USE OF SPATIAL AUTOCORRELATION TO BUILD REGRESSION MODELS OF TRANSACTION PRICES

Radosław Cellmer, PhD
Faculty of Geodesy and Land Management
University of Warmia and Mazury in Olsztyn
e-mail: rcellmer@uwm.edu.pl

Abstract

This paper presents the principles of studying global spatial autocorrelation in the land property market, as well as the possibilities of using these regularities for the construction of spatial regression models. Research work consisted primarily of testing the structure of the spatial weights matrix using different criteria and conducting diagnostic tests of two types of models: the spatial error model and the spatial lag model.

The paper formulates the hypothesis that the application of spatial regression models greatly increases the accuracy of transaction price prediction while forming the basis for the creation of cartographic documents including, among others, maps of land value.

Keywords: spatial autocorrelation, spatial models, market analysis.

JEL Classification: C21, R32.

Citation: Cellmer R., (2013), "Use of spatial autocorrelation to build regression models of transaction prices", Real Estate Management and Valuation, vol. 21, no. 4, pp. 65-74.

DOI: 10.2478/remav-2013-0038.

1. Introduction

The level of real estate transaction prices is shaped by many factors, among which location is of special importance. The qualitative character of this feature is a serious limitation in using classical regression models for price analysis. Moreover, these models do not take into consideration spatial correlations between variables, hence both model parameters and the residual are not a function of location. However, the effect of location can be included in modeling in the form of spatial interactions – understood as the degree of correlation of prices in a given location with prices in other locations.

The main aim of the conducted research (the results of which are presented in this study) was to analyze the use of spatial regression models for real estate market analyses using the phenomenon of spatial autocorrelation of transaction prices.

2. Spatial autocorrelation of transaction prices

Location plays a special role among the many factors shaping the value of land properties and, as a result, transaction prices. It is composed of all determinants connected, among others, with socioeconomic and spatial factors, factors relating to space use, accessibility, the distance from places of employment or those relating to local facilities as well as nuisances (BASU and THIBODEAU 1998). A similar location of real estate should, therefore, correspond to similar transaction prices, with the assumption that they do not differ significantly in terms of physical or legal characteristics. It is thus justified to adopt the thesis that land property prices are spatially correlated.

Spatial autocorrelation is defined as the degree of correlation of the observed variable value in a given location with the value of the same variable in another location (SUCHECKI 2010). Spatial autocorrelation can be employed when the occurrence of a phenomenon in a spatial unit causes an increase or decrease in the probability of the occurrence of this phenomenon in the neighboring units (JANC 2006).

The study of a spatial autocorrelation phenomenon requires, as a rule, determination of the spatial weights which represent the spatial relations presented in the form of a graph or matrix. In the case of a matrix notation, it is necessary to construct an adjacency matrix in which the relations between its elements are presented. The selection of the method of weight determination depends, to a large degree, on the nature of the analyzed phenomenon and on additional information not included in the set (LESAGE and PACE 2009). The obtained analysis results are highly dependent on the selection of weights, but it is not possible to provide objective patterns indicating in which case their individual types should be used (JANC 2006).

Spatial weights can be determined on the basis of contact analysis, distance or similarity (ANSELIN and BERA 1998). The most common neighborhood modeling method is the approach regarding the shared boundary of objects as the criterion of proximity (KOPCZEWSKA 2011). However, in the case of real estate properties which are the object of transactions, this is not appropriate because these objects relatively rarely neighbor each other (assuming a first-order neighborhood is intended). It seems more suitable to assume that these weights should be based on the criterion of distance. In the simplest case, one can create a matrix of neighbors within a *d* km radius with binary elements (w_{ij} = 1 when object *j* is *d* or less away from object *i* and in the remaining cases w_{ij} = 0). If we assume that distance is one of the key elements in real estate similarity, then we can use Euclidean measures applied for the coordinates of the centers of the analyzed real estate properties (particularly plots). In this case, the inverse distance or the square of the inverse distance can be used. Another solution is to base the weight matrix on the criterion of *k* nearest neighbors resulting from Delaunay's triangulation. Ligas (2006) proposes, moreover, to take into consideration the similarity of real estate properties with respect to their market characteristics, including the respective component in the matrix of the spatial structure.

Spatial autocorrelation measures may have both a global (determining the strength and character of spatial autocorrelation for the whole set of units) and local character. The most often used global spatial autocorrelation measures are I statistics, expressed by the following formula:

$$I = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{1}$$

where:

w_{ij} - weight of connections between unit i and j,
x_i, x_j - values of variables in the spatial unit i and j,
$\bar{x}$ - arithmetic mean of variable value for all units.

The value of Moran I statistics usually ranges from -1 to 1. The 0 value means a lack of spatial autocorrelation, negative values mean the occurrence of a varying level of the studied phenomenon in the neighborhood and positive values mean positive autocorrelation, i.e., the occurrence of a similar level of the studied phenomenon in the neighborhood.

The statistical significance of Moran I statistics is verified by normalized Z_I statistics, with a normal distribution at the expected value of 0 and variance of 1:

$$Z_I = \frac{I - \mathrm{E}(I)}{\sqrt{\mathrm{Var}(I)}} \tag{2}$$

using the following approximations of the expected value and variance (CLIFF, ORD 1973):

$$\mathrm{E}(I) = -\frac{1}{n-1} \qquad \mathrm{Var}(I) = \frac{n^2 S_1 - nS_2 + 3S_0^2}{(n^2-1)S_0^2} - \frac{1}{(n-1)^2} \tag{3}$$

where:

$$S_0 = \sum_{i=1}^{n} n \sum_{j=1} w_{ij} \ , \quad S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(w_{ij} + w_{ji}\right)^2 \ , \quad S_2 = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ji}\right)^2 \tag{4}$$

In the real estate market, it can be intuitively assumed that the spatial autocorrelation phenomenon should be clearly present. However, transaction prices are of course influenced by factors not connected with the location or connected with it indirectly (e.g., available utilities). The effect of these factors on transaction prices can be determined using statistical models and the use of the spatial autocorrelation phenomenon can improve their quality to a large degree.

2. Spatial regression models

The classical linear multiple regression model described extensively in literature (e.g., ISAKSON 1998; CZAJA 2001; BARAŃSKA 2002; BENJAMIN et al. 2004; SIRMANS et al. 2004; ADAMCZEWSKI 2006; BITNER 2007; SAWIŁOW 2010) is often used for the analysis of relations and dependencies between variables characterizing the real estate market. However, multiple regression models do not include spatial correlations between variables, hence both model parameters and the residual are not a function of the location. The solution to this problem can be the use of spatial regression models, including the effect of the location on the response variable.

Interdependencies or interactions in the real estate market can concern the response variable and explanatory variables, as well as a random component (SUCHECKI 2010). When these interactions concern an endogenous variable, spatial autoregression is involved. This means that the value of this variable from other locations influences the value of this variable at the analyzed location. If interactions concern the random component, the phenomenon of spatial autocorrelation of the random model component occurs. Of course, these interactions can also concern explanatory variables, i.e., a situation when the level of the explanatory variable from a given location is influenced by the values of exogenous variables from other locations is a case of cross spatial regression. Depending on the type of spatial interactions, two basic spatial regression models are most often used: the spatial lag model and the spatial error model (ANSELIN 1988; WILHELMSSON 2002; PÁEZ and SCOTT 2004; ARBIA 2006).

The general form of models including a spatial lag, also called spatial autoregressive models (SAR), is as follows (ANSELIN 1999; WALL 2004; ARBIA 2006):

$$y = \rho W y + X\beta + \varepsilon \tag{5}$$

where:

X - matrix of explanatory variables,

B - vector of coefficients (model parameters),

$\varepsilon \sim N(0,\sigma^2 I)$ – is the vector of model errors.

Wy is called the spatial lag response variable. The coefficient ρ is the spatial autocorrelation coefficient. If spatial autocorrelation does not occur ($\rho = 0$), then we obtain the classical linear multiple regression model.

If residuals in regression models are spatially correlated, the application of spatial models can improve the accuracy of the parameter's estimation. Some variables with a spatial character can then be included in the model and the effect of other variables, which cannot be included in the model, will be expressed in the form of residuals (OSLAND 2010). This, consequently, means that the global autocorrelation dependency will be included in the form of model errors. The general form of the spatial error model is as follows (ANSELIN 2003):

$$y = X\beta + \varepsilon \tag{6}$$

$$\varepsilon = \lambda W \varepsilon + \xi \tag{7}$$

where λ is the spatial autocorrelation coefficient and $W\varepsilon$ is the spatial lag error, which should be interpreted as the mean error from neighboring locations (KOPCZEWSKA 2006), and ξ is the independent model error.

Before starting spatial analyses, we do not always have knowledge of the spatial dependencies which should be included in this analysis. The effect of the spatial factor is difficult to measure and its representation in the form of a weight matrix depends largely on the arbitrary selection of the type of neighborhood criterion and the distance function. Before starting to test the functional form of the model, the Moran I test is used to find whether the residuals in the classical linear regression model show spatial autocorrelation. The aim of testing is to determine to which of two classes the model belongs: the class of spatial error models or spatial lag models. The tests require the estimation of an ordinary linear multiple regression model included in the null hypothesis. The alternative hypothesis is one of two forms of spatial models, established depending on the performed test. Two Lagrange multiplier tests are used to determine the type of spatial autocorrelation. They help to determine which case of autocorrelation occurs in a given model (ANSELIN 1988). The Lagrange multiplier test with the alternative hypothesis, including the spatial error model, has the χ^2 distribution with one degree of freedom. The test statistics have the following form (BURRIDGE 1980):

$$LM_{err} = \frac{1}{T_1}\left(\frac{e^T We}{\sigma^2}\right), \qquad \text{where } T_1 = tr\left(W^T W + W^2\right) \tag{8}$$

The Lagrange multiplier test with the alternative hypothesis assuming the functional form of the spatial lag model, is based on testing statistics with an χ^2 distribution with one degree of freedom (Anselin 1988):

$$LM_{lag} = \frac{1}{T_2}\left(\frac{e^T Wy}{\sigma^2}\right), \quad \text{where } T_2 = T_1 + \frac{(WX\beta)^T M(WX\beta)}{\sigma^2} \tag{9}$$

If both tests prove insignificant, spatial autocorrelation has not occurred and the results of linear regression should be assumed. If only one of these tests is significant, the spatial regression model should be selected depending on the type of autocorrelation. The situation may become complicated if both of these tests are significant. Test versions which are robust to bad model specifications (*Robust Lagrange Multiplier*) should then be used. The robust-to-spatial-displacement test and robust-to-spatial-error test can then be compared. The appropriate regression model is selected on the basis of their significance. If both robust tests are statistically significant, such a model should be chosen for which the significance is higher (ANSELIN and BERA 1998; JANC 2007). The problem of no clear indication of the form of the model by tests often remains unsolved also in this case.

The use of the least squares method to estimate spatial regression models will cause biased and inconsistent estimators to be obtained (ANSELIN 1999), hence the most frequently used method for these models is the maximum likelihood method. This method consists of determining such model parameters for which the likelihood function reaches its maximum.

For an N-element sample taken from a population with a continuous probability distribution, the likelihood function L has the form:

$$L(y,\beta) = \prod_{i=1}^{N} f(y_i \mid x_i, \beta) \tag{10}$$

The problem can be reduced to determining the maximum of the function which is the logarithm of the likelihood function (KUSZEWSKI 2007):

$$l(y,\beta) = \sum_{i=1}^{N} \ln\left[f(y_i \mid x_i, \beta)\right] \tag{11}$$

If the random component comes from an independent normal distribution with constant variance σ^2 and zero expected value, the logarithm of the likelihood function can be written as:

$$l\left(y \mid X, \beta, \sigma^2\right) = -\frac{N}{2}\ln\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}(y - X\beta)^T (y - X\beta) \tag{12}$$

For spatial error models, the random component has a multi-dimensional normal distribution with an expected value equal to zero and a variance-covariance matrix with the form:

$$V = \sigma^2\left[(I - \lambda W)^T (I - \lambda W)\right]^{-1} \tag{13}$$

The logarithm of the likelihood function for this type of models will be (MAGNUS 1978, BREUSCH 1980, ANSELIN 1999):

$$\ln L = -\frac{N}{2}\ln\left(2\pi\sigma^2\right) + \ln\left[\det(I - \lambda W)\right] - \frac{1}{2\sigma^2}(y - X\beta)^T (I - \lambda W)^T (I - \lambda W)(y - X\beta) \tag{14}$$

For spatial lag models, the maximum likelihood method is based on the same assumptions as the above. The logarithm of the likelihood function is (MAGNUS 1978, BREUSCH 1980, ANSELIN 1999):

$$\ln L = -\frac{N}{2}\ln\left(2\pi\sigma^2\right) + \ln\left[\det(I - \rho W)\right] - \frac{1}{2\sigma^2}(y - \rho Wy - X\beta)^T (y - \rho Wy - X\beta) \tag{15}$$

Detailed discussion on estimation of spatial regression models is presented in the extensive literature of the subject (among others, MAGNUS 1978; BREUSCH 1980, KELEJIAN and PRUCHA 1998; ANSELIN 1999; LIGAS 2006; SUCHECKI 2010; DEBARSY 2012).

3. Source data and the course of research

Research on the possibility of using the spatial autocorrelation phenomenon to construct regression models of transaction prices was carried out on the basis of the transaction prices of land properties designated for housing construction. These transactions were conducted within the city of Olsztyn in the years 2009-2011. The data on the transactions were obtained from the register of prices and values kept by the Department of Geodesy and Real Estate Management at the Olsztyn City Office. Before starting analyses, the data were selected and verified, removing data on transactions where the prices significantly diverged from the average value as well as data which indicated a non-market character of the transaction from the set used for analyses. Ultimately, data on 251 transactions were used.

Unit price was adopted as the response variable. The explanatory variables were the specific function (low- or high-intensity development), form of possession (ownership or perpetual usufruct), intensity of surrounding development (determined using the kernel density estimator), utilities, and geometric characteristics of the plot (shape and surface area). It should be added that the selection of explanatory variables and their measurement scales was the effect of thorough statistical analyses involving the construction of many experimental regression models. In the end, the author focused only on those variables whose effect on prices proved statistically significant.

A linear multiple regression model was constructed during the research and then, Moran I statistics were determined for both transaction prices and residuals from regression. Spatial weight matrices constructed on the basis of three criteria: k-nearest neighbors, a particular distance radius and inverse distance were used for that purpose. These matrices then served to construct spatial regression models in which the price depended not only on the adopted real property attributes, but also on prices or spatial lag residuals from regression. Six such models were constructed altogether.

If models estimated by the least squares method must be compared with spatial error and spatial lag models, not the coefficient of determination, but the values derived by maximizing the log-likelihood (LIK) function and information criteria: AIC (Akaike Information Criterion) or BIC (Bayesian Information Criterion) (ACQUAH 2010) are usually used. The respective formulas for the information criteria are as follows:

$$AIC = -\frac{2\ln L}{N} + \frac{2K}{N}, \qquad BIC = -\frac{2\ln L}{N} + \frac{K\log(N)}{N} \tag{18}$$

where K is the number of parameters in the model and N is the number of observations.

On the basis of the above criteria, individual models were evaluated and it was indicated which of them was best matched to the market data. R software was applied during the analyses using the package “spdep” and “sp”. GeoDa software was also used accessorily.

4. Research results

Estimation results of the linear multiple regression model by the least squares method are presented in Table 1. The scales of explanatory variable measurement were selected in such a way that the zero values corresponded to a typical real estate property. This allows the constant in the model to be interpreted as the average price of real estate properties with typical attributes.

Table 1

Estimation results of the linear multiple regression model

	Estimate	Std. error	t value	p-value
Intercept	255.601	6.647	38.454	0.0000
Designation	52.551	22.878	2.297	0.0225
Possession form	-72.881	24.720	2.948	0.0035
Intensity	182.395	35.293	5.168	0.0000
Public utility	-94.571	9.147	-10.339	0.0000
Geometry	-47.862	18.883	-2.535	0.0119

Source: own research.

The coefficient of determination R^2 was 0.536 and the standard estimation error 48.49. The multiple regression model was verified positively, both as to content and statistically, and proved quite well-matched to the data, although the location-related attribute was not included. However, it can be assumed that location is an important factor influencing prices, especially as the research area covered the whole city. This effect can manifest itself in the form of spatial interactions, expressed by spatial autocorrelation.

The study of spatial autocorrelation covered both transaction prices and residuals from the multiple regression model. Moran I statistics computed according to the formula (1) were used during the research. The method of selecting spatial weights, which can determine the results of analyses, deserves special attention. Of course, there are different methods for including neighborhood, with three selected methods adopted for the needs of the research:

1) neighborhood on the basis of k nearest neighbors,
2) neighborhood defined by the distance radius,
3) neighborhood defined as inverse distance.

Using the first method, it can be supposed that spatial autocorrelation decreases with an increasing number of nearest neighbors. A matrix of weights constructed on the basis of five nearest neighbors (real estate properties which were the subject of transactions) was used for the research. In the second method, a distance radius equal to 2,300 m was adopted, which is at the same time the minimum distance at which each transaction had at least one neighbor. Using the third method, it was assumed that the basis for the evaluation of neighborhood would be the inverse of distance determined on the basis of the coordinates of plot centroids. The results of spatial autocorrelation analysis are presented in Table 2.

Table 2

Value of Moran I statistics for prices and residuals from regression using different spatial weight matrices

Weights	Observation	Moran I	Z_I	p-value
nearest neighb.: k = 5	prices	0.593	16.257	0.0000
nearest neighb.: k = 5	residuals	0.378	10.396	0.0000
distance: d = 2300 m	prices	0.159	12.859	0.0000
distance: d = 2300 m	residuals	0.072	6.001	0.0000
inverse distance	prices	0.363	19.137	0.0000
inverse distance	residuals	0.269	14.226	0.0000

Source: own research.

Global spatial autocorrelation measured by Moran I statistics proved statistically significant in all the analyzed cases, with the highest autocorrelation for transaction prices, where the spatial weights were determined on the basis of the five nearest neighbors.

A dot diagram, which has the analyzed variable plotted on the x-axis and a spatial lag variable on the y-axis, may serve as a graphical interpretation of global Moran I statistics, (Figure 1). The slopes of the regression lines presented in Figure 1 are at the same time the value of Moran I statistics.

The occurrence of spatial autocorrelation both among transaction prices and residuals from the multiple regression model indicates the possibility of using spatial regression models. The selection of the model form (spatial lag or spatial error model) can be made with the application of the LM

(Lagrange Multiplier) test using formulas (8) and (9). The R-LM (robust LM) test was also used accessorily during the research. This test can help to select the appropriate model form if both LM tests prove statistically significant (table 2).

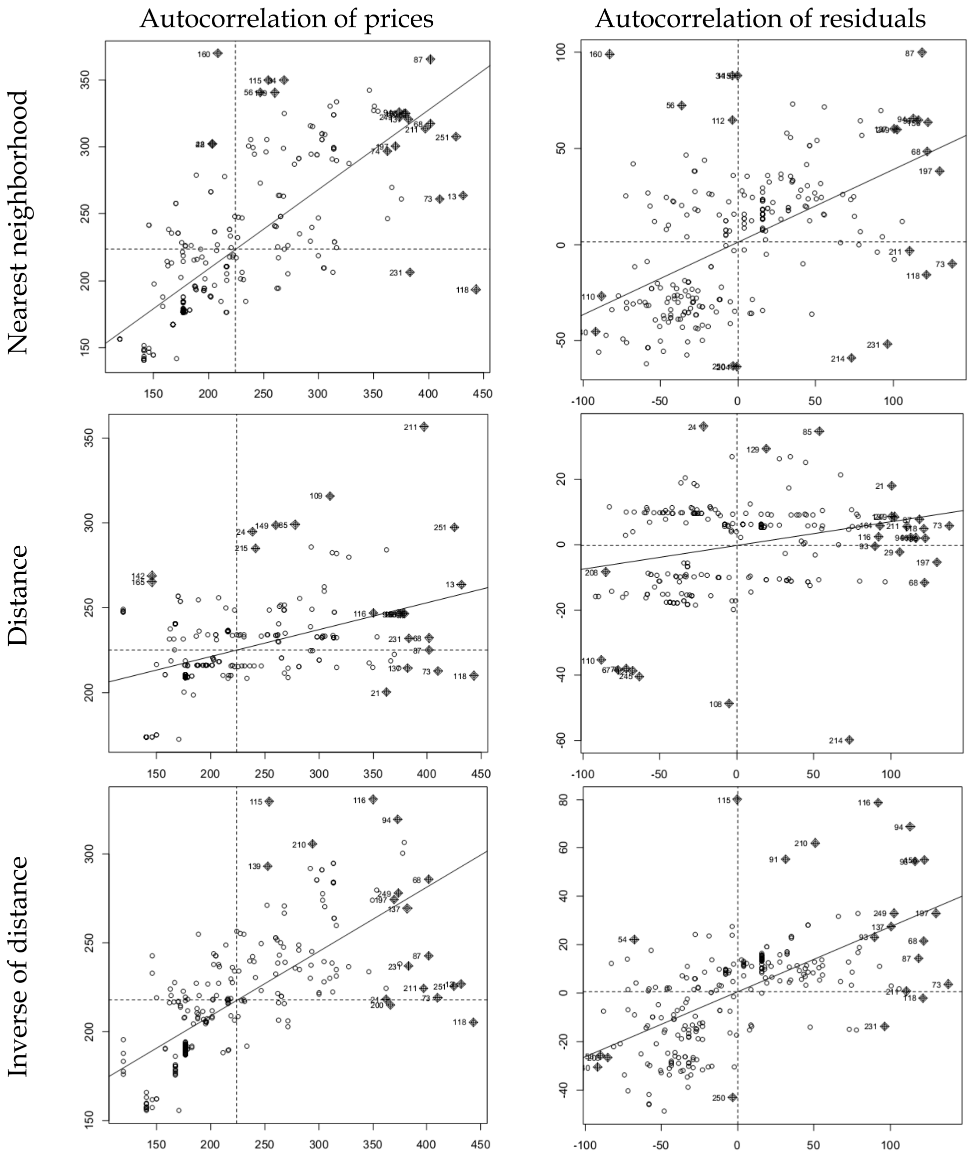

Fig. 1. Moran's dot diagrams for prices and residuals from regression created using different spatial weight matrices, *Source*: own research.

If LM_{err} is more significant than LM_{lag}, and $R\text{-}LM_{err}$ is statistically significant while $R\text{-}LM_{lag}$ is insignificant, then the spatial error model is more suitable. Otherwise, the spatial lag model is the more appropriate model. The spatial error model proves to be the more suitable model in the analyzed cases. In the spatial lag model, spatial interactions refer directly to transaction prices. In the spatial error model, it is assumed that these interactions refer to residuals, understood in this case as prices, in which the effect of non-spatial factors (explanatory variables in the model) has already been included. The hypothesis that, regardless of the adopted method of determining neighborhood in the spatial weights matrix, it is the spatial error model that better reflects the dependencies prevailing in the real estate market can therefore be adopted.

The parameters of individual models were estimated independently of the LM test results, at the same time determining the criteria for their evaluation: LIK, AIC and BIC. In regression models, the parameters of which are estimated by the least squares method, one of the basic evaluation criteria is the coefficient of determination R^2, which was also used during the research. Table 3 presents the evaluation results for the individual models.

Table 2

Results of the LM and R-LM test to select the appropriate model form

Weights	LM_{err} (p-value)	LM_{lag} (p-value)	$R\text{-}LM_{err}$ (p-value)	$R\text{-}LM_{lag}$ (p-value)
nearest neighborhood: k = 5	103.098 (0.000)	87.719 (0.000)	17.173 (0.000)	1.794 (0.180)
distance: d = 2300 m	26.889 (0.000)	5.032 (0.025)	21.860 (0.000)	0.002 (0.961)
inverse distance	180.078 (0.000)	100.140 (0.000)	79.990 (0.000)	0.052 (0.820)

Source: own research.

Table 3

Evaluation results for the analyzed spatial regression models

No.	Weights	Model	LIK	AIC	BIC	R^2
1	multiple regression		-1327.3	2668.7	2693.3	0.536
2	nearest neighb. k = 5	spatial lag	-1298.1	2612.1	2640.3	0.650
3		spatial error	-1294.0	2603.9	2632.1	0.668
4	distance d = 2300 m	spatial lag	-1324.4	2664.8	2693.0	0.548
5		spatial error	-1319.4	2654.8	2683.0	0.573
6	inverse distance	spatial lag	-1296.6	2609.2	2637.4	0.652
7		spatial error	-1285.7	2587.5	2615.7	0.690

Source: own research.

According to the LIK criterion, the spatial error model, in which spatial weights are determined on the basis of inverse distance, proves to be the best model. On the other hand, AIC and BIC information criteria indicate the classical linear multiple regression model. It should be noted, however, that differences between the values of the adopted criteria are small. The value of the coefficient of determination, similar to the criterion of the logarithm of the likelihood function, indicates that the last model is best matched to the data. The estimation results of the parameters of this model are presented in Table 4.

Table 4

Estimation results of the spatial error model using spatial weights as inverse distance

$\lambda = 0.917$				
	Estimate	Std. error	t value	p-value
Intercept	252.868	30.445	8.3057	0.0000
Designation	47.296	20.263	2.3341	0.0196
Possession form	-67.129	23.881	-2.8110	0.0049
Intensity	158.168	31.840	4.9676	0.0000
Public utility	-103.009	12.656	-8.1393	0.0000
Geometry	-35.646	17.539	-2.0323	0.0421

Source: own research.

The standard estimation error was 39.18 in this case, which is a much lower value than for the classical multiple regression model. This means that results which describe more accurately dependencies in the real estate market can be obtained by including spatial autocorrelation in the regression model.

5. Summary and conclusions

As a rule, multiple regression models applied for real estate market analysis and the determination of market value include location only indirectly. In such cases, location is expressed on an interval scale, though it should be noted that it is difficult to objectively fully evaluate the advantages and disadvantages of the location quantitatively. With the assumption that spatial interactions between transaction prices, i.e., the spatial autocorrelation phenomenon, occur in the real estate market, location can be included directly in regression models. As demonstrated by the conducted research, spatial regression models clearly show better matching to the market data than classical multiple regression models. The conducted tests also demonstrated that it is more justified to use error models than spatial lag models.

Besides the advantages of spatial regression models which are manifested in them matching the data better, these models also have important disadvantages. One of them is the difficulty in using them for the prediction of the response variable because its spatial lags are not known in this case. Prediction is, moreover, burdened with great error, which results from the structure of the models (HAINING 1990, KOPCZEWSKA 2011). Hence, such models can serve, above all, for diagnostics of processes occurring in space, which in this case consist of the quantitative effect of individual factors on transaction prices. They can, therefore, be a useful tool in the analysis of transaction prices on the real estate market.

6. Bibliography

ACQUAH H., 2010, *Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of an asymmetric price relationship*, Journal of Development and Agricultural Economics Vol. 2(1), pp. 1-6.

ADAMCZEWSKI Z., 2006, *Elementy modelowania matematycznego w wycenie nieruchomości. Podejście porównawcze*, Oficyna Wydawnicza Politechniki Warszawskiej.

ANSELIN L., 1988, *Lagrange multiplier test diagnostics for spatial dependence and spatial heterogenity*, Geographical Analysis 20, pp. 1-17.

ANSELIN L., 1999, *Spatial Econometrics*, Bruton Center School of Social Sciences, University of Texas at Dallas.

ANSELIN L., 2003, *Spatial externalities and spatial econometrics*. International Regional Science Review, vol. 26, no 2, pp. 153-166.

ANSELIN L., BERA A., 1998, *Spatial dependence in Linear Regression Models with an Introduction to Spatial Econometrics*, (in:) Ullah A., Giles D. (Eds.), Handbook of Applied Economic Statistic, Marcel Dekker, New York, pp. 237-289.

ARBIA G., 2006, *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*, Springer-Verlag, Berlin.

BARAŃSKA A., 2002, *Kryteria stosowania modeli stochastycznych w predykcji rynkowej wartości nieruchomości*, Zastosowania Metod Statystycznych w Badaniach Naukowych, www.statsoft.pl/czytelnia.

BASU S., THIBODEAU T., 1998, *Analysis of spatial autocorrelation in house prices*, Journal of Real Estate Finance and Economics, vol. 17:1, pp. 61-85.

BENJAMIN J. D., RANDALL S. GUTTERY R. S., SIRMANS C. F., 2004, *Mass Appraisal: An Introduction to Multiple Regression Analysis for Real Estate Valuation*, Journal of Real Estate Practice and Education, vol. 7, no 1, pp. 65-77.

BITNER A., 2007, *Konstrukcja modelu regresji wielorakiej przy wycenie nieruchomości*, Acta Scientiarum Polonorum, Administratio Locorum, No 6(4), 59-66.

BREUSCH T., 1980, *Useful invariance results for generalized regression models*, Journal of Econometrics 13, pp. 327–340.

BURRIDGE P., 1980, *On the Cliff-Ord test for spatial autocorrelation*, Journal of the Royal Statistical Society B 42, pp. 107-108.

CZAJA J., 2001, *Metody szacowania wartości rynkowej i katastralnej*, Komp-System, Kraków.

DEBARSY N., 2012, *The Mundlak Approach in the Spatial Durbin Panel Data Model*, Spatial Economic Analysis, vol. 7, no 1, pp. 109-131.

HAINING R., 1990, *Spatial data analysis in the social and environmental sciences*, Cambridge University Press.

ISAKSON H. R., 1998, *The Review of Real Estate Appraisals Using Multiple Regression Analysis*, Journal of Real Estate Research, vol. 15, issue 2, pp. 177-190.

JANC K., 2006, *Zjawisko autokorelacji przestrzennej na przykładzie statystyki I Morana oraz lokalnych wskaźników zależności przestrzennej (LISA) – wybrane zagadnienia metodyczne*, (w:) Komornicki T., Podgórski Z., (red.) Idee i praktyczny uniwersalizm geografii, Dokumentacja Geograficzna nr 33, IGiPZ PAN, Warszawa.

KELEIJAN H. H., PRUCHA I. R., 1998, *A generalized spatial two stage least squares procedure for estimating a spatial autoregressive models with autoregressive disturbances*, Journal of Real Estate Finance and Economics nr 17, pp. 99-121.

KOPCZEWSKA K., 2011, *Ekonometria i statystyka przestrzenna z wykorzystaniem programu R CRAN,* CeDeWu, Warszawa.

KUSZEWSKI W., 2007, *Jednorównaniowy model ekonometryczny – estymacja*, [w:] Gruszczyński M., Podgórska M. (red.), *Ekonometria*, Szkoła Główna Handlowa w Warszawie – Oficyna Wydawnicza, Warszawa, s. 33-34.

LESAGE J., KELLY PACE R., 2009, *Introduction to Spatial Econometrics*, Regional Research Institute, CRC Press/Taylor & Francis Group.

LIGAS M., 2006, *Przestrzenne modele autoregresji w zastosowaniu do wyceny nieruchomości,* Studia i Materiały Towarzystwa Naukowego Nieruchomości, vol. 14, nr 1, str. 123-136.

MAGNUS J., 1978, *Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix*. Journal of Econometrics 7, pp. 281–312.

OSLAND L., 2010, *An Application of Spatial Econometrics in Relation to Hedonic House Price Modelling*, Journal of Real Estate Research, vol. 32, issue 3, pp. 289-320.

PÁEZ A., SCOTT D. M., 2004, *Spatial statistics for urban analysis: A review of techniques with examples*, GeoJournal, *vol.* 61, pp. 53–67.

SAWIŁOW E., 2010, *Problematyka określania wartości nieruchomości metodą analizy statystycznej rynku*, Studia i Materiały Towarzystwa Naukowego Nieruchomości, Vol 8, No 1, s. 21-32.

SIRMANS G. S., MACPHERSON D. A., ZIETZ E. N., 2005, *The Composition of Hedonic Pricing Models*, Journal of Real Estate Literature, vol. 13, no 1, pp. 3-46.

SUCHECKI B. (red.), 2010, *Ekonometria przestrzenna. Metody i modele analizy danych przestrzennych*, C.H. Beck, Warszawa.

WALL M.M., 2004, *A close look at the spatial structure implied by the CAR and SAR models*. Journal of Statistical Planning and Inference, pp. 311-324.

WILHELMSSON, M., 2002, *Spatial Model in Real Estate Economics*. Housing Theory and Society, vol. 19, pp. 92-101.