Sigma J Eng & Nat Sci 38 (4), 2020, 1667-1680

Sigma Journal of Engineering and Natural Sciences
Sigma Mühendislik ve Fen Bilimleri Dergisi

Research Article

SPATIAL ANALYSIS OF THE ROAD TRAFFIC ACCIDENT STATISTICS IN THE PROVINCES OF TURKEY

Vural YILDIRIM[1], Yeliz MERT KANTAR*[2]

[1]*Institute of Earth and Space Sciences, Eskisehir Technical University, ESKISEHIR;*
ORCID: 0000-0002-6517-7849
[2]*Department of Statistics, Eskisehir Technical University, ESKISEHIR;* ORCID: 0000-0001-7101-8943

Received: 26.03.2020 Revised: 13.07.2020 Accepted: 23.07.2020

ABSTRACT

The aim of the study is to describe and model the spatial distribution of the road traffic accidents (RTAs) rate with the factors considering space–time relationship for the period between 2013–2018 in the provinces of Turkey. The RTA rate is modelled with the factors which are population, the number of different types of motor vehicles registered, lengths of three types of provincial roads and these factors measured in the neighbours. Firstly, spatial maps are used to demonstrate the spatial variability of RTAs in Turkey. Global and local spatial autocorrelation analyses are conducted to demonstrate whether the RTA in provinces with high rates of accidents show similar clusters. Spatial regression and panel models are considered a solution to examine the space–time relationship between the RTA and RTA's neighbourhood characteristics. We found that the RTA rate is not distributed randomly across Turkey. Spatial distribution of provinces with high rates of accidents is non-random (Moran's I changes between 0.52 and 0.59 with p <0.001). Moreover, while LISA analysis demonstrates the provinces determined as local clusters, the fixed effects models with different spatial structures show that the RTA rate is positively correlated with number of cars, vans, private vehicles and length of asphalt roads, other factors are negatively correlated and also non-asphalt road is not significant to explain the RTA rate. On the other hand, spatial parameters are significant in all models (p <0.1) and neighbouring region characteristics in terms of explanatory variables do not affect the explanation upon the RTA rate.
Keywords: Road traffic accident (RTA), spatial statistics, local indicators of spatial association, spatial panel econometrics, Turkey.

1. INTRODUCTION

Road traffic accidents (RTAs) have been identified as "human-caused natural disaster" by the World Health Organization. RTAs in most of the countries increase due to Human' mistakes (inattentive driving), traffic rule violations, increase of the vehicles in traffic, improper use of heavy vehicles on roads, environmental factors and unfavourable climatic conditions. RTAs in turkey often occur as a result of road infrastructure problems and traffic violations [1]. Turkey is a bridge between Asia and Europe [2-3]. Although the population was estimated as 72.75 million in 2010, and now it is over 82 million. These statistical results support that without any fall or

* Corresponding Author: e-mail: ymert@eskisehir.edu.tr, tel: (222) 321 35 50 / 4602

breakage the population of the country kept growing from the 19th century until now. This strong increase causes an increase in vehicle numbers and thus RTAs. The number of driver's license holders was about 30 million in 2017. This number increases every year. The parameters of RTA, which cause accidents , are numerous and have a complex structure and just because of this the researchers interest in analysing them for the purpose of understanding, comparing, modelling and demonstrating accidents. Particularly, spatial factors and also models are main interest recently. For example, Levine et al. [4] use a spatial lag model to investigate the spatial patterns of motor vehicle accidents in Honolulu, Hawaii, during the year 1990. Erdogan [3] describes the inter-province differences in traffic accidents and mortality on roads of Turkey by using spatial analysis. While Papadimitriou et al. [5] also consider another Spatial lag model to model the spatial variation of crash rates in Greece for the year 2002, LaScala et al. [6] propose a spatial error model to map locations of pedestrian injuries in San Francisco, California, for the year 1990 and the analysis supported that the geographical proximity has an impact on the traffic safety of spatial units. Network kernel density estimation with local Moran's I for hot spot detection (an area with higher concentration of RTAs) of traffic accidents is used in [7]. In the framework of spatial econometrics, crash rates analysis in China using a spatial panel model is conducted in Soro et al. [8] and it is found that freight traffic, the length of paved roads and the populations of age 65 and above are related to higher rates in RTA while the opposite trend is observed for the Gross Regional Product, the urban unemployment rate and passenger traffic. Lakes [9] studies a spatially explicit analysis of traffic accidents involving pedestrians and cyclists in Berlin and, the spatial distribution analysis shows, however, that there are significant spatial clusters (hot spots) of traffic accidents with a strong concentration in the inner city area. Daniel et al. [10] studies RTA in Nigeria with spatial analysis tools and they observe that there is significant clustering of RTA occurrence and death in the Federal capital territory and Nasarawa state. Hot spot analysis based on network spatial weights to determine spatial statistics of traffic accidents in Rize, Turkey are conducted in [1]. In [11], the RTAs statistics in Turkey show that drivers are responsible for the highest rate of defects. Factors influencing traffic accident frequencies on urban roads are studied with a spatial panel time-fixed effects error model in [12]. Akgungor and Dogan [13] use an artificial neural network model and a genetic algorithm model to estimate the number of accidents, fatalities and injuries in Ankara, Turkey, utilizing the data obtained between 1986 and 2005. While the number of people killed in RTAs in Turkey between 1990 and 2014 is modelled as the function of time by linear regression in [14], RTAs within last ten years in Turkey with an analysis of factors is described in [15]. Also, some methods are used to analyse the annual average daily traffic, the number of accident, injury and death, average velocity, distance, the number of links and junctions parameters in road between Erzincan and Agri (Gurbulak) cities in Turkey [16]. Dereli and Erdogan [17] aim to study traffic accidents in Turkey by means of GIS and statistical methods. GIS based surveillance of RTAs risk for Rawalpindi city is conducted in [18]. The occurrences and frequencies of RTAs in Hosanna town, Ethiopia, are discussed in [19]. As a result of the literature review, it can be claimed that RTA with multi-parameters, one of an important public health issue, should be studied with spatial statistics tools.

Thus, the study is to designed to describe and model the spatial distribution of the road traffic accidents (RTAs) rate with the factors considering space–time relationship for the period between 2013–2018 in the provinces of Turkey. The RTA rate is modelled with the factors which are *i.* population *ii.* the number of different types of motor vehicles registered, *iii.* Lengths of three types of provincial roads. Spatial maps show the spatial variability of RTAs in Turkey. It is a common practice to use the rate concerning the factor in order to compare regions or cities in terms of the certain factor. For this reason, the RTA rate, calculated by the number of RTAs /the population of cities, is used in analysis. Global (Moran's I) and local spatial autocorrelation (LISA) analyses applied to show whether provinces with high RTA show similar clusters. Spatial panel models are considered a solution to examine the space–time relationship between the RTA and RTA's neighbourhood characteristics. We find that the RTA rate is not distributed randomly across Turkey. The spatial distribution of provinces with high rates of accidents is non-random and it seen as globally clustered with significance of $p< 0.05$, where Moran's I changes between 0.52 and 0.59 with $p<0.001$). Moreover, LISA analysis shows the provinces determined as local clusters in terms of the RTA rate (with a level of significance of 95%). As a result, detecting the spatial pattern of RTA is necessary so that obtain reliable parameter estimates via panel models. Firstly, we estimated the pooled, random-effect and fixed-effects panel models and found that the fixed effects models should be preferred according to the Hausman specification tests. Thus, the fixed effects models with different spatial structures are operated to model the RTA rate. The results showed that the RTA rate is positively correlated with number of cars, vans, private vehicles and length of asphalt roads, other factors are negatively correlated and also non-asphalt road is not significant to explain the RTA rate. On the other hand, spatial parameters are significant usually in all models ($p<0.1$) and lagged-independent variables (neighbouring region characteristics in terms of explanatory variables) do not affect the explanation upon the RTA rate. For all computations and maps, In *R* software, *spdep* and *qqplot2* library are used for visualization of variables and spatial analyses, furthermore spatial panel analyses and econometrics analysis are conducted with *plm* and *splm* library.

2. DATA

RTA data is taken form Turkish Statistical Institute for the period of 2013-2018. Data is collected at province level. Map of Turkey showing the 81 provinces is provided in Fig. 1 to see the place of provinces. The maps displayed in Fig. 2 exhibit spatial distribution of the RTAs rate across all Turkish provinces. These maps enable us to clearly see the changes of the RTA rate in provinces between the years of 2013 and 2018. Specifically, while sub-regions in the west has the highest RTA rate, indicated with dark red area in the map, light red shows lowest rate. The higher RTA rate is still observed in Mugla in the south west part of Turkey for 2013-2014. As well as Mugla, Kilis has the highest RTA rate for 2014. These findings show the highest RTA rate of the most populated cities, Izmir, Istanbul, Bursa and, Ankara, which is the capital of Turkey, between the years 2015 and 2018. Istanbul is the world 8*th* largest city and has been inhabited since around 3000 BCE.

Figure 1. Turkey provinces maps

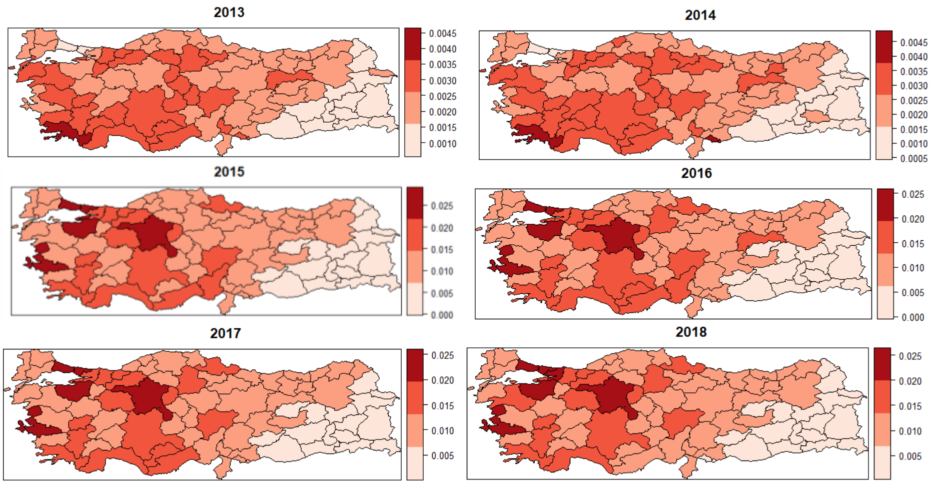

Figure 2. Natural breaks maps for traffic accidents for 2013-2018

Several variables proposed in the literature are considered, which affect the level of the RTA. In this study, the variables are selected according to the availability of data in Turkey at the level of 81 provinces for the entire period between 2013-2018. While RTA data used in the analysis are taken from Turkish Statistical Institute, data concerning lengths of three types of provincial roads is provided from General Directorate of Highways.

3. METHODS

Spatial econometrics include methods for spatial interaction (spatial autocorrelation or spatial dependency) and spatial structure (spatial heterogeneity) [20-22], while spatial statistics include both descriptive and inferential methods used for the analysis of georeferenced data. Both spatial statistics and spatial econometric models need quantifying the neighbour structure of the spatial units. This spatial structure is expressed by a spatial weights matrix W, whose element is w_{ij}. In other words, spatial structure between *n* spatial units (regions) is summarized by spatial weight matrix. Although there are many types of the spatial weight matrix W, rook, queen and *k* nearest neighbours continuity are well-known and mostly used matrix [23-29].

The spatial weight matrix, the queen matrix which defines all observations that shares common boundaries or vertices as neighbours, is given. The elements (w_{ij}) of the queen matrix are determined as

$$w_{ij} = \begin{cases} 1, & units\ i\ and\ j\ are\ neighbours \\ 0, & units\ i\ and\ j\ are\ not\ neighbours \end{cases}$$

The element w_{ij} $(i, j = 1, \dots, n)$ reflects the spatial influence of unit j on unit i, while w_{ii} always equals zero.

3.1. Spatial Autocorrelation

Two types of spatial autocorrelation are well-known: *i*. Global autocorrelation statistics, which demonstrate the spatial associations over the whole region, *ii*. Local autocorrelation measures, which are used traditional hotspot detection methods or finding local clusters. In the following, we briefly introduce both global and local spatial association measures.

Global spatial autocorrelation statistics

The well-known statistic to measure global spatial autocorrelations is the Moran's I, given as follows:

$$I = \frac{n}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \times \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

where n is the total number of spatial observations (i.e. districts), x_i and x_j are the values corresponding to the spatial districts i and j respectively. $\bar{x}$ is the mean value of x_i, i=1,…, n of all spatial districts, w_{ij} is an element of spatial weight matrix which shows spatial interaction between locations i and j. There is positive spatial autocorrelation when high values of one variable at one location are associated with high values at the neighbouring location. Generally, if the Moran's I is more/less than 0 and also 0, there is positive/negative autocorrelation and also no spatial autocorrelation.

Local Spatial autocorrelation analysis

While the global measures of spatial autocorrelation emphasize average spatial dependence over whole region or map, local measures of spatial association aim identifying patterns of spatial dependence within the sub-regions [26-28]. Although there are different measures for local spatial associations, the well-known one is the Local Moran's I proposed by [20] and its formula is given as follows:

$$I_i = n \times \frac{x_i \sum_{j=1}^{n} w_{ij}(x_j - \bar{x})}{\sum_{j=1}^{n}(x_j - \bar{x})^2}$$

While large positive I_i values point local cluster around the i-th location, large negative I_i values indicate that the data value sign at the th position is the opposite of its neighbors [26].

3.2. Local Spatial Clustering: LISA

Local Moran I is used to define local indicators of spatial relationship (LISA). LISA requires two conditions: (a) indicates the degree of significant spatial clustering for each location; (b) the sum of local statistics is proportional to a global indicator of spatial relationship [21-23]. The local spatial autocorrelation measure is defined as the presence of deviations from global spatial association patterns and "hotspots" such as local clusters or local outliers [30].

3.3. Spatial Panel Models

Spatial panel (SP) data contain time series observations for each geographical unit, they typically provide more information and variability as compared to one-dimensional data. Because of having N cross-section and T time series observations, panel data has many advantages such as giving more degrees of freedom and more information and having control between individual or time heterogeneity [31]. Thus, the SP model controls for both spatial and time effects. Three panel models (pooled, fixed effects, random effects) are extensively used in the literature. A fixed effect models check over whether intercepts change across group or time period, whereas a random effect models focus on differences in error variance components across individual or time period [32]. A two-way fixed model considers two sets of dummy variables. These commonly used panel models in applied researches are extended to include spatial error autocorrelation or a spatially lagged dependent or spatially lagged independent variables [33-35] which are briefly introduced as follows:

Spatial Pooled Model

The pooled linear regression model with spatial specific effects is compactly given as

$$Y_t = \rho W Y_t + X_t \beta + W X_t \theta + \alpha t_N + u_t$$
$$u_t = \lambda W u_t + \varepsilon_t$$

The model can be expressed with the time index (*i*) and the cross-sectional dimension (*t*),

$$y_{it} = \rho \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \sum_{i \neq j} w_{ij} x_{jt}\theta + \alpha + u_{it}$$
$$u_{it} = \lambda \sum_{i \neq j} w_{ij} y u_{jt} + \varepsilon_{it}$$

where *i* is an index for the cross-sectional dimension (spatial units), with *i*=1,...,*N*, and *t* is an index for the time dimension (time periods), with *t*=1,...,T and y_{it} is an observation on the dependent variable at *i* and *t*, y_{it} an vector of observations on the independent variables, and $\boldsymbol{\beta}$ is a vector of unknown parameters. ε_{it} is an independently and identically distributed error term for *i* and *t* with zero mean and constant variance, u_{it} is spatially auto-correlated error term.

Table 1 shows the spatial coefficient, spatial models, related information. The all considered spatial panel models (pooled, fixed and random) in this study are expressed with spatial effect or spatial models such as SLM, SEM, SDM and GNSM.

Table 1. The relationships different spatial dependence

Spatial Coefficient			Spatial Models		
Symbols	Definition	Meaning	Coefficients	Models	Abbreviation
ρ	Coefficients for spatially lagged dependent variable	Show endogenous interaction effects	$\rho \neq 0, \lambda = 0, \theta = 0$	Spatial lag model	SLM
λ	Spatial autoregressive coefficient	Show interaction effects among the error terms	$\rho = 0, \lambda \neq 0, \theta = 0$	Spatial error model	SEM
θ	Coefficients for spatially lagged dependent variable	Show exogenous interaction effects	$\rho = 0, \lambda = 0, \theta \neq 0$	Spatial Durbin Model	SDM
			$\rho \neq 0, \lambda \neq 0, \theta \neq 0$	General Nesting Spatial Model	GNSM
			$\rho = 0, \lambda = 0, \theta = 0$	Classic model	OLS

Spatial Fixed Model

In the fixed effects models, a dummy variable is included in the model. Fixed effects panel model with spatial effect are given as follows:

$$y_{it} = \rho \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \sum_{i \neq j} w_{ij} x_{jt}\theta + \alpha_i + u_{it}$$
$$u_{it} = \lambda \sum_{i \neq j} w_{ij} u_{jt} + \varepsilon_{it}$$
$$\alpha_i = (\alpha + \lambda_i)$$

Spatial and time effects in fixed effects models is provided by

$$y_{it} = \rho \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \sum_{i \neq j} w_{ij} x_{jt}\theta + \alpha_i + \alpha_t + u_{it}$$
$$u_{it} = \lambda \sum_{i \neq j} w_{ij} y u_{jt} + \varepsilon_{it.}$$

Spatial Random Effect Model

In the random effects model, it is assumed that the random variables u_i and ε_{it} are independent of each other.

Random effect panel model with spatial effects is presented as follows:

$$y_{it} = \rho \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \sum_{i \neq j} w_{ij} x_{jt}\theta + \alpha + u_{it}$$
$$u_{it} = \alpha_i + \lambda \sum_{i \neq j} w_{ij} y u_{jt} + \varepsilon_{it}$$

4. RESULTS

The spatial autocorrelation analysis of the RTA rate reveals global spatial clustering at the provincial level of Turkey, as seen Table 2. The queen weight matrix is used to define the spatial relationships between provinces. Table 2 summarizes the global Moran's I statistics and the related *p*-values for the corresponding years for the RTA. The RTAs exhibit significant spatial autocorrelation pattern for each year. All global Moran's I statistics are positive and significant at 0.001 level and suggests that the RTA rates have the clustered pattern in Turkey.

Table 2. Global Moran's I value for the spatial autocorrelation of the RTA rate in Turkey

Years	Moran's I	p-value
2013	0.5987499	< 0.001
2014	0.5201189	< 0.001
2015	0.5693021	< 0.001
2016	0.5724939	< 0.001
2017	0.5817596	< 0.001
2018	0.5744464	< 0.001

p-values in parentheses

The observed significant positive spatial autocorrelations from 2013 until 2018 indicated that the RTA rate remained unchanged at provincial level in Turkey. Thus, in order to further detect local autocorrelation (or the contribution of each province to global spatial autocorrelation), LISA (local correlation analyses) are performed [26]. Hot spots, in red color, and cold spots, in blue, is observed via LISA maps (Fig. 3). Pink and purple regions show spatial outliers (H-L and L-H), respectively. Provinces shown in white do not indicate absence or presence of any spatial dependency in figures (see Fig. 3). Sum up, the map usually contains four groups of observations. Whereas the HH indicates that high values are surrounded by high values, the LL group means that low values are surrounded by low values. Although, hot spots (HH) are seen on southwest of Turkey in 2013 and 2014 and in the following years, they are spread in the north. Outliers can be seen on east and southeast of Turkey.

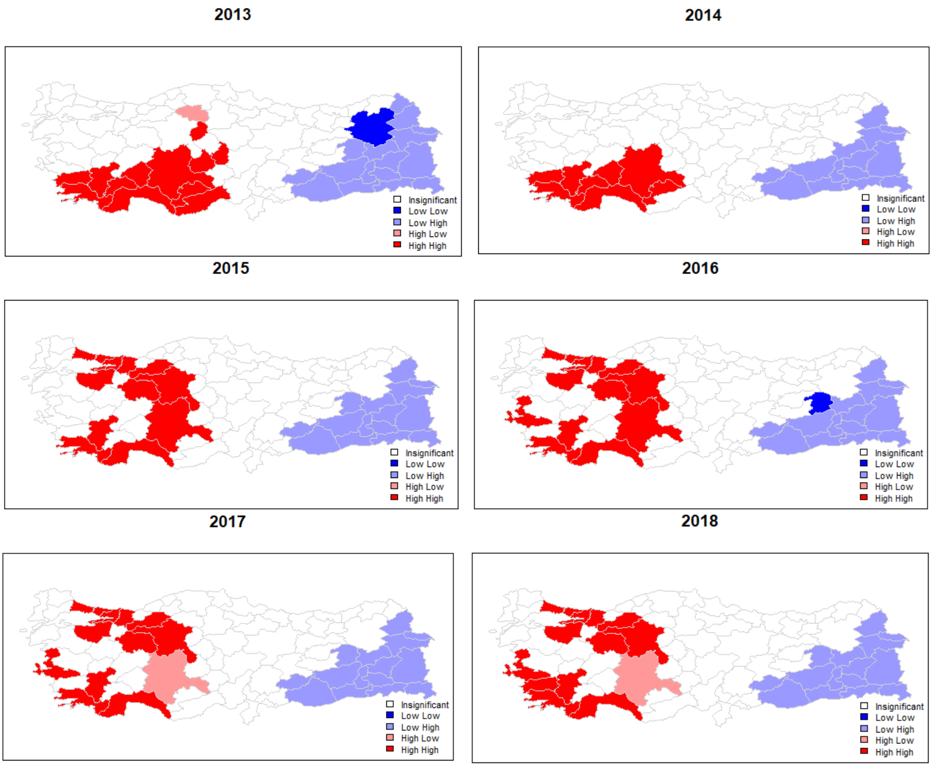

Figure 3. LISA cluster map for the RTA rate between 2013-2018 (5% significance level)

Since the RTA rate is proven to be spatially auto-correlated, which violates the assumption of OLS regression that the observations be independent, a spatial econometric model will be used to investigate the hidden influencing factors of RTA rate considering space–time relationship for the periods 2013-2018.

Table 3. Tests for non-spatial panel models

Tests	Test Value	p-value
F	4.8021	< 2.2e-16
LM	1.376	0.2408
Hausman	422.65	< 2.2e-16

Fixed effects are tested by the F test, while random effects are examined by the Lagrange multiplier (LM) test (Breusch and Pagan test). If the null hypothesis is not rejected, the pool regression can be suitable model. According to results of F, LM tests in Table 3, the pool model is not suitable to model the RTA rate. On the other hand, the Hausman specification test (Hausman test) compares a random effect models with fixed effect model. If the null hypothesis, stated that the individual effects are uncorrelated with the other repressors, is not rejected, a random effect model is favoured over its fixed counterpart [35]. As the Hausman test result suggests that the hypothesis of random effects must be rejected at % 5 per cent significance, thus, we adopt the fixed effects models and estimate the individual, time and individual-time fixed effects separately.

On the other hand, according to time-fixed effect tests (F test value=158.92, p-value < 2.2e-16 and LM test value=6005, *p*-value < 2.2e-16), time effects can be considered.

While the result of the Breusch-Pagan (BP) test (value=424.04, p-value < 2.2e-16) implies the presence of heteroscedasticity, the Breusch-Pagan LM test (value=10893, p-value < 2.2e-16) and the Pesaran CD test (value=95.419, p-value < 2.2e-16) indicate presence of cross-sectional dependence. Considering all panel test results and Moran's I values together, spatial models should be applied for modelling the RTA rate.

Table 4. Hausman Test for spatial panel models

Model	Test Value	p-value
SLM	1877.47487	< 0.001
SEM	66.46718	< 0.001
SDM	746.44437	< 0.001
GNSM	448.86711	< 0.001

As seen in Table 4, the results of Hausman test for spatial panel models shows that all fixed effects models are significant, thus, the fixed panel models with consideration of spatial effects are operated and the estimated results are presented in Tables 5 and 6 corresponding to spatial error autocorrelation, spatial dependent lag and spatial independent lag. Tables are divided in two parts, according to the fixed and fixed two ways models corresponding to spatial model, SEM, SLM, SDM, GNSM with different spatial structures, as in Table 1.

Table 5. Estimation results of the Fixed and Fixed two-way panel models with SEM and SLM

	SEM						SLM					
Models	Fixed Model			Fixed Two Ways			Fixed			Fixed Two Ways		
Variable	Coeff		p value	Coeff		p value	Coeff		p value	Coeff		p value
Number of cars	0,0685	***	2,59E-25	0,0684	***	2,99E-25	0,0680	***	5,19E-25	0,0679	***	4,34E-25
Number of minibuses	-0,4471	***	3,15E-06	-0,5264	***	3,20E-07	-0,4318	***	5,95E-06	-0,5262	***	3,01E-07
Number of buses	-1,0413	***	8,91E-05	-1,2401	***	5,26E-06	-0,9888	***	1,91E-04	-1,2290	***	5,80E-06
Number of vans	0,1160	***	9,57E-07	0,1440	***	8,39E-09	0,1113	***	2,69E-06	0,1438	***	8,18E-09
Number of trucks	-0,1399	**	1,63E-03	-0,1790	***	6,40E-05	-0,1311	**	3,20E-03	-0,1752	***	8,61E-05
Number of motorbikes	-0,0170	**	2,08E-03	-0,0179	**	1,20E-03	-0,0165	**	2,81E-03	-0,0181	***	9,82E-04
Number of private vehicles	1,5478	***	2,31E-04	1,6066	***	1,22E-04	1,6076	***	1,51E-04	1,6549	***	7,49E-05
Number of tractors	-0,0266	**	1,94E-03	-0,0252	**	3,26E-03	-0,0274	**	1,40E-03	-0,0252	**	3,10E-03
Length of asphalt roads	6,08E-06	***	1,62E-05	7,29E-06	***	4,12E-07	6,11E-06	***	1,53E-05	7,35E-06	***	3,13E-07
Length of surface coating roads	-3,05E-06	***	1,62E-06	-2,84E-06	***	1,03E-05	-3,20E-06	***	4,60E-07	-2,84E-06	***	9,43E-06
length of non asphalt roads	3,14E-06		3,84E-01	3,54E-06		3,23E-01	3,16E-06		3,81E-01	3,17E-06		3,75E-01
Spatial coefficient λ	0,1129	*1	8,42E-02	0,0258		7,03E-01 ρ	0,0883	*1	8,83E-02	0,0979	*1	6,59E-02

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '*1' 0.1 ' ' 1

Turning to the spatial panel models in Tables 5, the results demonstrate while the coefficient of spatially autocorrelation λ in the SEM in fixed model is significant (0.1129), λ in the fixed two-way panel model in the specifications of the SEM is not significant, and it has low value (0.0258). This suggests that the RTA rate is significantly influenced by the variation in neighbouring error term in fixed model. The results for SEM in Table 5 showed that the RTA rate is positively correlated with number of cars, private vehicle, vans and length of asphalt roads, other factors are negatively correlated. On the other hand, also non-asphalt road is not significant to explain the

RTA rate. The accident may be occurred on the asphalt due to high speed driving. Number of private vehicles have positively maximum effect and number of buses has negatively maximum effect way, and meanwhile length of asphalt roads and surface coating roads have minimum effect on SEM. Both of spatially autocorrelation parameter ρ are significant in SLM in fixed and fixed two ways models and have also almost same values (0.0883 and 0.0979). It means that, considering both of time and individual effect together neighbourhood effect is increased. Similar to SEM, number of private vehicles and number of buses have maximum effect, length of asphalt roads and surface coating roads have minimum effect on SLM. Number of buses has greater effect than number of minibuses in SEM and SLM. Number of motorbikes and tractors has naturally minimum effect among all vehicles.

On the other hand, the estimation results of the fixed and fixed two ways panel models with spatially lagged explanatory variables (SDM and GNSM) are given in Table 6. While spatial parameters are significant in all models ($p<0.1$) except GNSM fixed, the estimated coefficients of the lagged independent variables (neighbouring region characteristics in terms of independent variables) are not affect the RTA rate. The results obtained in the previous Table are also valid here. Although the results of fixed two-way panel model in the specifications of the SEM indicate that the RTA rate is positively affected by the number of cars, vans, private vehicles and length of asphalt roads, while length of surface coating roads are negatively correlated to explain RTA. Neighbours' explanatory variables, the observed indirect effects associated with explanatory variables are all statistically meaningless suggesting non- evidence of spillover effects. Therefore, these results provide justification for using spatial panel-data modelling approach to analyse the factors for RTA.

Considering significance, in GNSM in two ways effect, spatial parameters are higher than other models and other effects (ρ= -0.47 and λ=0.33) and also parameter of spatially lagged dependent variable is different than others as negative sign. Furthermore, they have smallest p-value in all of them. However, if it is taken into account significance of the lagged explanatory variables, it can be expressed that SEM or SLM are suitable for modelling the RTA rate.

Sum up, all the explanatory variables (except length of non-asphalt roads) are significant in all models and effects. As well as number of private vehicles, lagged number of private vehicles is only significant variable among all lagged variables (except in SDM in Two-Ways).

While spatial statistics have a potential to reveal the local or global spatial patterns of the considered factors, spatial econometric models allow us to find out spatial relationships and also can help to explain the factors behind observed spatial patterns. The applications of such spatial statistics tools to RTA will help to identification of provinces in terms of RTA with outstandingly spatial effects in Turkey. These results can be used by researchers working in road safety management.

Table 6. Estimation results of the Fixed and Fixed two-way panel models with SDM and GNSM

Models	SDM				GNSM			
	Fixed		Fixed Two-Ways		Fixed		Fixed Two-Ways	
Variables	Coeff	p value	Coeff	*p* value	Coeff	*p*-value	Coeff	*p*-value
Number of cars	0,0691 ***	3,48E-25	0,0681 ***	9,02E-25	0,0690 ***	4,01E-25	0,0617 ***	4,54E-22
Number of minibuses	-0,4969 ***	5,07E-07	-0,5411 ***	1,54E-07	-0,4951 ***	5,91E-07	-0,5465 ***	4,85E-08
Number of buses	-1,0770 ***	6,40E-05	-1,2281 ***	5,58E-06	-1,0634 ***	8,05E-05	-1,1282 ***	9,46E-06
Number of vans	0,1233 ***	2,90E-07	0,1484 ***	2,87E-09	0,1217 ***	4,90E-07	0,1499 ***	5,79E-10
Number of trucks	-0,1536 ***	7,07E-04	-0,1930 ***	2,40E-05	-0,1503 ***	9,41E-04	-0,1765 ***	3,88E-05
Number of motorbikes	-0,0163 **	3,28E-03	-0,0185 ***	8,41E-04	-0,0160 **	3,78E-03	-0,0174 ***	9,20E-04
Number of private vehicles	1,6465 ***	1,05E-04	1,6911 ***	5,58E-05	1,6709 ***	8,96E-05	1,9212 ***	2,45E-06
Number of tractors	-0,0257 **	2,90E-03	-0,0228 **	8,83E-03	-0,0260 **	2,58E-03	-0,0209 *	1,11E-02
Length of asphalt roads	5,65E-06 ***	7,24E-05	7,12E-06 ***	8,92E-07	5,65E-06 ***	7,39E-05	7,84E-06 ***	2,57E-08
Length of surface coating roads	-3,33E-06 ***	2,24E-07	-2,85E-06 ***	9,62E-06	-3,41E-06 ***	1,12E-07	-3,03E-06 ***	5,67E-07
length of non asphalt roads	2,99E-06	4,12E-01	3,10E-06	3,90E-01	3,00E-06	4,09E-01	2,52E-06	4,55E-01
Lagged Number of cars	-0,0004	9,73E-01	-0,0021	8,56E-01	-0,0004	9,75E-01	-0,0030	7,85E-01
Lagged Number of minibuses	-0,0912	6,80E-01	-0,1322	5,57E-01	-0,0885	6,89E-01	-0,1467	4,87E-01
Lagged Number of buses	0,1248	7,85E-01	0,0399	9,40E-01	0,1050	8,17E-01	-0,2213	6,45E-01
Lagged Number of vans	0,0276	5,78E-01	0,0337	5,03E-01	0,0276	5,79E-01	0,0358	4,53E-01
Lagged Number of trucks	-0,0928	3,51E-01	-0,1019	2,97E-01	-0,0934	3,48E-01	-0,1182	2,11E-01
Lagged Number of motorbikes	-0,0103	3,23E-01	-0,0081	4,30E-01	-0,0095	3,62E-01	0,0010	9,19E-01
Lagged_ Number of private vehicles	1,5995 *1	9,95E-02	1,5167	1,42E-01	1,6651 *1	8,56E-02	1,8736 *1	5,26E-02
Lagged Number of tractors	-0,0072	7,12E-01	-0,0037	8,45E-01	-0,0063	7,48E-01	0,0064	7,27E-01
Lagged_ Length of asphalt roads	1,50E-06	6,30E-01	2,07E-06	5,06E-01	1,50E-06	6,32E-01	2,24E-06	4,52E-01
Lagged_ surface coating roads	2,43E-12	1,00E+00	3,97E-08	9,75E-01	-6,94E-08	9,52E-01	-4,68E-07	6,85E-01
Lagged_ length of non asphalt roads	-3,05E-06	7,46E-01	-4,61E-06	6,19E-01	-2,99E-06	7,52E-01	-5,36E-06	5,49E-01
ρ	0,0883 *1	8,96E-02	0,0892 *1	9,34E-02	-0,0489	6,59E-01	-0,4713 ***	5,65E-06
λ					0,1187	1,48E-01	0,3324 ***	2,41E-07

Signif. codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘*1’ 0.1 ‘ ’ 1

5. CONCLUSION

The main results obtained from the presented study can be listed as follows:

- The RTA rate of 81 provinces in Turkey are examined by spatial statistics and spatial econometric tools over the period of 2013–2018.
- Since population is positively associated with RTA, the RTA rate is calculated as number of traffic accidents/ population in order to fairly compare the RTA in provinces.
- The Moran’s I value is found as around 0.5 and its $p < 0.001$ for the RTA rate confirms the clustered pattern and global spatial autocorrelation.
- The presence of neighbouring effects in the RTA rate is evidenced through the Moran's I statistic and natural breaking maps.
- LISA analysis show the significant provinces with a level of significance of 95% determined as a cluster with local Moran’s I. The result of LISA confirms that hot spots generally appear on western and southern regions which tend to have higher car speed/volume and more travel roads. The analysis also show that the majority of crashes occur in major provinces with large population and specific urban activity centres.
- According to the Hausman specification test, the fixed effects models should be preferred for modelling the RTA rate.

- The results showed that the RTA rate is positively correlated with number of cars, vans, private vehicles and length of asphalt roads, other factors are negatively correlated and also non-asphalt roads are not significant to explain RTA. The length of asphalt roads is unexpectedly one of the sources of higher RTA rate; speeding is probably the reason. It can be concluded from analysis if the number of private vehicles rise in provinces, it is triggering to increasing the RTA. Pavement maintenance is essential for ensuring good riding quality and avoiding traffic accidents. Improving road safety is one of the most important objectives for pavement management systems.
- Spatial parameters are mostly significant ($p<0.1$) and lagged independent variables (Neighbours' explanatory variables) do not affect the RTA rate, suggesting non- evidence of spillover effects. While a spatial spillover effect between Turkey's provinces does exist in terms of the RTA rate, only lagged number of private vehicles is significant to explain RTA in other provinces. In other words, the number of private vehicles in provinces has a significant positive spatial spillover effect on RTA in other provinces.

REFERENCES

[1] Colak, H.E., Memisoglu, T., Erbas, Y.S., Bediroglu, S., (2018) Hot Spot Analysis Based On Network Spatial Weights To Determine Spatial Statistics Of Traffic Accidents In Rize, Turkey, *Arabian Journal of Geosciences,* 11: 151.

[2] Işıldar, S., (2006) Road Accidents in Turkey 1995–2004, *IATSS Research*, 30(2), 115−118.

[3] Erdogan, S., (2009) Explorative Spatial Analysis of Traffic Accident Statistics and Road Mortality among the Provinces of Turkey, *Journal of Safety Research*, 40, 341–351.

[4] Levine, N., Kim, K.E., Nitz, L.H., (1995) Spatial Analysis of Honolulu Motor Vehicle Crashes: II. Zonal Generators, *Accid. Anal. Prev*., 27, 675–685.

[5] Papadimitriou E., Eksler V., Yannis G., Lassarre S., (2013) Modelling the spatial variation of road safety in Greece, *Proceedings of the Institution of Civil Engineers-Transport*, 49–58.

[6] LaScala, E.A., Gerber, D., Gruenewald, P.J., (2000) Demographic and Environmental Correlates of Pedestrian Injury Collisions: A Spatial Analysis, *Accid. Anal. Prev*., 32, 651–658.

[7] Zhixiao, X., Jun, Y., (2013) Detecting Traffic Accident Clusters with Network Kernel Density Estimation and Local Spatial Statistics: An Integrated Approach, *Journal of Transport Geography*, 31, 64-71.

[8] Soro, W.L., Zhou, Y., Wayoro, D., (2017) Crash Rates Analysis in China Using A Spatial Panel Model, *IATSS Research*, 41, 123–128.

[9] Lakes, T.A., (2017) Spatially Explicit Analysis of Traffic Accidents Involving Pedestrians and Cyclists in Berlin, *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz*, 60 (12), 1328-1335.

[10] Daniel, O., Adejumo, O., Oritogun, K., Mautin, G., Salako, A., (2017) Spatial Epidemiology of Road Traffic Crashes and Mortality in Nigeria, 2007-2015, C*urrent Journal of Applied Science and Technology,* 20(5), 1-11, DOI: 10.9734/BJAST/2017/33190.

[11] Unal, A., Atalay, A., Codur, M.Y., (2016) Evaluation of Traffic Accidents Happening in Recent Years in Turkey, *International Conference on Transportation and Traffic Engineering (ICTTE 2016),* https://www.researchgate.net/publication/320800768.

[12] Wang, W., Yuan, Z., Yang, Y., Yang, X., Liu, Y., 2019. Factors influencing traffic accident frequencies on urban roads: A spatial panel time-fixed effects error model. PLoS ONE 14(4): e0214539. https://doi.org/10.1371/journal.pone.0214539

[13] Akgungor, A.P., Dogan, E., (2009) An Artificial Intelligent Approach To Traffic Accident Estimation: Model Development and Application, *Transport*, 24:2, 135-142, DOI: 10.3846/1648-4142.2009.24.

[14] Akalin, K.B., Karacasu, M., Altın, A.Y., Ergul B., (2016) Curve Estimation of Number of People Killed In Traffic Accidents in Turkey. *IOP Conf. Ser.: Earth Environ. Sci. 44 052060*.

[15] Bicer, B.K., Unlu, H., Ozcebe, H., (2018) Traffic Accidents in Ten Years in Turkey: Advanced Analysis of National Data, *European Journal of Public Health*, 28(4), 319-320, cky214.032, DOI:10.1093/eurpub/cky214.032.

[16] Bayata, H.F., Bayrak, O.U., Pehlivan, H., (2018) Characteristics and Estimation of Traffic Accident Counts Using Artificial Neural Network and Multivariate Analysis: A Case Study in Turkey, *North Transit Interurban. Fresenius Environmental Bulletin*, 27(4), 2290-2298.

[17] Dereli, M.A., Erdogan, S., (2017), A New Model For Determining The Traffic Accident Black Spots Using GIS-Aided Spatial Statistical Methods, *Transportation Research Part A: Policy and Practice*, 103, 106-117.

[18] Butt, A., Ahmad, S.S., Sheikh, A.S., Shabbir, R., Erum, S., (2017) GIS Based Surveillance of Road Traffic Accidents (RTA) Risk for Rawalpindi City: A Geo-Statistical Approach, *Kuwait Journal of Science*, 44(4), 129-134.

[19] Hayidso, T.H., Gemeda, D.O., Abraham, A.M., (2019) Identifying Road Traffic Accidents Hotspots Areas Using GIS in Ethiopia: A Case Study of Hosanna Town, *Transport and Telecommunication*, 20(2), 123-132.

[20] Anselin, L., (1995) Local Indicators of Spatial Association-LISA, *Geographical Analysis*, 27, 93-115.

[21] Anselin, L., (2002) Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models, *Agricultural economics*, 27(3), 247–267.

[22] Anselin, L., Gallo, J.L., Jayet H., (2008) Spatial Panel Econometrics, *In: Mátyás L, Sevestre P (eds) The Econometrics of Panel Data*, Springer-Verlag, Berlin Heidelberg, 625-660.

[23] Aktas, S.G., Kumtepe, E.G., Kantar, Y.M., Ulukan, I.C., Aydin, S., Aksoy, T., Er, F., (2019) Improving Gender Equality in Higher Education In Turkey, *Applied Spatial Analysis and* Policy, 12(1), 167-189.

[24] Kantar, Y.M., Gunay, S.A., (2016) Spatial Correlation Analysis of Unemployment Rates in Turkey, *Journal of Eastern Europe Research in Business and Economics*, 1-9.

[25] Yidirim, V., Kantar, Y.M., (2020) Robust Estimation Approach for Spatial Error Model, *Journal of Statistical Computation and Simulation*, DOI:10.1080/00949655.2020.1740223.

[26] Scrucca L., (2005) Clustering Multivariate Spatial Data Based On Local Measures of Spatial Autocorrelation. An Application To The Labour Market Of Umbria, *Quaderni del Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia*, n. 20, Ottobre 2005.

[27] Lesage, J.P., Pace, R.K., (2009) Introduction to Spatial Econometrics. *CRC Press*, Boca Raton, FL.

[28] Getis, A. Ord, J.K., (1992) The Analysis Of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24, 189-206.

[29] Getis, A. Aldstadt, J., (2004) Constructing The Spatial Weights Matrix Using A Local Statistic. *Geographical Analysis*, 36, 90-104.

[30] Boots, B., (2002) Local Measures of Spatial Association, *Ecoscience*, 9(2), 168–176.

[31] Pillai N., Vijayamohanan., (2016) Panel Data Analysis with Stata Part 1: Fixed Effects and Random Effects Models, *Munich Personal RePEc Archive*. https://mpra.ub.uni-muenchen.de/76869/.

[32] Park, H.M., (2011) Practical Guides To Panel Data Modelling: A Step-By-Step Analysis Using Stata. *Tutorial Working Paper*. International University of Japan, Minami Uonuma.
[33] Elhorst, J.P., (2003) Specification and Estimation of Spatial Panel Data Models, *International Regional Science Review*, 26(3), 244-268.
[34] Elhorst, J.P., (2010) Spatial Panel Data Models, *In: Fischer M., Getis A., eds Handbook of Applied Spatial Analysis*. Springer, Berlin, Heidelberg, DOI: 10.1007/978-3-642-03647-7_19.
[35] Cliff, A. D. Ord, J. K., (1981) *Spatial Processes - Models and Applications*, London.