RESIDENTIAL REAL ESTATE PRICE MODELLING THROUGH THE METHOD OF THE GEOGRAPHICALLY WEIGHTED REGRESSION : GOMEL CITY CASE STUDY

Natalia Zhukovskaya, Olga Popko

Belarussian State University

Faculty of Geography and Geoinformatics

Contact: natazhuk@gmail.com, zhukousk@bsu.by

Abstract

One of the most challenging tasks in modelling of house pricing is to take into account the location factors. Geographically Weighted Regression (GWR) as a local regression model is an extremely effective instrument for spatial data analysis. Pricing modelling results in Gomel residential real estate market by the use of Geographically Weighted Regression are presented. The data of the Belarus' National Cadastral Agency on real estate transactions (apartments) in Gomel in 2019 are used as initial. The global Moran I index has been used to estimate a spatial autocorrelation of the dependent variable (price per square meter of residential real estate). Several factors having the impact on the apartment sale prices have been determined. Independent variables having been used in analysis can be divided into building characteristics and spatial characteristics. The building characteristics section includes the number of rooms within the property, property area (square meters), building age, number of floors in the building, floor of the property. The spatial characteristics group contains proximity to the city center, recreation areas, supermarkets, bus stops, healthcare and educational facilities. A regression model of housing price in Gomel has been developed. Mapping variable regression coefficients allows to explore spatial features of the impact of the different explanatory variables on the property price. Geographically weighted regression modelling has revealed the pricing peculiarities inherent for certain areas of the city.

Data and Methods

The real estate transactions (apartments) in Gomel in 2019 has been analyzed. The total number was 1,046 purchase and sale transactions.

Analysis steps:

- Identifying and modelling the spatial factors influence on a price
- Statistical data analysis
- Assessment of spatial autocorrelation
- Conducting Geographically Weighted Regression

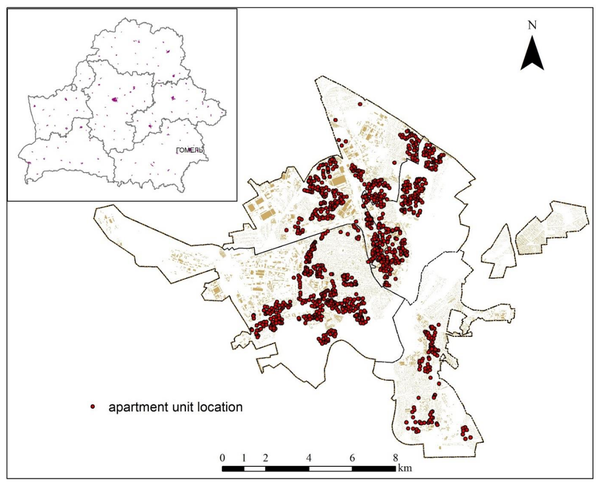

Fig.1. Apartment unit locations

The spatial factors models were constructed in the environment of ArcGIS 10.8 with using Euclidean distance tool. Statistical data analysis includes the variable distribution assessment with the help of histograms and fitting criterions (the Kolmogorov-Smirnov test, the Shapiro-Wilk test), calculating descriptive statistics, nonparametric correlation analysis. Spatial Autocorrelation Analysis was implemented with global and local Moran's I statistics (Anselin, 1995). The Geographically Weighted Regression analysis was used to model the relationship between an apartment price and some building, property and location characteristics.

Results and Discussions

Analysis of purchase and sale transactions in Gomel secondary housing market has showed that apartments prices (measured in USD per square meters) were variable, ranging from 250 to 750 USD, with an average of 544 USD.

Considering the spatial distribution of apartment unit prices (measured in USD per square meters), two types of spatial association are tested: global spatial autocorrelation and local spatial association.

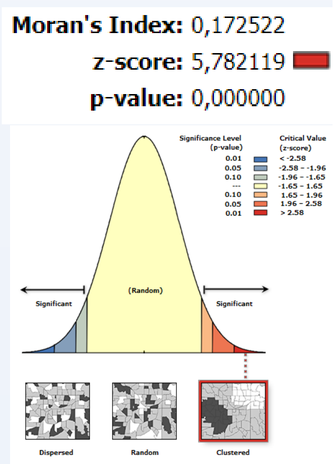

Fig.2. Global Moran's I analysis result

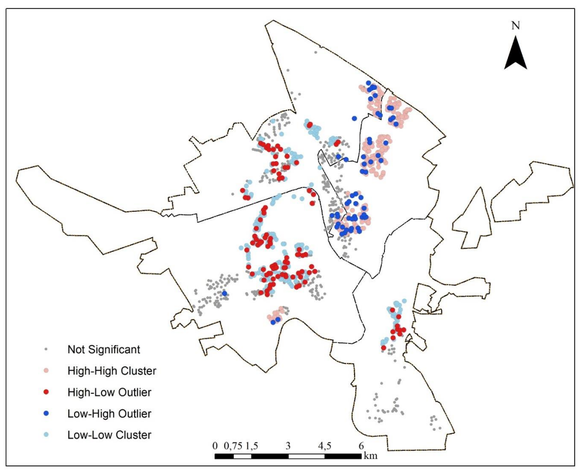

Fig.3. Spatial clusters and spatial outliers map of apartment unit prices in Gomel

Internal factors: property area, building age, number of floors in the building, floor of the property, wall material, number of property rooms.

Spatial factors: proximity to the city center, recreation areas, supermarkets, bus stops, healthcare and educational facilities.

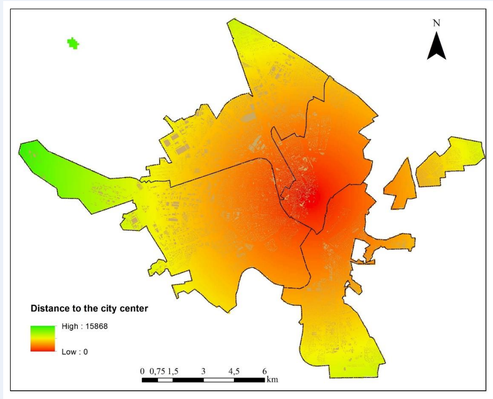

Fig.4. Distance to the city center

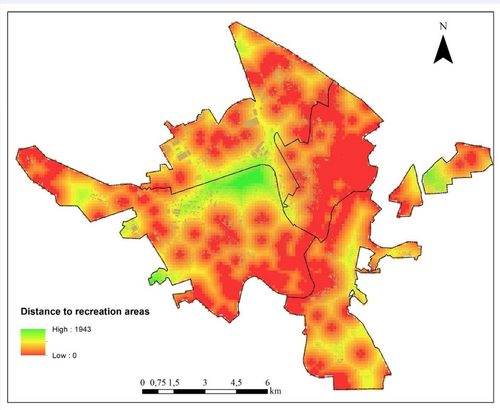

Fig.5. Distance to recreation areas

Variables	r
Property area (square meters)	**0.064**
Number of rooms	-0.039
Building age	**-0.327**
Number of floors in the building	**0.285**
Floor of the property	**0.175**
Wall material	**0.120**
Proximity to the city center	-0.030
Proximity to recreation areas	**-0.141**
Proximity to supermarkets	0.044
Proximity to healthcare facilities	**0.182**
Proximity to educational facilities	0.078
Proximity to bus stops	-0.005
Significant coefficients at the significance level 0.05 are determined in bold	

Fig.6. Spearman correlation coefficients between apartment prices per m² and building and location characteristics

Variables	Significant, %
Building age	100
Proximity to healthcare facilities	100
Proximity to recreation areas	100
Proximity to the city center	92
Proximity to supermarkets	86
Number of floors in the building	88
Wall material	82
Floor of the property	65
Number of rooms	41
Proximity to educational facilities	31
Property area (square meters)	30
Proximity to bus stops	7

Fig.7. Summary of Variable Significance (Exploratory Regression)

Diagnostic content	Value
Number of observations	1046
Bandwidth	945
Residual squares	10099.18
Sigma	111.23
AICc	10.06
R^2	0.77
Adjusted R^2	0.74

Fig.8. Basic diagnostic statistics of the GWR

Variable	Mean (min – max)
Intersept	525 (-982 – 1187)
Property area (square meters)	-0.68 (-2.65 - 1.49)
Building age	-1.00(-7.88 - 4.50)
Number of floors in the building	5.46 (-11.9 - 23.22)
Floor of the property	0.57 (-11.96 – 11.86)
Wall material	7.05(-50.68 - 46.69)
Proximity to the city center	0.001 (-0.094 - 0.198)
Proximity to recreation areas	-0.014(-0.151 - 0.329)
Proximity to supermarkets	0.009 (-0.491 - 0.184)
Proximity to healthcare facilities	0.010 (-0.134 - 0.242)

Fig.9. Results of the GWR model parameter estimation

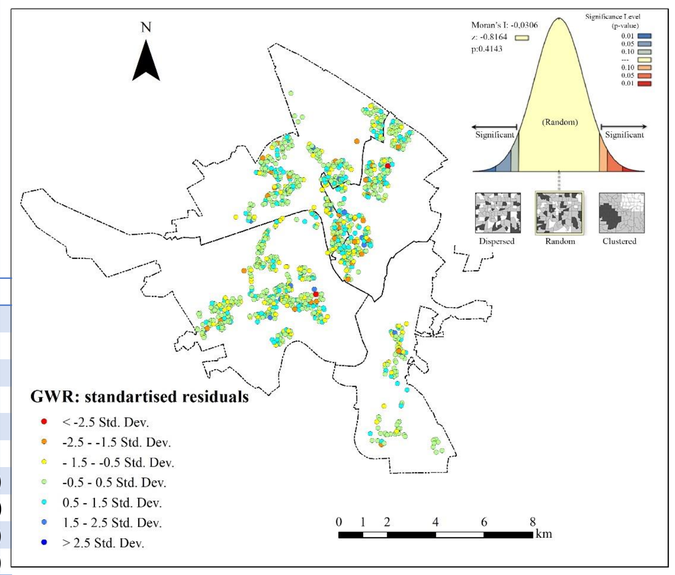

Fig.10. Standard residual of GWR for apartment unit prices in Gomel

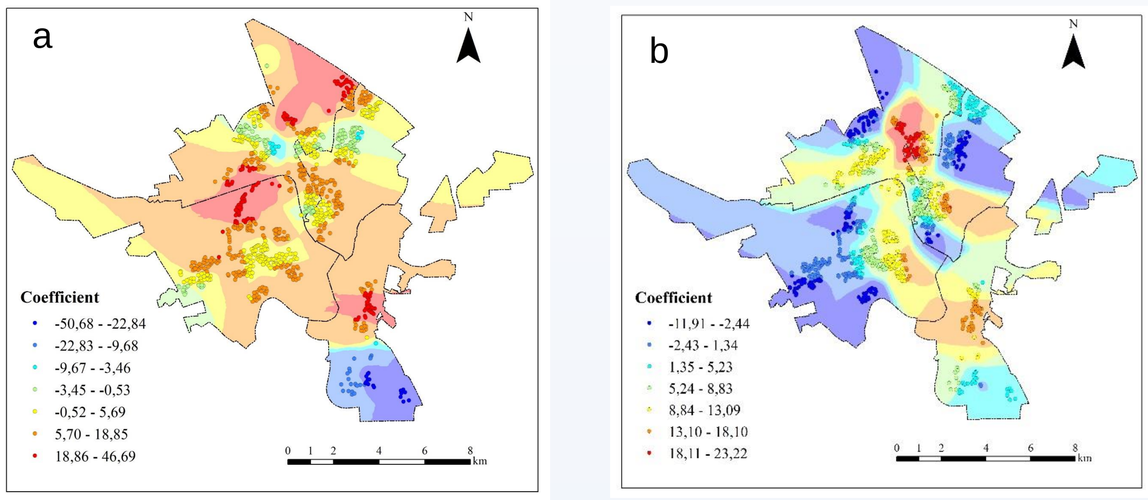

*Fig. 11. The spatial distribution of the GWR coefficients: (**a**) wall material; (**b**) number of floors in the building*

Conclusions

In this study, we investigated the relationships between residential real estate price (per sq.m) and both building and location characteristics for Gomel during 2019 using GWR. Geographically Weighted Regression is local regression technique which takes into account spatial heterogeneity. Several factors having the impact on the apartment sale prices have been determined. The final GWR model includes explanatory variables based on two categories: building and apartment characteristics (property area (square meters), building age, number of floors in the building, floor of the property, wall material) and location characteristics (proximity to city center, recreation areas, supermarkets, healthcare facilities). A regression model of housing price in Gomel has been developed. The main output from GWR is a set of localised parameter estimates and associated diagnostics. The GWR adjusted R^2 has been 0.74 (R^2 is 0.77). AICc has given a value of 10.06. Mapping the GWR model coefficients allows to explore the spatial features of relationships between residential real estate price (per sq.m) and explanatory variables. The spatial autocorrelation of GWR residuals for our model have resulted in a Moran's I value of −0.03 (p = 0.414), indicating a residuals spatial random distribution, which in turn suggests a good GWR fit.

References

1. Anselin L. (1995). Local Indicators of Spatial Association – LISA. Geographical Analysis, Volume 27, pp. 93–115.
2. Cellmer R., Cichulska A. Bełej M. (2020) Spatial Analysis of Housing Prices and Market Activity with the Geographically Weighted Regression. ISPRS International Journal of Geo-Information, Volume 9 (6), doi: 10.3390/ijgi9060380.
3. Cellmer R. (2012). The Use of the Geographically Weighted Regression for the Real Estate Market Analysis. Folia Oeconomica Stetinensia, Volume 11 (1), doi:10.2478/v10031-012-0009-6.
4. Fotheringham A.S., Brunsdon C., Charlton M. (2002). Geographically weighted regression: the analysis of spatially varying relationships. Chichester: Wiley. 269 p.
5. Fotheringham A.S., Charlton M.E., Brunsdon C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environment and Planning A, Volume 30 (11), pp.1905–1927.
6. Wing, Chau Kwong, Chin T. (2003). A Critical Review of Literature on the Hedonic Price Model. International Journal for Housing Science and Its Applications, Volume 27, pp. 145-165.