



EXAMINING SPATIAL VARIABILITY IN THE ASSOCIATION BETWEEN SOME FACTORS AND NUMBER OF PRISONERS IN TURKEY: A GWR ANALYSIS

Bahadır YÜZBAŞI, Department of Econometrics, Faculty of Economics and Administrative Sciences, Inonu University, Turkey, b.yzb@hotmail.com

( <https://orcid.org/0000-0002-6196-3201>)

Çetin GÖRÜR* Department of Econometrics, Inonu University, Social Sciences Institute, Turkey, gorurcetin@hotmail.com

( <https://orcid.org/0000-0002-9556-5068>)

Received: 27.04.2021, Accepted: 18.06.2021

*Corresponding author

Research Article

DOI: 10.22531/muglajsci.928734

Abstract

The crime phenomenon requires an economic as well as a sociological analysis in terms of its occurrence, its causes and its consequences. In this study, the effects of socio-economic factors on crime were investigated by using Multiple Linear Regression and Geographical Weighted Regression Methods. The aim of the study is to analyze the factors that have an effect on crime, as well as to analyze which of the regression methods used gives more effective results. Method comparisons were made using R square and Akaike's Information Criterion. As a result of the analysis, it was seen that the Geographic Weighted Regression gave more effective results than the Multiple Linear Regression. In the study, Turkey's 81 provinces were included in the analysis and used the data of the year 2019. The number of prisoners is dependent variable, migration, literacy rate and unemployment rate are independent variables. The data used in this study are taken from the official web address Turkey Statistical Institute. Analyses were made using the statistical R program.

Keywords: Geographic Weighted Regression, Multiple Linear Regression, Akaike's Information Criterion

TÜRKİYE'DEKİ BAZI FAKTÖRLER İLE MAHKUM SAYISI ARASINDAKİ İLİŞKİDE MEKANSAL DEĞİŞKENLİĞİN İNCELENMESİ: BİR GWR ANALİZİ

Özet

Suç olgusu, hem ortaya çıkış biçimi ve nedenleri ve hem de sonuçları açısından sosyolojik olduğu kadar ekonomik bir çözümlemeyi de gerekli kılmaktadır. Bu çalışmada, sosyo-ekonomik faktörlerin suç üzerindeki etkileri Çoklu Doğrusal Regresyon ve Coğrafik Ağırlıklı Regresyon Yöntemleri kullanılarak araştırılmıştır. Çalışmanın amacı, suç üzerinde etkili olan faktörleri analiz etmenin yanı sıra kullanılan regresyon yöntemlerinden hangisinin daha etkili sonuçlar verdiğini analiz etmektir. Yöntem karşılaştırmaları R² ve Akaike Bilgi Kriteri kullanılarak yapılmıştır. Analiz sonucu olarak Coğrafik Ağırlıklı Regresyonun Çoklu Doğrusal Regresyondan daha etkili sonuçlar verdiği görülmüştür. Çalışmada, Türkiye'nin 81 ili analize dahil edilmiş ve 2019 yılına ait veriler kullanılmıştır. Mahkûm sayısı bağımlı değişken, göç, okuryazar oranı ve işsizlik oranı bağımsız değişkenlerdir. Çalışmada kullanılan veriler Türkiye İstatistik Kurumu resmi web adresinden alınmıştır. Analizler istatistiksel R programı kullanılarak yapılmıştır.

Anahtar Kelimeler: Coğrafik Ağırlıklı Regresyon, Çoklu Doğrusal Regresyon, Akaike Bilgi Kriteri

Cite

Görür, Ç., Yüzbaşı, B., (2021). "Examining Spatial Variability in the Association between some Factors and Number of Prisoners in Turkey: A Gwr Analysis", *Mugla Journal of Science and Technology*, 7(1), 141-148.

1. Introduction

Since the first period of history, crime has attracted the attention of economists as well as other disciplines, as a problem that disrupts the institutional order of the society, breaks the ties that provide positive interaction in the society and radically affects this structure. Because the studies of scientists have revealed that many things will push people to commit crimes. With regard to social and economic policies, crime causes significant damage to society with its enormous social cost. Solving this problem primarily depends on the definition of the crime

phenomenon and the determination of the factors that reveal the crime. There are studies in the literature about the existence of many parameters for committing crime. In general, studies conducted towards the mid-1980s agreed that high unemployment rate would lead to an increase in crime. However, the relationship between unemployment and crime, when measured at the total level, has been shown to be both inconsistent and insignificant [1]. In [2], a decrease in the unemployment rate does not have any effect on the crime rate. In [3], in a study he conducted in Buenos Aires with the data of

1985-1997, found that there was no relationship between the unemployment rate and the crime rate, and the inequalities in the income distribution increased the crime rate. In [4], concluded that unemployment does not have a strong effect on crime rate in New Zealand. In [5], found the relationship between unemployment and all crime types to be negative and weak in their analysis, which investigated the socioeconomic determinants of the crime they made between provinces in Spain. In [6], using data from 1995 to 2002, in his analysis of the relationship between criminal activity and total unemployment in Turkey mentioned that it was weak but linear. He pointed out that during the recovery periods of the economy, the total number of crimes would decrease and increase in periods of recession and high unemployment. In [7], conducted a cross section analysis using the provincial data for the year 2000. In the analysis, it was concluded that unemployment positively affects the crime in 19 of the 36 models estimated. In [8], tried to determine the relationship between unemployment and crime by using the multiple regression method among provinces, and in their analysis results they stated that unemployment only increased the number of crimes of damaging property, which was statistically significant only at the level of 10%.

Studies examining the effect of education on crime generally emphasize that the increase in the education levels of individuals increases the skills and skills of individuals, thus their earnings from legal jobs increase and the opportunity cost of illegal activities increases. Many studies show that criminals have a less educated background than the innocent. The educational status of criminals is not good and they tend to drop out of school at an early age [9]. Since crime carries huge uncertain social costs, it has negative externalities for society. If the increase in education level decreases crime, there will be an enormous social benefit that cannot be calculated with schooling. In this case, the social gain of education will exceed its private gain [10]. After their analysis, concluded that schooling significantly reduces crime, 1% increase in high school completion rate of all men aged 20-60 in the USA reduces the cost of crime by \$ 1.4 billion per year. In fact, as [11] and [12] support, education increases skills and ability, then increases job opportunities and wage levels [13].

In [14], examined the effects of family background and environment on the behavior of young people living in the low-income Boston neighborhood. He stated that in the very poor environment in the city of Boston before 1989, "adult family behaviors are similar to the behaviors of young people, this connection is important for criminal activity". In [15] and [16], stated that population density is one of the important factors affecting crime formation. They argued that crime should not be encountered in places where the population is not dense [17]. There is a significant and expected relationship between the population, population density and urban population ratio of the provinces and most of

the crime types [8]. In [7], modeled population density and birth rate variables in a total of 13 combinations, it was found that the coefficients were statistically significant and these variables increased the crime rate.

When the studies on the relationship between crime and economic dynamics are examined, it is seen that the work of Becker and Ehrlich is the pioneer. In these studies, it was emphasized that the scarce conditions in the labor market can lead people to illegal activities or that the income earned by the person from illegal activities is more than the income from the labor market may be a reason for preference. The purpose of this study, Turkey's 81 provinces, migration, literacy rate and unemployment rate impact on the number of prisoners is to analyze with geographic weighted regression method.

2. Material and Method

Research material in 2019 and 81 provinces of Turkey, number of prisoner (NP), migration (IM), literacy rate (LR) and unemployment rate (UR) consists of data. The data were taken from the official web address of the Turkish Statistical Institute (TÜİK, 2021) and analyzed with the R statistics package program. Packages used in practice; sp, rgdal, rgeos, tmap, spgwr, grid, gridExtra and spdep.

In the study, Multiple Linear Regression (MLR) and Weighted Geographic Regression (GWR) methods were applied and thus the relationship between variables was determined. The existence of spatial relationships in variables was analyzed with the Moran's I test. The presence of multicollinearity problem between independent variables was tested with the Variance Inflation Factor (VIF) value. In addition, comparison of the results of MLR and GWR methods was made using Akaike's Information Criterion (AIC) and R Square (R^2) values.

2.1. Moran's I

Geographical data are in relation with each other due to spatial interaction [18]. When analyzing geographic data, the results are generally dependent on spatial relationship. The geographical relationship of the variables should be analyzed before doing spatial analysis. The existence of this relationship is made by the Moran's I test [19]. The Moran's I equation is given below:

$$I = \frac{n \sum_{i=1}^n \sum_{j \neq 1}^n \omega_{ij} (x_i - \bar{x})(x_j - \bar{x})}{n \sum_{i=1}^n \sum_{j \neq 1}^n \omega_{ij} (x_i - \bar{x})(x_i - \bar{x})} \quad (1)$$

$$Z_{score} = \frac{I - E(I)}{\sqrt{V(I)}}, \quad E(I) = \frac{-1}{n-1}, \quad V(I) = E(I^2) - E(I)^2$$

where the Moran's I is, n is number of research subjects, ω_{ij} are Weights representing spatial elements i and j , are Deviations of observations from the mean on the i , j spatial unit, Z_{score} is threshold for normalized statistics, $E(I)$ is Expected Autocorrelation, and $V(I)$ is variance. The Moran's I value is between -1 and 1. A Moran I value between -1 and 0 indicates a negative spatial relationship between geographic variables. If this value is between 0

and 1, it can be said that there is a positive spatial relationship.

2.2. Multiple Linear Regression Model

There are more than two variables in a multivariate geographic model. There is a linear regression relationship between these variables. This regression model is called as MLR. It is one of the most used methods to examine the relationship between dependent variables and independent variables. The Multiple Linear Regression model is as follows:

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \varepsilon_i. \quad (2)$$

In the equation, y_i is the fitted value of the land subsidence change, β_0 is a constant term, β_k is a regression coefficient of the k^{th} explanatory variable (the cumulative water level change of each aquifer), p is the number of observations, and ε_i is a random error term that fits the normal distribution.

2.3. Variance Inflation Factors

An example table is given by Table 1. Depending on the if there is a linear relationship between the independent variables, there is a multicollinearity problem. The existence of this problem is measured with VIF. It can be said that variables with values greater than 7.5 are unnecessary for the model. In order to obtain healthy results, the VIF value of the variables is considered to be less than 5. If the VIF value is high, high Std Error and low t-statistics values occur. However, for an accurate model estimation, small Std Error and high t-statistics values are preferred. The interpretation of the VIF value is given in the table below [20]:

Table 1. VIF interpretation

VIF value	Results
VIF=1	Not correlated
1 < VIF < 5	Moderately correlated
VIF > 5	Highly correlated

The equation below shows the VIF calculation:

$$VIF(\beta_i) = \frac{1}{1 - R_i^2}. \quad (3)$$

This *VIF* can be calculated for each predictor in the model, and the way is to regress the variable assume it is i^{th} variable against all other predictors. We obtain R_i^2 which can be used to find *VIF*, same thing can be applied to all other predictors [21].

2.4. Akaike's Information Criterion

In 1973, Akaike developed the AIC information criterion to see how different models exhibit an explanatory performance for a considered analysis. AIC comparison criteria are used in many scientific fields, especially in statistics and econometrics. In [22], showed that this

criterion is used in many different areas by using the marketing field. The AIC equation is as follows:

$$AIC = 2n \ln \sigma + n \ln 2\pi + n \left[\frac{n + tr(S)}{n - 2 - tr(S)} \right]. \quad (4)$$

In the equation, AIC is the value of AIC, n is the number of samples, σ is the standard deviation of the difference estimate, $tr(S)$ is the trace of the projection matrix S of the GWR model and is a function of the bandwidth. For the appropriate model, the model with low AIC value is selected [23].

2.5. R Square and Adjusted R Square

In a regression model, the explanation ratio of the variation in the dependent variable by the independent variables is determined by R^2 . The high value of the R^2 means that the independent variables are good at explaining the dependent variable. R^2 is considered to be a perfect explanation ratio if it is 1 [24]. The calculation of the R^2 is as follows [25]:

$$R^2 = 1 - \frac{RSS(Residual Sum of Squares)}{TSS(Total Sum of Squares)}. \quad (5)$$

Increasing the number of independent variables in a regression model causes a decrease in *TSS*. In this case, the R^2 value may not decrease. So no matter how many arguments are added to the model, the fit can look good. This situation really causes the question of whether the result is correct. The fact that the *Adj. R²* penalizes the arguments added to the model helps to clear up this suspicion. The Adjusted R^2 equation is as follows:

$$Adj. R^2 = \left\{ 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1} \right] \right\}.$$

2.6. Geographically Weighted Regression Model

It generates a single parameter for all spatial data in the Spatial Autoregression (SR) model. GWR, on the other hand, produces different parameters for each region, taking into account the differences of the regions of the data. In the literature, in the approximate solution studies for SR, exponential matrix properties, graph theory, Taylor series approximate solution, Chebyshev approximate solution, semi-parametric estimates, characteristic polynomial approximation, double bound likelihood estimation, lower-upper bound divide and conquer and SR local prediction were used ([26], [27], [28], [29] and [30]). Geographical weighted regression has been proposed by Fotheringham [31]. Unlike SR, it is a regression analysis that produces parameters within the neighborhood ratio according to the points in a geography and known reference points and gives results about the new point thanks to these parameters. Parameters of each point are different and parameter calculations can be made on a point basis or for all points on the area. In the SR model, the parameter is calculated to include the whole spatial area and is fixed.

The regression model defined in MLR can only be used in numerical analysis. In order to calculate the regressions of coordinates in geographical systems, the model must

be adapted to this. The following equation can be used for the GWR model:

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i)x_{ik} + \varepsilon_i. \quad (6)$$

In the equation, y_i is the fitted value of the land subsidence change for sample, i , (u_i, v_i) is the geographic center coordinate of sample space unit, $\beta_{i0}(u_i, v_i)$ is the constant term estimate of the i -th sample, $\beta_{ik}(u_i, v_i)$ is the k^{th} regression parameter of the i -th sample, which is a function of geographic location, $x_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{ip})^T$ are known vectors, and ε_i is an error that follows an independent normal distribution with a mean of zero.

Locally weighted parameters ($\beta_k(u_i, v_i)$) are calculated using the weighted least squares method. In order to calculate the parameters β , weights must be found. The weights are obtained from the continuous function of the coordinate $w_{ij}(u_i, v_i)$, the point for each value $i = 1, \dots, n$ and the distance between other points. The geographic equation with weighting is as follows:

$$\beta_i(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} (X^T W(u_i, v_i) X) Y. \quad (7)$$

Here $i = 1, \dots, k$ denotes the matrix row, $X = [x_1, x_2, \dots, x_n]^T$ the explanatory variables matrix, $Y = (y_1, \dots, y_n)$ the dependent variable and $W(u_i, v_i) n \times n$ the spatial weighting matrix. It takes the form of $W_i = \text{diag}[w_{i1}, w_{i2}, \dots, w_{in}]$ by calculating only the diagonal axis using Gaussian or bi-square functions.

Table 3. Results of MLS Analysis

Variables	Coefficient	S. Error	t value	VIF	p - value
Intercept	2.601	1.106	2.352		0.021 *
IM	5.446	2.303	3.233	1.038	0.001 *
LR	-2.396	1.165	-2.057	1.063	0.043 *
UR	1.587	4.184	3.794	1.025	0.001 *

* $p < 0.05$, $R^2 : 0.211$, Adjusted $R^2: 0.199$, F-statistic: 6.474, AIC: 1156.794

Table 3 shows the results of MLS regression. The level of significance was taken as 0.05 in the study. According to the MLS results, a significant relationship was found between the independent variables IM, LR and UR and the dependent variable NP ($p < 0.05$). When the relationship between the independent variables and the dependent variable was examined, a positive relationship was found between IM and NP. One unit of change in IM causes a change of 5.4 in NP. There is a negative relationship between LR and NP. One unit of change in LR causes a change of -2.3 in NP. On the other hand, a positive correlation was found between UR and NP. One unit of change in UR causes a change of 1.5 in NP. R^2 is 0.21, it means 21% of the variation in the dependent variable is explained by the independent variables. Likewise, the Adjusted R^2 value also carries a value close to the R^2 value. The AIC value of the model was found to be 1156.794. It is understood from the VIF values that there is no multiple linear correlation problem between the independent variables ($VIF < 5$). The distribution of

3. Analysis and Results

In this chapter; Information on study variables, MLS analysis results, Moran's I results, GWR analysis results and the comparison of MLS and GWR models are discussed in detail, and the results are given step by step below. In the study, NP are dependent variables, IM, LR and UR are independent variables.

Table 2. Descriptive of Summary of Study Variables

Variables	Mean	S. Deviation	Min.	Max.
NP	98.210	31.744	1.110	98.624
IM	103.541	36.678	89.000	101.451
LR	94.954	2.821	87.572	98.670
UR	1.232	1.903	0.096	15.323

Number of data points: 81

Table 2 shows the descriptive statistics from the variables used in the GWR modeling. The variables of the NP, IM, LR and UR were used in the study. Geographical analysis unit is composed of data from all provinces in Turkey. The result shows that the mean prevalence rate of NP was 98%. The average of IM (103%) is higher than LR (94%) as seen in the table. The last variable UR (1.2%) has the lowest average. The results of the MLS model are given in detail in the table below.

the coefficients of the variables by provinces is given in the map below.

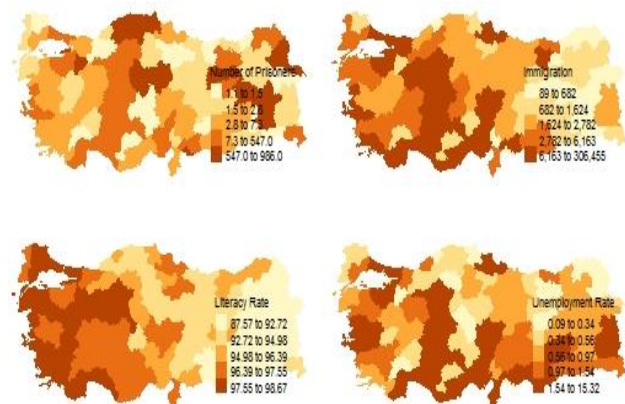


Figure 1. Distribution of Variables in Turkey by Provinces.

Figure 1 shows the distribution of NP, IM, LR and UR in Turkey by Provinces. When the figure is examined, it is

seen that the density ranges from light color to dark color for all variables. It is seen that the density is higher in dark colors. At the same time, it is clearly seen that provinces with border neighborhoods have values close to each other. This similarity is mostly seen in the literacy rate variable. In this case, it can be concluded that there is a relationship between provinces with border neighborhoods. These border neighborhoods were analyzed in detail with the Moran's I test. Below is the map drawn to calculate the border neighborhoods of the provinces.

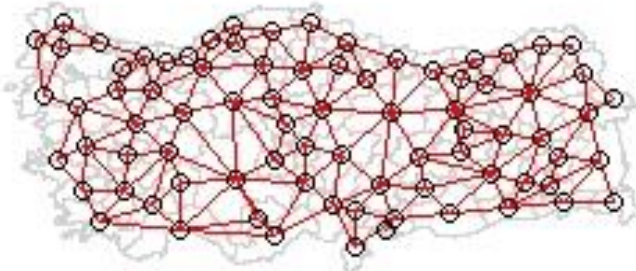


Figure 2. Determining neighborhood boundaries for the Moran I calculation

In order to analyze the existence of spatial autocorrelation, it is necessary to calculate Moran's I statistical values. Before calculating the Moran's I statistic value, the border neighborhoods of the provinces are determined. A detailed map of the provinces with border neighborhoods is given in Figure 2. The results of the Moran's I test are given in the table below.

Table 4. Moran's I statistical results.

Variables	Moran's I	Z Score	p - value
NP	0.168	2.533	0.005*
IM	-0.012	-0.203	0.049*
LR	0.807	11.454	0.001*
UR	0.099	1.618	0.048*

* p < 0.05

The global spatial autocorrelation index, with results shown in Table 4. The results showed that there was significant spatial autocorrelation for each variables (p<0.05). When the Moran's values are examined, it is observed that the spatial autocorrelation is positive in NP (0.16), LR (0.80) and UR (.09) variables, but this correlation is negative in the IM (-0.01) variable. Through the overall analysis of the Moran's I values, the global spatial autocorrelation was the lowest in IM (-0.01). The highest spatial autocorrelation belongs to the LR variable. The local spatial autocorrelation of the variables is given in detail in Figure 3 (a, b, c, d) below, according to the results of Moran's I values.

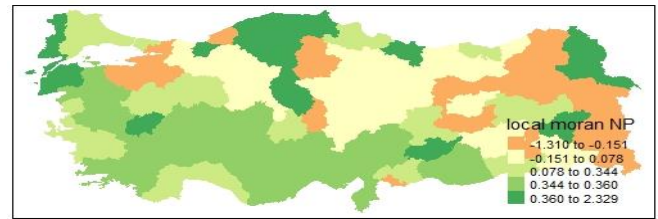


Figure 3a. Geographic distribution of Moran's I values

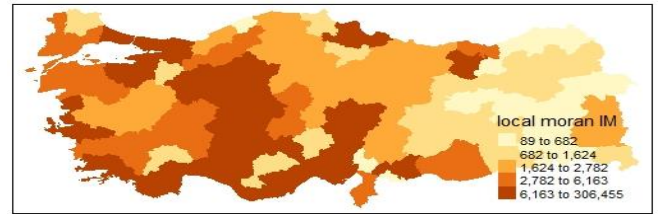


Figure 3b. Geographic distribution of Moran's I values

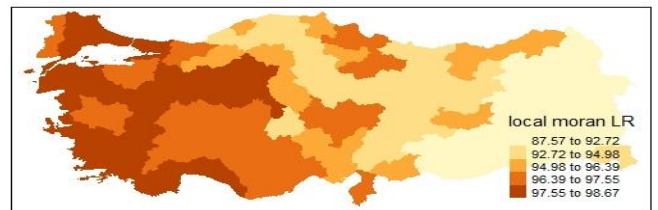


Figure 3c. Geographic distribution of Moran's I values

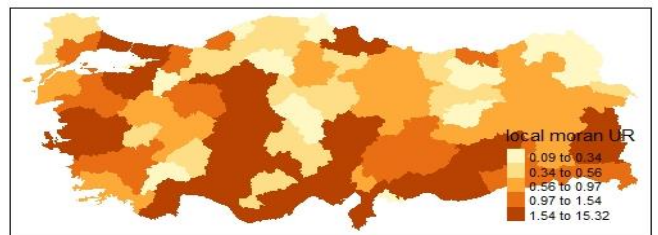


Figure 3d. Geographic distribution of Moran's I values

Figure 3 (a, b, c, d) shows the local spatial autocorrelation of the variables is given in detail according to the results of Moran's I values. For the NP (3a) variable, it is observed that the local spatial autocorrelation is negative in the regions with orange and light, and there is positive local spatial autocorrelation in the regions with green and dark green colors. There is a positive local spatial autocorrelation in all regions for the variable IM (3b). The degree of correlation is indicated from light to dark. It is observed that the local spatial autocorrelation is higher in regions with dark brown color. Local spatial autocorrelation is seen more clearly in the LR (3c) variable. It is clear that there is a positive local spatial autocorrelation between provinces. The degree of correlation likewise increases from light to dark, and it is observed that the local spatial autocorrelation is higher in the dark brown regions. There is a positive local spatial autocorrelation for the UR (3d) variable in the graph. The degree of correlation increases from light to dark color similar to other variables, and it is observed that the local spatial autocorrelation is higher in regions with dark brown color. The table below gives the results for the GWR model.

Table 5. Results of GWR Analysis

Variables	Min.	1st Qu.	Median	3rd Qu.	Max.
Intercept	2.288	2.424	2.507	2.617	2.691
IM	3.164	6.584	7.860	9.032	5.392
LR	-2.505	-2.429	-2.272	-2.185	-2.061
UR	1.348	1.891	1.660	1.412	1.919

R^2 : 0.249, Adjusted R^2 : 0.240, AIC: 1146.037, AICc: 1152.140

Table 5 gives the results for the GWR model. When the table is examined, it is seen that the relationship between independent variables and dependent variables are similar in the GWR model as in the MLS model. For example, in the GWR model, as in the MLS model, there is a positive relationship between IM and UR and NP, while there is a negative relationship between LR and NP. When the table for GWR is examined, one unit of change in IM causes change minimum 3.16 and maximum change 5.39 in NP. On the other hand, one unit of change in LR causes change minimum -2.50 and maximum change -2.06 in NP. Finally, one unit of change in UR causes change minimum 1.34 and maximum change 1.91 in NP. The explanation powers of the independent variables for the dependent variable are examined in detail in the local R^2 map below.

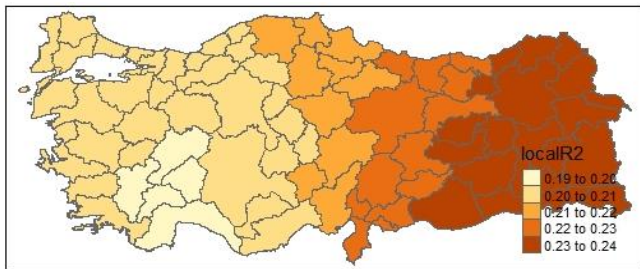


Figure 4. Local R^2 Estimates

In Figure 4, GWR local R-Squared values are mapped, yielding a number of interesting results. The local R-square was found to vary between 0.19 and 0.24 with the spatial variation. Figure 4 demonstrates an obvious regularity in the spatial distribution of R^2 : clearly, the east and southeast region have higher R^2 values. This illustrates that the relationships between IM, LR and UR and NP were better captured by the regression model in the east and southeast areas. It is observed that the power of independent variables to explain the dependent variable decreases as they move towards the west, as seen in the map.

Comparison of MLS and GWR models according to AIC and R^2 square values is given in table 6 below.

Table 6. Compare of MLS and GWR

Values	MLS	GWR
R^2	0.211	0.249
Adj. R^2	0.199	0.240
AIC	1156.794	1146.037
AICc		1152.140

Table 6 shows the comparison of MLS and GWR models. Comparison of MLS and GWR models was made according to R^2 and AIC values. Some of the model selection criteria are R^2 and AIC values. High R^2 and low AIC values are preferred for the model. When Table 5 is examined, it is observed that the R^2 value of the GWR model is higher than the R^2 value of the MLS model. This shows that the GWR model accounts for 24% of the change in NP, and the MLS model can only explain up to 21% of the change in NP. When the table for AIC values is examined, the AIC value of the GWR model is lower than the AIC value of the MLS model. The GWR model has a larger value of R^2 and a lower value of AIC. It can not only find the influence of each variables on number of prisoners, but also show obvious spatial heterogeneity. In addition, one of the advantages of the GWR model is that the spatial pattern generated under parameter estimation can be graphically represented, and the spatial pattern of the intensity of the impact of various attributes can be expressed in the form of a map. According to both the obtained values and the advantages of the GWR, it was deemed more appropriate to prefer the GWR model instead of the MLS model for this study.

4. Discussion and Conclusion

The crime phenomenon is handled and analyzed by many scientific disciplines (biology, psychology, psychiatry, law, economics, sociology, anthropology, etc.). The existence of many variables related to crime (poverty, unemployment, inequality, criminal peer group, problematic family structure, immigration, alcohol and drugs, character of the settlement, culture, weakness of formal and informal control elements, stigmatization, etc.) it makes it necessary to explain it in a widely perspective. In this context, theories analyzing the causes, nature and spread of crime from the past to the present have started to deal with the concepts of crime and criminal with the most comprehensive and systematic perspective.

When looking at the studies related to the factors affecting crime in the literature; in studies conducted in the 1980s, it is observed that unemployment increased crime rates. In the study conducted by [2], it was concluded that the decrease in unemployment had no effect on the crime rate. In [3], did not find any relationship between unemployment and crime in his study for New Zealand, but he concluded that income distribution increases crime. In [5], found a negative relationship between unemployment and crime rates in their study for Spain. In [8], study conducted in Turkey have found a positive correlation between crime rates and unemployment. In studies conducted in different regions to examine the relationship between education and crime rate, a generally negative relationship was found ([9], [10], [11] and [12]). In studies conducted to determine the relationship between population or immigration and crime, it was found that the rate of crime is high in regions with a high population, regions with high immigration and regions with a high population of poor people ([14], [15], [16], [17], [8] and [7]). In our study named "Examining Spatial Variability in the Association Between Some Factors and Number of Prisoners In Turkey: A Gwr Analysis", results similar to the studies in the literature were obtained among the factors discussed with unemployment.

In the analysis and findings section, after first giving statistical information about the variables, MLR analysis was performed and they were given on the map according to the densities of the variables. Later, the spatial autocorrelation of variables was analyzed according to the Moran's I test, and it was observed that the variables contained a significant autocorrelation according to the provinces. Finally, GWR analysis was performed and the results of MLS and GWR methods were compared. According to results, relationship between independent variables and dependent variables are similar in the GWR model as in the MLS model. The MLR and GWR models were compared according to the AIC and R^2 values, as a result, it was observed that the GWR model had stronger results. According to the GWR analysis results, one unit of change in IM causes change minimum 3.16 and maximum change 5.39 in NP. On the other hand, one unit of change in LR causes change minimum -2.50 and maximum change -2.06 in NP. Finally, one unit of change in UR causes change minimum 1.34 and maximum change 1.91 in NP. According to the GWR results, the variable that most affects the change in NP is IM. The least effective variable is UR. According to MLR and GWR model comparison results, R^2 value of the GWR model (0.2497107) is higher than the R^2 value of the MLS model (0.2314271). This shows that the GWR model accounts for 24% of the change in NP, and the MLS model can only explain up to 21% of the change in NP. When the table for AIC values is examined, the AIC value of the GWR model (1146.037) is lower than the AIC value of the MLS model (1156.794). The GWR model has a larger value of R^2 and a lower value of AIC.

In studies conducted in Turkey, the analysis of the factors affecting crime in general were examined by multiple linear regression method. In addition, it has been observed that the variables generally differ in the studies. At the same time, it has been observed that GWR analysis is mostly used in the field of geography. To measure the crime rate in Turkey, we believe that this work is differently due to the variables discussed and GWR method used. Therefore, the study is important and we anticipate it will contribute to the literature in Turkey.

5. References

- [1] Chiricos, T. G., "Rates of Crime and Unemployment; An Analysis of Aggregate Research Evidence", *Social Problems*, 34(2), .187-212, 1987.
- [2] Imrohoroğlu, A., Merlo, A. and Rupert P., "What Accounts for the Decline in Crime?", *International Economic Review*, 45(3), 707-729, 2004.
- [3] Hojman, D. E., "Explaining Crime in Buenos Aires: The Roles of Inequality, Unemployment and Structural Change", *Bulletin of Latin American Research*, 21(1), 121-128, 2002.
- [4] Papps, K. and Winkelman, R., "Unemployment and Crime: New Answer to an Old Question", *IZA Discussion Paper*, 25, 1- 25, 1998.
- [5] Buonanno, P. and Montolio, D., "Identifying the Socio-Economic and Demographic Determinants of Crime Across Spanish Provinces", *International Review of Law and Economics*, 28(2), 89-97, 2008.
- [6] Güvel, E. A., "*Suç ve Ceza Ekonomisi*", Roma Yayınları, Ankara, 2004.
- [7] Cömertler, N. and Kar, M., "Türkiye'de Suç Oranının Sosyo-ekonomik Belirleyicileri; Yatay Kesit Analizi", *Ankara Üniversitesi, Siyasal Bilgiler Fakültesi Dergisi*, 62(2), .37-57, 2007.
- [8] Durusoy, S., Köse, S. and Karadeniz, O., "Başlıca Sosyo Ekonomik Sorunlar Suçun Belirleyicisi Olabilir mi? Türkiye'de Deliller Arası Bir Analiz", *Elektronik Sosyal Bilimler Dergisi*, 7(23), ss.172-203, 2008.
- [9] Aytaç, M., Aytaç, S. and Bayram, N., "Suç Türlerini Etkileyen Faktörlerin İstatistiksel Analizi", *8. Türkiye Ekonometri ve İstatistik Kongresi*, İnönü Üniversitesi, Malatya, 24-25 Mayıs, 1-7, 2007.
- [10] Lochner, L. and Enrico, M., "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests and Self-Report", *American Economic Review*, 94(1), 155-189, 2004.
- [11] Freeman, R. B., *Crime and The Job Market*, NBER Working Paper Series, 4910, 1994.
- [12] Lochner, L., "Education, Work and Crime, Theory and Evidence", *Center for Economic Research Working Paper*, 465, 1-52, 1999.
- [13] Usher, D., "Education as Deterrent to Crime", *Canadian Journal of Economics*, 30(2), 367-384, 1997.
- [14] Case, A. C. and Lawrence F. K., "The Company You Keep the Effects of Family and Neighborhood on Disadvantaged Youth" *NBER Working Paper*, 3705, 1-25, 1991.
- [15] Meagan, E. C. and Gordon F. M., "The Determinants of Crime in Tucson, Arizona", *Urban Geography*, 24(7), 582-610, 2003.
- [16] William, V. A., "Socioeconomic Correlates of Increasing Crime Rates in Smaller Communities", *The Professional Geographer*, 50(3), 372-387, 1998.

- [17] Ayhan, İ. and Çubukçu, M, "Suç ve Kent İlişkisine Ampirik Bakış; Literatür Taraması", *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 3(5), .1-37, 2007.
- [18] Tobler, W.R., "A computer movie simulating urban growth in the Detroit region", *Econ. Geogr.* 46, 234–240, 1970.
- [19] Moran, P.A., "Notes on continuous stochastic phenomena" *Biometrika*, 37, 17–23, 1950.
- [20] Jamal I. Daoud 2017, "Multicollinearity and Regression Analysis", *J. Phys.: Conf. Ser.*, 2017.
- [21] Debbie, J. D. and Maria-Pia, V.F., "Robust VIF regression with application to variable Selection in large data sets", *The Annals of Applied Statistics*, 7(1), 319-341, 2013.
- [22] Andrews, R. and Currim, I., "Retention of latent segments in regression- based marketing models", *Int. J. Res. Mark.* 20(4), 315–321, 2003.
- [23] Hairuo Yu, Huili Gong, Beibei Chen, Kaisi Liu and Mingliang Gao, "Analysis of the influence of groundwater on land subsidence in Beijing based on the geographical weighted regression (GWR) model", *Science of the Total Environment*, 738, 2020.
- [24] Hintze, J. L., NCSS., *User's Guide III Regression and Curve Fitting*, NCSS Inc., Kaysville, Utah, 2007.
- [25] Newbold, P., Carlson, W. L. and Thorne, B. M., *Statistics for Business and Economics*, Eighth Edition, Pearson, England, 2013.
- [26] Pace, R. and LeSage J., "Closed-form maximum likelihood estimates for spatial problems (mess)", *Geographical Analysis*, 32(2), 157-172, 2000.
- [27] Pace, R. and LeSage, J., "Semiparametric maximum likelihood estimates of spatial de-pendence", *Geographical Analysis*, 34(1), 76–90, 2002.
- [28] Pace, R. and LeSage, J., "Chebyshev approximation of log-determinant of spatial weight matrices", *Computational Statistics and Data Analysis*, 45, 179-196 2004.
- [29] Martin, R., "Approximations to the determinant term in gaussian maximum likelihood estimation of some spatial models", *Statistical Theory Models*, 22(1):189–205, 1993.
- [30] Smirnov, O. and Anselin, L., "Fast maximum likelihood estimation of very large spatial auto-regressive models: A characteristic polynomial approach", *Computational Statistics and Data Analysis*, 35(3), 301–319, 2001.
- [31] Fotheringham, A.S., Brunson, C. and Charlton, M., *Geographically Weighed Regression the Analysis of Spatially Varying Relationships*, Wiley, West Sussex, UK, 2002.