DOI: 10.5604/01.3001.0014.0953

PRZEGLĄD STATYSTYCZNY
TOM LXVI – ZESZYT 4 – 2019

Krzysztof DMYTRÓW[1]

Calibration of attributes influence in the process of real estate mass appraisal by using decision-making methods[2]

Abstract. *There are situations in the real estate market in which a large number of properties have to be valued at the same time. In such cases it is advisable to use mass valuation methods. These methods involve estimating the value of a property on the basis of the values of the attributes defining it. The aim of the paper is to calibrate the influence of attributes on unit values of properties in mass appraisal in order to minimise the valuation error. The research was conducted for 318 residential properties located in Szczecin. The Szczecin Algorithm of Real Estate Mass Appraisal was used along with the econometric, statistical and expert approaches. The econometric approach is based on the ridge regression model, the statistical approach on the partial Kendall τ correlation coefficients, and the expert approach on the AHP method. The quadratic programming was co-employed with the statistical and expert approaches in order to minimise the mean square error (MSE) of the valuations. The econometric and statistical approaches with the minimisation of the MSE generated best results. The least accurate results were obtained by means of the statistical and expert approaches without the minimisation of the MSE. However, even though the optimisation of the MSE improves the quality of valuations, it also narrows down their volatility, which might make the valuation of properties from the outside of a given database more problematic.*

Keywords: real estate mass appraisal, real estate attributes, AHP method, statistical and econometric methods of real estate mass appraisal, quadratic programming.

JEL Classification: C34, C44, R30

Kalibracja wpływu atrybutów w procesie wyceny masowej nieruchomości za pomocą metod decyzyjnych

Streszczenie. *Zdarzają się sytuacje, gdy jednocześnie należy wycenić dużą liczbę nieruchomości. W takich przypadkach wskazane jest stosowanie metod masowej wyceny nieruchomości. Za ich pomocą wycenia się wartości nieruchomości na podstawie wartości definiujących je atrybutów. Celem artykułu jest taka kalibracja wpływu atrybutów na*

[1] University of Szczecin, Institute of Economics and Finance, ul. Mickiewicza 64, 71-101 Szczecin, e-mail: krzysztof.dmytrow@usz.edu.pl, ORCID: https://orcid.org/0000-0001-7657-6063.

[2] The research was conducted within the project financed by the National Science Centre, Project No. 2017/25/B/HS4/01813.

wartości jednostkowe nieruchomości w masowej wycenie, aby uzyskać jedynie minimalne błędy wycen. Badanie przeprowadzono dla 318 nieruchomości mieszkaniowych zlokalizowanych w Szczecinie. Zastosowano Szczeciński Algorytm Masowej Wyceny Nieruchomości wraz z podejściem ekonometrycznym, statystycznym i eksperckim. Podejście ekonometryczne oparto na modelu regresji grzbietowej, podejście statystyczne na cząstkowych współczynnikach korelacji τ Kendalla, natomiast podejście eksperckie na metodzie AHP. W celu zminimalizowania średniokwadratowego błędu wycen (MSE), zastosowano programowanie kwadratowe wraz z podejściem statystycznym i eksperckim. Najlepsze wyceny osiągnięto przy zastosowaniu podejścia ekonometrycznego i statystycznego z minimalizacją MSE. Najmniej dokładne wyniki uzyskano przy zastosowaniu podejścia statystycznego i eksperckiego bez minimalizacji MSE. Z jednej strony optymalizacja MSE poprawia jakość wycen, ale z drugiej zawęża ich zmienność, co może sprawić, że wycena nieruchomości dla innych danych może być problematyczna.

1. INTRODUCTION

Real estate appraisal can be performed individually, i.e. separately for each property, or by means of mass appraisal, i.e. when many properties are valued simultaneously. Each of the two above-mentioned methods has its advantages and disadvantages. The main advantage of the individual appraisal is that just one property or a limited number of properties are appraised by a real estate appraiser directly, which allows taking into account all the specific features that are sometimes unique and can be attributed only to one given property, therefore the appraisal can be very precise. On the other hand, however, the process of such valuation is time-consuming and the number of real estates that can be appraised individually in a defined period of time is limited. Individual appraisals are the most frequently performed procedures that use applicable valuation rules resulting from the law, a number of professional standards, basic and specialist valuation standards and interpretative notes (Źróbek and Bełej, 2000).

On the other hand, we can talk about real estate mass appraisal when the following conditions are met simultaneously (Hozer et al., 2002):

- the object of valuation is a large number of properties of the same type;
- the valuation is carried out using the same method for each property, which yields comparable results;
- all properties are valued simultaneously.

However, it should be noted that the two above-mentioned types of appraisal (individual and mass) are not considered as alternative or substitutional, but complementary. The individual real estate appraisal is usually carried out when there is a need for determining the value of a specific property, for the purpose of a purchase or sale of a property, for insurance purposes or to assess losses caused by random incidents, robberies, etc. The real estate mass appraisal, on the other hand, is adopted for the purposes of (Hozer et al., 1999):

- carrying out the revaluation of annual fees for the perpetual usufruct,
- estimation of the economic effects of adopting or amending local spatial development plans,
- monitoring the value of real estate that secures a bank's credit exposures in order to calculate LtV for the bank's credit portfolio,
- universal property taxation,
- other necessities, such as the sale of residential property from municipal resources, expropriation for linear investments, etc.

In order to perform mass real estate appraisal, all properties (of the same type) have to be definable by the same attributes. Therefore, it is very important to prepare a good, reliable database prior to the valuation process. As in the case of individual property appraisal, mass appraisal also requires the work of property appraisers, who have to determine reliable values of attributes. Real estate mass valuation should be carried out by means of quantitative methods. Literature distinguishes four main groups of quantitative methods used in real estate mass appraisal (Kauko and D'amato, 2008):

- model-driven methods;
- data-driven methods;
- methods based on machine learning;
- expert methods.

Model-driven methods constitute a set of classical, quantitative methods such as standard regression models, hedonic regression models or spatial regression models. Data-driven methods include non-parametric Geographically Weighed Regression (GWR) models. They are closely connected with machine learning techniques, such as the Artificial Neural Network, fuzzy sets, genetic algorithms, ridge regression, lasso regression, random forests, regression trees, etc. Nowadays, with the development of programming languages and increasing computing capability of computers, these methods are becoming increasingly popular (Ćetković et al., 2018). The first three groups of the above-mentioned quantitative methods require access to a large amount of data. In the case of incomplete databases (where some possible values of attributes are missing), it is either impossible or very difficult to obtain a reliable model. In such cases, it is necessary to consult experts about the influence of attributes on the value of a property. This can be done by means of the expert approach, based on multiple-criteria decision making (MCDM) methods. It would be hard to list all of the MCDM methods, because of their large number (Saaty and Ergu, 2015). They can be divided into two main groups (Hwang and Yoon, 1981):

- Multiple Attribute Decision Making (MADM);
- Multiple Objective Decision Making (MODM).

In the first group, the criteria are defined by the attributes, while in the second group – by the objectives. Also, in the first group, all the alternatives are explicitly known, while in the second, not necessarily (there could be an infinite number of them). Problems that arise while employing methods from the second group can be solved by means of the multiple objective mathematical programming procedures. As regards the methods from the first group, the decision-maker is usually interested in sorting, ranking or classifying alternatives. Due to the fact that the criteria are defined by attributes here, these methods are useful for the expert approach of specifying the influence of attributes on the value of properties. Methods used for solving problems within this group are called discrete MCDM methods or MADM methods. The best known of them are: AHP (Analytical Hierarchy Process), ANP (Analytic Network Process), REMBRANDT (Ratio Estimation in Magnitudes or deciBells to Rate Alternatives which are Non-DominaTed), DEMATEL (DEcision MAking Trial and Evaluation Laboratory), ELECTRE (ELimination Et Choix Traduisant la REalité), PROMETHEE (Preference Ranking Organization METHod for Enrichment of Evaluations), COPRAS (Complex Proportional Assessment of Alternatives), MAUT (Multi-attribute Utility Theory), MAVT (Multiattribute Value Theory), SAW (Simple Additive Weighting), VIKOR, TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) and many others (Saaty and Ergu, 2015). Also, there are MADM methods that originate from the multivariate statistical analysis, based on Hellwig's composite measure of development (Nermend, 2017).

If we want to employ experts to assess the influence of attributes on the value of real estate, it could be done by means of weights. In the expert approach, weights could be assessed directly by experts, but this would be highly subjective. Another method of obtaining weights of criteria (attributes) on the value of real estate is applying an appropriate multiple-criteria decision making technique. The question which technique to choose for this purpose could be answered by studying the problem closey. There are many publications which provide guidelines on how to choose the best-fitting decision-making technique. Guitouni and Martel (1998) presented the characteristics of almost 30 MCDM methods along with the recommendations as to which of them to use in which situation. Saaty and Ergu (2015) presented 16 criteria by which various MCDM methods were evaluated and compared. A comprehensive comparison of many MCDM methods was also carried out by Evangelos and Triantaphyllou (2000). If the aim of a study is to obtain the vector of weights of the criteria, the AHP method would be a good choice (Guitouni and Martel, 1998, p. 508; Trzaskalik, 2014, p. 274-275). The AHP is also a method most widely used in the real estate market of all the MCDM methods. It was applied, for example, to forecasting values of properties (Yalpir, 2014), establishing the weights of real estate attributes (Kozioł-Kaczorek, 2012), the appraisal of properties for purchase purposes (Ball and Srnivasan, 1994) and to making purchase-related decisions (Saaty, 1990). The main advantage of the expert methods over the other abovementioned ones is the fact that they can be used even if data is incomplete.

However, they also have several disadvantages. For example, instead of the real relationship between the property's attributes and its value, they reflect subjective (although supported by experience) estimation of this relationship by an expert. Such approach might be satisfactory in some cases, but generally it is not as good as when we can assess the above-mentioned relationship on the basis of full data. Also, estimations done by different experts tend to vary. It is assumed that real estate appraisers should assess the states and influence of attributes on the value of real estate in a similar way (this assumption is based on another assumption, namely that if the same real estate is evaluated by several appraisers, the obtained values should be approximately the same), but it is not always true. Therefore, expert methods should be regarded as a complement to other methods of mass real estate appraisal, rather than as an alternative.

The aim of the paper is to calibrate the influence of attributes on the unit value of properties in mass appraisal in order to minimise the valuation error. In the study presented below, properties were appraised by means of the Szczecin Algorithm of Real Estate Mass Appraisal (SAREMA) with the application of the econometric, statistical and expert (based on the AHP method) approaches. The quadratic programming was used in order to minimise the mean square error of valuations in the statistical and expert approaches.

2. RESEARCH METHODOLOGY

The following approaches of the calibration of influence of attributes were adopted in the research:

- econometric approach;
- statistical approach;
- AHP method;
- quadratic programming.

The starting point of each approach is the Szczecin Algorithm of Real Estate Mass Appraisal (SAREMA) (Hozer et al., 1999). In this algorithm, the unit value (or the transactional price) of a property is appraised by means of the following formula:

$$\hat{v}_{ji} = wwr_j \cdot v_b \cdot \prod_{k=1}^{K} \prod_{p=1}^{k_p} \left(1 + a_{kpi}\right), \tag{1}$$

where:

$\hat{v}_{ji}$ – unit market (or cadastral) value (value of 1 m^2) of the i-th property ($i = 1, 2, \dots, n$) in the j-th location attractiveness zone ($j = 1, 2, \dots, J$),

v_b – basic value of 1 m^2,
a_{kpi} – impact of the p-th state of the k-th attribute on the i-th property,
wwr_j – market value ratio in the i-th location attractiveness zone,
n – number of properties,
J – number of attractiveness zones,
K – number of attributes,
k_p – number of states of the k-th attribute.

The basic value (v_b) can be estimated in several ways. It can be the theoretical value of 1 m^2 of a property with the worst states of attributes in the cheapest location attractiveness zone. It can also be the value of 1 m^2 of the cheapest property in the appraised area. However, this approach is not suitable if we want to extend the algorithm to real estate from beyond the database currently being analysed. The reason is that it is strictly connected with the market value ratio (wwr_j). This measure informs us about the influence of the widely understood location on the value of a property. The location attractiveness zones are spatial units where properties of the same purpose with the same values of attributes should obtain similar values. In order to estimate the market value ratios, we need to have a set of representative properties that have to be appraised by real estate appraisers. Next, the values of the same properties have to be calculated using the algorithm. It is done by the following formula:

$$\hat{v}_{hji} = v_b \cdot \prod_{k=1}^{K} \prod_{p=1}^{k_p} (1 + a_{kpi}), \tag{2}$$

where $\hat{v}_{hji}$ – unit value of the i-th representative property in the j-th location attractiveness zone.

The value of wwr_{ji} for each representative property is calculated in the following way:

$$wwr_{ji} = \frac{v_{rji}}{\hat{v}_{hji}}, \tag{3}$$

where v_{ri} – unit value of the i-th representative property in the j-th location attractiveness zone.

Finally, market value ratios for each location attractiveness zone are calculated using the following formula:

$$wwr_j = \sqrt[l_j]{\prod_{i=1}^{l_j} wwr_{ji}}, \tag{4}$$

where l_j is the number of properties in the j-th attractiveness zone. Because the market value ratio is understood as the so-called 'location premium', its value for each location attractiveness zone should be no less than 1. The algorithm has a multiplicative form, therefore the geometric mean is used.

The most important and difficult part of applying the algorithm described by Formula (1) is the estimation of the impact of the p-th state of the k–th attribute on the i-th property (a_{kpi}). It can be done using various approaches. The econometric approach, being the modification of Formula (1), was proposed by Doszyń (2018). In this approach, logarithms of both sides of Equation (1) were calculated, and the error term was added:

$$ln(v_{ji}) = \alpha_0 + \sum_{k=1}^{K}\sum_{p=2}^{k_p} \alpha_{kp} x_{kpi} + \sum_{j=2}^{J} \alpha_j te_j + u_i, \tag{5}$$

where:

α_0 – intercept parameter (logarithm of the basic unit value),
α_{kp} – impact of the p-th state of attribute k,
x_{kpi} – dummy variable for the p-th state of attribute k for the i-th property ($i = 1, 2, \dots, n$),
α_j – (the logarithm of) the market value coefficient for the j-th location attractiveness zone,
te_j – dummy variable equal to 1 for the j-th location attractiveness zone (and zero for others),
u_i – error term.

By means of Formula (5), we represent the SAREMA in the econometric form. Real estate attributes are measured on the ordinal scale, therefore dummy variables were used. Variables determining the location attractiveness zones are categorical, therefore they are also represented as dummy variables. The main advantage of applying the econometric model is that the influence of each state of every attribute and location attractiveness zone on the unit value is analysed separately and independently. The main disadvantage of this approach is that there are many explanatory variables, which may cause problems with collinearity.

The next approach used in the study is the statistical one. It is based on the relationship between the real estate attributes and the value of 1 m^2 of a piece of real estate, for the representative properties. Because real estate attributes are measured on the ordinal scale, the rank correlation coefficient was applied. The two rank correlation coefficients that are most widely used are the Spearman and the Kendall coefficients. It is possible to use both of them; however, the Spearman coefficient calculates the differences between ranks, which is methodologically incorrect for the ordinal scale. Therefore, the Kendall rank coefficient is a better choice for this kind of data (Doszyń, 2017; Foryś and Gaca, 2016). Due to the fact that the number of cases is higher than the number of variants for each attribute, the Kendall rank coefficients with tied ranks (τ_B) was used (Parker et al., 2011):

Since the attributes may be correlated, partial τ_B Kendall coefficients were calculated:

$$\tau_{yx.z} = -\frac{R_{\mathrm{yx}}}{\sqrt{R_{\mathrm{yy}} \cdot R_{\mathrm{xx}}}}, \tag{6}$$

where:

R_{yx} – determinant of a matrix cofactor obtained by removing the row corresponding to the explained variable y and the column corresponding to the explanatory variable x,
R_{yy} – determinant of a matrix cofactor obtained by removing the row and column corresponding to the explained variable y,
R_{xx} – determinant of a matrix cofactor obtained by removing the row and column corresponding to the explanatory variable x.

Having calculated partial correlation coefficients, we can proceed to the estimation of weights of each attribute in the algorithm (Kolenda, 2006):

$$w_k = \frac{\left|\tau_{yx_k.z}\right|}{\sum_{k=1}^{K}\left|\tau_{yx_k.z}\right|} \cdot 100\%, \tag{7}$$

where:

k – number of the analysed attribute,
K – number of attributes.

The expert approach is based on the AHP method. It was developed by Thomas L. Saaty (1980). It involves pairwise comparisons of the decision criteria, which in this study were the attributes of properties. Pairwise comparisons between the attributes were performed on the basis of a survey, addressed to four real estate appraisers. Instead of applying the 9-point Saaty scale (Brunelli, 2015), the appraisers suggested using a shorter, 4-point scale, which measured the dominance of attribute 1 over attribute 2:

- attribute 1 has a definitely greater impact (4);
- attribute 1 has a noticeably greater impact (3);
- attribute 1 has a slightly greater impact (2);
- the attributes are indifferent (1);
- attribute 2 has a slightly greater impact (1/2);
- attribute 2 has a noticeably greater impact (1/3);
- attribute 2 has a definitely greater impact (1/4).

The results of the pairwise comparisons c_{kl}, where k, l – compared attributes, $(k, l = 1, 2, \dots, K; k \neq l)$ between the attributes are placed in the AHP matrix.

If the obtained AHP matrix is consistent, then the weights of attributes can be estimated. The weights obtained in the statistical and expert approaches are used to assess the impact of the p-th state of the k-th attribute $\left(1 + a_{kp}\right)$ in Formulas (1) and (2):

$$1 + a_{kp} = e^{\ln\left(\frac{v_{\max}}{v_b}\right) \cdot u_{kp}} = \left(\frac{v_{\max}}{v_b}\right)^{u_{kp}} \tag{8}$$

where:

v_{max} – maximum, the theoretical value of 1 m^2 of a property with the best states of attributes in the most expensive location attractiveness zone,

u_{kp} – influence of the weight of the p-th state of the k-th attribute, calculated in the following way:

$$u_{kp} = \frac{w_k}{k_p - 1}(p - 1). \tag{9}$$

Formula (15) is the author's approach to the estimation of the impact of attributes on a unit value of real estate. The ratio $\frac{v_{\max}}{v_{\mathrm{b}}}$ is used in order to transfer the range of unit values of properties onto the valuated ones. The main assumption of Formulas (8) and (9) is that firstly, for the poorest state of a property's attribute ($p = 1$), the value of $1 + a_{kp} = 1$, and secondly, the higher state of the attribute, the higher the value of Formula (8). Formula (9) assumes that the transitions between the states of attributes are (relatively) linear.

The main advantage of the statistical approach is that it reflects the true relationships between the attributes and the unit value of a property. Its main drawback, however, is that in the first stage, the impacts of the p-th state of the k-th attribute $(1 + a_{kp})$ are estimated, while the market value ratios (wwr_j) are calculated in the second stage. It might lead to a situation where values $1 + a_{kp}$ would depend not only on the relationships between the attributes and the unit value, but also on location. The influence of location is considered in the second stage. Therefore the influence of location on the unit value of a property might be biased. The expert approach should be free from this disadvantage, but the assessment of the relationship between the attributes and the unit value of a property in this approach is subjective. The main advantage of the expert approach is that it can be used regardless of the availability of data.

The statistical and expert approaches are the bases for the application of the quadratic programming to the minimisation of the valuation errors. It is assumed that the highest value of a property is obtained for the best states of attributes, and the lowest value of a property – for the poorest states of attributes. For a perfect database (where properties appraised by experts have values corresponding to the states of attributes), the valuation error should have a minimal value without the need of optimisation (because the algorithm always satisfies the assumption that better states of attributes generate higher values of properties). However, databases (especially from the real estate market) are usually far from being perfect. Therefore, the ratio $\frac{v_{\max}}{v_{\mathrm{b}}}$, obtained by the values set by appraisers, does not necessary reflect the real variability of the values of properties. This ratio, then, will be the subject of optimisation in quadratic programming. The values of the representative properties were appraised both by the appraisers and by the algorithm. The quadratic programming was used to minimise the Mean Square Error (MSE):

$$1 + a_{kp} = e^{\ln\left(\frac{v_{\max}}{v_b}\right)\cdot u_{kp}} = \left(\frac{v_{\max}}{v_b}\right)^{u_{kp}} \tag{10}$$

where:

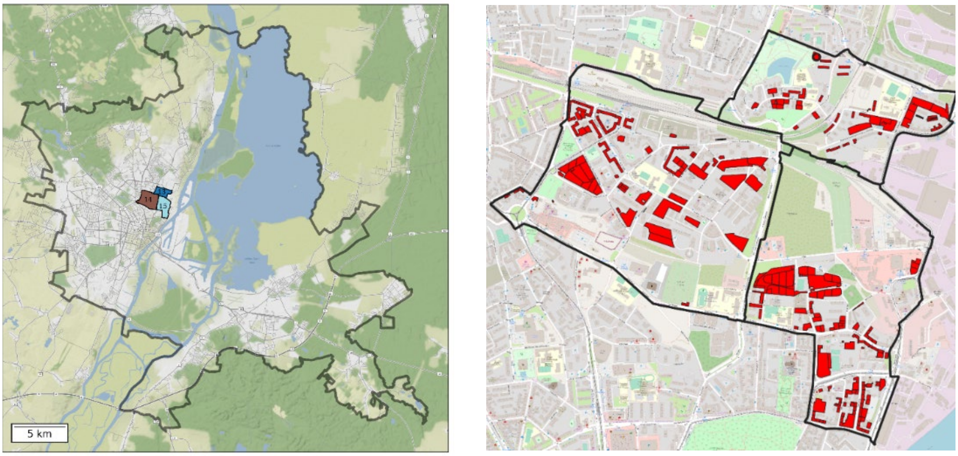
v_{ji} – real unit value of the property determined by the appraiser,
$\hat{v}_{ji}$ – theoretical unit value of the property determined by the algorithm (1).

The quadratic programming was not used in the econometric approach, because the procedure of estimation of the econometric model assumes the minimisation of the sum of squares, which is virtually the same as the *MSE*.

3. DATA USED IN THE STUDY

The research was based on 318 properties located in Szczecin. All of them served housing purposes and were located in three location attractiveness zones (numbered 13, 14 and 15). Their spatial distribution is presented in Figure 1.

Figure 1. Spatial distribution of valued properties



S o u r c e: author's work based on data from the Central Office of Geodesy and Cartography.

All properties were described by five attributes:

- area: 1 – small (up to 500 m^2), 2 – average (500 – 1200 m^2), 3 – large (over 1200 m^2);
- access to utilities: 1 – none, 2 – partial, 3 – full;
- accessibility of public transport: 1 – poor, 2 – average, 3 – good;
- quality of surroundings: 1 – onerous, 2 – unfavourable, 3 – average, 4 – favourable;
- attractiveness of the shape of a plot: 1 – low, 2 – average, 3 – high.

The interpretation of the first four attributes seems relatively straightforward, thus not requiring further explanation. However, the last attribute – 'the attractiveness of the shape of a plot' needs an explanation. It is assumed that that the optimal shape of a plot is a rectangle with the side length ratio of 3:2 (Dmytrów et al., 2018). All the plots were characterised by their area and circumference, among other features. For the circumference, the hypothetical area assuming the optimal rectangle was calculated and compared with the real area. When the ratio of the real area and the hypothetical one was higher than 0.9, the value of the attribute was 3. When it was between 0.5 and 0.9, the value of the attribute was 2, and when it was lower than 0.5, the value of the attribute was 1.

All the 318 properties had full access to utilities, therefore this attribute was removed from the econometric and statistical approaches, as it was impossible to measure its impact on the unit value. It was retained for the expert approach, though, because this approach allows measuring the influence of access to utilities on the value of a property.

Out of the set of 318 properties, 30 representative ones were selected. It should be explained here that the term 'representative property' is not understood in the sense of the representative method. Representative properties are selected in order to ensure that the whole range of attributes in every location attractiveness zone are taken into account in the study. In the analysis presented in this paper it was done by means of stratified sampling from sets of properties with the same values of attributes and in the same attractiveness zones. The borders of location attractiveness zones were determined by experts. The assumption behind this process is that the location attractiveness zones should consist of similar properties, i.e. for the given values of attributes, the values of properties should be roughly the same. But, one might ask, if the goal was to appraise 318 properties (a relatively small population), what was the purpose of selecting 30 of them? It is because of the method – the researcher who uses the SAREMA, has at his/her disposal only these representative properties on the basis of which the market value ratios are estimated. The number of representative properties is usually small, because individual appraisal is time-consuming and expensive. In this study, the representative properties were valued by appraisers and used to estimate the market value ratios in the three analysed location attractiveness zones, by means of equations (2) – (4) (in the statistical and expert approaches). In the statistical approach, they were also used to estimate the values of the partial Kendall τ_B correlation coefficients and the impact of states of attributes on the unit value of a property. In the econometric approach, these representative properties were used to estimate the parameters of the model (5).

The results obtained for the representative properties were then used to appraise all the 318 properties. In order to verify the effectiveness of the SAREMA, all the 318 properties were also valued by real estate appraisers, therefore it was possible to calculate valuation errors for the whole population. The basic error measure was the percentage error (PE):

$$PE_{ji} = \frac{v_{ji} - \hat{v}_{ji}}{v_{ji}} \cdot 100\%. \tag{11}$$

The optimal value of the percentage error is 0%. For all properties, the mean percentage error (MPE) was calculated:

$$MPE = \frac{\sum_{i=1}^{n} PE_{ji}}{n}. \tag{12}$$

The mean percentage error shows if valuations are biased. Its optimum value is 0%. If it is positive, it means that the valuations are on the average underestimated, and if it is negative, it means they are overestimated.

The relative accuracy of valuations was measured using the mean absolute percentage error ($MAPE$):

$$MAPE = \frac{\sum_{i=1}^{n} \left| PE_{ji} \right|}{n}. \tag{13}$$

The closer the value of $MAPE$ to 0%, the better – it means that valuations are closer to the real values of properties. The value of $MAPE$ indicates the mean percentage deviation of values of properties obtained by means of the algorithm from their real unit values.

The absolute accuracy of valuations was measured by means of the root mean square error ($RMSE$):

$$RMSE = \sqrt{MSE}. \tag{14}$$

The smaller the value of $RMSE$, the better. It carries information about the mean absolute deviation of the values of properties obtained by means of the algorithm from their real unit values.

The final measure of accuracy of the valuations were the shares of properties for which the percentage errors (PE) fell within the range of ±5%, ±10% and ±15%.

4. EMPIRICAL RESULTS

In the first step of the analysis, 30 representative properties were valued. The following approaches were adopted:

- econometric (Ec);
- statistical (St);
- expert (Ex);
- statistical with the minimisation of the MSE (St_{MSE});
- expert with the minimisation of the MSE (Ex_{MSE}).

The basic value of 1 m^2 (v_b) was estimated by property appraisers at 279 PLN. Theoretical maximum value of 1 m^2 (v_{max}) was estimated at 708 PLN. Therefore the $\frac{v_{\max}}{v_{\mathrm{b}}}$ ratio equalled 2.54; it means that the most expensive property that had the best states of attributes and was located in the most expensive location attractiveness zone was expected to be 154% more expensive than the cheapest one that had the poorest states of attributes and was located in the cheapest location attractiveness zone. After the optimisation of the *MSE* for statistical and expert approaches, this ratio was set at the levels of 1.37 and 1.24, respectively.

In order to apply the algorithm (1) to the statistical and expert approaches, weights of attributes had to be calculated. For the expert approach, four real estate appraisers did pairwise comparisons of attributes in the AHP method. The results (AHP matrixes – C_1, C_2, C_3, C_4) are presented in equations (15) – (18) (the order of rows/columns is the same as the order of attributes: area, access to utilities, accessibility of public transport, quality of surroundings and attractiveness of the shape of a plot):

$$C_1 = \begin{bmatrix} 1 & \frac{1}{3} & 1 & \frac{1}{3} & 1 \\ 3 & 1 & 3 & 2 & 3 \\ 1 & \frac{1}{3} & 1 & \frac{1}{3} & \frac{1}{2} \\ 3 & \frac{1}{2} & 3 & 1 & 3 \\ 1 & \frac{1}{3} & 2 & \frac{1}{3} & 1 \end{bmatrix}, \tag{15}$$

$$C_2 = \begin{bmatrix} 1 & 2 & 2 & \frac{1}{2} & 1 \\ \frac{1}{2} & 1 & 4 & 3 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & 1 & \frac{1}{2} & \frac{1}{4} \\ 2 & \frac{1}{3} & 2 & 1 & 2 \\ 1 & 2 & 4 & \frac{1}{2} & 1 \end{bmatrix}, \tag{16}$$

$$C_3 = \begin{bmatrix} 1 & \frac{1}{4} & 2 & \frac{1}{4} & 2 \\ 4 & 1 & 4 & \frac{1}{2} & 4 \\ \frac{1}{2} & \frac{1}{4} & 1 & \frac{1}{3} & 2 \\ 4 & 2 & 3 & 1 & 4 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 1 \end{bmatrix}, \tag{17}$$

$$C_4 = \begin{bmatrix} 1 & 1 & 1 & \frac{1}{3} & 2 \\ 1 & 1 & 1 & \frac{1}{3} & 2 \\ 1 & 1 & 1 & \frac{1}{3} & 2 \\ 3 & 3 & 3 & 1 & 3 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix}. \tag{18}$$

The weights of attributes obtained on the basis of matrixes C_1, C_2, C_3, C_4 are presented in Table 1.

TABLE 1. WEIGHTS OF ATTRIBUTES FOR THE AHP METHOD FOR ALL APPRAISERS

Appraiser	Attributes				
	area	access to utilities	accessibility of public transport	quality of surroundings	attractiveness of the shape of a plot
1	10.91%	37.73%	9.73%	28.73%	12.91%
2	20.27%	24.75%	7.30%	24.33%	23.35%
3	12.36%	31.39%	10.17%	39.11%	6.97%
4	16.09%	16.09%	16.09%	42.26%	9.47%

S o u r c e: author's calculation.

As presented in Table 1, the weights obtained from the four experts were diverse. The highest weights were obtained for the access to utilities and the quality of surroundings, whereas the lowest for the accessibility of public transport and the attractiveness of the shape of a plot. The AHP matrix for the second expert was not consistent (matrixes obtained for the three other appraisers were), despite the fact that this person's assessments were repeated twice (the *consistency ratio* for this expert equalled 1.17). However, the author decided to leave it in the study, because some researches demonstrated that vectors of weights obtained for participants with the consistency ratio higher than 0.1 were not significantly different from these obtained for the consistency ratio equal to or lower than 0.1 (Apostolou and Hassell, 1993). Subsequently, mean weights for each attribute were used in the expert approach. The weights of attributes for the statistical approach (based on the partial Kendall τ_B coefficients) and for the expert approach are presented in Table 2.

TABLE 2. WEIGHTS OF ATTRIBUTES FOR STATISTICAL AND EXPERT APPROACHES

Approach	Attributes				
	area	access to utilities	accessibility of public transport	quality of surroundings	attractiveness of the shape of a plot
Statistical	7.57%	—	51.42%	31.78%	9.24%
Expert	14.91%	27.49%	10.82%	33.61%	13.17%

S o u r c e: author's calculation.

Weights obtained by the statistical and expert approaches are quite different from one another. Besides the fact that it was impossible to estimate the weight of the ‘access to utilities’ attribute for the statistical approach (in the expert approach, the weight of this attribute was the second largest, at 27.5%), the most significant difference can be observed in the case of the ‘accessibility of public transport’ attribute. The experts associated the least weight to this attribute (10.8%), whereas from the point of view of the strength of the relationship between this attribute and the unit value of the representative properties, calculated by the statistical approach, its weight (over 51.0%) surpassed the sum of the weights of all the other attributes. The ‘quality of surroundings’ attribute had a similar weight in both approaches. The smallest weights in the statistical approach were obtained for the ‘area’ and the ‘attractiveness of the shape of a plot’ attributes. They also assumed relatively small weights in the expert approach. These weights were transformed into the impact of attributes on the unit value of a property using equations (8) and (9).

In the econometric approach, it was impossible to estimate the classical regression model (5) because of the collinearity of attributes. In order to overcome this difficulty, the ridge regression (Tikhonov regularisation) was applied (Calvetti et al., 2000). The regularisation parameter was set at the level of 0.0001. Table 3 presents the vectors of the impact of attributes on the unit value of properties $(1 + a_{kp})$ for the econometric, statistical and expert approaches.

TABLE 3. VECTORS OF IMPACT OF ATTRIBUTES ON UNIT VALUE OF PROPERTIES $(\mathbf{1} + \boldsymbol{A_{KP}})$

Attributes	States of attributes	Approach				
		Ec	St	Ex	St_{MSE}	Ex_{MSE}
Area	1	1.000	1.000	1.000	1.000	1.000
	2	0.987	1.036	1.072	1.012	1.016
	3	1.003	1.073	1.149	1.024	1.033
Access to utilities	1	—	—	1.000	—	1.000
	2	—	—	1.137	—	1.030
	3	—	—	1.292	—	1.061
Accessibility of public transport	1	1.000	1.000	1.000	1.000	1.000
	2	0.969	1.270	1.052	1.085	1.012
	3	1.032	1.614	1.106	1.176	1.024
Quality of surroundings	1	1.000	1.000	1.000	1.000	1.000
	2	1.014	1.104	1.110	1.034	1.025
	3	1.039	1.218	1.232	1.069	1.050
	4	1.097	1.344	1.367	1.106	1.075
Attractiveness of the shape of a plot	1	1.000	1.000	1.000	1.000	1.000
	2	1.083	1.044	1.063	1.015	1.014
	3	1.058	1.090	1.131	1.030	1.029

S o u r c e: author”s calculation.

Analysing the vectors of impact of the states of attributes on the unit value of a property, we can see that the highest values of $1 + a_{kp}$ in the statistical and expert approaches (both with and without the minimisation of the MSE) were obtained for the attributes with the highest weight and the lowest otherwise (compare Table 2). Because the $\frac{v_{\max}}{v_{\mathrm{b}}}$ ratio for non-optimised approaches was higher than the ratio for the optimised ones, the values of $1 + a_{kp}$ in the former approaches were higher. A higher $\frac{v_{\max}}{v_{\mathrm{b}}}$ ratio will certainly cause more significant dispersion of results. When we compare the results of the statistical and expert approach with the econometric one, it turns out that the latter has most in common with the statistical approach with the minimisation of the MSE, but anyway the differences between the two approaches are relatively significant. The first observation is that in the econometric approach, higher states of attributes do not always have to have a stronger impact on the unit value of a property. In fact, this is only true in the case of one attribute, namely the quality of surroundings. This results from the fact that the values of properties estimated by experts have not always been higher when the states of attributes were higher. Moreover, the econometric approach considers all the states of each attribute separately (as dummy variables do), while in the statistical approach, the correlation strength and direction are reflected for the whole range of the attribute.

After the application of the above-presented results to the whole set of 318 properties, the measures of the accuracy of valuations were obtained (Table 4).

TABLE 4. MEASURES OF ACCURACY OF VALUATIONS (THE MOST ACCURATE VALUES ARE BOLDED)

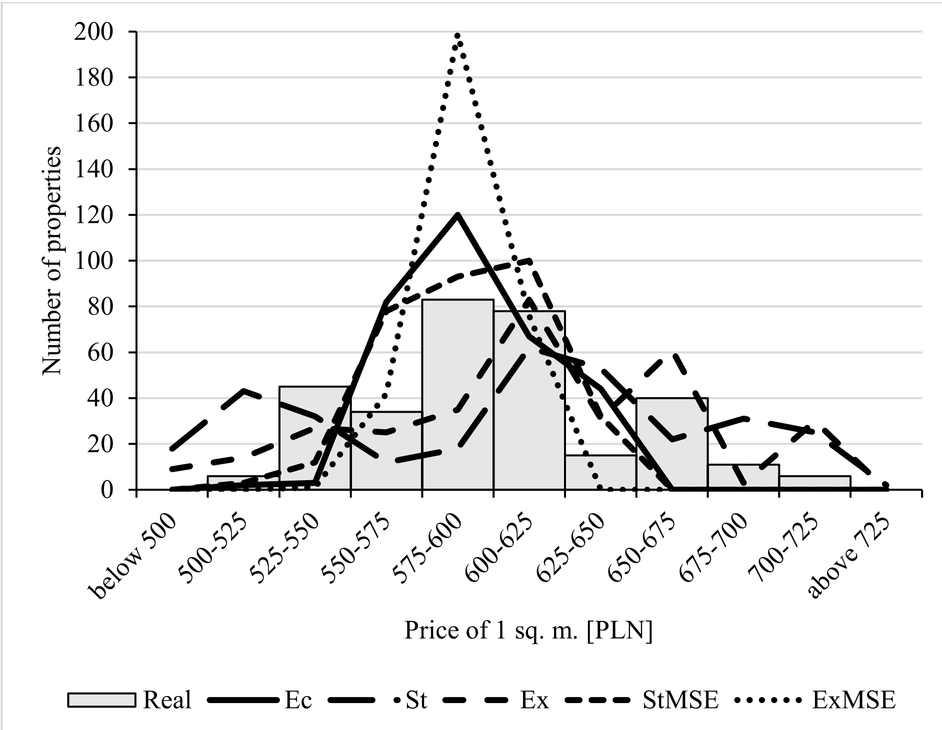
Measures of accuracy	Approach				
	Ec	St	Ex	St_{MSE}	Ex_{MSE}
MPE	**0.77%**	–0.87%	–2.18%	0.92%	0.79%
$MAPE$	4.74%	7.96%	7.82%	**4.51%**	5.21%
$RMSE$	36.39	59.55	56.53	**35.94**	40.19
Minimum appraised value	**520.19**	427.25	419.70	522.69	543.98
Maximum appraised value	644.97	729.09	**723.90**	631.46	619.53
Maximum underestimation	–14.54%	–23.10%	–21.32%	**–9.71%**	–14.24%
Maximum overestimation	16.72%	25.27%	30.19%	**14.45%**	15.53%
Share of properties with errors between ±5%	57.23%	38.37%	36.16%	**58.49%**	53.15%
Share of properties with errors between ±10%	88.37%	70.76%	66.35%	**93.08%**	84.59%
Share of properties with errors between ±15%	99.37%	86.16%	87.42%	**100.00%**	99.69%

S o u r c e: author’s calculation.

It can be seen that as regards the accuracy of valuations, we can divide applied methods into two groups. The first, with relatively high accuracy, consists of the econometric approach and statistical and expert approaches with the minimisation of the MSE. The second group, presenting lower accuracy, comprises statistical and expert approaches without the minimisation of the MSE. In the first group, the valuations tend to be slightly underestimated (the smallest mean percentage error – 0.767%, was obtained for the econometric approach). In the second group, the situation is opposite – valuations are overestimated, with the mean overestimation equal to 0.87% for the statistical approach and 2.18% for the expert approach. The smallest mean absolute percentage and standard errors were obtained for the statistical approach with the minimisation of the MSE. If we compare the extreme (minimum and maximum) valuations with the real values (502.11 and 701.43 PLN), the minimum valuation closest to the real valuation was obtained for the econometric approach, while the maximum valuation closest to the real one – for the expert approach. All the remaining measures of accuracy indicate that the most accurate mass real estate appraisal approach is the statistical approach with the minimisation of the MSE. In this approach, the maximum underestimation of a unit value of the appraised properties was below 10%, and the highest overestimation – less than 15%. Almost 58.5% of all the valued properties had the valuation percentage error equal to or lower than ±5%. Valuation errors lower than ±10% were obtained for over 93% of all the properties, and the highest valuation error equalled ±15%. As expected, higher range of valuations was obtained for the statistical and expert approaches without the minimisation of the MSE than for the variants with the minimisation of the MSE (because of the higher $\frac{v_{\max}}{v_{\mathrm{b}}}$ ratio in the former). It is worth noting that even in the group of approaches yielding the least accurate results (statistical and expert approaches without the minimisation of the MSE), the results are still relatively accurate. The maximum relative errors oscillate around 30%, over 86-87% of all properties had relative valuation errors smaller than ±15%, and the mean relative error did not exceed 8%.

The results presented in Table 4 demonstrate that in the analysed case, the most effective approach among the methods for mass real estate appraisal was the statistical approach with the minimisation of the MSE, closely followed by the econometric approach. The expert approach with the minimisation of the MSE was the third most effective method, whereas the statistical and expert approaches without the minimisation of the MSE were the least effective ones. However, assessing the quality of methods on the basis of only a few parameters might be misleading. If we look at the distribution of unit values of properties obtained by means of the analysed approaches and the real values, we can observe that the expert approach with the minimisation of the MSE caused strong narrowing of the value range (Figure 2).

Figure 2. Distributions of the real and the estimated unit values of properties



S o u r c e: author's calculation.

The distributions of unit values of properties obtained using the econometric and statistical approaches with the minimisation of the MSE were relatively close to the real ones, but the extreme values tended to be over- or underestimated. The extreme values, on the other hand, turned out to be estimated relatively precisely by the statistical and expert approaches without the minimisation of the MSE. This demonstrates that the minimisation of the MSE is a compromise: its application decreases average valuation errors, but at the same time causes flattening of the results. This particular behaviour of the results was useful for this study, but it will not necessarily be useful in all the cases. Therefore, it should always be decided whether the advantages of the optimisation of valuation errors overweight the disadvantages (narrowing of the results, underestimation of high values and overestimation of low values of properties).

5. CONCLUSIONS

The aim of the article is to perform real estate mass appraisal using decision-making methods, and then to compare the obtained results with the results of the classical (econometric, statistical and expert) approaches. The outcome of the study demonstrates that adopting decision-making methods (in this case, the quadratic programming) greatly improves the accuracy of estimations calculated by means of the Szczecin Algorithm of Real Estate Mass Appraisal (SAREMA). The statistical approach, based on the partial Kendall τ_B correlation coefficients with the minimisation of the MSE, turned out to be the most accurate one, followed closely by the econometric approach. The expert approach with the minimisation of the MSE, although yielding relatively precise results, was noticeably less accurate than the two above-mentioned approaches, which just confirmed the earlier expectations. It is because both the statistical approach with the minimisation of the MSE and the econometric approach are based on the real relationships between the attributes of a property and its value per 1 m^2. The expert approach, based on real estate appraisers' experience, on the other hand does not always reflect the real relationships in the real estate market. Therefore it could be recommended as a complementary approach, adopted when the statistical or econometric approaches could not be used (e.g. because of small or incomplete data sets).

Although the statistical approach with the minimisation of the MSE is the most effective method in the presented research, it cannot be universally acclaimed as the best one. Relatively much depends on the quality of the used database, which, in turn, is determined by the quality of work of experts (real estate appraisers). If they define the attractiveness zones correctly (properties belonging to each attractiveness zone are alike, i.e. for given states of attributes their values are similar), if the real estate attributes are correctly identified and the representative properties are properly appraised (their value is higher for better states of attributes and lower for poorer states of attributes), then the quality of database could be regarded as good. In such a case, the base value would reflect the value of a property in the cheapest location attractiveness zone with the poorest values of attributes, and the maximum value would represent the value of a property in the most expensive location attractiveness zone and with the best values of attributes. A database prepared with such a degree of diligence would not even need the optimisation of valuation errors, because the original base and maximum values would reflect the variability of real estate values accurately.

It should also be stressed that even though the above findings apply to the analysed 318 properties, they cannot be automatically employed to every situation. Before the application of any of the above-mentioned methods to other pieces of research, it should first be checked which approach would generate the best results for that particular case. It is generally hard to find a method which would be universally applicable to the real estate market, because every territorial unit, i.e. every city, town or village has its own specificity, to which the applied methods should be adapted.

Although the SAREMA has been used for 20 years, this study introduces new approaches to the calibration of the influence of attributes for this algorithm (statistical and expert approaches with the minimisation of valuation errors), presents their advantages and disadvantages, and facilitates the selection of the most relevant approach for a particular study. Further research into this subject should cover other methods of calibrating the influence of attributes (where transitions between the states would not necessarily be linear) and other approaches to the applications of the SAREMA.

REFERENCES

Apostolou B., Hassell J. M., (1993), An empirical examination of the sensitivity of the analytic hierarchy process to departures from recommended consistency ratios, *Mathemmatical. Computer Modelling*, 17(4/5), 163–170. DOI: 10.1016/0895-7177(93)90184-Z.

Ball J., Srnivasan V. C., (1994), Using the Analytic Hierarchy Process in House Selection, *The Journal of Real Estate Finance and Economics*, 9, 69–85. DOI: 10.1007/BF01153589.

Brunelli M., (2015), Introduction to the Analytic Hierarchy Process, Springer, Cham Heidelberg, New York, London. DOI: 10.1007/978-3-319-12502-2.

Calvetti D., Morigi S., Reichel L., Sgallari F., (2000), Tikhonov regularization and the L-curve for large discrete ill-posed problems, *Journal of Computational and Applied Mathematics*, 123(1–2), 423–446. DOI: 10.1016/S0377-0427(00)00414-3.

Ćetković J., Lakić S., Lazarevska M., Žarković M., Vujošević S., Cvijović J., Gogić M., (2018), Assessment of the Real Estate Market Value in the European Market by Artificial Neural Networks Application, *Complexity*, (special issue, article ID 1472957), 1–10. DOI: 10.1155/2018/1472957.

Dmytrów K., Gnat S., Kokot S. (2018), Próba uwzględnienia kształtu nieruchomości gruntowych jako atrybutu w procesie zalgorytmizowanej wyceny, *Problemy Rynku Nieruchomości*, 2(50), 4–10.

Doszyń M., (2017), Statistical Determination of Impact of Property Attributes for Weak Measurement Scales, *Real Estate Management and Valuation*, 25(4), 75–84. DOI: 10.1515/remav-2017–0031.

Doszyń M., (2018), Ekonometryczna wersja szczecińskiego algorytmu masowej wyceny nieruchomości, *Problemy Rynku Nieruchomości*, 50(2), 11–17.

Foryś I., Gaca R., (2016), Application of the Likert and Osgood Scales to Quantify the Qualitative Features of Real Estate Properties, *Folia Oeconomica Stetinensia*, 16(2). 7–16. DOI: 10.1515/foli-2016-0021.

Guitouni A., Martel J. M., (1998), Tentative guidelines to help choosing an appropriate MCDA method, *European Journal of Operational Research*, 109(2), 501–521. DOI: 10.1016/S0377-2217(98)00073-3.

Hozer J., Kokot S., Foryś I., Zwolankowska M., Kuźmiński W., (1999), *Ekonometryczny algorytm masowej wyceny nieruchomości gruntowych*, Uniwersytet Szczeciński, Szczecin.

Hozer J., Kokot S., Kuźmiński W., (2002), *Metody analizy statystycznej rynku w wycenie nieruchomości*, Polska Federacja Stowarzyszeń Rzeczoznawców Majątkowych, Warszawa.

Hwang C. L., Yoon K., (1981), *Multiple Attribute Decision-Making: Methods and Applications*, Springer-Verlag, New York.

Kauko T., D'Amato M., (eds.), (2008), *Mass Appraisal Methods: An International Perspective for Property Valuers*, Blackwell Publishing, Oxford.

Kolenda M., (2006), *Taksonomia numeryczna. Klasyfikacja, porządkowanie i analiza obiektów wielocechowych*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Kozioł-Kaczorek D., (2012), Hierarchizacja cech nieruchomości z zastosowaniem analitycznego procesu hierarchicznego, *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 20(1), 165–174.

Nermend K., (2017), *Metody analizy wielokryterialnej i wielowymiarowej we wspomaganiu decyzji*, Wydawnictwo Naukowe PWN, Warszawa.

Parker R. I., Vannest K. J., Davis J. L., Sauber S. B., (2011), Combining Nonoverlap and Trend for Single-Case Research: Tau-U, *Behavior Therapy,* 42(2), 284–299. DOI: 10.1016/j.beth. 2010.08.006.

Saaty T. L., (1980), *The Analytic Hierarchy Process*, McGraw-Hill, New York.

Saaty T. L., (1990), How to make a decision: The analytic hierarchy process, *European Journal of Operational Research*, 48(1), 9–26. DOI: 10.1016/0377-2217(90)90057-I.

Saaty T. L., Ergu D., (2015), When is a Decision-Making Method Trustworthy? Criteria for Evaluating Multi-Criteria Decision-Making Methods, *International Journal of Information Technology & Decision Making*, 14(6), 1171–1187. DOI: 10.1142/S021962201550025X.

Triantaphyllou E., (2000), *Multi-criteria decision-making methods: A comparative study*, Kluwer Academic Publishers, London. DOI: 10.1007/978-1-4757-3157-6.

Trzaskalik T., (ed.), (2014), *Wielokryterialne wspomagania decyzji. Metody i zastosowania*, Polskie Wydawnictwo Ekonomiczne, Warszawa.

Yalpir S., (2014), Forecasting residential real estate values with AHP method and integrated GIS, in: *Conference proceedings of People, Buildings and Environment, an international scientific conference,* Kroměříž, Czech Republic, 694–706.

Źróbek S., Bełej M., (2000), *Podejście porównawcze w szacowaniu nieruchomości*, Educaterra, Olsztyn.