitative attributes and to formulate the appropriate algorithm. As Bruneau and Cherfouh [8] and McCartney [9] claim in their papers, long and short-term analysis of both commercial real estate prices and rents is also an important element of theoretical considerations.

This research paper presents the principle of determining appropriate quantitative and qualitative attributes for particular types of premises and buildings as well as for a plot of land, which will be directly related to the transaction price of such real properties. For such attributes, a market analysis and valuation algorithms for land with commercial buildings have been presented and which are referred directly to the total transaction prices of similar properties.

2. Development of Quantitative and Qualitative Attributes to Describe Volatility of Transaction Prices of Commercial Real Estate

A set of sale transactions of similar real estate that may include usable floor spaces for housing, commerce, services or production, is understood as the market of land with commercial buildings, subject to their location being in the same city or in neighboring communes. The sale of such real estate always consists of a total transaction price, including a plot of land and all buildings as components of this land.

The transaction price of a land property with commercial buildings depends on its quantitative as well as qualitative attributes. Quantitative attributes include surface areas of plots of land and usable floor spaces of premises and buildings with various intended purposes, and they may also comprise values of rents and capitalization rates. Qualitative attributes are represented by global attributes of these properties, such as their location, technical condition, or their standard.

If lands with commercial buildings are located in different districts of the city or in different neighboring communes, then the attribute "location" should be formulated for them, according to the following scale:

- 4 – very advantageous,
- 3 – advantageous,
- 2 – average,
- 1 – disadvantageous.

The market value of the components of a land property is, to a large extent, dependent on the technical condition of the buildings. If the considered land properties include components constructed in a different time period or using building materials of different durability, then the attribute “technical condition of the building” must be defined, according to the following scale:

- 4 – very good,
- 3 – good,
- 2 – satisfactory,
- 1 – unsatisfactory.

Transaction prices of the analyzed properties in the database for market analysis should be adjusted for the date of the analysis (valuation). In order to determine the change rate of transaction prices in time, the best solution is to take commercial premises with a homogeneous intended purpose or land properties earmarked for commercial development. Such groups of real estate allow for specifying unit transaction prices, which will constitute the basis for determining the unit price change rate within one month.

Adjustment of total transaction prices of land with commercial buildings as for the analysis date, should be performed according to the following formula:

$$C_{TK} = C_T [1 + B(t_K - t_T)],$$

where:

- C_{TK} – transaction price is adjusted as for the analysis date,
- C_T – transaction price obtained in the sale,
- B – unit price change rate within 1 month,
- t_K – number of months from first transaction to adjustment (analysis) date,
- t_T – number of months from first transaction to considered transaction.

Having taken the above assumptions into account, the database of land with commercial buildings for market analysis should be presented in the appropriate table, the general form of which is illustrated in Table 1.

Based on the exemplary real estate database contained in Table 1, it is noticeable that it is possible to create seven pairs associating transaction price with the specific seven attributes. The size of these pairs is different, therefore, the specific analyzed attributes have different representativeness, and thus, with varying probability, they describe the volatility of transaction prices. All attributes with a minimum size of 2 will be used for the market analysis. The table above demonstrates that the pairs; price – land, price – warehouse and price – location are represented by six properties, while the pairs; price – office and price – commerce are represented by five properties. The pairs price – flat and price – production are represented by three properties.

Table 1. General form of database for six land properties with commercial buildings

Adjusted transaction price C_{TK}	Surface area of land $[a_G]$ $[m^2]$	Residential floor space $[a_{MK}]$ $[m^2]$	Office space $[a_B]$ $[m^2]$	Commercial space $[a_H]$ $[m^2]$	Production floor space $[a_{PR}]$ $[m^2]$	Warehouse floor space $[a_M]$ $[m^2]$	Attribute location $[a_L]$
C_{T1}	S_{G1}	x	S_{B1}	x	S_{P1}	S_{M1}	a_{L1}
C_{T2}	S_{G2}	S_{MK2}	S_{B2}	S_{H2}	x	S_{M2}	a_{L2}
C_{T3}	S_{G3}	x	x	S_{H3}	S_{P3}	S_{M3}	a_{L3}
C_{T4}	S_{G4}	x	S_{B4}	S_{H4}	x	S_{M4}	a_{L4}
C_{T5}	S_{G5}	S_{MK5}	S_{B5}	S_{H5}	S_{P5}	S_{M5}	a_{L5}
C_{T6}	S_{G6}	S_{MK6}	S_{B6}	S_{H6}	x	S_{M6}	a_{L6}

The market analysis of land with commercial buildings will consider all the pairs that associate transaction price with individual attributes. These pairs will represent a two-dimensional random variable, in which the dependent variable is the transaction price, and the independent variable determines the considered attribute. For each analyzed pair as a two-dimensional random variable, it is necessary to specify the characteristic parameters which include: the average value and standard deviation for both analyzed variables, as well as the total correlation coefficient. Due to the different size of the studied pairs, the probability of representativeness of individual attributes should be taken into account, the value of which will be determined as the ratio of the number of elements in the analyzed attribute and the number of all real properties in the database. The value of this probability should be taken into account when putting together the percentages of individual attributes contributing to the market value of the entire property.

Based on the performed market analysis, according to the principles set out above, the market value of any property can be predicted in each pair considered, i.e. in any two-dimensional random variable. The regression analysis for the two-dimensional random variable represented by the price and the determined attribute is used for this forecast. The ultimate market value of the analyzed property is a weighted average of the market values forecast for each attribute. The formulas for determining these weights are derived using the analysis of variance and probability theory.

To implement the above-mentioned principles of market analysis and real estate valuation, formulas are developed and substantively justified for the estimation of the parameters of a two-dimensional random variable and for the prediction of a dependent variable and its inaccuracy. All theoretical considerations constitute the subject of the next part of the research paper, and their practical illustration, in the form of numerical examples, is presented in the last part of this study.

3. Correlation and Regression of Two-dimensional Random Variable with Respect to Market Analysis and Property Valuation

In the statistical analysis of the market, the price is a dependent variable, whereas the specified property attribute represents an independent variable. Relationships between the price and the individual attributes of real estate can be described using independent models of a two-dimensional random variable.

In order to establish a regression model for two random variables, which are represented by the attribute a and by the transaction price C , appropriate scales must be determined for them. The graphic image of these two variables in the orthogonal coordinate system is called the correlation diagram. Based on the correlation diagram, it is possible to initially conclude what the relationship between the two variables is.

The parameter defining the mutual dependence of the random variables is the covariance between a variable a and a variable C , which is denoted by $\text{cov}[a, C]$. Depending on the units of the analyzed random variables, the quantity $\text{cov}[a, C]$ may take different values. Therefore, the value of the covariance can be standardized using the values of standard deviations of both random variables in boundary distributions, i.e. $\sigma[a]$ and $\sigma[C]$.

The standardized $\text{cov}[a, C]$ is a measure of linear interdependence of the random variables a and C , and it is called (Pearson's) total correlation coefficient, i.e.:

$$r = \frac{\text{cov}[a, C]}{\sigma[a] \cdot \sigma[C]} \quad (1)$$

The value of the total correlation coefficient can be determined based on results from the sample, i.e. the database of real estate for market analysis, according to the following alternative formulas:

$$r = \frac{\sum_{i=1}^n (a_i - \bar{a})(C_i - \bar{C})}{\sqrt{\sum_{i=1}^n (a_i - \bar{a})^2} \cdot \sqrt{\sum_{i=1}^n (C_i - \bar{C})^2}} \Leftrightarrow r = \frac{\frac{1}{n} \sum_{i=1}^n a_i C_i - \bar{a} \cdot \bar{C}}{\sigma[a] \cdot \sigma[C]} \quad (2)$$

where $\bar{a}$, $\bar{C}$ and $\sigma[a]$, $\sigma[C]$ denote the average value and standard deviations of the analyzed attribute and transaction price respectively.

Based on the total correlation coefficient and other characteristic parameters of a two-dimensional random variable, it is possible to determine the functional relationship between the dependent variable and the independent variable, which is called the linear regression model.

Determining the regression model for a two-dimensional random variable (a, C) involves a selection of such a function:

$$C = g(a) \quad (3)$$

which will represent (replace) the set of values of both variables. The estimation of the parameters of this function is always carried out according to the ordinary least squares (OLS) approach, i.e. according to the principle:

$$\Psi(a_i) = \sum_i [C_i - g(a_i)]^2 = \sum_i \delta_i^2 = \min \quad (4)$$

If the function $g(a)$ is adopted in the following linear form:

$$C = A + B \cdot a \quad (5)$$

the condition (4) is expressed by the following relationship:

$$\Psi(a_i) = \sum_i [C_i - A - B \cdot a_i]^2 = \min \quad (6)$$

After raising the expression in square brackets to the power of two, a function is obtained that represents the following quadratic form:

$$\Psi(a_i) = \sum_i C_i^2 + B \sum_i a_i^2 + A \cdot n + 2B \sum_i a_i C_i + 2A \sum_i C_i + 2A \cdot B \sum_i a_i = \min \quad (7)$$

Function (7) will satisfy the minimum condition when its partial derivatives are equal to zero, i.e.:

$$\begin{aligned} \frac{\partial \Psi}{\partial A} &= 2A \cdot n + 2B \sum_i a_i - 2 \sum_i C_i = 0 \\ \frac{\partial \Psi}{\partial B} &= 2A \sum_i a_i + 2B \sum_i a_i^2 - 2 \sum_i a_i C_i = 0 \end{aligned} \quad (8)$$

After solving the above system of equations, the final formulas for A and B are as follows:

$$B = r \cdot \frac{\sigma(C)}{\sigma(a)} \quad (9)$$

$$A = \bar{C} - r \frac{\sigma(C)}{\sigma(a)} \cdot \bar{a} \quad (10)$$

After substituting formulas (9) and (10) into the expression (5), we obtain an association of all characteristic parameters of a two-dimensional random variable (a, C) , in the form of the following expression:

$$C - \bar{C} = r \frac{\sigma(C)}{\sigma(a)} (a - \bar{a}) \quad (11)$$

The quantities occurring in this relationship are defined by formulas (1) to (2), which means that the linear regression model is defined by the average values and standard deviations in boundary distributions determined subject to the database of sold real estate, and taking into account the total correlation coefficient of the variable with respect to the variable C .

Based on the regression model in formula (11), with a predetermined value of one variable, it is possible to predict (calculate) the value of the second variable. In the case of a property value forecast, the value of the attribute (a_p) will be determined, and the predicted unit price of the property (C_p) will be calculated.

Using the analysis of variance for a two-dimensional random variable, it is possible to prove that the absolute correlation coefficient squared (r^2) is closely related to the reliability of explaining transaction prices by the attribute under consideration. For the analysis of the variance of a two-dimensional random variable represented by the value of the attribute a and the value of the transaction price C , their average values will be used.

In the analysis of variance, the sums of squares of differences of a random variable in relation to its average value, and the sum of squares of differences of a random variable with respect to its predicted value from the regression line, are always considered. A measure of the total dispersion of the value of a random variable C in relation to its average value $\bar{C}$ is the sum of the squares of their differences, called the total sum of squares (TSS), i.e.:

$$TSS = \sum_{i=1}^n (C_i - \bar{C})^2 \quad (12)$$

If the considered values of the random variable are described by the regression line, then the difference $C_i - \bar{C}$ can be replaced with two components:

$$C_i - \bar{C} = (C_i - C_{p_i}) + (C_{p_i} - \bar{C}) \quad (13)$$

The first component of this sum relates to the part of the random variable C_i which is not explained by the linear regression model, while the second component relates to the part of this variable explained by the linear regression prediction.

The explained sum of squares (EES) of the variable C_i in the regression model is defined based on the difference between the predicted value of C_{p_i} from the

regression line and the average value of $\bar{C}$, hence it is expressed by the following formula:

$$ESS = \sum_{i=1}^n [g(a_i) - \bar{C}]^2 \quad (14)$$

The residual part of the sum of squares (RSS) of the variable C_i by the regression lines is the difference between TSS and ESS , i.e.:

$$RSS = TSS - ESS \quad (15)$$

The aim of the presented analysis of variance is to show that the ratio of the explained sum of squares (EES) to the total sum of squares (TSS) is equal to the square of the absolute correlation coefficient (r^2) for the considered attribute relative to property prices considered in the database, i.e.:

$$\frac{ESS}{TSS} = \frac{\sum_{i=1}^n (C_{Pi} - \bar{C})^2}{\sum_{i=1}^n (C_i - \bar{C})^2} = r^2 \quad (16)$$

Based on the regression line equation (11), it is possible to calculate the difference between the predicted value of a random variable and its average value ($C_{Pi} - \bar{C}$), i.e.:

$$C_{Pi} - \bar{C} = r \frac{\sigma(C)}{\sigma(a)} (a_{Pi} - \bar{a}) \quad (17)$$

Having taken into account formula (17) and the definition for the variance of the variable C , the expression (16) can be written in the following form:

$$\frac{ESS}{TSS} = \frac{\sum_{i=1}^n \left[r \frac{\sigma[C]}{\sigma[a]} (a_{Pi} - \bar{a}) \right]^2}{n \cdot \sigma^2[C]} = \frac{r^2 \frac{\sigma^2[C]}{\sigma^2[a]} \cdot n \cdot \sigma^2[a]}{n \cdot \sigma^2[C]} = r^2 \quad (18)$$

Dependence (18) proves that r^2 determines the measure of adjustment of the straight line of regression of formula (11) to the set of points representing a two-dimensional random variable represented by a predetermined attribute and the price of individual properties. The quantity r_j^2 can be a measure of explanation of the transaction price by a linear regression model for a predetermined j -th attribute a_{jp} , hence its value can be taken as the weight of accuracy of the prediction C_{jp} based on formula (17). In the analysis of the real estate market, it can be assumed that the value of r_j^2 is the contribution of the considered attribute to the explanation of the price volatility in the database. If these contributions are standardized, then these values

will represent the weighting factors of the attributes. Determination of weighting factors for specific attributes is an essential stage of the market analysis and forms the basis for the valuation of each real property.

4. Algorithm for Estimating Market Value of Land with Commercial Buildings

In the first stage of the algorithm, each pair of “attribute – transaction price” from the real estate database must be considered independently, i.e. specify the following parameters for a two-dimensional random variable:

- $\bar{a}_j$ – average value for the attribute,
- $\bar{C}_j$ – average value for the corresponding prices,
- $\sigma(a_j)$ – standard deviation for the attribute,
- $\sigma(C_j)$ – standard deviation for the corresponding prices,
- r_j – total correlation coefficient,
- P_j – probability of representativeness of the pair ($P_j = n_j/n$).

Based on the above parameters, for each specified j -th pair, a model (11) of linear two-dimensional regression should be determined, with the following denotations:

$$C_j = \bar{C}_j + r_j \frac{\sigma(C_j)}{\sigma(a_j)} (a_j - \bar{a}_j) = A + B \cdot a_j \quad (19)$$

Equation (19) is the basis for forecasting the market value of a similar property, for each analyzed attribute in the database for market analysis. For this purpose, it is necessary to determine the value of the considered attribute a_{jp} , and the value of the price forecast, which represents the predicted market value, should be calculated for this value. Thus, for each j -th attribute, the representativeness of which is determined by the ratio of the number of its value to the number of all transaction prices, the predicted market value ($M\tilde{V}_j$) is determined.

$$M\tilde{V}_j = A + B \cdot a_{jp} \quad (20)$$

According to the considerations contained in (18), the reliability of the market value prediction according to the relationship (20) is determined using r_j^2 , and therefore the weights of the forecast must be determined for the average value from the predicted market values for each attribute. Having adopted the earlier denotations, the weight of reliability of the market value prediction from the j -th attribute should be determined according to the following formula:

$$p_j = \frac{n_j}{n} \cdot r_j^2 \quad (21)$$

The estimated market value (EMV) of a property is calculated as the weighted average of forecasts determined from individual attributes, i.e.:

$$EMV = \frac{\sum_j M\tilde{V}_j \cdot p_j}{\sum_j p_j} \quad (22)$$

The compliance of the valuation model with the database of comparable real estate may be assessed according to the following procedure. For each real property in the database, formula (22) is implemented and, consequently, its market value (EMV_i) is calculated. The standard deviation (σ_n) from differences in transaction prices (C_i) and estimated market values (EMV_i) of each real estate from the database is an average measure of inconsistency of the valuation model with the database of comparable real estate, i.e.

$$\sigma_n = \sqrt{\frac{1}{n} \sum_{i=1}^n (C_i - EMV_i)^2} \quad (23)$$

The non-compliance ratio of the valuation model with the database of comparable real estate will be expressed by the coefficient of dispersion which has the following form:

$$\lambda = \frac{\sigma_n}{\bar{C}} \quad (24)$$

where $\bar{C}$ means the average price of real estate in the database.

5. Numerical Example of Market Analysis and Property Valuation

Subject to six notarial deeds analyzed in a given commune, the following market information was obtained:

- 1) Rep. 1 – Real estate developed with multifunctional pavilions with usable floor space comprising: office space 80 m², commercial space 420 m², warehouse space 800 m², on the plot with the area of 4,800 m². Transaction price for the entire property amounted to PLN 4,300,000.
- 2) Rep. 2 – Real estate developed with an office pavilion with usable floor space comprising: office space 640 m², commercial space 60 m², on the plot with the area of 3,200 m². Transaction price for the entire property amounted to PLN 1,550,000.
- 3) Rep. 3 – Real estate developed with multifunctional pavilions with usable floor space comprising: office space 240 m², commercial space 620 m², warehouse space 960 m², on the plot with the area of 7,450 m². Transaction price for the entire property amounted to PLN 5,800,000.

- 4) Rep. 4 – Real estate developed with multifunctional pavilions with usable floor space comprising: office space 180 m², commercial space 360 m², warehouse space 780 m², on the plot with the area of 5,200 m². Transaction price for the entire property amounted to PLN 4,300,000.
- 5) Rep. 5 – Real estate developed with multifunctional pavilions with usable floor space comprising: office space 120 m², commercial space 540 m², warehouse space 580 m², on the plot with the area of 5,700 m². Transaction price for the entire property amounted to PLN 3,680,000.
- 6) Rep. 6 – Real estate developed with an office-commercial pavilion with usable floor space comprising: office space 880 m², commercial space 120 m², on the plot with the area of 2,200 m². Transaction price for the entire property amounted to PLN 1,910,000.

The real estate described by the notarial deeds has the following identical attributes:

- city zone – outskirts,
- earmarked in the land use plan – areas for services and commerce,
- technical condition and standard of materials used – good.

Therefore, qualitative attributes will not be taken into account in the market analysis and property valuation.

The above information from notarial deeds forms the basis for comparing the values of individual quantitative attributes and transaction prices for the analyzed real estate, as illustrated in Table 2.

Table 2. Database for six land properties with commercial buildings for market analysis and valuation

Transaction no.	Land surface area a_G [m ²]	Office space a_B [m ²]	Commercial space a_H [m ²]	Warehouse space a_M [m ²]	Transaction price C_T [PLN]
1	4,800	80	420	800	4,300,000
2	3,200	640	60	–	1,550,000
3	7,450	240	620	960	5,800,000
4	5,200	180	360	780	4,300,000
5	5,700	120	540	580	3,680,000
6	2,200	880	120	–	1,910,000

Based on the database of sold real estate contained in Table 2, all statistical parameters for two-dimensional random variables, represented by the selected attribute and the transaction price, were calculated. The values of these parameters are summarized in Table 3.

Table 3. Statistical parameters for individual two-dimensional random variables

Attribute	Average value for attribute	Average value for corresponding prices	Standard deviation for attribute	Standard deviation for corresponding prices	Total correlation coefficient	Probability of representativeness
Land surface area	4,758	3,590,000	1,960	1,605,266	0.9278	1
Office space	357	3,590,000	326	1,605,266	-0.7865	1
Commercial space	353	3,590,000	224	1,605,266	0.9169	1
Warehouse space	780	4,520,000	156	901,998	0.9347	4/6 = 0.67

The parameters summarized in Table 3 form the basis for each attribute and transaction price, a comparison of the linear regression equation that will be used to predict the market value of the real estate being valued.

The valuation was carried out for the land with commercial buildings, with qualitative attributes identical to the ones included in the database, but with the following quantitative attributes:

- land surface area – 5,600 m²,
- office space – 220 m²,
- commercial space – 520 m²,
- warehouse space – 720 m².

All the calculated parameters of the regression line and the predicted market value for real estate with the above-listed attributes are demonstrated in Table 4.

Table 4. Summary for each attribute and transaction price of regression equation and predicted market value for real estate being valued

Name of attribute	Regression equation (general form) $C_{jp} = A_j + B_j \cdot a_{jp}$	Value of attribute of real estate being valued a_{jp}	Predicted market value $M\tilde{V}_j$	Weight of reliability $p_j = P_j \cdot r_j^2$
Land surface area	$C_G = -219,921 + 800.6839 \cdot \tilde{a}_G$	5,600	4,263,909	0.86
Office space	$C_B = 4,971,082 - 3,872.194 \cdot \tilde{a}_B$	220	4,119,200	0.62
Commercial space	$C_H = 1,268,645 + 6,569.8725 \cdot \tilde{a}_H$	520	4,684,979	0.84
Warehouse space	$C_M = 298,571 + 5,412.088 \cdot \tilde{a}_M$	720	4,195,275	0.58
$EMV = \frac{4,263,909 \cdot 0.86 + 4,119,200 \cdot 0.62 + 4,684,979 \cdot 0.84 + 4,195,275 \cdot 0.58}{0.86 + 0.62 + 0.84 + 0.58} =$ $= \frac{12,589,507}{2.9} = 4,341,209 \text{ PLN}$				

In order to compare the obtained property valuation result with the estimated value according to the parametric model, appropriate matrices were defined using the values contained in Table 2. Having applied the matrix calculations, according to Subsection 4.6 of the research paper Parzych and Czaja (2015), estimation of the parameters of the Gauss-Markov model was performed, and then, on their basis, the market value of the analyzed property was calculated. As a result of this valuation, the market value of $EMV = \text{PLN } 4,424,600$ was obtained. When compared to the value contained in Table 5, it appears that the difference between the estimated market values according to the two analyzed models does not exceed 2%.

If the weights of the reliability of the attributes are standardized, i.e. $\hat{p}_j = p_j / \sum_j p_j$, and these values are multiplied by $EMV = \text{PLN } 4,341,209$ and then divided by the value of the analyzed attribute (surface area), then the unit market value for specific real estate components is obtained. All the calculations aimed at obtaining the unit market value have been carried out in Table 5.

Table 5. Calculation of unit market values for individual components of real estate being valued

Attribute – real estate component	Standardized weights of attributes $\hat{p}_j = \frac{p_j}{\sum_j p_j}$	Estimated market value of component of real estate being valued $EMV_j = EMV \cdot \hat{p}_j$ [PLN]	Value of attribute of real estate being valued a_{jp} [m ²]	Estimated unit market value of components of real estate being valued $EM\tilde{V}_{junit}$ [PLN/m ²]
Land surface area	0.296354	1,263,625	5,600	225.65
Office space	0.212929	877,096	220	3,986.80
Commercial space	0.289376	1,355,718	520	2,607.15
Warehouse space	0.201492	845,313	720	1,174.05

In order to assess the conformity of the valuation model with the database of comparable real estate, for each real estate in the database its market value (EMV_i) was calculated. Standard deviation (σ_n) from differences in transaction prices (C_i) and estimated market values (EMV_i) is a measure of the inaccuracy of estimating the market value of the analyzed property. The non-compliance ratio of the valuation model with the database of comparable real estate was calculated according to formula (24) and summarized in Table 6.

The value of the coefficient $\lambda = 0.12$ demonstrates that the market value estimated according to the multiple two-dimensional linear regression procedure may be determined with an inaccuracy of less than 12%. The value of this coefficient leads to an interval $0.90 \geq (1 - \lambda) > 0.85$, which means that the compliance of the valuation model with the database of comparable real estate is at a fairly high level.

6. Numerical Example of the Valuation of a Land Property Developed with a Building Intended for Trade and Services

The land properties selected for the market analysis were those developed with buildings intended for trade and services, located within the valued property and being the subject of trading in the last six months. The analysis was made of the information on the nine market transactions, in which the land properties were described using six price-setting attributes. The scales for optional attributes were established on the following four levels:

- 3 – very advantageous,
- 2 – advantageous,
- 1 – average,
- 0 – disadvantageous.

Table 7 contains information on the attributes (features) of land properties developed with buildings intended for trade and services as well as on their rents and unit transaction prices.

Table 7. Database of sold properties developed with buildings intended for trade and services

Transaction no.	Location	Transport access	Surrounding	Monthly unit rent [PLN/(m ² ·m-c)]	Land surface area [m ²]	Usable floor space of premises [m ²]	Transaction price [C_i] [PLN]	Unit price [c_i] [PLN/m ²]
1	2	2	1	270	600	130	702,000	5,400
2	3	3	2	290	610	120	672,000	5,600
3	2	2	1	230	540	175	805,000	4,600
4	2	2	1	215	520	160	688,000	4,300
5	2	2	1	270	630	140	756,000	5,400
6	3	2	2	280	650	150	840,000	5,600
7	3	3	2	310	780	180	1,080,000	6,000
8	2	2	1	260	630	130	676,000	5,200
9	3	2	2	295	650	110	627,000	5,700
Valuation	2	2	2	251	580	146	?	?

The subject of the valuation is a land property developed with buildings intended for trade and services, which has the attributes listed in the last row of Table 8. It is more beneficial to determine the unit transaction price for each analyzed property rather than consider the total transaction price as then the further analysis of the attribute may be omitted [usable floor space of premises].

After using the Excel application, all the statistical elements necessary for the implementation of formula (11) were determined and their values included in Table 8.

Table 8. Summary of statistical elements
used to estimate the unit market value of the analyzed property

Designations	Location	Transport access	Surrounding	Monthly unit rent [PLN/(m ² ·m-c)]	Land surface area [m ²]	Unit price [c _i] [PLN/m ²]
1	2	2	1	270	600	5,400
2	3	3	2	290	610	5,600
3	2	2	1	230	540	4,600
4	2	2	1	215	520	4,300
5	2	2	1	270	630	5,400
6	3	2	2	280	650	5,600
7	3	3	2	310	780	6,000
8	2	2	1	260	630	5,200
9	3	2	2	295	650	5,700
Average value	2.44	2.22	1.44	253.78	623.33	5,311.11
Standard deviation	0.50	0.42	0.50	5.75	70.24	510.87
Total correlation coefficient	0.7246	0.5115	0.7246	0.9919	0.8784	–
Absolute correlation coefficient squared (r_j^2)	0.5251	0.2617	0.5251	0.9838	0.7716	$\Sigma = 3.07$
Attribute weight $p_j = \frac{r_j^2}{\sum_j r_j^2}$	0.1712	0.0853	0.1712	0.3207	0.2516	–
Weighting factor of the attribute $\left[p_j \cdot r_j \cdot \frac{\sigma(c)}{\sigma(a_j)} \right]$	127.54	53.62	127.54	28.27	1.61	–
Attribute difference ($a_{jw} - \bar{a}_j$)	–0.44	–0.22	0.56	–2.78	–43.33	–
Product $\left[p_j \cdot r_j \cdot \frac{\sigma(c)}{\sigma(a_j)} \right] \cdot (a_{jw} - \bar{a}_j)$	–56.69	–11.92	70.86	–78.52	–69.64	Total –145.91

The final formula for the unit market value has the following form:

$$MV_j = \bar{c} + \sum_j \left[p_j \cdot r_j \frac{\sigma(c)}{\sigma(a_j)} \right] \cdot (a_{jw} - \bar{a}_j).$$

The estimated unit market value of the analyzed property was determined after taking into account the average value of unit transaction prices $\bar{c} = 5,311.11$ and the sum of the products of weighting factors of the attributes and their differences ($a_{jw} - \bar{a}_j$), the value of which equals to -145.91 :

$$MV_j = 5,311.11 - 145.91 = 5,165.20 \text{ PLN/m}^2.$$

According to the considered algorithm, the estimated market value of the analyzed usable floor space of 146 m^2 is:

$$MV = 5,165.20 \cdot 146 = 754,120 \text{ PLN}.$$

All the calculations presented above were performed with the Excel application.

7. Conclusions

The procedure presented for estimating market value of land developed with commercial buildings is based on the assumptions of statistical analysis of the market. The derived formulas and substantively justified algorithms may form the basis for market analysis and estimation of the market value of commercial real estate. The database for market analysis and valuation should contain information regarding at least six properties similar to the one being valued, which were subject to real estate transactions on the market.

The proposed procedure for estimating the market value of real estate is based on multiple models of two-dimensional regression, defined by the pairs: attribute – transaction price, the parameters of which are estimated according to the method of least squares. The algorithms of this procedure allow us to take into account the lack of some components of quantitative attributes in the considered pairs. The result of estimating the market value of a property can be assessed in terms of its inaccuracy, and then the assessment of the conformity of the valuation model with the database of comparable real estate can be performed.

The analysis of the presented algorithms made on the basis of the two numerical examples demonstrated that the procedure proposed for estimating market value will always yield results that are consistent with the Gauss–Markov model. All the calculations occurring in this procedure may be performed with the widely available Excel software application.

References

- [1] Czaja J., Parzych P.: *Szacowanie rynkowej wartości nieruchomości w aspekcie Międzynarodowych Standardów Wyceny*. Stowarzyszenie Naukowe im. Stanisława Staszica, Kraków 2007.
- [2] d'Amato M.: *Cyclical Capitalization*. [in:] Lorenz D., Dent P., Kauko T. (eds.), *Value in a Changing Built Environment*, Wiley Blackwell, 2017, pp. 151–172.
- [3] d'Amato M., Amoruso P.: *Application of a Cyclical Capitalization Model to the London Office Market*. *International Real Estate Review*, Global Social Science Institute, vol. 21(1), 2018, pp. 113–143.
- [4] Fehribach F., Rutherford R., Eakin M.: *An Analysis of the Determinants of Industrial Property Valuation*. *Journal of Real Estate Research*, vol. 8, no. 3, 1993, pp. 365–376. <https://doi.org/10.1080/10835547.1993.12090717>.
- [5] Saderion Z., Smith B., Smith Ch.: *An Integrated Approach to the Evaluation of Commercial Real Estate*. *Journal of Real Estate Research*, vol. 9, no. 2, 1994, pp. 151–167. <https://doi.org/10.1080/10835547.1994.12090743>.
- [6] Parzych P., Czaja J.: *Szacowanie rynkowej wartości nieruchomości*. Wydawnictwa AGH, Kraków 2015.
- [7] Epley D.: *The Need to Reference Automatic Valuation Models to the Valuation Process*. *Journal of Real Estate Literature*, vol. 25, no. 1, 2017, pp. 237–251. <https://doi.org/10.1080/10835547.2017.12090450>.
- [8] Bruneau C., Cherfouh S.: *Long-run equilibrium for the Greater Paris office market and short-run adjustments*. *Journal of Property Research*, vol. 32(4), 2015, pp. 301–323. <https://doi.org/10.1080/09599916.2015.1089310>.
- [9] McCartney J.: *Short and long-run rent adjustment in the Dublin office market*. *Journal of Property Research*, vol. 29(3), 2012, pp. 201–226. <https://doi.org/10.1080/09599916.2012.689990>.