Romanian Journal of Regional Science
Vol. 17, No. 2, Winter 2023

DOI: https://doi.org/10.61225/rjrs.2023.07

UNRAVELLING THE ROAD SAFETY CHALLENGES: A SPATIAL ANALYSIS OF ROAD FATALITIES RATES

Elena-Maria Prada[a,*], Erika Marin[a], Monica Roman [a]

[a] Faculty of Economic Cybernetics, Statistics and Informatics, Bucharest University of Economic Studies, Romania

* Corresponding author

Address: 15-17 Calea Dorobanților, Bucharest, Romania

E-mail: elena.prada@csie.ase.ro

Biographical Notes

Elena Prada, PhD is Lecturer at the Bucharest University of Economic Studies, Department of Statistics and Econometrics. Her research interests focus on economic statistics, macroeconomic analysis, regional statistics, and spatial analysis. ORCID ID: https://orcid.org/0000-0001-7215-3207

Erika Marin, PhD is Full Professor at the Bucharest University of Economic Studies, Department of Statistics and Econometrics. Her research interests focus on economic statistics, macroeconomic analysis and forecasting, regional statistics, higher education, labour markets, and econometrics. ORCHID ID: https://orcid.org/0000-0002-3611-0151

Monica Roman, PhD is Full Professor at the Bucharest University of Economic Studies, Department of Statistics and Econometrics. Her research interests focus on topics linking the fields of labour and demographic economics, regional studies and transition economies. Her research includes empirical works about international migration, regional development, transition economies, research and higher education and policy evaluation. ORCID ID: https://orcid.org/0000-0002-9347-0639

Abstract

Road safety is a critical concern for all countries worldwide. Due to the lockdown measures imposed during the COVID-19 crisis in the European Union (EU), road fatalities have slightly decreased. We address the road safety from a regional perspective by selecting those NUTS 2 regions of the EU that are contiguous and have no missing data. Three steps make up our analysis: first, we analyse spatial autocorrelation globally; next, we analyse it locally; and, finally, we use spatial regression analysis to look at the relationship between the variables. Through the utilization of Moran's I, a spatial autocorrelation statistic, we seek to unveil the hidden patterns surrounding road accident death and injury rates. Also, our analysis includes a series of maps that show the general situation of regional road fatalities as reflection of the results obtained. The findings highlight the need for region-specific interventions and policy adjustments to mitigate road accidents and promote safer transportation systems. By implementing targeted strategies, policymakers and authorities not only can reduce accident rates but also enhance the overall quality of transportation experiences for residents and visitors alike, contributing to a safer, more sustainable future.

Keywords: spatial analysis, spatial autocorrelation, spatial regression, spatial lag, spatial error, road fatalities
JEL Classification: C21; C31; C40; L91

1. Introduction

Pollution road fatalities are important in many ways, particularly in terms of the human and societal costs they imply. These costs encompass not only the lives lost, but also the social impact in terms of traffic jams, cost of insurance, and law enforcement costs (Ghadi et al., 2018).

The primary focus of this study is to identify patterns in road fatalities across various NUTS 2 regions, investigating whether accident rates exhibit spatial clustering or dispersion. Secondly, our goal was to examine the macroeconomic factors that determine road fatalities in detail, acknowledging the complexity of the variables affecting road safety. Even though we had originally planned to include a larger number of variables, we were forced to depend only on the number of cars and NEETs (Not in Education, Employment, or Training), both available at NUTS 2 level. Despite this challenge, the study uncovered the indirect impact of the NEET (Not in Education, Employment, or Training) population on road safety via various channels. Understanding these dynamics could provide invaluable insights into reducing road risks and improving safety measures.

Our analysis includes Austria, Switzerland, Czechia, Germany, Estonia, Finland, Hungary, Italy, Lithuania, Latvia, Malta, Poland, Romania, Sweden, and Slovakia, offering a relevant image for the European countries. The accessibility of data and the coherence of neighbouring areas guided our methodical selection of these regions. We aimed for a large but manageable dataset, prioritising regions with extensive data on traffic accidents and fatalities. We sought to capture meaningful spatial relationships and patterns by emphasizing coherence among neighbouring regions, taking advantage of geographic proximity for a more informative examination of road accident dynamics.

The literature currently lacks extensive exploration into spatial analysis of traffic accidents, indicating a gap that needs to be filled. This research gap provides an excellent opportunity to delve deeper into the spatial complexities of traffic accidents, such as patterns, clustering, and dispersion.

In this regard, we consider that such research could yield critical insights for developing specific policies and measures intended for reducing road fatalities and injuries.

More spatially focused research could provide insights into the localised nature of road accidents, potentially revealing region-specific risk factors or issues that could inform more effective preventive measures.

This paper is organized into several sections, beginning with an introductory overview of existing literature in the field, followed by methodology and data collection process and an extensively examinations of our findings. Finally, a thorough discussion and concluding observations are made.

2. Literature review

Most studies that delves into the determinants of road fatalities do not account for spatial interactions (Ali et al, 2019; Dhibi, 2019; Saeednejad, et al., 2020; Pammer, et al 2021; Prada, 2021).

Erdogan (2009) used Geographic Weighted Regression (GWR) to investigate road mortality across Turkish provinces in 2003, examining death rates alongside vehicle counts separated by their type, such as cars and buses. GIS-assisted spatial analysis was used in the investigation to pinpoint provinces with safety deficiencies so that targeted preventive actions could be taken. The accident and fatality distributions at the province level were revealed by this method, which combined GIS and spatial statistics. The study underlined the need for additional preventive actions, focusing on the areas that the research explicitly identified.

Osayomi and Areola (2015) explored road traffic accidents, injuries, and fatalities in Nigeria from 1980 to 1984. Using spatial regression techniques, they investigated variables such as traffic density and highway availability.

Alkhadour et al. (2021) studied traffic accidents in Amman, Jordan, between 2017 and 2019. The variables in their study included accident dates, locations, types of injuries sustained, and the vehicles involved, and they used hotspot distribution analysis to identify accident clusters. The results showed concentrated hotspots in industrial, commercial, and residential zones, mostly in and near Amman's downtown areas. These densely populated regions with public amenities also have higher traffic volumes and slower speed restrictions, which add to the concentration of incidents in these places.

The spatial analysis of traffic accidents is a relatively unexplored area in the literature. This gap provides an opportunity for additional research and exploration in the field. Understanding the spatial patterns, clustering, and dispersion of traffic accidents can help to develop targeted interventions and policies to reduce fatalities and injuries.

3. Method

Our comprehensive analysis includes variables for the year 2020 across all NUTS 2 in several European countries: Austria, Switzerland, Czechia, Germany, Estonia, Finland, Hungary, Italy, Lithuania, Latvia, Malta, Poland, Romania, Sweden, and Slovakia.

The NUTS 2 regions from specific European countries were carefully chosen based on several factors, the most important being data availability and the upkeep of coherent neighbouring regions. We aimed to create a comprehensive yet manageable dataset by focusing on regions with reliable data on traffic accidents, fatalities, and relevant variables. Therefore, our dataset included 152 observations in total, providing reliable and complete data on traffic incidents, fatalities, and other relevant covariates presented in Table 1.

Table 1. The variables of the model

Variable	Notation	Definition	Source
Death rate	DR	The variable is computed as a crude rate, represented as the number of deaths resulting from traffic accidents per 1,000 individuals within the population.	Eurostat
Injury rate	IR	The variable is computed as a crude rate, represented as the number of injuries resulting from traffic accidents per 1,000 individuals within the population.	Eurostat
Cars Rate	CarsR	The variable is computed as a percentage rate, represented as the number of cars as percentage of total number of cars at national level.	Eurostat
NEET rate	NEET	Represents the people between 18 and 24 years old who are not currently involved in any formal educational activities, are unemployed, and are not enrolled in any kind of job training or apprenticeship programmes.	Eurostat

Source: Own representation.

Our primary objective is focused on capturing meaningful spatial relationships and patterns. With this aim in view, we took into account geographic proximity and similarities between adjacent regions in order to facilitate a more coherent and insightful examination of road accident dynamics across these areas.

In our analytical approach, we employed both global and local Moran's I statistics. To explore the neighbourship, we employed a contiguity matrix based on Euclidean distance constraints to map out neighbourhood relationships across such a large geographical area. This method enabled us to create spatial connections between regions based on their proximity, allowing us to investigate spatial patterns and interactions between these disparate areas. Moran's I can take values from -1 to +1, being negative shows dissimilarities among territorial units and being positives shows similarities between territorial units, as detailed in table 2. below:

Table 2. Moran's I interpretation

Strength of Association	Positive	Negative
	Similarity (high-high or low-low)	**Dissimilarity (high-low or low-high)**
Small	0.1 to 0.3	-0.1 to- 0.3
Medium	0.3 to 0.5	-0.3 to -0.5
Large	0.5 to 1.0	-0.5 to -1.0

Source: Own representation based on Moura and Bráulio (2020).

In particular, according to the positive spatial autocorrelation indicates that the relationship between the reference spatial entity and its neighbours tends to typically displays either a high-high or a low-low pattern. This pattern shows the presence of similarities between neighbouring regions and the reference region. As for the negative spatial autocorrelation, which shows dissimilarities, the relationship between the reference spatial entity and its neighbours tends to exhibit either a high-low or a low-high pattern.

Equation (1) represents the Moran's I formula used, allowing us to assess spatial autocorrelation.

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

Where:

- n is the number of territorial units;
- x_i and x_j represent the values of the employed variable in territorial units i and j;
- $\bar{x}$ is the mean of variable x across the n territorial units;
- w_{ij} is a binary weighting factor with a value of $w_{ij} = 1$ if territorial units i and j are neighbouring, and zero otherwise.

Subsequently, equation (2) describes the significance testing of Moran's I, aiding in understanding the strength and significance of spatial patterns observed in the data.

$$z_I = \frac{I - E(I)}{S_{error\ I}} \quad (2)$$

Where:

- I is the Moran's I value;

- $S_{error\ I}$ is the standard error of the Moran's I value;
- E(I) is the expected value of Moran's I.

To test the significance of Moran's I, the null hypothesis typically assumes spatial randomness (no spatial autocorrelation). To generate a distribution under the null hypothesis, Monte Carlo simulation (MC simulation) or permutation test is frequently employed. To determine the statistical significance of the observed Moran's I value, this distribution is compared to it. Thus if the $z_I > Z_{\alpha/2}$ the null hypothesis is rejected, and we can conclude the presence of spatial autocorrelation.

Additionally, our methodology included the use of the spatial regression model. This model, as opposed to the classic regression, incorporates spatial dependencies between nearby regions into account in addition to correlations between variables. We applied the Spatial Autoregressive Model (SAR) and the Spatial Error Model (SEM), two spatial regression effects available in GeoDa.

$$\text{SEM: } y = X\beta + \mu;\ \mu = \lambda W\mu + \varepsilon \text{ (3)}$$

$$\text{SAR: } y = \rho Wy + X\beta + \varepsilon \text{ (4)}$$

Where:

- X are the independent variables;
- Y is the dependent variables;
- B are the regression coefficients;
- λ is the spatial error coefficient;
- P is the spatial lag coefficient.

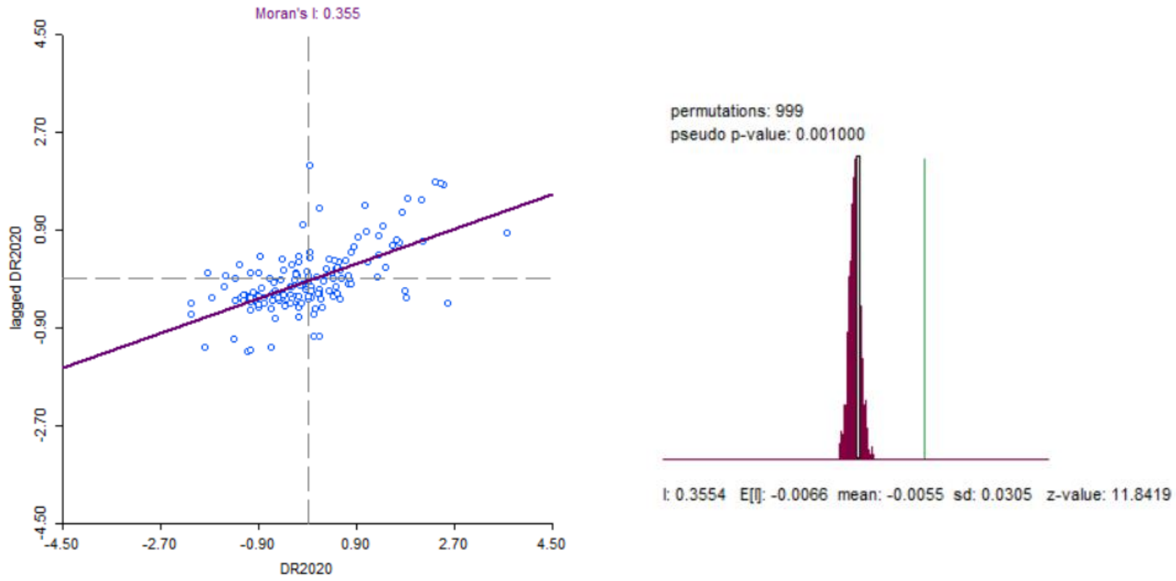
When examining the effects of various factors on the dynamics of traffic accidents throughout the areas, these two models serve to comprehend and consider spatial interactions. We consider that by incorporating the spatial effects, the results Our expectation is that the dependent variable will present spatial lag.

4. Results and Discussion

Our analysis unfolds in three stages: global analysis of spatial autocorrelation, local analysis of spatial autocorrelation, and examining the relationship between variables through spatial regression analysis.

First, we apply Moran's I to do a global spatial autocorrelation assessment. For every variable employed in the analysis, the results show a statistically significant presence of spatial autocorrelation. The following charts (Figures 1 through 4) show the relevant significance test of the statistics on the right and Moran's I diagram on the left, which also contains the Moran's I value. The value of Moran's I obtained for all variables is positive, so we can conclude that the observations present similarities within the neighbouring regions.

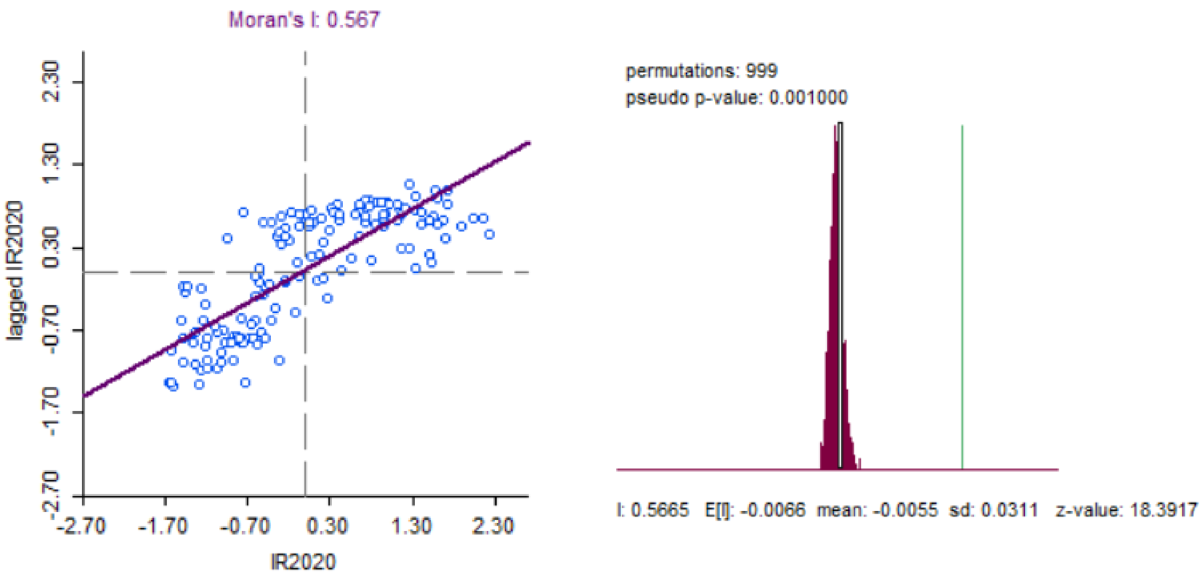
Chart 1. The Global Moran's I for Death Rate



Source: results estimated by the authors using GeoDa

The death rate presents a moderate positive spatial autocorrelation (Moran's I = 0.355), with the permutation test indicating a statistically significant presence of spatial autocorrelation.

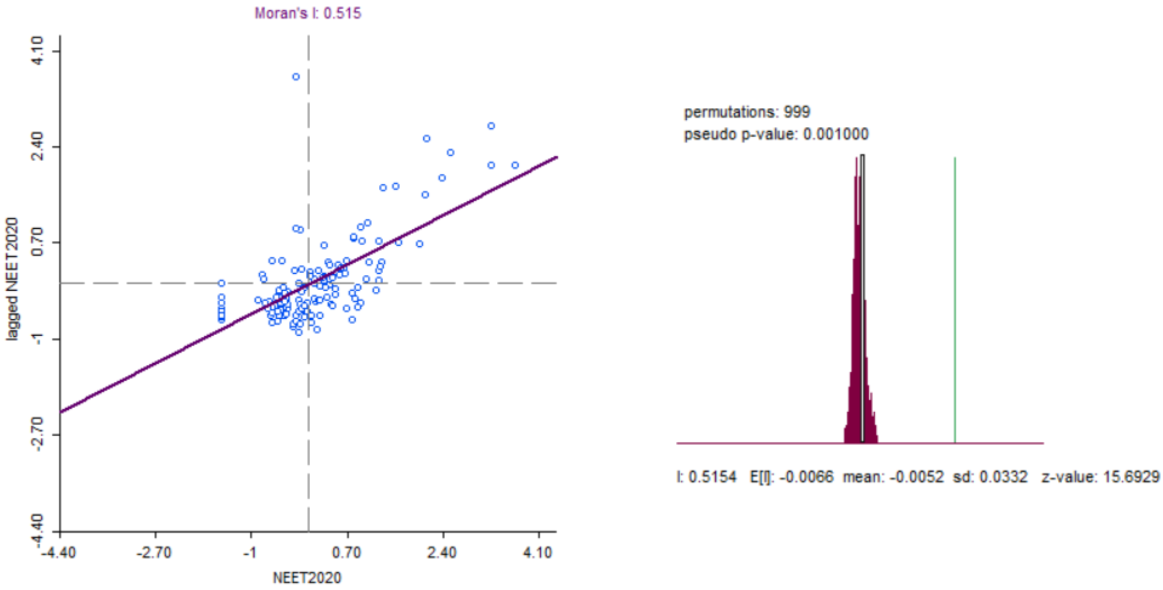
Chart 2. The Global Moran's I for Injury Rate



Source: results estimated by the authors using GeoDa

The injury rate presents a large positive spatial autocorrelation (Moran's I = 0.567), with the permutation test indicating a statistically significant presence of spatial autocorrelation.

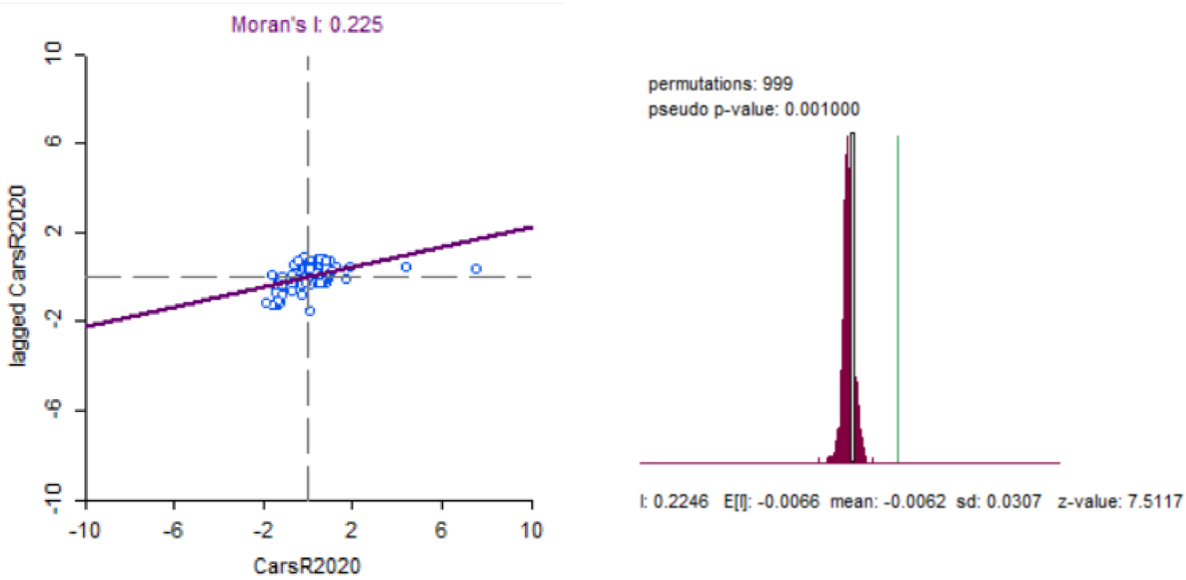
Chart 3. The Global Moran's I for NEETs



Source: results estimated by the authors using GeoDa

The NEETs injury rate shows a large positive spatial autocorrelation (Moran's I = 0.567), with the permutation test indicating a statistically significant presence of spatial autocorrelation. As for the cars rate we observe a low positive spatial autocorrelation (Moran's I = 0.225), also the permutation test indicating a statistically significant presence of spatial autocorrelation.

Chart 4. The Global Moran's I for Cars Rate



Source: results estimated by the authors using GeoDa

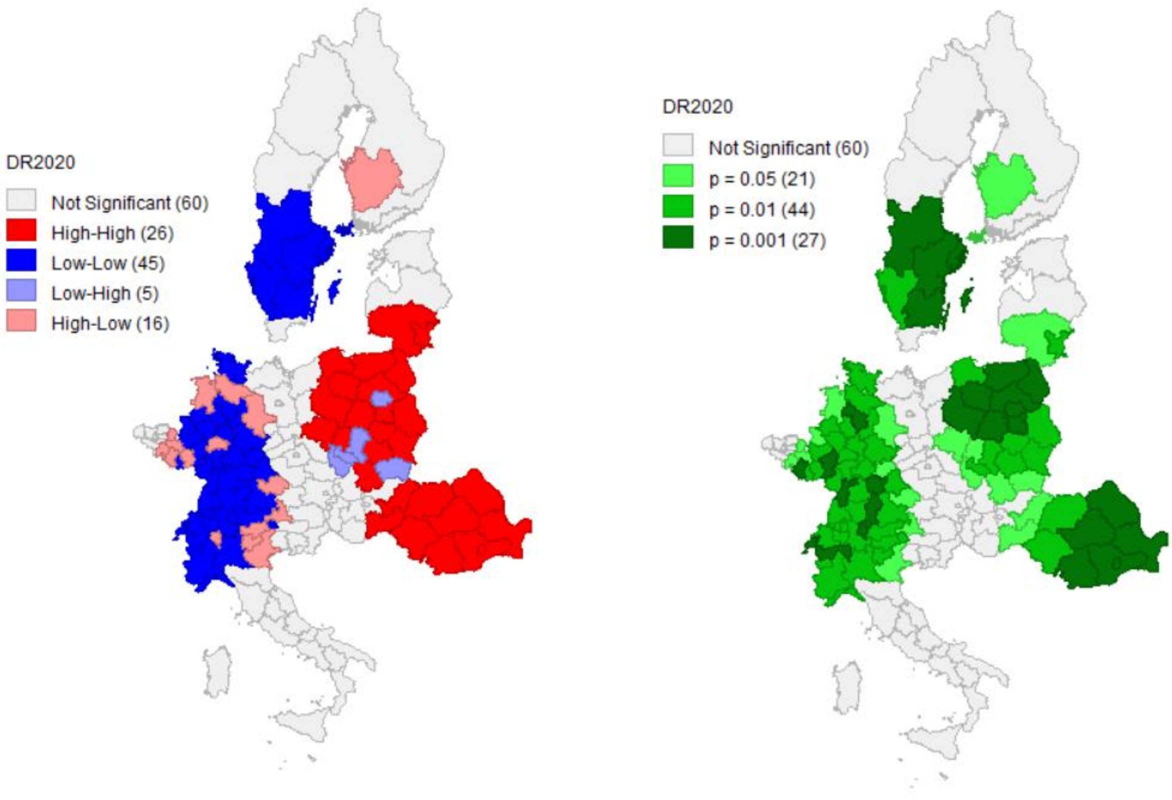
The global spatial autocorrelation can reveal spatial patterns of the variables employed, by clustering the regions based on their similarities. Also, it can show interesting dissimilarities between NUTS 2 regions. Considering the value obtained of the global spatial autocorrelation in this analysis was excluded the variable number of car. Additionally, assuming that purchasing a car can be a personal decision, a global autocorrelation at the NUTS 2 level cannot present meaningful result.

The following charts (from 5 to 7) are divided as follows: on the left side, it's the type of spatial relationship between the statistical units, while on the right side, the meaning of the relationship is presented.

The maps of the Local Moran's I value for death rate and injury rate expose interesting patterns, indicating two different clusters across the countries considered. We can distinguish two different parts between the western and the eastern parts of the territorial units which present similarities. The fact that these clusters are partially mirrored is an intriguing aspect. For the mortality rate, the countries in the west present a low-low interaction, those in the east present a high-high interaction; however, for the injury rate, these clusters are reversed.

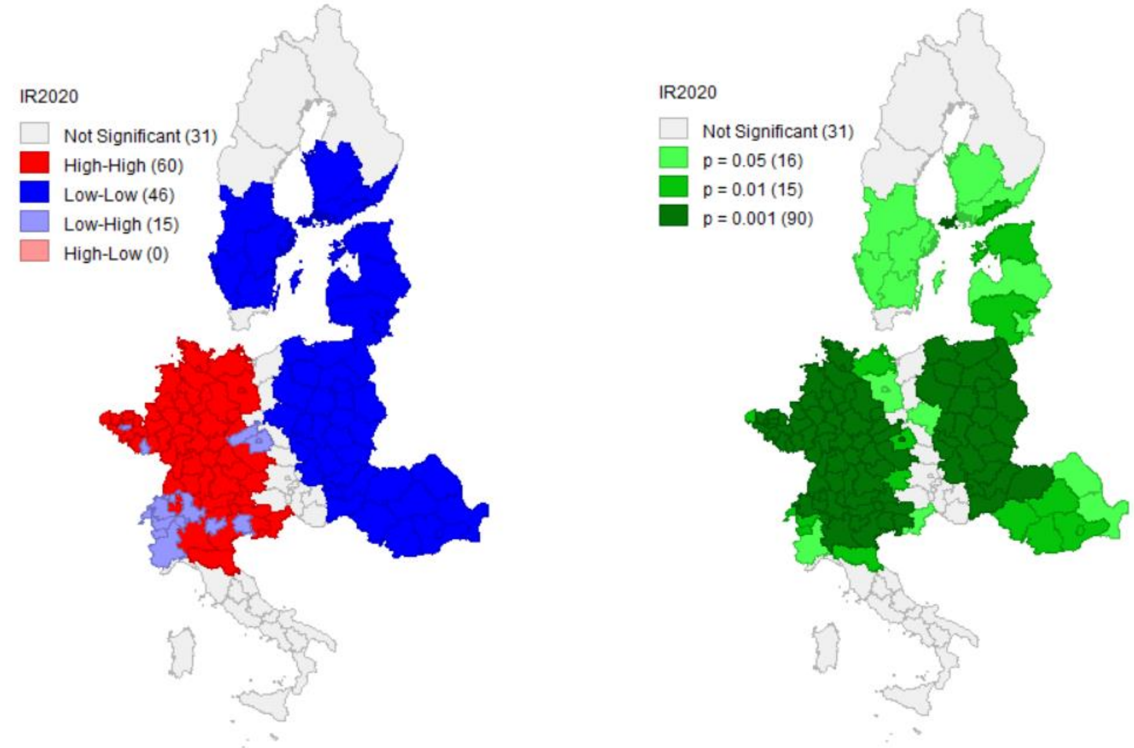
As for NEETs variable, even the pattern is comparable with the death rate variable, we can observe that are regions with a low NEETs rate, but their neighbours have a high NEETs rate.

Chart 5. The Local Moran's I for Death Rate



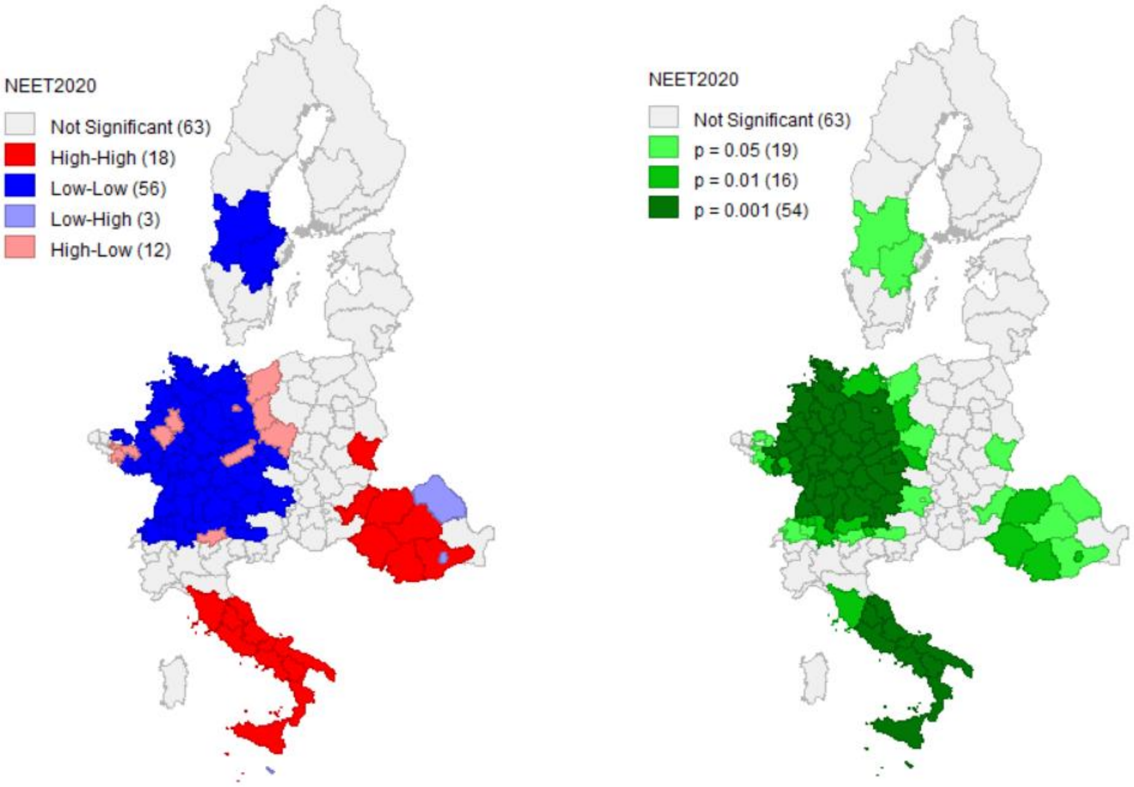
Source: results estimated by the authors using GeoDa

Chart 6. The Local Moran's I for Injury Rate



Source: results estimated by the authors using GeoDa

Chart 7. The Local Moran's I for NEETs



Source: results estimated by the authors using GeoDa

The final stage of the paper is to study the spatial relationship through the multiple linear regression by estimating the spatial effects. We used in turn as dependent variable the death rate (DR) and the injury rate (IR). The results presented in Table 3. are the spatial dependence diagnostics for both models, provided by GeoDa. Lagrange Multiplier and robust Lagrange Multiplier (LM) tests are provided for the two types of spatial models: lag(SAR) and error (SEM).

To decide in favour of a model, we first examine the significance of the tests by comparing the Prob value with 0.05 significance level. In case of death rate variable model SAR outperformed the SEM model.

Table 3. Diagnostics for Spatial Dependence

Variable/Test	DR		IR	
	Value	Prob	Value	Prob
Lagrange Multiplier (lag)	**67.03**	**0.00**	185.39	0.00
Robust LM (lag)	**3.52**	**0.06**	1.98	0.16
Lagrange Multiplier (error)	64.79	0.00	**207.99**	**0.00**
Robust LM (error)	1.28	0.26	**24.58**	**0.00**

Significance level: (***) 0.001; (**) 0.01; (*) 0.05
Source: results estimated by the authors using GeoDa

In the case of death rate variable model SAR outperformed the SEM model. This conclusion was drawn considering that Lagrange Multiplier for lag has Prob value lower than 0.05 and is larger than the Lagrange Multiplier for error. For injury rate as dependent variable the SEM model outperformed the SAR model, even though we considered suggestive to present both results.

The estimates are presented in Table 4. The spatial lag coefficient ρ is denoted as W and it shows the "dependence in observation value of dependent variable in a spatial unit and the corresponding neighbouring units" (Saputro, et al, 2019, p2). The spatial error coefficient lambda (λ) shows the "spatial dependence in the error term of a spatial unit and the corresponding neighbouring unit" (Saputro, et al, 2019, p2).

Model 1 shows the spatial regression results obtained for the dependent variable mortality rate. It can be seen that as the mortality rate increases in the reference region it will also lead to an increase in the mortality rate in the neighbouring regions, thus generating a spill over effect. Of the independent variables included in the model, only NEETs is significant and shows a direct relationship between it and the death rate (DR). This result presents an interesting aspect, given the nature of this indicator. We cannot conclude that there is a causal effect, but the result shows that there is an influence on road fatalities generated by young people and NEETs.

For the injury rate (IR) variable, we estimated both the model with spatial lag, in order to compare the results with those obtained for the DR variable, and the model with spatial error, which according to the spatiality tests was the best fit. The results for the two spatial models are shown in the second model for the spatially lagged model and in the third model for the spatially lagged model.

Table 4. Spatial regression results for death rate and injury rate dependent variables

Model	Model 1 SAR		Model 2 SAR		Model 3 SEM	
Variable	DR	Probability	IR	Probability	IR	Probability
W	0.72(***)	0.00	0.88(***)	0.00	-	-
Constant	0.01	0.06	1.01(***)	0.00	3.172(***)	0.00
CarsR	-7.07E-06	0.42	-7.7E-04	0.06	-0.001(**)	0.01
NEET	4.39E-4(**)	0.03	-0.03(***)	0.00	-0.03(**)	0.01
lambda	-	-	-	-	0.89(***)	0.00
R-Squared	0.355		0.628		0.629	

Significance level: (***) 0.001; (**) 0.01; (*) 0.05
Source: results estimated by the authors using GeoDa

Comparing the model with spatial lag for the DR variable and that for the IR variable, there is no significant change in the spatial lag coefficient W, which maintains its sign and significance in the case of the model with the dependent IR variable. Interestingly, in the case of NEETs for the IR dependent variable, although the coefficient remains significant, the sign becomes negative, indicating the presence of an indirect relationship. Therefore, if NEETs increases then IR will decrease. In the case of the model with spatial error, the one preferred following spatial diagnostics, we observe that all coefficients are statistically significant. The NEETs variable preserves the indirect reality with the dependent variable IR.

The results obtained are even more interesting as the analysis of road fatalities does not only involve variables that intuitively might influence either death rate or injury rate. Taking into account variables at macroeconomic level may show a different perspective for policy-makers.

Considering the results obtained, we trust that the paper outlines how important it is to understand the complex dynamics of road safety based on macroeconomic data to develop successful policy interventions and reduce fatalities.

5. Conclusion

The main constraint faced in this study is the lack of available data for NUTS levels, in particular at the NUTS 2 level for all the EU member states. Given that our dataset comprises only 16 of the 27 EU member states, this limitation highlights the critical need for improved data availability at the NUTS 2 level. It is critical to expand the breadth and depth of available data across all NUTS levels in order to contribute to a comprehensive understanding of the intricate dynamics of road safety across different regions. The issues of data availability is an ongoing concern and has been raised by many researchers (Panzera Postiglione, 2014; Awan and Pettenella, 2017).

Access to data at all NUTS levels would significantly improve the outcome of policy development and intervention strategies focused on reducing road fatalities and improving overall road safety throughout the European Union. With an expanded dataset, policymakers and stakeholders would be able to create more targeted and nuanced initiatives that are customised for the specific situations and challenges that exist in various regions. This, in turn, will open the path for more effective and impactful measures aimed at reducing traffic fatalities and ensuring safer roads across the EU countries.

We acknowledge certain limitations in our data analysis and modelling, even though our study explores a largely uncharted territory. Going forward, investigating different kinds of spatial analysis is a good direction for this field's future study. This method might provide more information and deepen our comprehension of the intricate dynamics surrounding this subject.

References

Ali, Q., Yaseen, M. R. and Khan, M. T. I., 2019. The causality of road traffic fatalities with its determinants in upper middle income countries: a continent-wide comparison. *Transportation research part A: policy and practice*, 119, pp. 301-312. https://doi.org/10.1016/j.tra.2018.12.002

Alkhadour, W., Zraqou, J., Al-Helali, A. and Al-Ghananeem, S., 2021. Traffic accidents detection using geographic information systems (GIS). *International Journal of Advanced Computer Science and Applications*, 12(4), pp. 484-494. https://dx.doi.org/10.14569/IJACSA.2021.0120462

Awan, H. U. M. and Pettenella, D., 2017. Pine nuts: a review of recent sanitary conditions and market development. *Forests*, 8(10), pp. 367-384. https://doi.org/10.3390/f8100367

Dhibi, M., 2019. Road safety determinants in low and middle income countries. *International journal of injury control and safety promotion*, 26(1), pp.99-107. https://doi.org/10.1080/17457300.2018.1482926

Erdogan, S., 2009. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of safety research*, 40(5), pp. 341-351. https://doi.org/10.1016/j.jsr.2009.07.006

Ghadi, M., Török, Á. and Tánczos, K., 2018. Study of the economic cost of road accidents in Jordan. *Periodica Polytechnica Transportation Engineering*, ***46***(3), pp. 129-134. http://dx.doi.org/10.3311/PPtr.10392

Moura, A. C. M. and Bráulio M. F., 2020. ESDA (Exploratory Spatial Data Analysis) of Vegetation Cover in Urban Areas—Recognition of Vulnerabilities for the Management of Resources in Urban Green Infrastructure. *Sustainability* 12(5), pp.1933-1955. https://doi.org/10.3390/su12051933

Osayomi, T. and Areola, A. A., 2015. Geospatial analysis of road traffic accidents, injuries and deaths in Nigeria. *The Indonesian Journal of Geography*, 47(1), pp. 88-98. https://doi.org/10.22146/ijg.6749

Pammer, K., Freire, M., Gauld, C. and Towney, N., 2021. Keeping safe on Australian roads: overview of key determinants of risky driving, passenger injury, and fatalities for indigenous populations. *International journal of environmental research and public health*, 18(5), 2446. https://doi.org/10.3390/ijerph18052446

Panzera, D. and Postiglione, P., 2014. Economic growth in Italian NUTS 3 provinces. *The Annals of Regional Science*, 53(1), pp. 273-293. https://doi.org/10.3390/su12176717

Prada, E. M., 2021. One for the road: the determinants of the road fatalities in the European Union. In *Proceedings of the International Conference on Business Excellence*, 15(1), pp. 228-234. https://doi.org/10.2478/picbe-2021-0022

Saeednejad, M., Sadeghian, F., Fayaz, M., Rafael, D., Atlasi, R., Houjaghan, A. K., Salamati, P., 2020. Association of social determinants of health and road traffic deaths: a systematic review. *Bulletin of Emergency & Trauma*, 8(4), pp. 211-217. https://doi.org/10.30476/beat.2020.86574