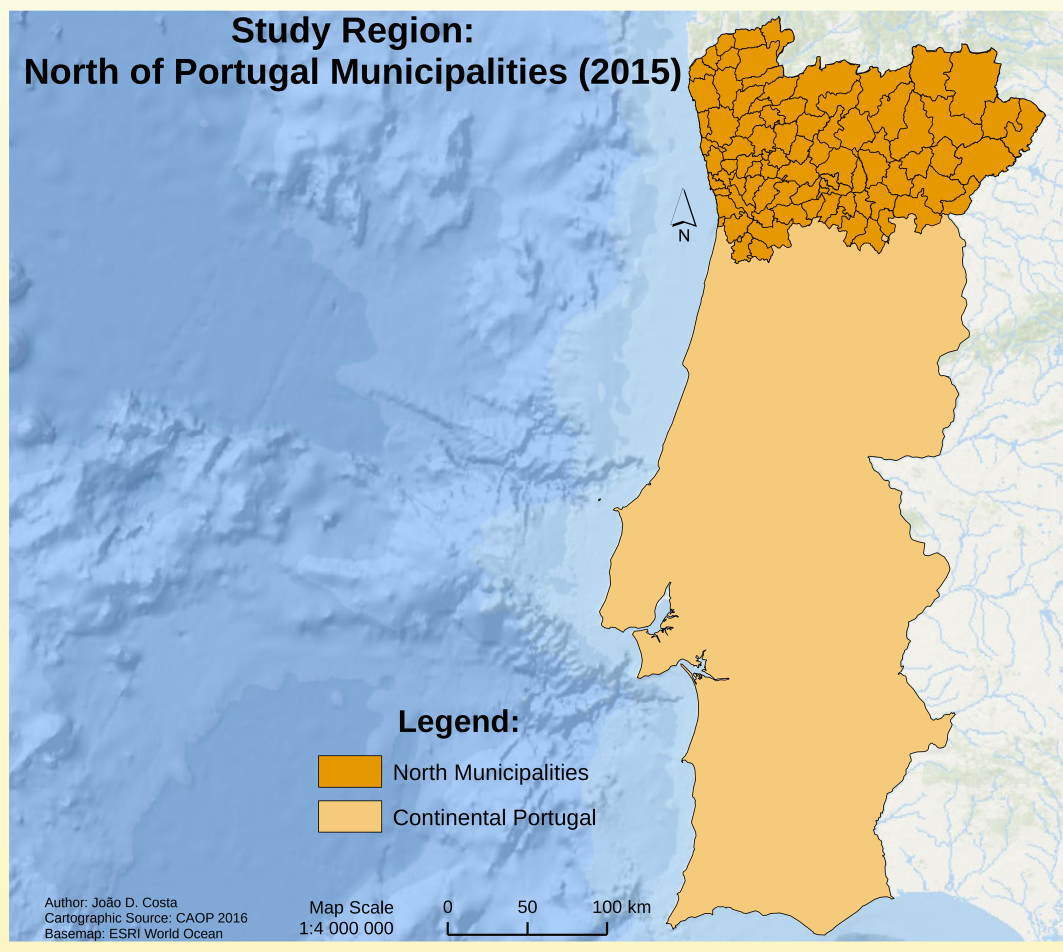
Application of spatial regression to investigate current patterns of crime in the north of Portugal

João Diogo Costa, Ana Cristina Costa
NOVA IMS, Universidade Nova de Lisboa

Motivation

Crime is a phenomenon that accompanies societies from the earliest civilization and has a dynamic presence in time and space. Although individual incidences are unpredictable and difficult to anticipate, geographic studies have shown that crime is often concentrated in clusters (Wang et al. 2013), thus the phenomenon is neither random nor homogeneous in space, especially when considering urban areas (Nezami & Khoramshahi 2016). Researchers recognize the importance of considering the non-stationarity of the spatial process, and so they focus on the study of crime at the local level (Cahill & Mulligan 2007). Even though many geographers are interested in crime research, there are few attempts in the community to support and standardize this issue, and the "geography of violence" is still an emerging field of research (Springer & Le Billon 2016).

This investigation seeks to identify the current patterns of crime in the northern region of mainland Portugal. Additionally, we investigate potential covariates of crime.



Methodology and Data

EXPLORATORY SPATIAL DATA ANALYSIS

- Summary statistics
- Data posting
- Regional histogram
- IDW
- Local Moran's I statistics
- Global Moran's I statistic
- Hotspot analysis
- Scatterplot graphs

OLS MODEL DIAGNOSTICS

- Multicollinearity
- T-tests
- Jarque-Bera test
- Koenker test
- Global Moran's I statistic
- Joint F-test
- Adjusted R-Squared
- Adjusted Akaike's Information Criterion (AICc)

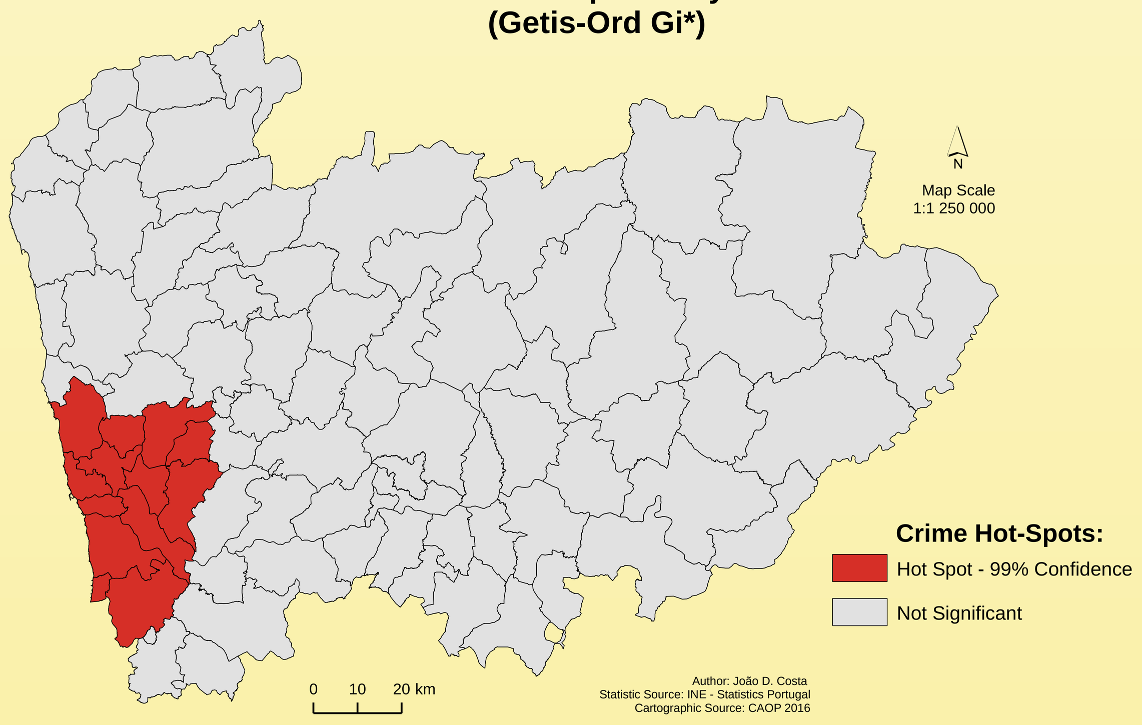
GWR MODEL DIAGNOSTICS

- Local multicollinearity
- Global Moran's I statistic
- Local R-Squared
- Local Standard errors
- Adjusted Akaike's Information Criterion (AICc)

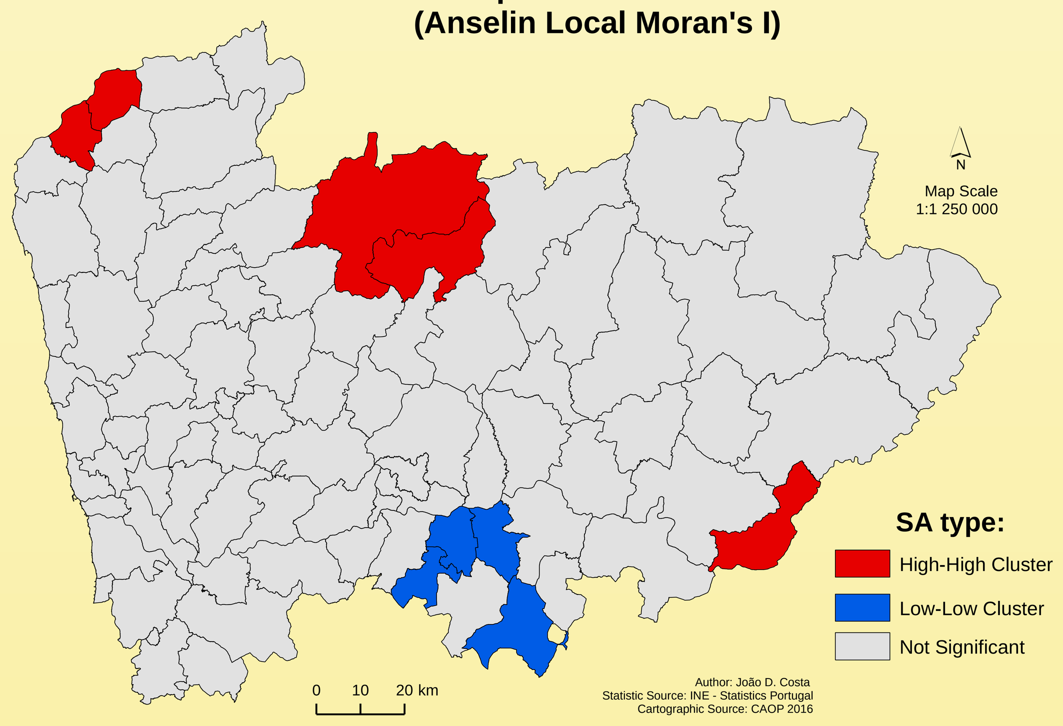
Predictor variables investigated
1.Distance to district capital (meters)
2.Area (meters²)
3.Population density
4.Resident population
5.Youth population (nº of inhabitants aged by 15 to 29)
6.Foreign population
7.Low level of schooling (nº of inhabitants from age 15 who has less than the secondary school level)
8.Job offers (annual average number)
9.Remuneration (monthly average salary in €)
10.Buying power per capita (%)
11.Unemployement rate (%)
12.Social integration income (SII) beneficiaries

Results

Crime hot-spot analysis (Getis-Ord Gi*)



Crime spatial autocorrelation (Anselin Local Moran's I)

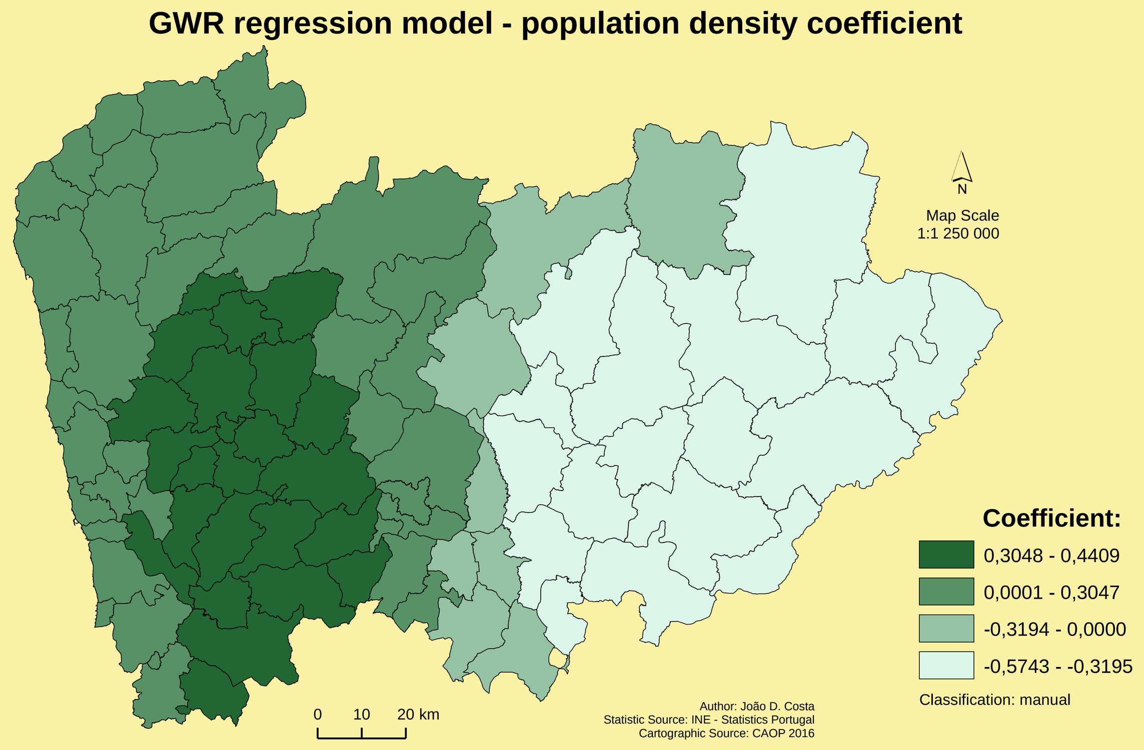


Model variable	Coefficient	t-value
Intercept	4.3857	6.6877
Population Density	0.2575	4.0011*
SII Beneficiaries	0.4540	5.8528*
Distance	–0.0737	–3.0488*

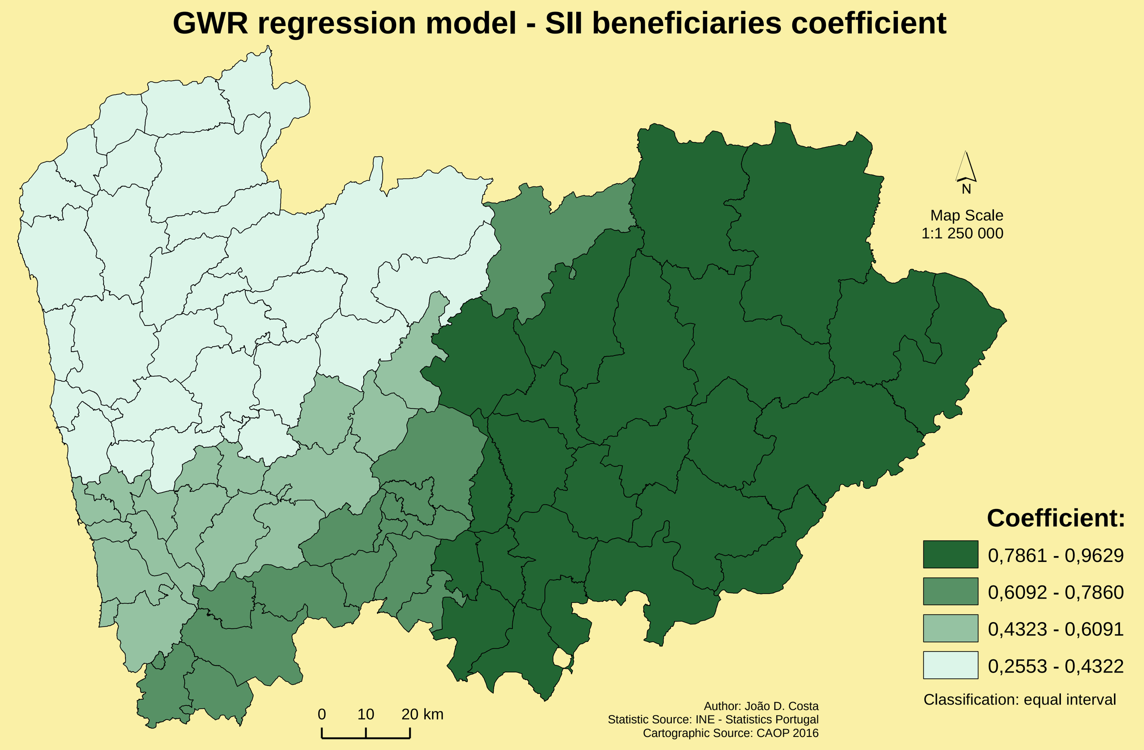
* Significant at 1% level

OLS Diagnostics	Statistic	p-value
Adjusted R-Squared	0.7887	
AICc	206.9527	
Joint-F statistic	106.7370	0.0000
Jarque-Bera statistic	2.3559	0.3079
Koenker statistic	5.4979	0.1388
Global Moran's I statistic	0.5096	0.0000

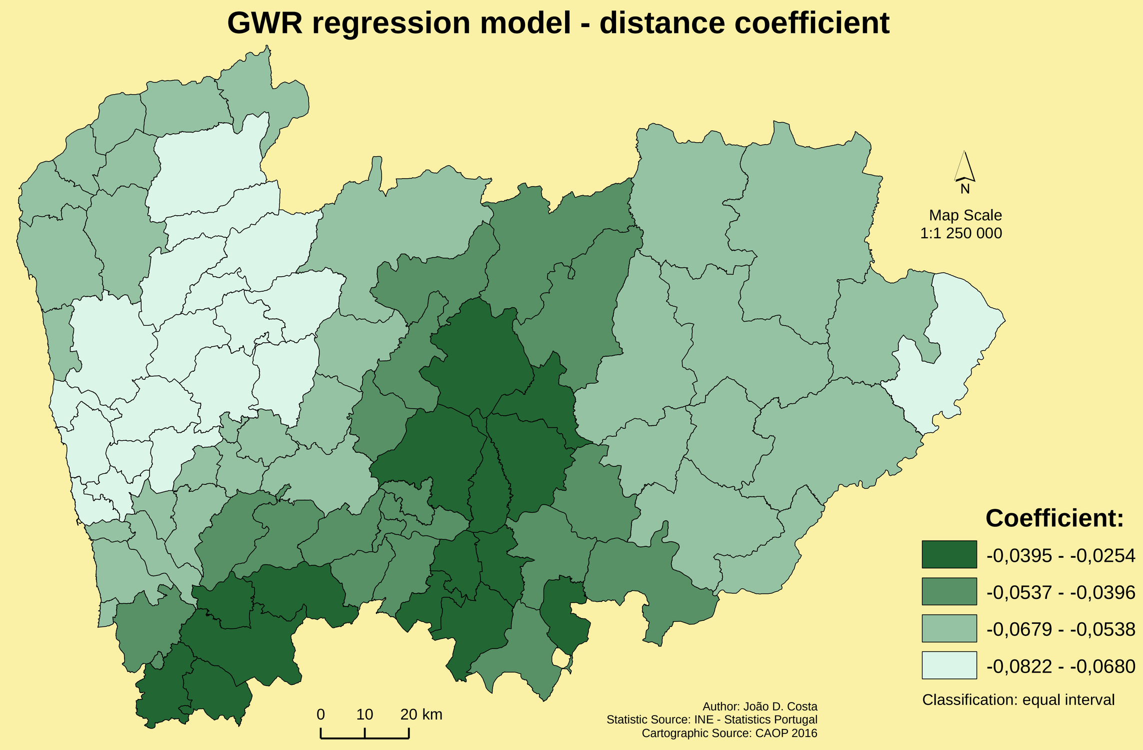
GWR regression model - population density coefficient



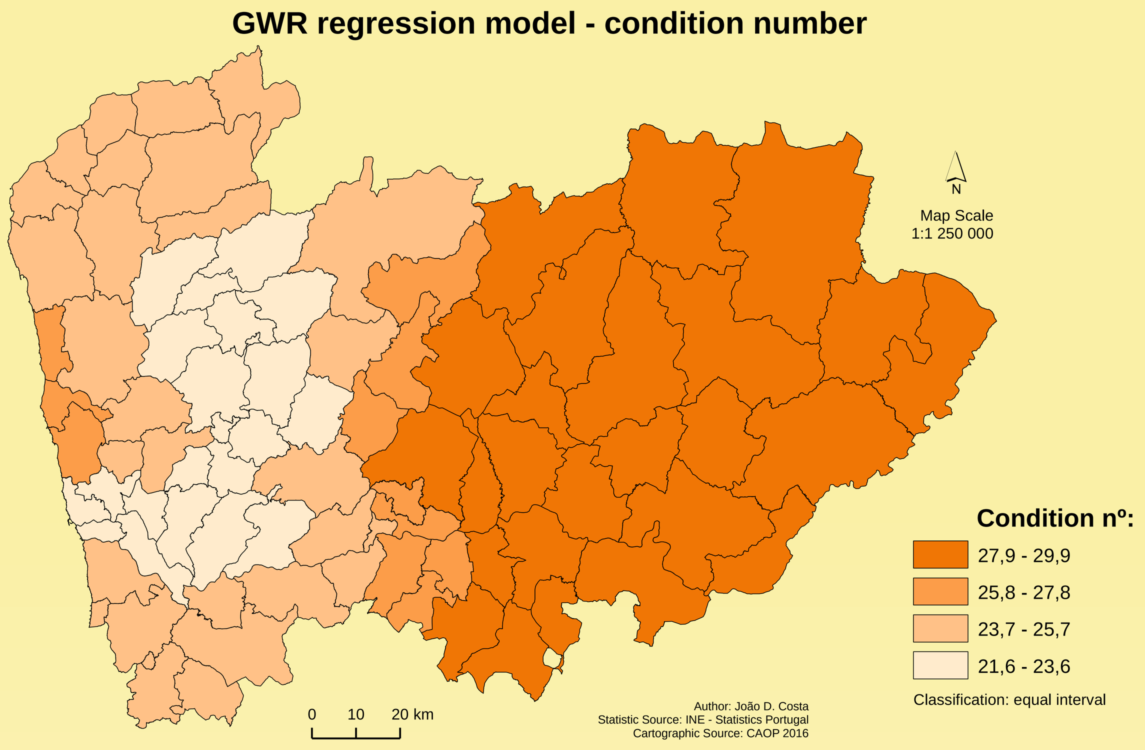
GWR regression model - SII beneficiaries coefficient



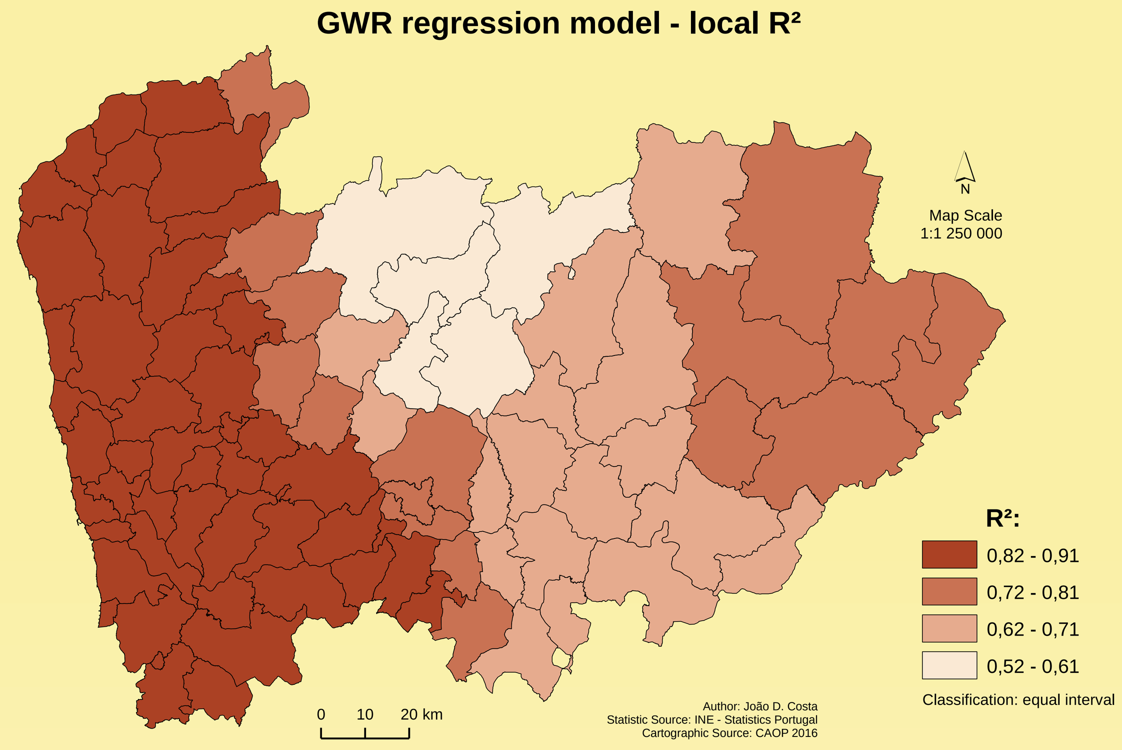
GWR regression model - distance coefficient



GWR regression model - condition number



GWR regression model - local R²



Conclusion

Comparing with the OLS model, the GWR model improved the goodness of fit results: the AICc decreased to 155, and the GWR model now explains 90% of the crime variability (Adjusted R-Squared). Unlike the OLS model, the spatial distribution of the GWR residuals exhibits a random pattern, which was confirmed by a non-significant value of the Global Moran's I statistic.

GWR provided further insights about the regional variation of the explanatory variables: population density has more impact in criminality on the west; distance to the district's capital is more relevant for municipalities located in the centre and south; and the coefficients of the SII beneficiaries have increasing values from west to east.

Even though the GWR model provided good diagnostic results, the interpretation of map patterns for individual coefficients should be done with caution, given the methodological limitations of GWR (e.g., Wheeler 2014).

References

- Cahill, M., & Mulligan, G. (2007). Using geographically weighted regression to explore local crime patterns. *Social Science Computer Review*, 25(2), 174–193.
- Nezami, S., & Khoramshahi, E. (2016). Spatial Modeling of Crime by Using of GWR Method. In: *2016 Baltic Geodetic Congress (BGC Geomatics)*, Gdansk, June 2016, pp. 222– 227. doi: 10.1109/BGC.Geomatics.2016.47.
- Springer, S., & Le Billon, P. (2016). Violence and space: An introduction to the geographies of violence. *Political Geography*, 52, 1–3.
- Wang, D., Ding, W., Lo, H., Morabito, M., Chen, P., Salazar, J., & Stepinski, T. (2013). Understanding the spatial distribution of crime based on its related variables using geospatial discriminative patterns. *Computers, Environment and Urban Systems*, 39, 93–106.
- Wheeler, D. C. (2014). Geographically weighted regression. In: *Handbook of Regional Science*, M. M. Fischer & P. Nijkamp (eds.), Springer Berlin Heidelberg, pp. 1435–1459. doi: 10.1007/978-3-642-23430-9_77