Spatial Econometrics – Usage in Transportation Sciences: A Review Article

Zsombor Szabó[1]*, Árpád Török[2]

[1] Department of Transport Technology and Economics, Faculty of Transportation Engineering and Vechicle Engineering, Budapest University of Technology and Economics, H-1521 Budapest, P.O.B. 91, Hungary
[2] Department of Automotive Technologies, Faculty of Transportation Engineering and Vechicle Engineering, Budapest University of Technology and Economics, H-1521 Budapest, P.O.B. 91, Hungary
* Corresponding author, e-mail: szabo.zsombor@mail.bme.hu

Abstract

Nowadays the spatial econometrics is became widely used in transportation sciences. In order to know which method can be used or how they should be used, the review articles give answers. In this recent paper the goal is to collect all of the methods in one article which can be used in further researches. The improvement in this article is that beside the spatial econometric methods, other necessary techniques are also introduced. With this fact a whole analysis can be applied.

Keywords

spatial econometrics, transportation geography, transportation economics

1 Introduction

In the field of spatial sciences, it is a basic objective to handle interdependences of input data, which may have disadvantageous effect on estimation reliability of classical regression models. In case of the above mentioned types of interdependencies related to the spatial characteristic of the basic dataset, spatial methods are proposed to be applied. Accordingly, the aim of the article to introduce a selection of these kinds of models, which are applicable to handle spatial related interdependencies.

The main aspect of the authors in selecting spatial methods to be introduced is their connection to transportation sciences (Levulytė et al., 2017). As it was mentioned above it is not a goal to give a full picture about the topic, but to give an insight of the applicability of the referred group of methods. In the first step the article (Ord, 1975) focuses on which introduced a newly developed spatial index for all of counties of Ireland related to road accessibility. In the article a spatial model has been introduced where the dependent variable has been the percentage of gross agricultural output of each county consumed, while the independent variable has been being the index for road accessibility (Ord, 1975).

The paper (Sipos, 2017) has concluded local spatial tests can be applied to identify hot and cold points from a traffic safety point of view. Spatial autocorrelation related to Hungary's accident data has been proven between 2010-2012 with global Moran-test (Sipos, 2017).

There are other well-structured and applicable review and summary articles in the international literature (Anselin, 2001; Getis, 1991; LeSage, 2008; Varga, 2002), however our main goal is to collect those methodological components which can be used in our further researches (Szabó et al., 2017).

2 Spatial Autocorrelation

The first demands for spatial econometrics based evaluations are rooted to the underlying development of the demographic analysis of the 40's. Waldo Tobler assumed that according to his model, Ann Arbor's population in 1940 depends form the city's population in 1930. Furthermore, Ann Arbor's population depends on the population of all the other settlements of the World in 1930. Accordingly, Tobler introduced his first law of geography: "*Everything is related to everything else, but near things are more related than distant things.*" (Tobler 1970: p. 236).

The basic idea of spatial econometrics is very similar to the classic approach of econometrics, however in this case instead of time-lag, a space lag is used. So spatial econometric model is desirable for use, if spatial autocorrelation

exists. In case of a classic econometric model there is a one-way temporal dependency between the measured units: the latter value of two neighboring units depends on the value of the earlier unit. In spatial sense the dependency can be bilateral, so while a property of an investigated unit depends on another unit, then the other unit's property depends also on the original measured unit.

According to another approach, in land use models, the type of the land depends on the time (t), and also on the properties of the neighborhoods (the model was built in a square lattice, so the place depends from i and j). If the type of the land only depends on the same property of the neighboring areas (according to the first law of geography), then that is a geographical model, where there is a possibility to search for spatial autocorrelation (Tobler, 1979).

There are two types of the spatial autocorrelation: positive and negative. In case of positive autocorrelation similar elements make groups in space. In case of negative autocorrelation, units' pattern is shaped like a chessboard. The causes of the positive autocorrelation are well-researched; however, the causes of negative autocorrelation are not detailed in the same deepness in the literature. The positive autocorrelation can have two main reasons (Varga, 2002):

1. Technical reason: in this case the autocorrelation is always based on the specification of the element identification. For example, it is likely to find autocorrelation among political, economic or legal characteristic of cities located in the same geographical area.

2. The substantive reason: in this case the spatial autocorrelation is based on some kind of social interaction. The main reason of it is mostly the first law of the geography.

3 Econometric Usage of Spatial Autocorrelation

As it was mentioned before the spatial autocorrelation's best analysis is based on the econometric models. Because of the existence of two-way dependency, weight matrices are proposed to be used. There are three possible ways of it. The first is the binary neighborhood matrix, the second is the inverse distance matrix, while the third is the mixture of the first two, the Getis matrix. The weight matrix is signed by W, while w_{ij} denotes its elements, which represents the relationship between elements i and j (Varga, 2002).

3.1 Binary Neighborhood Matrix

The precondition of setting up a binary neighborhood matrix is the definition of the degree of neighborhood. Two points of interest which have a common border are called direct (first order) neighbors. For example, in case
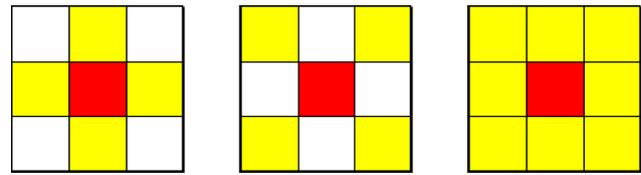

Fig. 1 The neighbor patterns (rook, bishop and queen)
(Source: authors)

of a 3 × 3 square lattice there are three types of neighborhood, as it is shown in Fig. 1.

The first type is called rook, the second is called bishop while the third is called queen type of association based upon the rules of chess. Fig. 1 shows the neighboring elements (with yellow) from the chosen point (with red) in each case.

Based on the previously shown definition, if one element is given, the direct neighbor of a direct neighbor (which is not the investigated fields), is its second degree neighbor, and so on. Based on this fact, the binary neighborhood matrix could only be referred only to one neighbor degree (which can be referred first degree neighbors, second degree neighbors, or else, but only one of them). In this case $w_{ij} = 1$ only if the neighbor degree between i and j are fitted to the actual order what the matrix represents (in case of a first degree binary neighborhood matrix, W, if $w_{ij} = 1$ then there is a first degree neighboring relationship between i and j) (Varga, 2002).

3.2 Inverse Distance Matrix

In case of inverse distance matrix three different matrices can be set up. The first one is the classic distance matrix where the actual elements show the distance of the two investigated points. The distances between the points are based on the spherical cosine rules, because in case of geographical locations, especially in case of long distances the classical Pythagorean theorem is not applicable. The formula of the spherical cosine rule is Eq. (1)

$$\cos c = \cos a \cos b + \sin a \sin b \cos \gamma \qquad (1)$$

where:
a, b, c the three sides of the spherical triangle
γ the angle opposite of the investigated side (c).

The first step of the distance calculation is to find a spherical triangle of which two points are the investigated locations, and the third is the North Pole. The points are defined by their coordinates, however the south latitude, and the west longitude coordinates will be negative numbers. In this case the elements of the Eq. (1) is given below, if $A(A_1; A_2)$ and $B(B_1; B_2)$ Eqs. (2)-(4)

$$a = 90° - B_1 \tag{2}$$

$$b = 90° - A_1 \tag{3}$$

$$\gamma = \left| A_2 - B_2 \right| \ . \tag{4}$$

In this case the distance between the two actual points will be $d(i,j) = cR$, where R is the radius of the Earth (6380 km).

The next matrix is the weight matrix; elements of which can be calculated with Eq. (5)

$$w_{ij} = \frac{1}{d(i,j)^a} \ . \tag{5}$$

In this case a is a parameter chosen arbitrary. The value of a is usually 2, but the previously carried out researches demonstrated that its value should be chosen between 0.4 and 3.3 w ith a median of 1.94 (Haggett, 2001). The other assumption is, that in case of $i = j$ the $w_{ij} = 0$ will be true.

As it will be shown in the next part of the article, sometimes it is proper to use row-standardized matrices. In this case, all of the elements of the matrix must be divided by the sum of the actual row. Because of this step, the sum of the rows will be equal to 1, and $\sum_{i,j} w_{ij} = N$, where N is the number of the investigated points.

3.3 Getis Matrix

The weight matrix founded by Getis is a fusion of the neighborhood and distance matrices. The values of the matrix elements can be 0 or 1 similarly to the neighborhood matrices, but an element of the matrix can only be 1 if the distance between the two investigated points is below a predefined distance (Getis, 1991; Getis and Ord, 1992).

4 Spatial Statistics

In spatial econometrics it is important to find out whether there is spatial autocorrelation in our data or not. To evaluate this, it is proper to use cross-product statistics. The general form of this statistic is given as Eq. (6) (Getis, 1991)

$$\Gamma = \sum_{i,j} W_{ij} Y_{ij} \tag{6}$$

where:
W_{ij}: elements of the weight matrix between i and j
Y_{ij}: measure of association between i and j .

The statistic test of the autocorrelation has two different types: the global and the local statistics. In case of the global evaluation the overall dataset is taken into consideration during the analysis. In case of the local investigation the statistic is calculated for only one geographical location (Varga, 2002).

4.1 Moran I-test

The most widely used form of the cross-product statistics for autocorrelation is the Moran's I-test (Moran, 1948). The statistic has a global form and a local form too. In this case the cross-product is given as Eq. (7) (Getis, 1991)

$$Y_{ij} = (x_i - \bar{x})(x_j - \bar{x}) \ . \tag{7}$$

The formula of the global Moran statistics is defined by Eq. (8) (Anselin, 2001)

$$I = \frac{N}{S_0} \frac{\sum_{i,j} \left(w_{ij} (x_i - \mu)(x_j - \mu) \right)}{\sum_i (x_i - \mu)^2} \tag{8}$$

where:
N : number of the investigated points
x_i , x_j : the observed value of two points of interest
μ : the expected value of x
w_{ij} : the elements of the spatial weight matrix
S_0 : normaliser $-$ $S_0 = \sum_{i,j} w_{ij}$.

Because of the previously mentioned considerations, in case of a row-standardized weight matrix $S_0 = N$, it is possible to use I^* instead of I. Its formula is given as Eq. (9) (Varga, 2002)

$$I^* = \frac{\sum_{i,j} \left(w_{ij} (x_i - \mu)(x_j - \mu) \right)}{\sum_i (x_i - \mu)^2} \ . \tag{9}$$

Based upon I^* it is possible to decide whether there is a positive or a negative autocorrelation between the data. It is enough to analyze its relation with its expected value. If $I^* > E(I^*)$, then there is a positive autocorrelation in our data, if it is not true then negative autocorrelation can be defined. The expected value of I^* can be expressed by Eq. (10) (Varga, 2002)

$$E(I^*) = \frac{-1}{N-1} \ . \tag{10}$$

There are two possible ways to define the significance of the result of the Moran test. The methods analyze the given dispersion to random dispersions, and if the actual dispersion seems to be extreme compared to the random ones, then the spatial autocorrelation's presence is significant (Varga, 2002).

In case of the first method the null hypothesis is that the values tend to normal distribution. Then, for I value a u-statistics can be calculated. However, in most of the cases this statistic tends to be too strict, so it is advised to test the results with other methods as well (Varga, 2002).

A possibly applicable other way to test the results of the Moran test is to apply permutation method. In this case the

null hypothesis assumes that the probability of any permutation of the data (the observed values) has the same result. To carry out the analyses, the first step is to order a random permutation of the values to the points; then the alternate Moran-I statistic is need to be calculated. These alternate I values are need to be compared to the original one. Finally, the pseudo significance value can be calculated based on Eq. (11) (Bajmócy and Szakálné Kanó, 2009; Varga, 2002).

$$\frac{T+1}{M+1} \tag{11}$$

where:

T : number of the cases, where the alternate I value is higher than the original one

M : number of the permutations.

The low pseudo significance value indicates that the null hypothesis can be rejected, which signs the presence of the spatial autocorrelation (Varga, 2002).

Among the local statistics Anselin's local Moran statistic is the most widely used. In this case, the used formula is Eq. (12) (Varga, 2002).

$$I_i = z_i \sum_j w_{ij} z_j, \text{ where } z_i = x_i - \mu \tag{12}$$

where:

x_i , x_j : the observed value of two points of interest

μ : the expected value of x

w_{ij} : the elements of the spatial weight matrix.

To calculate the significance level, the previously shown methods are appropriate.

4.2 Geary c-test

The Geary c-test is another test to evaluate the presence of spatial autocorrelation. The c contiguity ratio is a generalization of von Neumann's ratio (von Neumann, 1941) for time-series. The benefit of the given parameter is that it results normally distributed dataset for a low amount of points. Considering the above mentioned advantageous characteristics of c contiguity ratio, to test the significance of the results a u-test is adequate (Geary, 1954). This is another type of cross-product statistics, so in this case formula can be defined as Eq. (13) (Getis, 1991)

$$Y_{ij} = \left(x_i - x_j\right)^2 . \tag{13}$$

The formula of the contiguity ratio is defined as Eq. (14) (Geary, 1954)

$$c = \frac{N-1}{K_1} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_i - x_j\right)^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)} \tag{14}$$

where $K_1 = \sum_{i=1}^{n} k_i$ where k_i stands for the number of the connexions (Geary, 1954).

If Geary's c value is equal to 1, then there is no autocorrelation between the data (Getis, 1991).

4.3 Getis-Ord G-test

The G-test is slightly newer than the other test, because it was introduced in 1991. The main idea of the test, that this statistic makes a connection between the spatial interaction and spatial autocorrelation models (Getis, 1991). This is also a cross-product type statistic, as it can be seen in Eq. (15):

$$Y_{ij} = x_i x_j . \tag{15}$$

The Getis-Ord G-test has also local and global form. The local form has also two types differentiated by the existence of $i = j$ relationship. The formula of the two statistics are given as Eqs. (16) and (17) (Getis and Ord, 1992)

$$G_i = \frac{\sum_{j=1}^{n} w_{ij} x_j}{\sum_{j=1}^{n} x_j} \forall j \neq i \tag{16}$$

$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij} x_j}{\sum_{j=1}^{n} x_j} . \tag{17}$$

The difference between the two formula is the restriction $j \neq i$. Because there are only small differences between the two tests, hence the significance test will only be present for the first one (the other one is similar). For the significance test a simple u-test can be used (Getis and Ord, 1992).

As it was mentioned before there is a global form of the G-test. It is quite similar to the local form, hence only its formula is presented at Eq. (18) (Getis and Ord, 1992)

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} x_i x_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j} \forall j \neq i . \tag{18}$$

5 Estimation

In spatial econometric models it can be necessary to define the effects of the independent variables. To do this, other kind of models – differing significantly from the above mentioned methods – are need to be introduced. The basis of these models is the classic linear regression model. Because it is usually estimated by the ordinary least squares (OLS) model hence they are usually called OLS models. The estimation formula is given as Eq. (19) (Elhorst, 2010)

$$y = \alpha + \beta x + \varepsilon \ . \tag{19}$$

All of the introduced spatial models are based on the previously presented formula, however these formulas contain different spatial parameters (Elhorst, 2010).

5.1 Spatial Lag Model

The essence of the spatial lag model can be concluded as follows: the time lag used in the classic econometrics is used combined with the weight matrix. The formula of the model is the Eq. (20) (Varga, 2002)

$$\boldsymbol{y}_{(N\times1)} = \rho \boldsymbol{W}_{(N\times N)} \boldsymbol{y}_{(N\times1)} + \boldsymbol{X}_{(N\times K)} \boldsymbol{\beta}_{(K\times1)} + \boldsymbol{\varepsilon}_{(N\times1)} \tag{20}$$

where:

$\boldsymbol{y}$: vector of the dependent variables

ρ : autoregressive parameter

$\boldsymbol{W}$: weight matrix

$\boldsymbol{\beta}$: the coefficient vector

$\boldsymbol{X}$: the matrix of the independent variables

ε: vector of errors ($E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2$)

N : number of points of interest

K : number of independent variables.

In spatial econometrics, during the estimation process, the classic least square method cannot be used, so other methods are need to be applied, (e.g. maximum likelihood approach).

Thanks to the properties of the likelihood estimation, it is enough to find the maximum of the concentrated likelihood function. For the spatial lag model's concentrated likelihood function, the only unknown variable is the ρ, so it can be found after the maximization. If the ρ is calculated, then the $\boldsymbol{\beta}$ and the σ^2 are also can be calculated. The formula of the spatial lag model's concentrated likelihood function is given as Eq. (21) (Varga, 2002)

$$L_c = \sum_i \ln(1 - \rho\omega_i) - \frac{N}{2} \ln\left[\frac{(\boldsymbol{e}_0 - \rho\boldsymbol{e}_L)^T (\boldsymbol{e}_0 - \rho\boldsymbol{e}_L)}{N} \right] \tag{21}$$

where:

L_c : the value of the concentrated likelihood function

ω_i : the eigen values of the $\boldsymbol{W}$ row standardized weight matrix

$\boldsymbol{e}_0$: the residual vector of the estimation, where the $\boldsymbol{y}$ is the vector of the dependent variables, while $\boldsymbol{X}$ is the matrix of the independent variables

$\boldsymbol{e}_L$: the residual vector of the estimation, where the $\boldsymbol{Wy}$ is the vector of the dependent variables, while $\boldsymbol{X}$ is the matrix of the independent variables.

As it was mentioned before after finding the maximum of the likelihood function, the $\boldsymbol{\beta}$ and the σ^2 can be calculated from the ρ value Eqs. (22)-(23) (Anselin, 1988)

$$\boldsymbol{\beta} = \boldsymbol{b}_0 - \rho\boldsymbol{b}_L \tag{22}$$

$$\sigma^2 = \frac{(\boldsymbol{e}_0 - \rho\boldsymbol{e}_L)^T (\boldsymbol{e}_0 - \rho\boldsymbol{e}_L)}{N} \ . \tag{23}$$

5.2 Spatial Error Model

The other model is the spatial error model, where the spatial effects are taken into consideration as an error. So in this case the final model is exempt from the spatial autocorrelation. In this case the model's mathematical representation is Eqs. (24) and (25) (Varga, 2002)

$$\boldsymbol{y}_{(N\times1)} = \boldsymbol{X}_{(N\times K)} \boldsymbol{\beta}_{(K\times1)} + \boldsymbol{\varepsilon}_{(N\times1)} \tag{24}$$

$$\boldsymbol{\varepsilon}_{(N\times1)} = \lambda \boldsymbol{W}_{(N\times N)} \boldsymbol{\varepsilon}_{(N\times1)} + \boldsymbol{\zeta}_{(N\times1)} \tag{25}$$

where:

$\boldsymbol{\zeta}$: vector of spatial dependent errors

λ : autoregressive error parameter.

As it was mentioned before in spatial econometrics the classic ordinary least squares method cannot be used. So the maximum likelihood estimation is ought to be used. As it was mentioned above it is enough to use the concentrated likelihood function for the maximization, and after finding the maximum value the $\boldsymbol{\beta}$ and the σ^2 parameters can be calculated from λ . The formula of the likelihood function in this case is Eq. (26)

$$L_c = \sum_i \ln(1 - \lambda\omega_i) - \frac{N}{2} \ln\left[\frac{\boldsymbol{e}^T (\boldsymbol{I} - \lambda\boldsymbol{W})^T (\boldsymbol{I} - \lambda\boldsymbol{W}) \boldsymbol{e}}{N} \right] \tag{26}$$

where $\boldsymbol{e}$ is the residual vector of the general least square model.

5.3 Other Spatial Models

There is a main disadvantage of the previously shown methods as there is only one spatial parameter in the models. After 2007 several methods were introduced with more than one spatial parameters (Elhorst, 2010). However, in this paper, these methods are not the subject of the investigation.

5.4 Choosing the proper model

In order to choose the proper model, Lagrange-Multiplier (LM) tests are ought to be used. There are tests for the

presence of spatial error and spatial lag too. The first test is for the spatial error Eq. (27) (Varga, 2002)

$$LM - ERR = \frac{\left(\frac{e^T W e}{s^2}\right)^2}{T} \tag{27}$$

where Eqs. (28) and (29):

$$s = \frac{e^T e}{N} \tag{28}$$

$$T = tr\left(W^T W + W^2\right). \tag{29}$$

In this case the tr stands for the trace of a matrix.

As it was mentioned before there are tests for spatial lag too. The most frequently used is the Eq. (30) (Varga, 2002)

$$LM - LAG = \frac{\left(\frac{e^T W y}{s^2}\right)^2}{RJ_{\rho-\beta}} \tag{30}$$

where Eqs. (31) and (32):

$$RJ_{\rho-\beta} = \frac{T + (WX\beta)^T M (WX\beta)}{s^2} \tag{31}$$

$$M = I - X\left(X^T X\right)^{-1} X^T. \tag{32}$$

All of the statistics are χ^2 statistics with $df = 1$. If the test seems to be significant then the given method cannot be used (Varga, 2002).

6 Goodness of the models

To determine the goodness of the actual model, there are two possibilities. Firstly, the square of the Pearson like correlation parameter can be used, while the other possibility is to apply the likelihood based information criteria.

In case of Pearson like correlation parameters, when the model is spatially lagged then the well-known formula can be used Eq. (33) (Anselin, 1988; Bolla and Krámli, 2005)

$$R^2 = \frac{\hat{y}^T \hat{y} - N\bar{Y}^2}{y^T y - N\bar{Y}^2}, \text{ where } y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \tag{33}$$

where $\bar{Y}$ is the mean of the elements of y (Bolla and Krámli, 2005).

However, if there is a spatial autoregressive error term (λ) in the model, then the previous formula cannot be used. In this case a pseudo R^2 value should be calculated, its formulas are given by Eqs. (34) and (35) (Anselin, 1988)

$$R^2 = 1 - \frac{e^T (I - \lambda W)^T (I - \lambda W) e}{(y - \iota y_w)^T (I - \lambda W)^T (I - \lambda W)(y - \iota y_w)} \tag{34}$$

$$y_w = \frac{\iota^T (I - \lambda W)^T (I - \lambda W) y}{\iota^T (I - \lambda W)^T (I - \lambda W) \iota} \tag{35}$$

where: ι is an N by 1 vector of ones.

In case of the likelihood based values, four types of statistics can be applied. These are the L_c value of the concentrated likelihood function and the AIC, BIC and $CAIC$ statistics. The L_c value can be calculated based upon Eqs. (21) and (26), and in general, if the value is lower, the model seems to be better. However, the value of the L_c depends from the number of the independent variables, and the sample size, so an index of goodness should be applied. These are the information criteria (Cameron and Trivedi, 2005).

The first information criterion is the AIC (Akaike Information Criterion), the simplest among all. In case of using this criterion the sample size is not taken into consideration, only the number of the independent variables. Its formula is the Eq. (36) (Cameron and Trivedi, 2005)

$$AIC = -2\ln L + 2q \tag{36}$$

where:

L : value of the likelihood function

q : number of the independent variables.

Because the L_c is the value of the logarithmic likelihood function, in case of the information criteria the logarithmic transformation needs not be used. While the L_c value is a negative number of which minimum value is $-N$, where N is the sample size, the information criterion will be a positive number. According to the AIC, the model will be better if the value is lower. Because the sample size is not taken into consideration and the q value's weight is not proven to be so emphatic, so that it is not used frequently (Cameron and Trivedi, 2005).

In order to improve the AIC index KLIC (Kullback–Liebler Information Criterion) should be used. The basis of this criteria is that the second part of the formula should be replaced by a formula where the N and q values are taken into consideration, and they are more than $2q$. In this paper two of them will be introduced, the BIC (Bayesian Information Criterion), and the $CAIC$ (consistent AIC) ones. Their formulas are the Eqs. (37) and (38) (Cameron and Trivedi, 2005)

$$BIC = -2\ln L + (\ln N)q \tag{37}$$

$$CAIC = -2\ln L + (1 + \ln N)q \tag{38}$$

where:

L : value of the likelihood function

q : number of the independent variables

N : number of the elements in the sample.

7 Conclusion

This article has summed up the used methods in the topic of spatial econometrics. There are numerous articles about this field, but the aim of our article was to introduce a new view point in the overview of the topic. The goal has been to identify those methods which can be used in our further research (Szabó et al., 2017). For other researchers,

these tools could be really useful, when territorial units are compared to each other, for example in case of road safety like (Török, 2017) or (Pauer, 2017).

The article mentioned all steps of a spatial econometric analysis and examples for the applicable methods have also been presented. In the first step a weight matrix is needed to be found which was introduced in Section 3. The second step is to find out the presence of spatial autocorrelation (Section 4), and finally a proper model is need to be set up (Sections 5 and 6).

References

Anselin, L. (1988) "Spatial Dependence in Regression Error Terms", In: Spatial Econometrics: Methods and Models, Studies in Operational Regional Science, 1st ed., Springer, Dorddrecht, The Netherlands, pp. 100–118.
https://doi.org/10.1007/978-94-015-7799-1_8

Anselin, L. (2001) "Spatial Econometrics", In: Baltagi, B. H. (ed.) A Companion to Theoretical Econometrics, 1st ed., Blackwell Publishing Ltd., Padstow, United Kingdom, pp. 310–330.
https://doi.org/10.1002/9780470996249.ch15

Bajmócy, Z., Szakálné Kanó, I. (2009) "Hazai kistérségek innovációs képességének elemzése" (Measuring the Innovation Performance of Hungarian Subregions), Tér és társadalom, 23(2), pp. 45–68. (in Hungarian)
https://doi.org/10.17649/TET.23.2.1234

Bolla, M., Krámli, A. (2005) "Statisztikai következtetések elmélete" (Theory of statistical inference), 2nd ed., Typotex Kft., Budapest, Hungary. (in Hungarian)

Cameron, A. C., Trivedi, P. K. (2005) "Microeconometrics: Methods and Applications", 1st ed., Cambridge University Press, New York, USA.
https://doi.org/10.1017/CBO9780511811241

Elhorst, J. P. (2010) "Applied Spatial Econometrics: Raising the Bar", Spatial Economic Analysis, 5(1), pp. 9–28.
https://doi.org/10.1080/17421770903541772

Geary, R. C. (1954) "The Contiguity Ratio and Statistical Mapping", The Incorporated Statistician, 5(3), pp. 115–146.
https://doi.org/10.2307/2986645

Getis, A. (1991) "Spatial Interaction and Spatial Autocorrelation: A Cross-Product Approach", Environment and Planning A: Economy and Space, 23(9), pp. 1269−1277.
https://doi.org/10.1068/a231269

Getis, A., Ord, J. K. (1992) "The Analysis of Spatial Association by Use of Distance Statistics", Geographical Analysis, 24(3), pp. 189−206.
https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Haggett, P. (2001) "Geography: A Global Synthesis", 4th ed., Pearson Education, Harlow, UK.

LeSage, J. P. (2008) "An Introduction to Spatial Econometrics", Revue d'économie industrielle, (3), pp. 19−44.
https://doi.org/10.4000/rei.3887

Levulytė, L., Baranyai, D., Sokolovskij, E., Török, Á. (2017) "Pedestrians' Role in Road Accidents", International Journal for Traffic and Transport Engineering, 7(3), pp. 328−341.
https://doi.org/10.7708/ijtte.2017.7(3).04

Moran, P. A. P. (1948) "Some Theorems on Time Series: II. The Significance of the Serial Correlation Coefficient", Biometrika, 35(3-4), pp. 255–260.
https://doi.org/10.2307/2332344

Ord, K. (1975) "Estimation Methods for Models of Spatial Interaction", Journal of the American Statistical Association, 70(349), pp. 120–126.
https://doi.org/10.1080/01621459.1975.10480272

Pauer, G. (2017) "Development Potentials and Strategic Objectives of Intelligent Transport Systems Improving Road Safety", Transport and Telecommunication, 18(1), pp. 15–24.
https://doi.org/10.1515/ttj-2017-0002

Sipos, T. (2017) "Spatial Statistical Analysis of the Traffic Accidents", Periodica Polytechnica Transportation Engineering, 45(2), pp. 101–105.
https://doi.org/10.3311/PPtr.9895

Szabó, Z., Sipos, T., Török, Á. (2017) "Spatial Econometric Analysis of the Hungarian Border Crossings", MATEC Web of Conferences, 134, article number: 00057.
https://doi.org/10.1051/matecconf/201713400057

Tobler, W. R. (1970) "A Computer Movie Simulating Urban Growth in the Detroit Region", Economic Geography, 46, Supplement: Proceedings. International Geographical Union. Commission on Quantitative Methods, pp. 234–240.
https://doi.org/10.2307/143141

Tobler, W. R. (1979) "Cellular Geography", In: Gale, S., Olsson, G. (eds.) Philosophy in Geography, Theory and Decision Library, 1st ed., Springer, Dordrecht, The Netherlands, pp. 379–386.
https://doi.org/10.1007/978-94-009-9394-5_18

Török, Á. (2017) "Comparative Analysis between the Theories of Road Transport Safety and Emission", Transport, 32(2), pp. 192–197.
https://doi.org/10.3846/16484142.2015.1062798

Varga, A. (2002) "Térökonometria" (Spatial econometrics), Statisztikai Szemle, 80(4), pp. 354–370. (in Hungarian)

von Neumann, J. (1941) "Distribution of the Ratio of the Mean Square Successive Difference to the Variance", The Annals of Mathematical Statistics, 12(4), pp. 367–395.
https://doi.org/10.1214/aoms/1177731677