Article

Spatial Econometric Analysis of Road Traffic Crashes

Tibor Sipos *, Anteneh Afework Mekonnen and Zsombor Szabó

Department of Transport Technology and Economics, Faculty of Transportation and Vehicle Engineering, Budapest University of Technology and Economics, H-1111 Budapest, Hungary; anteneh.mekonnen@kjk.bme.hu (A.A.M.); szabo.zsombor@kjk.bme.hu (Z.S.)

* Correspondence: sipos.tibor@kjk.bme.hu

Abstract: Keeping the basic principles of sustainable development, it must be highlighted that decisions about transport safety projects must be made following expert preparation, using reliable, professional methods. A prerequisite for the cost–benefit analysis of investments is to constantly monitor the efficiency of accident forecasting models and to update these continuously. This paper presents an accident forecasting model for urban areas, which handles both the properties of the public road infrastructure and spatial dependency relations. As the aim was to model the urban environment, we focused on the road public transportation modes (bus and trolley) and the vulnerable road users (bicyclist) using shared infrastructure elements. The road accident data from 2016 to 2018 on the whole road network of Budapest, Hungary, is analyzed, focusing on road links (i.e., road segments between junctions) by applying spatial econometric statistical models. As a result of this article, we have developed a model that can be used by decision-makers as well, which is suitable for estimating the expected value of accidents, and thus for the development of the optimal sequence of appropriate road safety interventions.

Keywords: spatial dependence; econometric models; autocorrelation; road traffic accidents; spatial lag; spatial error

Citation: Sipos, T.; Afework Mekonnen, A.; Szabó, Z. Spatial Econometric Analysis of Road Traffic Crashes. *Sustainability* **2021**, *13*, 2492. <https://doi.org/10.3390/su13052492>

Academic Editors: Victoria Gitelman and Elżbieta Macioszek

Received: 15 December 2020

Accepted: 19 February 2021

Published: 25 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Major investments in the field of transportation have several unique characteristics that differentiate them from investments in other fields of economy. Consequently, professional preparation of decisions is crucial, as investment costs are typically very high. Such projects create infrastructure with several decades of service life, and thus problems arising due to poor preparation can only be amended with a considerable amount of resources. Road safety is one of the pillars in the Sustainable Development Goals, and it does not only explicitly address the issues related to good health and safety of the people, but it is also associated with the targets towards the development of sustainable cities [1]. Because of the scarce resources available, the implementation of the most effective road safety improvements is of paramount importance, so an analysis of each project's effectiveness cannot be avoided [2]. Efficiency analysis ought to apply a complex approach, including a socio-economic study [3,4]. Together with the financial analysis, parameters affecting a wide range of economic actors need to be involved and evaluated from several points of view. Such a determining feature is the incidence and socio-economic impact of traffic accidents.

The scarcity of widely available methodologies for defining and managing these effects in developed countries shows that the analysis of these effects is complicated, though mandatory. However, it is of utmost importance to detect areas where a transportation safety improvement will have a significant impact.

Since accidents are believed to be discrete, random, and non-negative, Poisson and negative binomial models were generally implemented (e.g., [5]). The problem with those models is that they assume that accidents are independent of the space they occur in,

although this is not the case in reality [6]. Generally, the safety performance functions (SPF) used in international practice usually modeled by negative binomial distribution due to the typically occurring overdispersion problem, as the equations can be solved, but there is no evidence that the accident frequency indeed follows a negative binomial distribution. Thus, we aimed to take the spatial effect into account when setting up the estimator model.

While road traffic crashes tend to occur at specific times, they are also affected by a comprehensive interaction of spatial factors [7]. The heterogeneity among traffic crashes on roads with the same geometric conditions is because of the unlikeliness of independence across space [8].

Accidents are not randomly distributed in space but represent spatial autocorrelation [7]. Spatial autocorrelation is the degree of influence by the value of a variable at a certain location on the value of the same variable at a contiguous location [9].

In the literature, 36 articles have been identified that also take spatiality into account in some way when analyzing traffic accidents. Table 1 summarizes how the literature can be classified, what explanatory variables that are relevant to us are used, and which of the most commonly used spatial econometric models (SAR—Spatial Autoregression, SEM—Spatial Error Model) appear in it.

Table 1. Summary of Literature.

Articles	Handling of Spatial Interactions	Analyzed Units	Speed (KPH/MPH)	AADT (veh/day)	Infrastructure Design Attributes	SAR	SEM	Ref.
Shankar et al. 2015	(i) the delimitation of spatial units appears, but the modeling of the interactions between them is omitted	(a)	-	-	×	-	-	[10]
Briz-Redón et al. 2019		(a)	-	LN	×	-	-	[11]
Ng et al. 2002		(b)	-	-	-	-	-	[12]
Noland and Quddus 2005		(b)	-	-	×	-	-	[13]
Lu et al. 2016		(b)	×	-	×	-	-	[14]
Casando-Sanz et al. 2010		(b)	-	×	×	-	-	[15]
Cantillo et al. 2015		(b)	-	×	×	-	-	[16]
Quddus 2008		(b)	×	LN	×	NS	×	[17]
Wang et al. 2009		(a)	-	LN	×	-	-	[18]
Castro et al. 2013		(a)	-	-	×	×	-	[19]
Xie et al. 2014	(ii) comparing models that handle spatial interaction with models that do not take spatial interaction into account	(a)	-	-	-	×	×	[20]
Rhee et al. 2016		(b)	-	×	×	×	×	[8]
Jia et al. 2018		(b)	-	-	-	×	×	[21]
Pljakić et al. 2019		(b)	-	×	×	×	×	[22]
Soro et al. 2017		(b)	-	×	×	×	×	[23]
Castro et al. 2012b		(a)	-	LN	×	×	-	[24]
Ha & Thill 2011		(b)	-	-	-	×	-	[25]
Ryder et al. 2019		(b)	-	×	×	-	-	[26]
Lee & Gim 2020		(b)	-	-	×	×	×	[27]
Wang et al. 2019a		(b)	-	-	×	×	×	[6]
Xu et al. 2016		(a)	-	-	×	-	-	[28]
Al-Hasani et al. 2019		(b)	-	-	-	×	×	[29]
Saeed et al. 2020		(b)	-	-	-	-	-	[30]
Zhang et al. 2020		(b)	-	-	×	-	-	[31]
Lee et al. 2018		(b)	-	-	-	×	×	[32]
Wang et al. 2019b		(b)	-	-	×	×	×	[33]

Table 1. Cont.

Articles	Handling of Spatial Interactions	Analyzed Units	Speed (KPH/MPH)	AADT (veh/day)	Infrastructure Design Attributes	SAR	SEM	Ref.
Chiou et al. 2014	(iii) only models that handle spatial interactions are set up	(a)	-	×	×	-	×	[34]
Liu and Sharma 2017		(b)	-	×	-	-	-	[35]
Simoes et al. 2015		(b)	-	-	-	-	-	[36]
Black 1991		(a)	-	-	-	-	-	[37]
Karaganis and Mimis 2006		(a)	-	-	-	-	×	[38]
Azimian and Pyrialakou 2020		(b)	-	-	-	-	-	[39]
Moons et al. 2008		(a)	-	-	-	-	-	[40]
Kortter 2016		(b)	-	-	-	×	-	[41]
Steenberghen et al. 2004		(b)	-	-	-	-	-	[42]
Li et. al. 2020		(a)	-	-	-	-	-	[43]

× the given model contains that parameter, - the given model does not contain that parameter, (a) elements of line road network, (b) spatial units, AADT: Annual average daily traffic, LN: the given parameter is logarithmic, NS: the given parameter is not significant.

Spatial traffic accident models that appear in the literature can be divided into three major groups according to how spatial interactions are handled: (i) the delimitation of spatial units appears, but the modeling of the interactions between them is omitted; (ii) they compare models that handle spatial interaction with models that do not take spatial interaction into account; (iii) only models that handle spatial interactions are set up. Within these, one of two approaches was used in each case: (a) the analyses are based on the elements of the line road network, within which it is delimited whether the given accident occurs at an intersection or between two intersections; (b) the analyses are based on spatial units. Based on this, the articles published in the literature can be classified as shown in Table 2.

Table 2. Classification of Articles.

	(a) Elements of Line Road Network	(b) Spatial Units
(i) no spatial interactions considered	[10,11]	[12–16]
(ii) compare models with and without spatial parameters	[18–20,24,28]	[6,8,17,21–23,25–27,29–33]
(iii) models considering spatial interactions	[34,37,38,40,43]	[35,36,39,41,42]

In the literature, 36 spatial models were identified. As Tables 1 and 2 illustrate, for 31 of these, either the spatial units are the observed units (b); or they do not take into account spatial interactions (i); or they only analyze spatial models (iii). This article follows the five studies enumerated in the highlighted cell of Table 2, as we have compared spatial and non-spatial models (ii), based on the elements of the linear road network (a), including the sections between intersections. Two of the mentioned articles deal with intersections [24,28]. The remaining three discuss accidents on motorways. The article by Wang et al. [18] examines the accident frequency of the M25 motorway around London, on which they build Poisson-lognormal, Poisson-Gamma, and Poisson-lognormal CAR (conditional autoregressive) [44] models. The range of explanatory variables includes infrastructure characteristics as well as traffic characteristics. Out of these, the Congestion Index is highlighted, but traffic volume and speed are also displayed. The article by Castro et al. [19] deals with motorways around Austin, TX, USA, and the range of explanatory variables includes crash characteristics, highway design attributes, characteristics of drivers and vehicles involved in the crash, and environmental factors. The article by Xie et al. [20] examines accidents caused by Hurricane Sandy according to the type of accident, where it occurred, and at what time of day.

The articles detailed above, which most closely resemble ours, take into account highway sections; thus, our analysis of urban sections fills a gap. The range of explanatory variables appears in many of the articles examined, such as traffic volume, speed, number of lanes, or length of sections (e.g., [18]), the existence of bus lanes (e.g., [11]), or ratio of truck traffic (e.g., [34]). Another difference is that in addition to the traditionally used SAR and SEM approaches, Spatial Autocorrelation (SAC) [45] models are also used. Although several articles in the literature have taken other approaches in addition to the SAR and SEM models, such as the Spatial Durbin Model (e.g., [6,30]) or the Besag–York–Mollié model [46] (e.g., [35,36]), SAC models have not been used by anyone in the literature known to us yet. In summary, we focused on an area that is missing from the literature: joint usage of four techniques: (i) we used spatial econometric models; (ii) we used real network sections instead of other spatial units; (iii) we analyzed urban roads instead of rural area; (iv) we take into account the spatial autocorrelation (SAC) modeling technique.

The main objective of this article is to identify a proper model for estimating the accident frequencies on different sections of the road infrastructure with different properties in Budapest, Hungary (for Hungary, there are numerous traffic safety models, such as [47–49]). Our assumption that a spatial econometric point of view could help us to create such a model.

The paper is organized as follows. The dataset and its preprocessing are described in the next chapter. Section 3. explains the methodology employed to diagnose spatial dependence among the data. The test and model results are discussed in Section 4, and the paper is concluded by Section 5.

2. Data

Data Collection

Detailed data were collected on accidents with personal injury and fatality from the Center for Budapest Transport (BKK) for the analysis. The accident data from 2016 to 2018 were used for the analysis. The road network data comes from the EFM (official traffic model for Budapest and the agglomeration), where modeling of the current state for Budapest is available. This database contains a huge amount of data about the road segments, including average daily traffic of passenger cars, buses, bicycles, and HGVs (Heavy Good Vehicle); properties of the road segments (number of lanes, and the existence of bus lane); and free flow speed and current speed.

This line network was segmented into junctions and links. Because there is not a real barrier between the junctions and the links, it was assumed that if the accident is less than 50 m away from a junction, then that accident belongs to the junction. With this process segments are occurred, divided by the junction's 50-m radius area. Only accidents on the road links are considered in this study.

In general, accidents are not completely randomly distributed, but their occurrence is highly associated with the traffic condition and geometric features of the road. This postulate leads researchers to conduct segment-based traffic safety analysis [50]. The total number of road segments in Budapest that are considered in this study are 4701. The total length of the road network in the dataset is 1310 km. The average link length is 0.28 km. 1586 accidents on the road links for three years from 2016 to 2018 in Budapest were considered in the analysis. Figure 1 is a schematic map from the QGIS, version 3.10.12 (QGIS Association) software, showing the distribution of accidents on the road links for the city of Budapest for the three years surveyed.

Figure 1. Spatial Distribution of Accidents in Budapest, 2016–2018. (Source: Authors).

Accidents were assigned to each corresponding road link (accidents varying from 0 to 28 per link), and every possible parameter that could be related to the accidents from the EFM was described. The most important properties, which will be our decision variables during the modeling process, are listed in Table 3. In order to reach better modeling results, two variables (number of lanes, and free-flow speed) were merged together. The reason behind this is, that these two variables can specify clearly the main type of the road. Thus, a new set of explanatory variables occurred, which is denoted by CAT_XY, where X denotes the free-flow speed category, while Y represents the number of lanes. For this, we have to categorize the roads according to the available free-flow speed by identifying four clusters in harmony with the Hungarian laws:

1. Residential streets (free-flow speed is less than or equal with 20 kph);
2. Urban roads (free-flow speed is between 20 kph and 50 kph);
3. Main roads (free-flow speed is between 50 kph and 90 kph);
4. High-speed roads (free-flow speed is above 90 kph).

Table 3. Descriptions of Variables Considered in the Study (Source: Authors).

No.	Symbol	Variable Description
0	CAT_XY	Merged category variable of the free-flow speed, and the number of lanes
1	BUS_LANE	Dummy variable for the bus lane existence
2	BUS_VOL	Average daily bus and trolleybus traffic
3	HGV_12_D	Dummy variable whether the HGVs above 12 tons are allowed or not
4	BIKE_VOL	Average daily bicycle traffic
5	AADT	Annual average daily traffic of the passenger cars and LGVs (HGVs below 3.5 tons)
6	Ratio_HGV	Ratio of the HGVs above 3.5 tons compared to the motorized traffic
7	LENGTH	Length of the given road segment

For the dummy variables, further discussion is needed in order to analyze the results. Previously two dummy variables are introduced; the first belongs to the bus lanes' existence, while the second belongs to the availableness of the heavy good vehicles above 12 tons. The rules for the dummy variables are denoted by Equations (1) and (2).

$$HGV_12_D_i = \begin{cases} 1, & \text{if HGVs above 12 tons are allowed on the given section} \\ 0, & \text{if HGVs above 12 tons are not allowed on the given section} \end{cases} \quad (1)$$

$$BUS_LANE_i = \begin{cases} 1, & \text{if there is a bus lane at least in one direction} \\ 0, & \text{if there is no bus lane along the section} \end{cases} \quad (2)$$

Accidents that occurred on a given link over the examined period are aggregated. Based on this aggregation method, it is assumed that the contributing factors for accidents that occurred on the same link are similar, which may not be right all the time [51].

All insignificant variables were eliminated based on likelihood ratio analysis and standard errors of the parameter estimate values. The correlation matrix was analyzed, and if two variables were correlated to each other strongly, one of them was eliminated from the model. This helps avoid the confounding effect of the variables. Even if a variable is found to be significant but has less contribution to the decrease in deviance, it was eliminated from the analysis.

Accidents events and the frequency of road links are summarized in Table 4.

Table 4. Accidents and Frequency of Road Links.

No. of Accidents	No. of Road Links	%	No. of Accidents	No. of Road Links	%
0	3115	66.3	9	3	0.1
1	909	19.3	10	3	0.1
2	350	7.4	11	2	0.0
3	141	3.0	12	1	0.0
4	75	1.6	15	1	0.0
5	49	1.0	16	1	0.0
6	19	0.4	17	1	0.0
7	19	0.4	18	1	0.0
8	10	0.2	28	1	0.0
Total				4701	100.0

3. Methodology

The applied methodology is mostly based on the research of Luc Anselin [52], and Attila Varga [53]. The main idea behind the applied model environment is based on the phenomenon that traditional linear regression models estimated by ordinary least squares methods cannot take into consideration the fact that panel data based upon spatial specifications is not independent of its spatial location [52]. This means, models in which spatial autocorrelation can be found in the linear regression errors cannot be used for further investigation; instead, a spatial econometric model ought to be used.

The spatial econometric analysis applied here follows the methodology described by [54]. The first step is to set up the Ordinary Least Squares (OLS) models with proper explanatory variables (3).

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_m x_{im} + \varepsilon_i \quad \forall i = 1..N \quad (3)$$

where

- y : dependent variable;

- α : constant term;
- β_j : estimated coefficients of the linear regression model ($\forall j = 1..m$);
- x_i : explanatory variables ($\forall j = 1..m$);
- ε : error term ($E\varepsilon_i = 0$, $V(\varepsilon_i) = \sigma^2$);
- N : number of observations.

The second step is the definition of the weight matrices. In accordance with the methodology proposed in the literature [54], queen type binary (B) matrices have been applied in our research, which are transformed into row standardized weight matrices (W). In the case of binary matrices, the spatial units have an effect on each other, depending on whether they are neighbors to each other (1) or not (0). A detailed description of this method can be found in [52–56]. Two conditions of neighborhood are introduced, one is where two points of interest are neighbors to each other if they are closer than a given distance. The second is when the given point neighbors are the k closest points (k can be chosen arbitrarily) or k nearest neighbors.

The third methodological step is to prove the existence of spatial autocorrelation in the dataset. The presence of spatial autocorrelation depends on the weight matrix. The existence of spatial autocorrelation and the proper weight matrix can be tested by the Moran- I test [57]. If the test results in a significant effect, a spatial econometric model can be applied [53].

$$I = \frac{N}{S_0} \frac{\sum_{i,j} (w_{ij}(x_i - \mu)(x_j - \mu))}{\sum_i (x_i - \mu)^2} \quad (4)$$

where

- N : number of 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 : normalizer— $S_0 = \sum_{i,j} w_{i,j}$

If the existence of spatial autocorrelation is proven, three types of models can be set up: the SAR, the SEM, and the SAC. In order to decide which model should be used, the Lagrange–Multiplier test is available [54,58]. If it can be assumed that the spatial lagged dependent variable also affects the dependent variable, the SAR model should be chosen. In this case, the following regression formula can be applied [53]:

$$\mathbf{y}_{(N \times 1)} = \rho \mathbf{W}_{(N \times N)} \mathbf{y}_{(N \times 1)} + \mathbf{X}_{(N \times K)} \boldsymbol{\beta}_{(K \times 1)} + \boldsymbol{\varepsilon}_{(N \times 1)} \quad (5)$$

where

- $\mathbf{y}$: vector of the dependent variables;
- ρ : autoregressive parameter;
- $\mathbf{W}$: weight matrix;
- $\boldsymbol{\beta}$: coefficient vector;
- $\mathbf{X}$: matrix of the independent variables;
- $\boldsymbol{\varepsilon}$: vector of errors ($E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2$);
- N : number of points of interest;
- K : number of independent variables.

If spatial dependence can be eliminated from the model, and the spatial effects can be transferred to the error term, the SEM model should be used. In this case, the formulas presented below can be used [53]:

$$\mathbf{y}_{(N \times 1)} = \mathbf{X}_{(N \times K)} \boldsymbol{\beta}_{(K \times 1)} + \boldsymbol{\varepsilon}_{(N \times 1)} \quad (6)$$

$$\boldsymbol{\varepsilon}_{(N \times 1)} = \lambda \mathbf{W}_{(N \times N)} \boldsymbol{\varepsilon}_{(N \times 1)} + \boldsymbol{\zeta}_{(N \times 1)} \quad (7)$$

where

- ζ : vector of spatial dependent errors;
- λ : autoregressive error parameter.

The third option is to use the two approaches (the modeling of spatial lag, and spatial error) together. There are several models for this [45], out of which the SAC was used, because this is the one which can handle two different weight matrices at the same time. Its formula is given in Equation (8).

$$\mathbf{y}_{(N \times 1)} = \rho \mathbf{W}_1 \mathbf{y}_{(N \times 1)} + \mathbf{X}_{(N \times K)} \boldsymbol{\beta}_{(K \times 1)} + \lambda \mathbf{W}_2 \boldsymbol{\varepsilon}_{(N \times 1)} + \boldsymbol{\zeta} \quad (8)$$

After building the OLS, SAR, and SEM models, the results were compared by Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) [59]. In addition to AIC and BIC, we used the likelihood test ratio (LR) as well. According to this approach, two hierarchical models (the superior one has one or more surplus estimated parameter) can be compared to each other based upon the ratio of their likelihood value. The test formula is given in Equation (9) [60]:

$$\text{LR} = 2 \ln \frac{L_2}{L_1} = 2(\ln L_2 - \ln L_1) \sim \chi_{df}^2 \quad (9)$$

where

- L_1 : likelihood value of the inferior model;
- L_2 : likelihood value of the superior model;
- df : degree of freedom for the chi-square distribution, equal to the number of the surplus estimated variables.

Spatial econometric analyses were performed in the R 3.4.0 environment [61]. The maptools [62], sp [63,64], spdep [63,65] and spatialreg [63,66,67] libraries were used during the analysis.

4. Results and Discussions

The following two hypotheses were tested to imply whether there is a spatial dependence or not.

H_0 : Accident frequencies by spatial unit are not spatially correlated.

H_1 : Accident frequencies by spatial unit are spatially correlated.

4.1. Diagnosis of Spatial Dependence

The Moran's I test was run for different bandwidths and k nearest neighbor values. As Table 5, shows all the p -values were statistically significant and pass the five percent significance level criterion, and z -scores are positive and significantly higher than 1.96 (z -value for 95% confidence). Therefore, the null hypothesis is rejected, and the dataset is proven to be spatially clustered. In addition to that, since " I " exceeds $E(I) = -1/N$, it attests that there is a positive spatial autocorrelation. Figure 2 presents the Global Moran's I test summary and it also confirms that given the z -score of 4.888, there is less than 1% likelihood that this clustered pattern could be the result of random chance. Thus, it can be clearly substantiated that in the examined sample, a clear positive autocorrelation can be detected regarding personal injury road traffic accidents that occurred on the line network in 2016–2018 in the urban area of Budapest. Therefore, the result of the traditional OLS procedure cannot be considered due to the non-fulfillment of the Gaussian–Markov conditions [68], since autocorrelation can be detected in the error terms. Therefore, it is expedient to move to the spatial econometric approach, as it is suitable for dealing with spatial dependencies.

Table 5. Moran's I Test Result. (Source: Authors based on [54,63,66,67])

Type		Moran I	$E(I)$	$V(I)$	z-Value	p-Value
D (km)	1	3.66×10^{-2}	-2.12811×10^{-4}	6.67×10^{-5}	10.0090	$<2.2 \times 10^{-16}$
	0.75	3.66×10^{-2}	-7.77000×10^{-4}	2.41×10^{-5}	7.6060	2.828×10^{-14}
	0.5	4.95×10^{-2}	-9.75400×10^{-4}	5.04×10^{-5}	7.1088	1.17×10^{-12}
	0.25	5.05×10^{-2}	-1.30000×10^{-3}	1.72×10^{-4}	3.9444	8.00×10^{-5}
k nearest neighbor	1	8.71×10^{-2}	-1.17000×10^{-3}	3.09×10^{-4}	5.0199	5.17×10^{-7}
	10	5.30×10^{-2}	-7.35700×10^{-4}	3.62×10^{-5}	8.9820	$<2.2 \times 10^{-16}$
	25	3.59×10^{-2}	-2.12766×10^{-4}	1.45×10^{-5}	9.6163	$<2.2 \times 10^{-16}$
	40	2.93×10^{-2}	-6.35600×10^{-4}	9.01×10^{-6}	9.9616	$<2.2 \times 10^{-16}$

Figure 2. Global Moran's I Summary. (Source: Authors).

Since all Moran's I for regression residuals are significant and positive, the highest I (8.71×10^{-2}) based on k -nearest neighbor with k value of 1 is finally selected, and the spatial weight matrix is set up, which could give a better model. The matrix is row standardized. Figure 3 shows the Moran's I scatterplot for Euclidean distance and k nearest neighbors, respectively.

Figure 3. Moran's I Scatterplot for Two Weight Matrices: (a) Euclidean Distance; (b) k Nearest Neighbor. (Source: Authors).

Based on the Lagrange Multiplier test, the following outputs were recorded (Table 6). LMerr and LMlag show which spatial econometric model is ought to be used, RLMerr and RLMlag are the robust tests for the same, while SARMA indicates whether the SAC model could be used.

Table 6. Lagrange Multiplier Test Output. (Source: Authors based on [54,63,66,67]).

LM Test	Statistic	p -Value	Significance Codes
LMerr	59.64	1.14×10^{-14}	***
LMlag	55.02	1.19×10^{-13}	***
RLMerr	4.91	0.02664	*
RLMlag	0.29	0.58851	
SARMA	59.93	9.68×10^{-14}	***

The significance level is indicated in the tables by the following notations: stands for $p < 0.1$, * and *** stands for $p < 0.001$.

4.2. Model Outputs

Since both the LMerr and the LMlag are significantly different from zero, it is required to check their robust versions. The robust counterparts of the tests imply that the error model is the more likely to be used.

The F statistics (44.55) of the OLS model is highly significant, with p -value $< 2.26 \times 10^{-16}$ which infers that the model is acceptable.

Multiple models with different weight matrices for each technique (OLS, SAR, SEM, and SAC) were run. Finally, the champion models' parameter estimates, and goodness of fit results are summarized in Table 7.

Table 7. The Model's Parameter Estimates. (Source: Authors based on [54,63,66,67]).

Parameter	OLS		SAR		SEM		SAC	
	Estimate	t-Value	Estimate	z-Value	Estimate	z-Value	Estimate	z-Value
(Intercept)	-4.33×10^{-1}	(−0.956)	-4.67×10^{-1}	(−3.8636) ***	-4.71×10^{-1}	(−1.0454)	-4.72×10^{-1}	(−1.0462)
CAT_12	9.77×10^{-1}	(2.368) *	9.67×10^{-1}	(10.8092) ***	1.02×10^0	(2.5033) *	1.02×10^0	(2.4917) *
CAT_21	1.26×10^0	(3.07) **	1.25×10^0	(13.7577) ***	1.30×10^0	(3.1897) **	1.29×10^0	(3.1796) **
CAT_22	1.01×10^0	(2.501) *	1.00×10^0	(11.6113) ***	1.06×10^0	(2.6346) **	1.05×10^0	(2.624) **
CAT_23	1.47×10^0	(3.217) **	1.46×10^0	(6.4268) ***	1.51×10^0	(3.3261) ***	1.51×10^0	(3.3161) ***
CAT_24	1.02×10^0	(2.429) *	1.01×10^0	(7.3039) ***	1.05×10^0	(2.533) *	1.05×10^0	(2.5264) *
CAT_25	9.04×10^{-1}	(1.769).	9.01×10^{-1}	(2.826) **	9.40×10^{-1}	(1.852).	9.38×10^{-1}	(1.8466).
CAT_26	1.54×10^0	(3.188) **	1.51×10^0	(5.3031) ***	1.51×10^0	(3.1471) **	1.51×10^0	(3.1491) **
CAT_28	8.57×10^{-1}	(0.643)	8.37×10^{-1}	(0.6308)	9.89×10^{-1}	(0.7529)	9.78×10^{-1}	(0.7436)
CAT_31	8.51×10^{-2}	(0.101)	7.42×10^{-2}	NA	1.42×10^{-1}	(0.1683)	1.37×10^{-1}	(0.162)
CAT_32	7.73×10^{-1}	(1.847).	7.63×10^{-1}	(5.0947) ***	8.14×10^{-1}	(1.9556).	8.10×10^{-1}	(1.9459).
CAT_33	1.14×10^0	(2.571) *	1.12×10^0	(5.5661) ***	1.17×10^0	(2.6534) **	1.16×10^0	(2.6448) **
CAT_34	8.47×10^{-1}	(2.079) *	8.38×10^{-1}	(7.0821) ***	8.82×10^{-1}	(2.1846) *	8.79×10^{-1}	(2.1758) *
CAT_35	8.18×10^{-1}	(1.647).	8.24×10^{-1}	(2.803) **	8.78×10^{-1}	(1.7773).	8.74×10^{-1}	(1.7692).
CAT_36	4.62×10^{-1}	(1.087)	4.60×10^{-1}	(2.6302) **	4.79×10^{-1}	(1.1386)	4.78×10^{-1}	(1.1353)
BUS_LANE	5.30×10^{-1}	(5.344) ***	5.27×10^{-1}	(5.4219) ***	5.30×10^{-1}	(5.3421) ***	5.30×10^{-1}	(5.3438) ***
BUS_VOL	5.16×10^{-4}	(6.556) ***	5.20×10^{-4}	(6.662) ***	5.46×10^{-4}	(6.8316) ***	5.44×10^{-4}	(6.8093) ***
HGV_12_D	2.22×10^{-3}	(0.047)	4.95×10^{-3}	NA	1.55×10^{-2}	(0.3211)	1.46×10^{-2}	(0.3041)
BIKE_VOL	9.10×10^{-5}	(5.283) ***	8.86×10^{-5}	(5.1538) ***	8.61×10^{-5}	(4.9792) ***	8.63×10^{-5}	(4.9947) ***
LN_AADT	1.44×10^{-1}	(7.437) ***	1.43×10^{-1}	(8.0805) ***	1.43×10^{-1}	(7.305) ***	1.43×10^{-1}	(7.3103) ***
Ratio_HGV	-5.77×10^{-1}	(−2.318) *	-5.41×10^{-1}	(−2.2386) *	-5.22×10^{-1}	(−2.0826) *	-5.24×10^{-1}	(−2.0897) *
LN_LENGTH	9.48×10^{-1}	(24.499) ***	9.34×10^{-1}	(24.3398) ***	9.51×10^{-1}	(24.3371) ***	9.50×10^{-1}	(23.9841) ***
ρ	-	-	0.0638	(5.1107) ***	-	-	0.0058	(0.15)
λ	-	-	-	-	0.071	(5.3861) ***	0.0654	(1.6359)
AIC	15,601		15,576		15,574		15,576	
BIC	15,750		15,732		15,729		15,738	
Log-Likelihood	−7777.686		−7764.688		−7763.103		−7763.089	

The significance level is indicated in the tables by the following notations: stands for $p < 0.1$, * stands for $p < 0.05$, ** stands for $p < 0.01$, and *** stands for $p < 0.001$.

The robust counterparts of the Lagrange Multiplier test infer that the error model is preferable over the lag model. This is now proved with the *AIC*, *BIC*, and *LR* value (Table 8). *AIC* and *BIC* values of the SEM are less than those values of the SAR and SAC. Furthermore, the Log-Likelihood values (L_c) of the SEM are higher than those of the SAR and SAC. The SEM was finally selected to be the best one of the three spatial econometric models. *LR* values prove that SEM (and in general all of the spatial econometric models) is significantly better than the standard linear regression model. In coherence with the *AIC* and *BIC* values, the SAC model's significant difference cannot be proven in this case.

Table 8. The *LR* values of the different models. (Source: Authors).

	Inferior Model				Superior Model			<i>LR</i>	<i>df</i>	
	L_c	<i>AIC</i>	<i>BIC</i>		L_c	<i>AIC</i>	<i>BIC</i>			
OLS	−7777.686	15,601	15,750	SAR	−7764.688	15,576	15,732	25.996	1	***
OLS	−7777.686	15,601	15,750	SEM	−7763.103	15,574	15,729	29.166	1	***
OLS	−7777.686	15,601	15,750	SAC	−7763.089	15,576	15,738	29.194	2	***
SAR	−7764.688	15,576	15,732	SAC	−7763.089	15,576	15,738	3.198	1	
SEM	−7763.103	15,574	15,729	SAC	−7763.089	15,576	15,738	0.028	1	

The significance level is indicated in the tables by the following notations: *** stands for $p < 0.001$.

5. Discussion

For the SEM model, which is considered to be the best, the sign of all explanatory variables (except for the ratio of HGVs) is positive and significantly different from zero. This can be considered as obvious for the length of the examined sections or the average daily traffic because it means that the longer the given section, or the higher the traffic on it, the higher the accident frequency. For bicycle traffic volume, the same observation can be applied, as if cycling traffic is higher, the frequency of accidents also increases. This is incoherent with the models observed in Denmark or the Netherlands due to the fact that in Budapest, there is a few physically separated infrastructure for cyclists [69], and the volume of bicycle traffic has not reached the critical level where the trend is reversed, and the risk of an accident decreases as bicycle traffic increases.

In the case of the category variables describing the characteristics of the infrastructure, it can be stated that the minimum number of accidents occurred on the high-speed seven-lane roads (this was chosen as a reference among the category variables). Generally, the frequency of accidents is lower on the road developed for higher speed, which can be derived from that these sections have a higher level of safety (guardrail, and divided road).

In the number of lanes: between 20 and 50 kph, there is no clear correlation between the number of lanes and the accident frequency; at higher speed designed sections (above 50 kph), the accident frequency decreases for more than four lanes.

In the case of parameters describing the quality of road public transport, it can be observed that the increase in bus traffic increases the risk of accidents. This is due to the fact that buses in traffic perform maneuvers that other road users do not expect (approaching and leaving a bus stop, handling overhead lines in the case of trolleybuses, and taking into account insulated sections).

The same is right for bus lanes, which, with their special rules, contribute to an increase in the risk of accidents, typically due to conflict situations resulting from pre-turn lane changes.

In the case of HGV traffic, the properties of the parameters are different from the previous ones. On the one hand, whether HGVs (over 12 tons) can use a given road section has no significant effect on the frequency of accidents. On the other hand, the ratio of HGVs negatively affects the accident frequency, i.e., the higher the proportion of HGV traffic in total traffic, the lower the frequency of accidents. This is due to the fact that lorries are less dynamic in traffic, perform less unforeseen maneuvers, and likely to use more protected main roads.

If the spatial parameters are taken into consideration, both the SAR and SEM model's parameters are significant; however, this cannot be said about the SAC model. Both SAR and SEM models have a positive spatial parameter, which means if a section's accident frequency is high, then the closest one's frequency will also be increased. This indicates that improvements for longer road sections seem to be beneficial in order to avoid traffic accidents. In the case of SAC models, both spatial parameters are not significantly differentiated from zero, which can be due to the fact that the two approaches (spatial error and delay) effect cancel each other out. Since the spatial parameters are not significant in this case, the effects described above are not fulfilled in this case.

The best model with a significantly positive spatial parameter (λ) is the SEM approach. This indicates that the closest infrastructure elements have a positive effect on each other, which could be eliminated from the model. These effects can be derived from the urban structure and the geographical properties affecting the city. These parameters are hard to take into consideration in a classic linear regression; thus, it can be stated that the spatial econometric approach is a more effective technique for modeling road accident frequency.

6. Conclusions

In this study, road traffic accidents occurring on the road links of Budapest over a three-year-long period were surveyed. The accidents on the same link were aggregated, and centroids of each road link were considered for the formulation of three different types of models. The global Moran's I test results with significant p -value and positive z -score confirm the presence of spatial dependence of road traffic accidents. OLS, SAR, SEM, and SAC models were formulated in this study. The results proved that incorporating spatial effect in the model gives a better result than the traditional multiple linear regression model. The Lagrange–Multiplier test, in conjunction with the AIC , BIC , and Log-Likelihood values proved that the spatial error model gives the best output over the remaining two spatial econometric models in this study. Confirmation of spatial dependence among the data; illustration of better performance of the spatial models over traditional OLS models; and identification of significant contributors or predictors of the road traffic accidents are the main findings of this study.

The spatial econometric approach allows the development of a universal and simple model that will significantly assist decision-makers in the development of road safety investments. The role of spatial econometrics is to include difficult or non-descriptive parameters in the model, which allows for a more accurate prediction of accident frequencies. However, it is essential to update these models every 3–5 years, as there may be changes in the urban structure that will fundamentally change traffic safety conditions.

This study's main limitation is that the analysis only takes into consideration the city of Budapest. Furthermore, it considered accident events on the road links only; those at the intersections are not considered. The third limitation is that accident severity was not considered; instead all accidents that occurred on the road links over a period are aggregated.

These limitations define the future research directions. Three main directions could be identified: (i) choosing a different environment rather than Budapest, such as an analysis on the whole Hungarian network; (ii) take into consideration the accidents of the junctions in a separate model, or a joint one as well; (iii) taking into consideration the accident types and the accident results. Developing such spatial econometric models of good performance and discovering the most significant parameters would help authorities identify the specific interventions in improving road safety and ultimately ensure sustainable transport and urban development.

Author Contributions: Conceptualization, T.S., A.A.M. and Z.S.; methodology, T.S., A.A.M. and Z.S.; software, T.S., A.A.M. and Z.S.; validation, T.S., A.A.M. and Z.S.; formal analysis, T.S., A.A.M. and Z.S.; investigation, T.S., A.A.M. and Z.S.; resources, T.S., A.A.M. and Z.S.; data curation, T.S., A.A.M. and Z.S.; writing—original draft preparation, T.S., A.A.M. and Z.S.; writing—review and editing, T.S., A.A.M. and Z.S.; visualization, T.S., A.A.M. and Z.S.; supervision, T.S., A.A.M. and Z.S.; project

administration, T.S., A.A.M. and Z.S.; funding acquisition, T.S., A.A.M. and Z.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from BKK Zrt. (Budapesti Közlekedési Központ—Centre for Budapest Transport) and are available from the authors with the permission of BKK Zrt.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jones, S.; Lidbe, A.; Hainen, A. What can open access data from India tell us about road safety and sustainable development? *J. Transp. Geogr.* **2019**, *80*, 102503. [\[CrossRef\]](#)
2. Llewellyn, R.; Cowie, J.; Maher, M. Active Road Studs as an Alternative to Lighting on Rural Roads: Driver Safety Perception. *Sustainability* **2020**, *12*, 9648. [\[CrossRef\]](#)
3. Skoufas, A.; Basbas, S.; Salanova Grau, J.M.; Aifadopoulou, G. Analysis of In-Vehicle Warning System for Rail-Road Level Crossings: Case Study in the City of Thessaloniki, Greece. *Period. Polytech. Transp. Eng.* **2020**, *49*, 42–59. [\[CrossRef\]](#)
4. Lu, Q.; Tettamanti, T.; Hörcher, D.; Varga, I. The impact of autonomous vehicles on urban traffic network capacity: An experimental analysis by microscopic traffic simulation. *Transp. Lett.* **2020**, *12*, 540–549. [\[CrossRef\]](#)
5. Astarita, V.; Caliendo, C.; Giofrè, V.P.; Russo, I. Surrogate Safety Measures from Traffic Simulation: Validation of Safety Indicators with Intersection Traffic Crash Data. *Sustainability* **2020**, *12*, 6974. [\[CrossRef\]](#)
6. Wang, S.; Chen, Y.; Huang, J.; Chen, N.; Lu, Y. Macrolevel Traffic Crash Analysis: A Spatial Econometric Model Approach. *Math. Probl. Eng.* **2019**, *2019*, 1–10. [\[CrossRef\]](#)
7. Hong, J.; Lee, S.; Lim, J.; Kim, J. Application of Spatial Econometrics Analysis for Traffic Accident Prediction Models in Urban Areas. In Proceedings of the Eastern Asia Society for Transportation Studies, Taipei, Taiwan, 9–12 September 2013; Volume 8.
8. Rhee, K.-A.; Kim, J.-K.; Lee, Y.; Ulfarsson, G.F. Spatial regression analysis of traffic crashes in Seoul. *Accid. Anal. Prev.* **2016**, *91*, 190–199. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Griffith, D. *Spatial Autocorrelation: A Primer*; Association of American Geographers: Washington, DC, USA, 1987.
10. Shankar, V.; Mannering, F.; Barfield, W. Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accid. Anal. Prev.* **1995**, *27*, 371–389. [\[CrossRef\]](#)
11. Briz-Redón, Á.; Martínez-Ruiz, F.; Montes, F. Spatial analysis of traffic accidents near and between road intersections in a directed linear network. *Accid. Anal. Prev.* **2019**, *132*, 105252. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Ng, K.; Hung, W.; Wong, W. An algorithm for assessing the risk of traffic accident. *J. Saf. Res.* **2002**, *33*, 387–410. [\[CrossRef\]](#)
13. Noland, R.B.; Quddus, M.A. Congestion, and safety: A spatial analysis of London. *Transp. Res. Part A Policy Pract.* **2005**, *39*, 737–754. [\[CrossRef\]](#)
14. Lu, J.J.; Xing, Y.; Wang, C.; Cai, X. Risk factors affecting the severity of traffic accidents at Shanghai river-crossing tunnel. *Traffic Inj. Prev.* **2016**, *17*, 176–180. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Casado-Sanz, N.; Guirao, B.; Attard, M. Analysis of the Risk Factors Affecting the Severity of Traffic Accidents on Spanish Crosstown Roads: The Driver's Perspective. *Sustainability* **2020**, *12*, 2237. [\[CrossRef\]](#)
16. Cantillo, V.; Garcés, P.; Márquez, L. Factors influencing the occurrence of traffic accidents in urban roads: A combined GIS-Empirical Bayesian approach. *DYNA* **2016**, *83*, 21–28. [\[CrossRef\]](#)
17. Quddus, M.A. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. *Accid. Anal. Prev.* **2008**, *40*, 1486–1497. [\[CrossRef\]](#)
18. Wang, C.; Quddus, M.A.; Ison, S.G. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. *Accid. Anal. Prev.* **2009**, *41*, 798–808. [\[CrossRef\]](#)
19. Castro, M.; Paleti, R.; Bhat, C.R. A spatial generalized ordered response model to examine highway crash injury severity. *Accid. Anal. Prev.* **2013**, *52*, 188–203. [\[CrossRef\]](#)
20. Xie, K.; Ozbay, K.; Yang, H. Spatial analysis of highway incident durations in the context of Hurricane Sandy. *Accid. Anal. Prev.* **2015**, *74*, 77–86. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Jia, R.; Khadka, A.; Kim, I. Traffic crash analysis with point-of-interest spatial clustering. *Accid. Anal. Prev.* **2018**, *121*, 223–230. [\[CrossRef\]](#)
22. Pljakić, M.; Jovanović, D.; Matović, B.; Mičić, S. Macro-level accident modeling in Novi Sad: A spatial regression approach. *Accid. Anal. Prev.* **2019**, *132*, 105259. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Soro, W.L.; Zhou, Y.; Wayoro, D. Crash rates analysis in China using a spatial panel model. *IATSS Res.* **2017**, *41*, 123–128. [\[CrossRef\]](#)

24. Castro, M.; Paleti, R.; Bhat, C.R. A latent variable representation of count data models to accommodate spatial and temporal dependence: Application to predicting crash frequency at intersections. *Transp. Res. Part B Methodol.* **2012**, *46*, 253–272. [\[CrossRef\]](#)
25. Ha, H.-H.; Thill, J.-C. Analysis of traffic hazard intensity: A spatial epidemiology case study of urban pedestrians. *Comput. Environ. Urban Syst.* **2011**, *35*, 230–240. [\[CrossRef\]](#)
26. Ryder, B.; Dahlinger, A.; Gahr, B.; Zundritsch, P.; Wortmann, F.; Fleisch, E. Spatial prediction of traffic accidents with critical driving events—Insights from a nationwide field study. *Transp. Res. Part A Policy Pract.* **2019**, *124*, 611–626. [\[CrossRef\]](#)
27. Lee, J.; Gim, T.-H.T. A spatial econometrics perspective on the characteristics of urban traffic accidents: Focusing on elderly drivers' accidents in Seoul, South Korea. *Int. J. Inj. Control Saf. Promot.* **2020**, *27*, 520–527. [\[CrossRef\]](#)
28. Xu, X.; Xie, S.; Wong, S.C.; Xu, P.; Huang, H.; Pei, X. Severity of pedestrian injuries due to traffic crashes at signalized intersections in Hong Kong: A Bayesian spatial logit model. *J. Adv. Transp.* **2016**, *50*, 2015–2028. [\[CrossRef\]](#)
29. Al-Hasani, G.; Asaduzzaman, M.; Soliman, A.-H. Comparison of Spatial Regression Models with Road Traffic Accidents Data. In Proceedings of the International Conference on Statistics: Theory and Applications (ICSTA'19), Lisbon, Portugal, 13–14 August 2019.
30. Saeed, T.U.; Nateghi, R.; Hall, T.; Waldorf, B.S. Statistical Analysis of Area-wide Alcohol-related Driving Crashes: A Spatial Econometric Approach. *Geogr. Anal.* **2020**, *52*, 394–417. [\[CrossRef\]](#)
31. Zhang, Y.; Lu, H.; Qu, W. Geographical Detection of Traffic Accidents Spatial Stratified Heterogeneity and Influence Factors. *IJERPH* **2020**, *17*, 572. [\[CrossRef\]](#)
32. Lee, D.; Guldman, J.-M.; von Rabenau, B. Interactions between the built and socio-economic environment and driver demographics: Spatial econometric models of car crashes in the Columbus Metropolitan Area. *Int. J. Urban Sci.* **2018**, *22*, 17–37. [\[CrossRef\]](#)
33. Wang, W.; Yuan, Z.; Yang, Y.; Yang, X.; Liu, Y. Factors influencing traffic accident frequencies on urban roads: A spatial panel time-fixed effects error model. *PLoS ONE* **2019**, *14*, e0214539. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Chiou, Y.-C.; Fu, C.; Chih-Wei, H. Incorporating spatial dependence in simultaneously modeling crash frequency and severity. *Anal. Methods Accid. Res.* **2014**, *2*, 1–11. [\[CrossRef\]](#)
35. Liu, C.; Sharma, A. Exploring spatio-temporal effects in traffic crash trend analysis. *Anal. Methods Accid. Res.* **2017**, *16*, 104–116. [\[CrossRef\]](#)
36. Simões, P.; Shrubbsall, S.; Natário, I. A spatial econometrics analysis for road accidents in Lisbon. *Int. J. Bus. Intell. Data Min.* **2015**, *10*, 152–173. [\[CrossRef\]](#)
37. Black, W.R. Highway Accidents: A Spatial and Temporal Analysis. *Transp. Res. Rec.* **1991**, *1318*, 75–82.
38. Karaganis, A.N.; Mimis, A. A Spatial Point Process for Estimating the Probability of Occurrence of a Traffic Accident. In Proceedings of the 46th Congress of the European Regional Science Association: “Enlargement, Southern Europe and the Mediterranean”, Volos, Greece, 30 August–3 September 2006.
39. Azimian, A.; Pyrialakou, D. Exploratory Spatial Data Analysis in Traffic Safety. *Int. J. Geospat. Environ. Res.* **2020**, *7*, 4.
40. Moons, E.; Brijs, T.; Wets, G. Hot Spot Analysis: Improving a Local Indicator of Spatial Association for Application in Traffic Safety. In *Computational Science and Its Applications—ICCSA 2008*; Gervasi, O., Murgante, B., Laganà, A., Taniar, D., Mun, Y., Gavrilova, M.L., Eds.; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5072, pp. 221–231. [\[CrossRef\]](#)
41. Korter, G.O. A Spatial Analytical Framework for Examining Road Traffic Crashes. *J. Mod. App. Stat. Meth.* **2016**, *15*, 570–583. [\[CrossRef\]](#)
42. Steenberghen, T.; Dufays, T.; Thomas, I.; Flahaut, B. Intra-urban location and clustering of road accidents using GIS: A Belgian example. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 169–181. [\[CrossRef\]](#)
43. Li, X.; Dadashova, B.; Yu, S.; Zhang, Z. Rethinking Highway Safety Analysis by Leveraging Crowdsourced Waze Data. *Sustainability* **2020**, *12*, 10127. [\[CrossRef\]](#)
44. Besag, J. Spatial Interaction and the Statistical Analysis of Lattice Systems. *J. R. Stat. Soc. Ser. B (Methodol.)* **1974**, *36*, 192–236. [\[CrossRef\]](#)
45. Zhukov, Y.M. Applied Spatial Statistics in R, Section 6. Geostatistics 2010. Available online: <http://www.people.fas.harvard.edu/~zhukov/Spatial6.pdf> (accessed on 5 December 2019).
46. Besag, J.; York, J.; Mollié, A. Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Stat. Math.* **1991**, *43*, 1–20. [\[CrossRef\]](#)
47. Ágoston, G.; Madleňák, R. Road Safety Macro Assessment Model: Case Study for Hungary. *Period. Polytech Transp. Eng.* **2020**, *49*, 89–92. [\[CrossRef\]](#)
48. Kiss, Á.O.; Sesztakov, V.; Török, Á. Road Safety Analysis in Győr. *Period. Polytech Transp. Eng.* **2013**, *41*, 51–56. [\[CrossRef\]](#)
49. Rudolph, F.; Mátrai, T. Congestion from a Multimodal Perspective. *Period. Polytech. Transp. Eng.* **2018**, *46*, 215–221. [\[CrossRef\]](#)
50. Zhao, X.; Lord, D.; Peng, Y. Examining Network Segmentation for Traffic Safety Analysis with Data-Driven Spectral Analysis. *IEEE Access* **2019**, *7*, 120744–120757. [\[CrossRef\]](#)
51. Ziakopoulos, A.; Yannis, G. A review of spatial approaches in road safety. *Accid. Anal. Prev.* **2020**, *135*, 105323. [\[CrossRef\]](#)
52. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988.
53. Varga, A. Térökonometria. *Stat. Szle.* **2002**, *80*, 354–370.
54. Sarmiento-Barbieri, I. An Introduction to Spatial Econometrics in R. 2016. Available online: http://www.econ.uiuc.edu/~lab/workshop/Spatial_in_R.html (accessed on 11 July 2018).

-
55. Szabó, Z.; Török, Á. Spatial Econometrics—Usage in Transportation Sciences: A Review Article. *Period. Polytech Transp. Eng.* **2019**, *48*, 143–149. [[CrossRef](#)]
 56. Getis, A. Spatial interaction and spatial autocorrelation: A cross-product approach. *Environ. Plan. A* **1991**, *23*, 1269–1277. [[CrossRef](#)]
 57. Moran, P.A.P. Some Theorems on Time Series: II the Significance of the Serial Correlation Coefficient. *Biometrika* **1948**, *35*, 255–260. [[CrossRef](#)]
 58. Anselin, L.; Bera, A.K.; Florax, R.; Yoon, M.J. Simple diagnostic tests for spatial dependence. *Reg. Sci. Urban Econ.* **1996**, *26*, 77–104. [[CrossRef](#)]
 59. Cameron, A.C.; Trivedi, P.K. *Microeconometrics: Methods and Applications*; Cambridge University Press: New York, NY, USA, 2005.
 60. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **1981**, *17*, 368–376. [[CrossRef](#)] [[PubMed](#)]
 61. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Wien, Austria, 2017.
 62. Bivand, R.S.; Lewin-Koh, N. *Maptools: Tools for Reading and Handling Spatial Objects*; R Foundation for Statistical Computing: Wien, Austria, 2017.
 63. Bivand, R.S.; Pebesma, E.J.; Gómez-Rubio, V. *Applied Spatial Data Analysis with R*, 2nd ed.; Springer: New York, NY, USA, 2013.
 64. Pebesma, E.J.; Bivand, R.S. Classes and methods for spatial data in R. *R News* **2005**, *5*, 9–13.
 65. Bivand, R.S.; Wong, D.W.S. Comparing implementations of global and local indicators of spatial association. *TEST* **2018**, *27*, 716–748. [[CrossRef](#)]
 66. Bivand, R.S.; Hauke, J.; Kossowski, T. Computing the Jacobian in Gaussian spatial autoregressive models an illustrated comparison of available methods. *Geogr. Anal.* **2013**, *45*, 150–179. [[CrossRef](#)]
 67. Bivand, R.S.; Piras, G. Comparing Implementations of Estimation Methods for Spatial Econometrics. *J. Stat. Softw.* **2015**, *63*, 1–36. [[CrossRef](#)]
 68. Maddala, G.S. *Introduction to Econometrics*, 3rd ed; John Wiley&Sons Ltd.: Chichester, UK, 2001.
 69. Wang, H.; Chang, S.K.J.; de Backer, H.; Lauwers, D.; de Maeyer, P. Integrating Spatial and Temporal Approaches for Explaining Bicycle Crashes in High-Risk Areas in Antwerp (Belgium). *Sustainability* **2019**, *11*, 3746. [[CrossRef](#)]