Traffic & Transportation

Exploring and Comparing Spatial Clusters of Pedestrian Night-Time Crashes Based on Different Street Lighting Conditions at County Spatial Unit

František KEKULA[1], Bernard KOSOVEC[2], Darko BABIĆ[3], Pavel HRUBEŠ[4]

Original Scientific Paper
Submitted: 17 June 2024
Accepted: 9 Oct 2024

[1] Corresponding author, kekulfra@fd.cvut.cz, Czech Technical University in Prague, Faculty of Transportation Sciences
[2] bkosovec@fpz.unizg.hr, University of Zagreb, Faculty of Transport and Traffic Sciences
[3] dbabic@fpz.unizg.hr, University of Zagreb, Faculty of Transport and Traffic Sciences
[4] pavel.hrubes@cvut.cz, Czech Technical University in Prague, Faculty of Transportation Sciences

Publisher:
Faculty of Transport and Traffic Sciences, University of Zagreb

ABSTRACT
This paper attempts to determine the role of street lighting in the spatial clustering of night-time crashes involving pedestrians in the Republic of Croatia. Five-year (2018–2022) night-time pedestrian crash data were used in conditions with and without street lighting. First, distance-based statistical methods were used to assess the spatial clustering and deviations from complete spatial randomness (CRS) of the crash patterns. Second, the global Moran's I analysis was conducted to investigate a degree of spatial autocorrelation of the annual crash counts aggregated in 21 counties of Croatia. Finally, the local indicators of spatial association (LISA) were used to identify the locations of the crash count hotspots. The results of the ANND analysis confirm a significant clustering of crashes for both street lighting conditions. However, different global Moran's I values for both conditions were obtained with a high and statistically significant positive value for the crash counts without street lighting. Local Moran's I analysis reveals that the High-High (H-H) county clusters are located in coastal regions of Croatia, while the Low-Low (L-L) county clusters appear in the East continental part, next to Slavonia. The results suggest that inadequate lighting conditions have an impact on the clustering of pedestrian crashes at night.

KEYWORDS
night-time pedestrian crashes; street lighting; distance-based statistical methods; spatial autocorrelation; global Moran's I; local Moran's I.

1. INTRODUCTION

Road crashes constitute a substantial part of the global socio-economic burden, particularly in low- and middle-income countries [1]. Compared to other types of road crashes, pedestrian crashes create almost 20% [2] of all types of crashes in the European Union (EU). Essentially, pedestrians are unprotected traffic participants who interact with other, more protected road users of higher speed and mass and a better protective structure from whom they can hardly protect themselves. Therefore, when a crash occurs, they often suffer fatal or severe injuries [3].

Among others, the low lighting level of the road environment is one of the main factors that contributes significantly to the increase in the risk of severe or fatal crashes [4, 5]. Low visibility conditions reduce the visual function, affecting the visual acuity and perception of drivers, and therefore negatively impacting their visual performance [6]. Furthermore, these conditions contribute to more severe crashes involving pedestrians and cyclists [7-10]. Ferenchak et al. [11] found that the presence of street lighting evinces a positive correlation with better safety outcomes. They used linear regression to investigate the relationships between the per cent change in pedestrians killed (or injured) as the dependent variable and the year as the independent variable. This analysis was divided into dark and daylight categories to account for lighting conditions. The results revealed that the death and injury rates increased more significantly in the dark conditions compared to the daylight conditions with an increase of 43.3%. Kemnitzer et al. [12] used generalised estimating equations with a logit link to estimate the odds of pedestrian injury. They found that in dark conditions without road lighting, the odds of injury in pedestrian strikes increase by 49%. Furthermore, Kim et al. [13] investigated three-year pedestrian-vehicle crash data using a mixed logit model and found that darkness without street lighting significantly increases the probability of fatal injuries and very serious injuries from 137% to 325% and from 12% to 19%, respectively.

The role of street lighting in pedestrian crashes at night has so far been at the margins of scholarly interest. Although some authors reported that street lighting was one of the major contributing factors [14, 15], others reported it as a factor that had a minor influence [16, 17]. However, the interpretation of these findings can vary between countries for various reasons, including pedestrian and traffic volume at night, pedestrian infrastructure [18], road network, speed limits [18, 19], land use [20], extent of education on the importance of the pedestrian conspicuity problem [18] in addition to the spatial socio-economic and socio-demographic characteristics [21]. Ćosić et al. [22] analysed the relationships between the external factors and pedestrian crash blackspots in the city of Zagreb based on a sample of 1,333 pedestrian road crashes. To identify the black spots of the crash, they used the KDE method which revealed that the highest density was observed in the city centre and on the main roads. Further analysis of the possible impacts of external factors at the identified black spots revealed only a minor increase in the probability of pedestrian crashes at night [22]. However, this study was aimed at the capital city of Croatia, where street lighting installations are supposed to be of a better quality compared to the locations on the outskirts. Kučinić et al. [23] investigated pedestrian crashes in the Republic of Croatia during an observation period of 5 years and reported no correlations between injured and fatally injured pedestrians and lighting conditions. However, this study did not consider any technique for exploratory spatial analysis to discover patterns of spatial clusters of road crashes. Moreover, none of these studies were focused primarily on pedestrian crashes at night under lighted and unlighted conditions. Therefore, it is not yet clear where more crashes are likely to be concentrated under these conditions. To overcome the limitations mentioned above, the authors investigated the clustering of night-time pedestrian crashes with respect to different street lighting conditions at the county spatial scale, giving valuable information on the distribution of spatial clusters of the crashes.

Therefore, the objective of this paper is to assess the spatial pattern and degree of spatial clustering of night-time pedestrian crashes in the Republic of Croatia with respect to different street lighting conditions in order to identify areas where the crashes concentrate. To examine the spatial clustering tendency of crashes, ANND analysis and nearest neighbour distance function G were used. Second, the univariate global Moran's I analysis was conducted to investigate statistically significant spatial autocorrelation of pedestrian crash counts under conditions with and without street lighting. Finally, LISA was employed to detect localised clustering. To perform these analyses, the night-time pedestrian crashes with and without street lighting from 2018 to 2022 in Croatia were gathered and analysed. This paper consists of five sections, which are structured as follows. Section two explains our methodology, comprising two subsections, namely, the data filtering and preprocessing, and exploratory spatial analysis. Additionally, this section also covers the description of the case study area and data utilised for the purpose of the analysis. In the third section, the results of the exploratory spatial analysis are summarised. The fourth section addresses and discusses the findings from the analysis. The last section presents the main conclusions of this paper.

2. DATA AND METHODOLOGY

In this section, we describe the utilised data, and our research methodology, which comprises two main methodological steps: (1) data filtering and preprocessing; (2) exploratory spatial data analysis. Detailed description of these steps will be given in the following sub-sections.

2.1 Study area and GIS data

The Republic of Croatia was selected as the case study area for this paper. It is located at the crossroads of Central and Southeastern Europe, covering an area of 56,594 km with a population of 4,000,000. Furthermore, only three cities in Croatia have more than 100,000 inhabitants, but about half of the population lives in urban settlements. It is fair to mention the seasonal tourist characteristics of Croatia since, during the warm period of the year, a significant number of tourists visit Croatia, especially the coastal counties.

The Republic of Croatia is administratively divided into 21 counties interconnected by motorways (primary roads; approximately 1,480 km) and, dominantly, by state roads (secondary roads; approximately 7,390 km). The spatial data of administrative boundaries, specifically county boundaries, were obtained from the Geoportal of the State Geodetic Administration of the Republic of Croatia and are shown in *Figure 1*.

Croatia faces severe road safety challenges, as the number of people killed in crashes has stagnated in recent years. The latter is especially emphasised when it comes to vulnerable road users. That is, a total of 244 pedestrians died on Croatian roads during the 2018–2022 period, of which 127 were over 65 years of age. Although before 2020, the number of pedestrian deaths was descending, that number grew from 38 in 2020 to 43 pedestrians in 2022. Another challenge is the maintenance and renewal of the road infrastructure in a way that would be safer for road users, especially considering the rising volume of motorised traffic. Reconstruction and maintenance of roads and road equipment, in a responsible manner, remains a challenge from a variety of perspectives, such as terrain obstacles, financial sources, property-legal relations, etc.

The authors used GIS data on road crashes and spatial borders of the study area counties for spatial aggregation of pedestrian crashes. The crash data were collected from the Ministry of Interior of Croatia. These data covered the period from 2018 to 2022, including 1,624 night-time crashes, of which 1,449 (89%) occurred with the presence of street lighting and 175 (11%) occurred without street lighting. The geo-referenced raw data were imported in used GIS software, and the next chapter will explain detailed data filtering and a preprocessing protocol to utilise the data in further analysis.

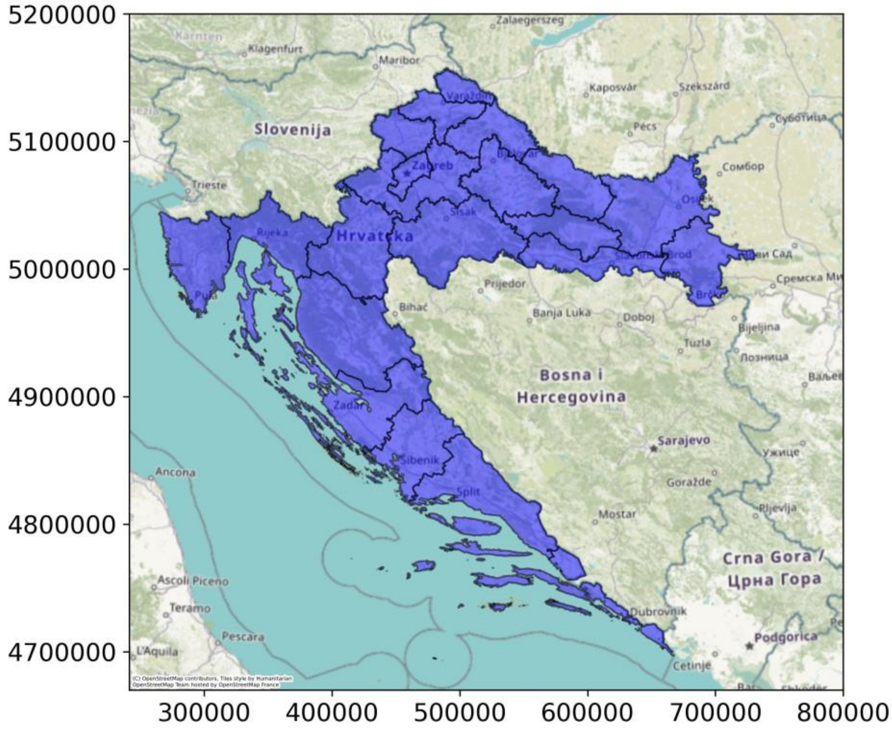

Figure 1 – Case study area (Republic of Croatia) including 21 counties

2.2 Data filtering and preprocessing

In this paper, we used geo-referenced crash data covering the period from 2018 to 2022 and geo-referenced boundary data including 21 counties of the Republic of Croatia. These data were further filtered and pre-processed in order to obtain relevant data sets for the purpose of exploratory spatial data analysis. The steps of the filtering and preprocessing are given in the following sub-sections.

Road crashes filtering

The crash data include several attributes for each crash record such as the unique identification crash number, crash occurrence date and time, geographic coordinates (WGS84), type of crash, crash consequences, as well as other attributes regarding the crash record. In this paper, only relevant attributes were considered for filtering. Two input crash data sets that we dealt with were prepared by applying a three-step filtering criterion performed in Quantum GIS (QGIS) environment with the algorithms "Extract by attribute" and "Extract by expression" [24]. As mentioned above, we filtered the records from the input original crash database according to our filter criterion to determine only the pedestrian crashes at night under the different street lighting conditions. The original crash database contains information on the visibility conditions (2 = night), type of crash (10 = pedestrian collision), and street lighting conditions (1 = functional, 2 = non-functional, 3 = not present) according to which the researchers obtained the two data sets of pedestrian crashes at night for two different street lighting conditions as shown in *Figure 2*. It is important to note here that the values of 2 and 3 of the attribute determining the street lighting condition were considered the condition without street lighting.

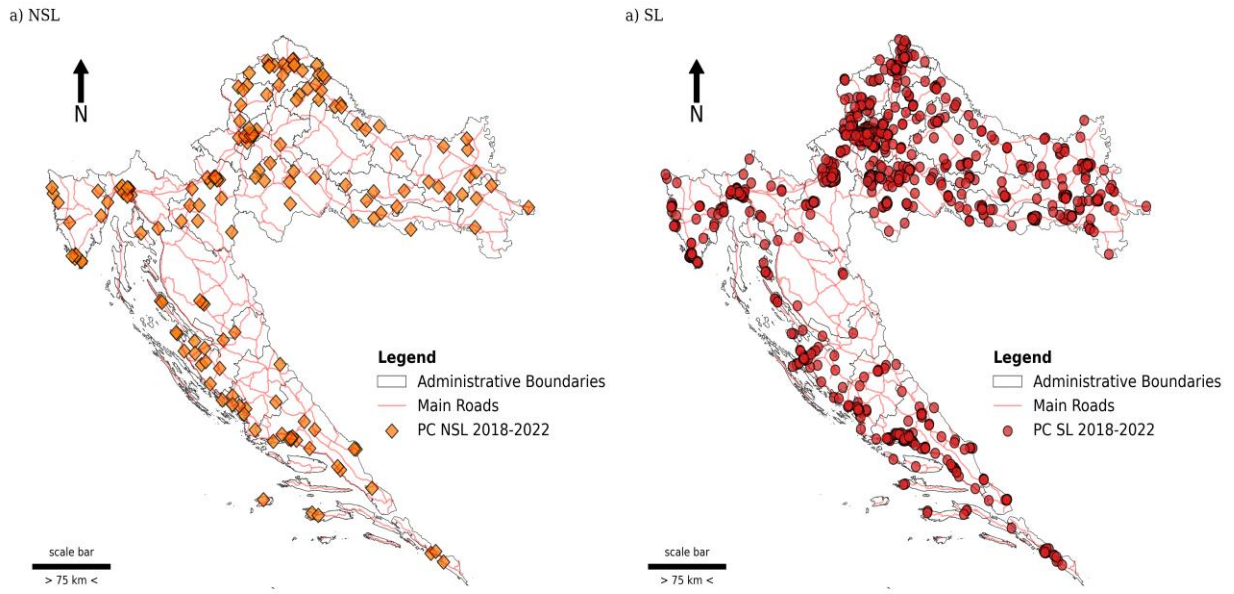

Figure 2 – Distribution of the night-time pedestrian crashes
(a) pedestrian crashes under the NSL condition; b) pedestrian crashes under the SL condition

Table 1 summarises the number of pedestrians involved in crashes with respect to the severity of the injury. For simplicity, we refer to the condition without street lighting as "No Street Lighting (NSL)" and the condition with street lighting as "Street Lighting (SL)".

Spatial level of crashes aggregation

Spatial aggregation is prone to the effects of the modifiable areal unit problem (MAUP), which has been thoroughly investigated and described by many researchers [25, 26]. According to Openshaw [29] and Fotheringham and Wong [30], the MAUP effect can be decomposed into two effects, namely the scale effect and the zoning effect. Both effects affect the statistical results in the analysis of the same data set in different ways. The scale effect occurs when the size of the spatial unit changes. On the contrary, the zoning effect comes into play when there is a change in the boundaries of spatial elements. The outcomes depend on the spatial aggregation through the scale and the shape of a spatial unit. Various published studies have used different scales and shapes of spatial aggregating units for geographical road crashes and road safety studies. These units range from different administrative units, such as the country unit [31, 32], the city unit [33] or the settlement unit [34], to traffic analysis zones (TAZs), such as census blocks [35] or census tracts [36]. Due to the fact that there is no analytical solution to the MAUP [37], the effect of the MAUP on statistical results can only be investigated by simulations of many different spatial units [38 p. 139]. In this paper, we considered the county spatial scale, as we believe that the use of this administrative-based spatial unit provides a valuable comparison of the crash clusters and allows for a better comparison with other demographic and economic characteristics.

Table 1 – Numbers of pedestrians involved in the night-time crashes under different street lighting conditions with respect to the severity of the injury

Injury severity	SL condition	NSL condition	Total
Fatality	83 (72%)	32 (28%)	115
Severe injury	466 (88%)	63 (12%)	529
Mild injury	907 (92%)	78 (8%)	985
No injury	157 (86%)	25 (14%)	182
Not determined	27 (90%)	3 (10%)	30

Night-time crashes aggregation

Subsequent exploratory spatial data analysis utilised the aggregates of filtered night-time pedestrian crashes based on the spatial unit and selected crash characteristics. First, we imported filtered crashes from comma-separated values files as the point layers into QGIS and saved them as point ESRI shapefiles. Second, the crash aggregation was carried out using the algorithm "Count points in polygon" within the QGIS environment [24], where the point layers of the crashes were laid over the polygon layer of the county and the total number of crashes was calculated for each county in the polygon. This method aggregated the points to the counties in the study region. In total, two aggregates per county level of night-time pedestrian crash counts for two street lighting conditions were generated and exported as polygon ESRI shapefiles.

2.3 Exploratory spatial analysis

In this paper, exploratory spatial analysis employs four spatial statistics to investigate spatial clustering and autocorrelation of night-time pedestrian crashes under two different street lighting conditions. Firstly, an ANND analysis of the spatial distribution of the crashes was performed to test whether the crashes were randomly distributed or not. Then the nearest neighbour distance function G was used to measure the distribution of the nearest neighbour distances. If randomness is present, the exploratory spatial analysis process will stop. In contrast, if there is no randomness, the univariate global Moran's I statistic was calculated to measure the overall degree of spatial clustering of the crashes. Finally, if the crashes show statistically significant global spatial autocorrelation, the LISA was performed as the univariate local indicator for the detection of clustering at the local level. CrimeStat 4.02 was used for ANND analysis, the spatstat package in R was used to perform nearest neighbour distance function G, and the Python Spatial Analysis Library (PySAL) was used to carry out both the global and local Moran's I analyses.

Average nearest neighbour analysis

Several studies have been published using different methods to examine the clustering of road crashes. Amiri et al. [39] compared 5 different techniques for mapping clusters of road crashes and detecting high-risk locations, including the average nearest neighbour, Getis-Ord Gi*, global Moran's I, kernel density estimation (KDE) and mean centre. He found that the average nearest neighbour technique is one of the techniques showing the best performance after the Moran's I in terms of accuracy. Anderson [40] used KDE to study the spatial patterns of injury-related crashes in London. However, he pointed out that the main drawback of the KDE is related to determining the statistical significance of the resulting clusters. Additionally, the density patterns generated by KDE are influenced by the choice of bandwidth. Furthermore, Nicholson [41] found that the nearest neighbour distance approach is a powerful and robust technique for detecting crash patterns and is still popular for use in exploratory spatial analyses of crashes [42, 43, 44]. Therefore, this paper considered the ANND technique in the analysis of the distribution of the point patterns of the crashes.

This method is based on the nearest neighbour index (NNI) calculation obtained as the division of the observed mean distance and random mean distance. The observed mean distance represents the minimum average distance between each data point and its nearest neighbour data point. In contrast, the random mean distance is the minimum average distance between neighbours in a randomly spatially generated point pattern. Based on the outcome value of the NNI it could be determined if a given point pattern tends to be clustered, dispersed or randomly distributed over a delimited space. If the observed mean distance is less than the random mean distance, the NNI is less than 1 indicating spatial clustering of the given point pattern. On the other hand, if the observed mean distance is greater than 1, then the pattern tends to be spatially dispersed. The index value of 1 represents the example of a randomly distributed point pattern.

To explore the significance of the calculated value of NNI, the measure of p-values and z-scores is used to indicate whether the null hypothesis can be rejected or not. The null hypothesis is that all points in the given pattern are randomly distributed. In other words, whether the given point pattern is statistically clustered or not. The nearest neighbour index is calculated according to the following formula:

$$NNI = \frac{D_O}{D_E} \tag{1}$$

where D_O is the mean distance between each point and its nearest neighbour calculated from the input data set and D_E is the expected mean distance for the points generated within a random pattern process. The observed mean distance D_O represents the average of minimum distances of each data point to its nearest data point and is calculated as follows:

$$D_O = \frac{\sum_{i=1}^{n} \sum_{i \neq j=1}^{n-1} Min(D_{ij})}{n} \tag{2}$$

where D_{ij} represents the minimum distance of a given data point to its nearest neighbour, and n is the number of points in the input data set.

The expected random mean D_E, calculated from the completely spatially random point pattern, is calculated by:

$$D_E = 0.5 \cdot \sqrt{\frac{A}{n}} \tag{3}$$

where A is the area of the study region. This area can be represented as a minimum rectangle enclosing all features, or it can be defined as a user-specified area [45]. The average nearest neighbour z-score to test the significance of the NNI is given by [46]:

$$z = \frac{D_O - D_E}{SE} \tag{4}$$

where SE is the standard error of the mean random distance which is approximately given by:

$$SE = \frac{0.26136}{\sqrt{\frac{n^2}{A}}} \tag{5}$$

The results of the ANND analysis of the pedestrian night-time crashes for two street lighting conditions are summarised in *Table 2*.

Nearest neighbour distance function G

The nearest neighbour distance function G is a cumulative distribution function that was used for describing the distribution of the nearest neighbour distances within two crash patterns and as the distance-based summary function to test if these patterns departure from the null hypothesis. This distance-based statistical method is based on the comparison of empirical and theoretical functions. The empirical function $\hat{G}(d)$ expresses the proportions of distances that are less than a particular distance and is estimated from the observed pattern [47]:

$$\hat{G}(d) = \frac{\sum_{i=1}^{n} \mathbf{1}(d_i \leq d)}{n} \tag{6}$$

where n is the number of points in the pattern, $\mathbf{1}$ is the indicator function, d_i is the nearest neighbour distance and d is the distance variable (inter-point distance). The indicator function takes a value:

$$\mathbf{1}(d_i) = \begin{cases} 1 & d_i \leq d \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where d_i is defined as:

$$d_i = min_{j \neq i}\{d_{ij}\}. \tag{8}$$

The theoretical function $G(d)$ expresses the proportions of the distances under the assumption of the null hypothesis [47]:

$$G(d) = 1 - e^{-\lambda \pi d^2} \tag{9}$$

where λ represents the mean number of events per unit area. To explore the crash patterns over different distances between points, the respective empirical and theoretical functions were calculated and compared. Additionally, to assess the departure of the observed crash patterns from the null hypothesis, the two-sided Monte Carlo tests based on global envelopes were performed. First, the empirical functions for the observed patterns and the theoretical functions were calculated under the assumption of the null hypothesis. Then 2 sets of 39 random simulated point patterns were generated for both lighting conditions using the estimated parameters from the observed patterns. Consequently, we obtained estimates for these random patterns. Finally, we calculated the test statistics T as the absolute maximum deviation over the range of distances d as follows [48]:

$$T = max_d \left| \widehat{G}_J(d) - G_{theo}(d) \right| \tag{10}$$

where $\hat{G}_j(d)$ represents estimated function of the j-th simulation in a set and $G_{theo}(d)$ represents theoretical function. In total, we calculated two statistics T_{NSL} and T_{SL} for the crashes under NSL and SL conditions, respectively. Then, the upper limit UL and lower limit LL of the envelopes were obtained as:

$$LL = G_{theo}(d) - T, \tag{11}$$

$$UL = G_{theo}(d) + T \tag{12}$$

From these limits, the global envelopes for the crashes under NSL and SL conditions were plotted. If the empirical function $\hat{G}(d)$ lies outside the envelope, the test rejects the null hypothesis. Otherwise, one can conclude that the crashes are randomly distributed. This test was performed with the significance level α of 0.05. Simulations and calculations were performed using the function “envelope” of the spatstat package [49].

Global spatial autocorrelation

Global measures of spatial autocorrelation are based on a single statistic that provides a general measure of the similarity between neighbours throughout the study region. In our case, these neighbours are represented by county spatial units of the Republic of Croatia. The most common statistic for global measures of spatial autocorrelation is the global Moran’s I statistic [50]. This statistic can assess a spatial relationship with the combination of spatial proximity captured by spatial weights and the attribute similarity captured by variable covariance. This statistic can be calculated for a given variable according to the following formula [51]:

$$I = \frac{n}{W_o} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij} \cdot (z_i - \bar{z}) \cdot (z_j - \bar{z})}{\sum_{i=1}^{n}(z_i - \bar{z})^2} \tag{13}$$

where n is the number of observations, W_o is the normalising factor, W_{ij} is the element in the spatial weights matrix that corresponds to the observation pair i and j, z_i and z_j are the observations for the areas i and j with the mean value $\bar{z}$. For the normalising factor we can write the following formula [51]:

$$W_o = \sum_{i=1}^{n} \sum_{j \neq i}^{n} W_{ij} \tag{14}$$

The spatial weights matrix was used in the row-standardised form, thus the $W_o = n$, and therefore *Equation 13* can be simplified to the following formula:

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \cdot (z_i - \bar{z}) \cdot (z_j - \bar{z})}{\sum_{i=1}^{n} (z_i - \bar{z})^2} \tag{15}$$

The value of global Moran's I ranges between -1 and +1, indicating spatial clustering of similar values and spatial spreading of dissimilar values, respectively. Values close to 0 suggest a random spatial pattern, thus removing spatial autocorrelation. It is important to note that the resulting value of global Moran's I is highly dependent on the chosen spatial weight matrix [52]. There are various methods to represent spatial relationships between features, namely distance-based, contiguity-based, K-nearest neighbour, kernel and Delaunay triangulation methods [53]. Gedamu et al. [54] reported that the distance-based weight matrix is very sensitive to the choice of distance and produces too many neighbours in the case of road crashes. Moreover, Saeed et al. [55] found that the fixed distance-based spatial weights matrix was statistically outperformed by the contiguity-based matrix. In this paper, we considered both types of spatial weights, the queen contiguity weight matrices based on first- and second-order neighbour, and the distance-based weights matrix.

For distance-based spatial weights, as Gedamu et al. [54] stated, the choice of distance has a great impact on the degree of spatial autocorrelation due to the different number of neighbours included in the calculation. To find the distance most appropriate for this situation, we performed a sensitivity analysis of several distances from which the distance value that maximises the degree of clustering phenomenon was chosen [56, 57]. First, we calculated the initial distance of 98,366.526522 m, at which any point has at least one neighbour. Then we determined the maximum band distance as the average of maximum distances from each data point to its farthest neighbour, which was 339,943.002017 m. Finally, we chose the distance increment of 10 m to obtain more accurate values. In total, 48,316 global Moran's I values for 24,158 various distances for both street lighting conditions were computed.

Inferencing was based on a random permutation procedure, which recalculates the statistic many times to generate a reference distribution. The obtained statistic is then compared to this reference distribution, and the pseudo-significance level of the p-value is computed. The permutation procedure was carried out with the default value of 999 permutations, as the effect of a higher number of permutations is marginal to this default value [52].

Local spatial autocorrelation

LISA was used as the univariate local indicator of spatial clusters and outliers in the night-time pedestrian crashes for both street lighting conditions. This statistic can represent two types of spatial clusters and also outliers (High-High, Low-Low, Low-High, High-Low) [58, 59]. The local Moran's I statistic for an area unit sample i in a given study area can be calculated according to the following formula [50]:

$$I_i = (z_i - \bar{z}) \cdot \sum_{j \epsilon J_i}^{n} W_{ij} \cdot (z_j - \bar{z})^2 \tag{16}$$

where J_i denotes the neighbourhood set of area i, and the summation in j runs only over those areas belonging to J_i, $\bar{z}$ denotes the average of these neighbouring observations [59]. The inference from this statistic was based on a conditional random permutation test, which was used to produce the so-called pseudo-significance levels. In that case, we used 999 permutations.

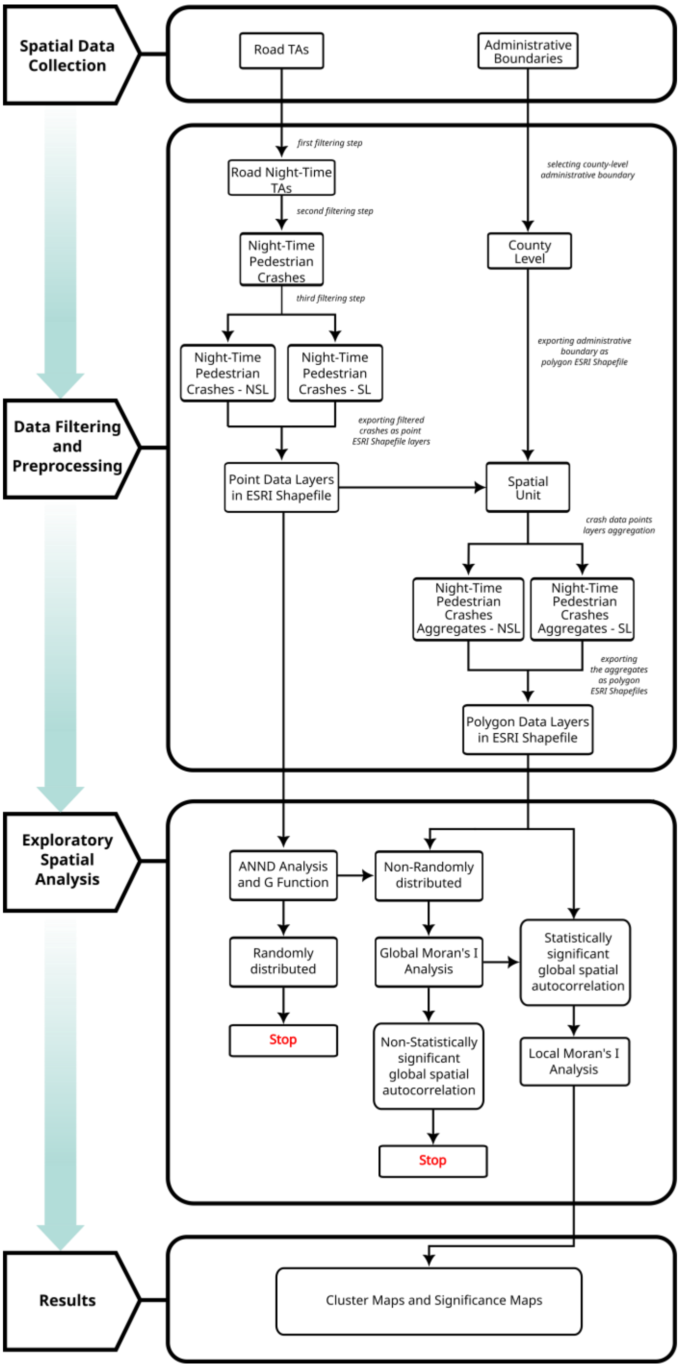

Figure 3 – Flow chart of the methodology

3. RESULTS

In this section, we present the results from four employed spatial statistics, such as the ANND, nearest neighbour distance function G, univariate global Moran's I and univariate local Moran's I. First, we present the results from the distance-based spatial statistics, and then the results from the global and local Moran's I are provided.

3.1 Average nearest neighbour analysis

Two NNIs of pedestrian night-time crashes under two different street lighting conditions were calculated. The results of the ANND analysis are summarised in *Table 2*. According to the NNI values in *Table 2* and their respective z-scores and p-values, it is evident that the point patterns of the crashes show statistically significant spatial clustering under both conditions. However, the NNI for the SL condition is by more than a half lower than for the NSL condition, indicating that the crashes in the SL condition tend to be more concentrated. Furthermore, the z-score (-55.8171) is almost five times higher in the SL condition compared to the NSL condition (-12.9703), demonstrating the higher magnitude of clustering. However, the ANND analysis does not prove the existence of hotspots. Therefore, it is essential to perform further investigation concerning the potential presence of spatial clusters.

Table 2 – Results of the ANND analysis of the crashes for both street lighting conditions during the study period (NSL - No Street Lighting, SL - Street Lighting)

Lighting condition	D_O (m)	D_E (m)	NNI	SE (m)	z-score	p-value
NSL	8072.94	16560.19	0.4875	654.36	-12.9703	0.0001
SL	1380.55	5911.97	0.2335	81.18	-55.8171	0.0001

3.2 Nearest neighbour distance function G

As can be seen in *Figure 4*, both empirical functions $\hat{G}(d)$ are mainly greater than their theoretical functions $G(d)$, suggesting that the crashes for both conditions (NSL and SL) tend to be much closer to each other compared to the null landscape (CRS). This proximity is most apparent for the crashes under the SL condition, where the very steep positive slope of the empirical function reveals that about 60% of the crashes appear to be the most clustered at distances less than approximately 1,000 m. This is consistent with the fact that there is a high concentration of crashes in the city centre. After that, the function decreases and reaches the theoretical function at distances greater than approximately 12,500 m, indicating that the crashes tend to be randomly located. The empirical function for crashes under the SL condition also indicates clustering. However, for distances less than 1,250 m, the crashes are slightly dispersed. For distances greater than this distance, the clusters appeared to be separated and scattered over larger distances, as the waveform of the empirical function resembles the staircase waveform. This behaviour is especially noticeable between distances of 1,250 and 7,500 m. Maximum clustering occurs at a distance of about 20,000 m. Additionally, both envelopes represented by the grey areas allow one to assess whether the point patterns depart from the null hypothesis. Because the empirical functions are above the envelopes, the crashes are more clustered than would be expected at random.

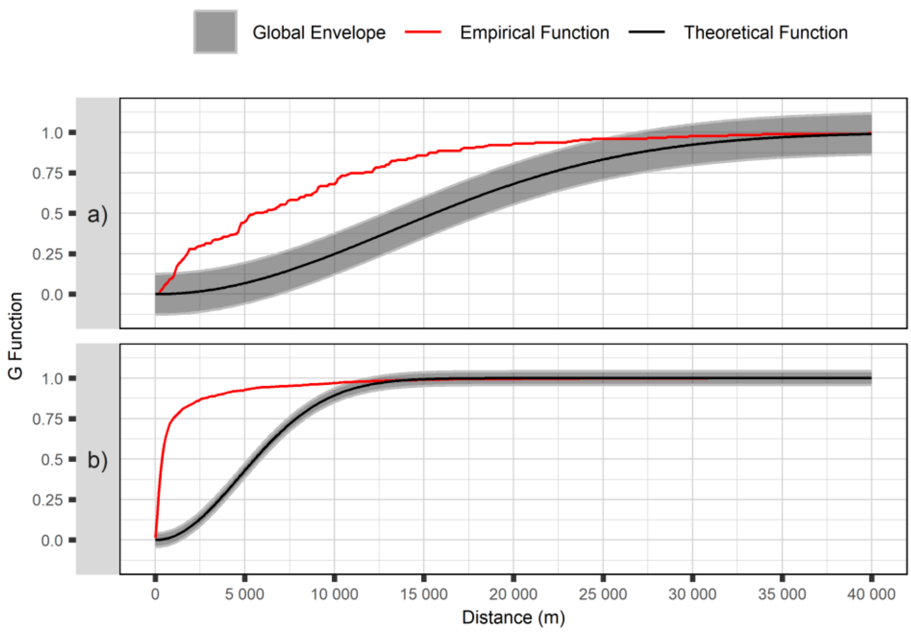

Figure 4 – The global envelopes from Monte Carlo simulations and respective empirical and theoretical functions; the grey areas represent the envelopes which enclose the G functions of 39 simulations of random point patterns (a) – NSL, (b) – SL

3.3 Global Moran's I analysis

Univariate global Moran's I analysis and statistical test were used to identify statistically significant spatial clustering of pedestrian night-time crashes under two different street lighting conditions. *Table 3* and *Table 4* summarise the results of global Moran's I analysis of the crashes for considered spatial weights and for both street lighting conditions.

The values of Moran's I for the crash counts under the NSL condition are almost positive for all the types of spatial weights indicating positive spatial clustering, except for the 1st order queen contiguity matrix. In that case, the Moran's I negative value reaches zero (-0.033236), suggesting a weak negative spatial autocorrelation. However, its very low z-score value (0.099520) and high pseudo-p-value (0.455) classify this result as statistically non-significant. For the 2nd order queen contiguity matrix, the Moran's I value is positive and noticeably higher, with a z-score value of 1.542897 reaching the threshold of 1.96 at a 95% confidence interval. However, the pseudo-p-value (0.077) is still above the value of 0.05, and thus this result cannot be classified as statistically significant.

On the other hand, the use of a distance-based spatial weights matrix yields a statistically significant positive autocorrelation with the Moran's I value of 0.181927, accompanied by the highest z-score (2.091764) and the lowest pseudo-p-value (0.032). This result indicates a high degree of spatial clustering. The bandwidth at which the maximum magnitude of clustering was found according to the distance sensitivity analysis is 124,687 km. In contrast, the values of Moran's I for the crash counts under the SL condition are negative for all spatial weight matrices. Their values hover above zero accompanied by low z-score values and high pseudo p-values. This result suggests a weak and non-statistically significant negative spatial autocorrelation. However, it is interesting to note that the trend of the values of Moran's I for both conditions is increasing as the spatial weight matrices change.

*Table 3 – Summary of the results of the global Moran's I analysis for night-time pedestrian crash counts under the NSL condition (*124,687 km)*

Spatial weight	Moran's I	Expected index	Variance	z-score	pseudo-p-value
Queen-1	-0.033236	-0.050463	0.029965	0.099520	0.455
Queen-2	0.110977	-0.047951	0.010610	1.542897	0.077
Fixed distance*	0.181927	-0.060779	0.013463	2.091764	0.032

Table 4 – Summary of global Moran's I analysis results for night-time pedestrian crash counts under the SL condition (†286,752 km)

Spatial weight	Moran's I	Expected index	Variance	z-score	pseudo-p-value
Queen-1	-0.072226	-0.048209	0.017173	-0.183263	0.452
Queen-2	-0.064434	-0.050549	0.005858	-0.181423	0.475
Fixed distance†	-0.029332	-0.051325	0.000368	1.146721	0.110

3.4 Univariate local Moran's I

Since crash counts under the SL condition showed significant spatial autocorrelation, these counts were not considered in the subsequent analysis of local clustering. The authors focused only on crash counts under the NSL condition with the distance-based spatial weights matrix, where the clustering phenomenon was maximised.

Global Moran's I analysis might not detect localised clusters; therefore, we employed LISA using the univariate local Moran's I index to calculate the values of local Moran's I and generate cluster maps and their respective significance maps. LISA represents the indicator of local spatial autocorrelation, providing additional information on local clustering [59] and can identify the locations of these local patterns within the study area.

Results from the examination of the local spatial autocorrelation of the crash counts using LISA are summarised in the form of the cluster map and significance map in *Figure 5*. The cluster map (*Figure 5a*) shows the spatial distribution of the local indicators of High-High (H-H), Low-Low (L-L), Low-High (L-H) and High-Low (H-L) county clusters with crash count values. The significance map (*Figure 5b*) shows the statistical significance of each county cluster in the cluster map at three different levels of significance pseudo-p-value. Significant H-H indicators are represented as dark-red counties where the high crash count value is surrounded by another high crash count value.

On the contrary, significant L-L indicators are represented by dark-blue counties that show the locations where the low crash count value is surrounded by another low crash count value. The significant indicators H-L and L-H represent the locations where a high crash count value is surrounded by a low crash count value and vice versa, respectively. Significant H-H and L-L indicators are dominant over the H-L and L-H clusters and are separated by the statistically non-significant counties. Clusters of significant H-H indicators were found in the coastal regions of Croatia. On the other hand, significant L-L indicators are dominant in the Slavonia region, in the East close to the border with Serbia, and the West close to the centre of Croatia.

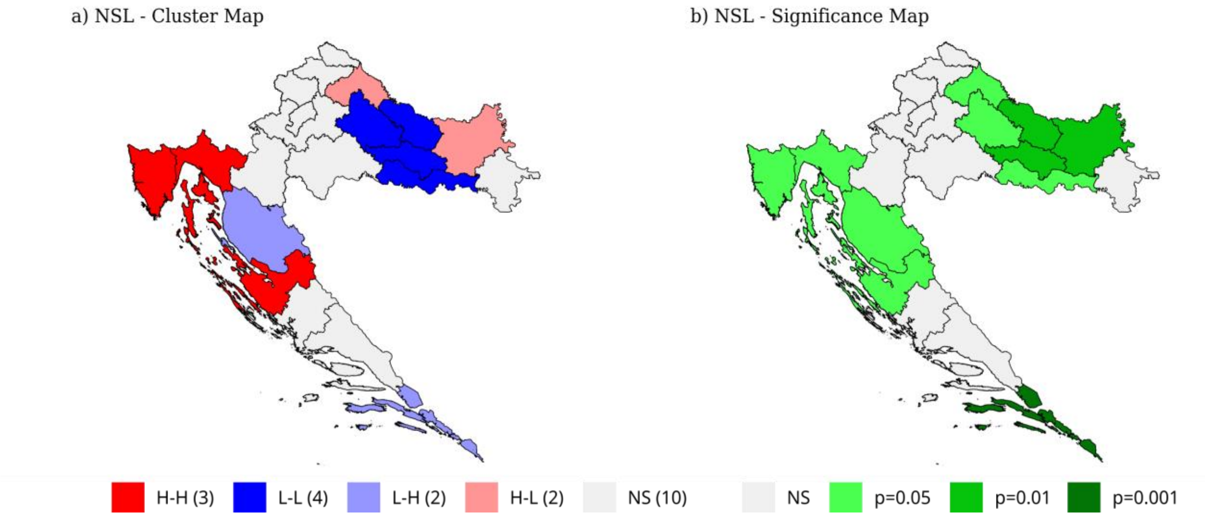

Figure 5 – Cluster and significance maps for night-time pedestrian crash counts under the NSL condition (a) cluster map; b) significance map

Additionally, *Table 5* summarises the descriptive characteristics of the county clusters by calculating the average values of the number of pedestrians who are fatally, seriously and mildly injured. The average values in the H-H clusters typically stand out in all injury severities relative to the average values of the L-L, H-L and L-H clusters or of the non-significant clusters. The average value of pedestrian fatalities in the H-H clusters is almost three times higher than in the L-L clusters. Furthermore, the average value of severely injured pedestrians in the H-H clusters is four times higher than in the L-L clusters. Although the L-L clustered counties have a significantly lower average value of severely injured pedestrians than the H-H clustered counties, the H-L clustered counties have the same average value as the H-H clusters. On the other hand, the average values of pedestrian fatalities and severe injuries are the lowest in the L-L clusters.

Table 5 – Descriptive characteristics of clustered and non-clustered counties with respect to average values of fatally, severely and mildly injured pedestrians (H-H: High-High clusters, L-L: Low-Low clusters, H-L: High-Low clusters, L-H: Low-High clusters, NS: Non-Significant clusters)

Injury severity	H-H	L-L	H-L	L-H	NS	All
Fatally injured (average value)	2.00	0.75	1.00	1.00	1.90	1.52
Severely injured (average count)	5.00	1.25	5.00	2.00	2.90	3.00
Mildly injured (average count)	5.33	1.00	4.00	2.00	4.60	3.71

4. DISCUSSION

The results from the ANND analysis suggest that pedestrian nigh-time crashes with the presence of street lighting tend to be more clustered than crashes without street lighting. This finding was confirmed by using the nearest neighbour distance function G and Monte Carlo tests based on global envelopes, where the clustering was most apparent at distances less than 1,000 m. One potential explanation for this result may lie in the different number of crashes since the calculation of the NNI is dependent on the total number of points and the selected area. However, in both conditions, the NNIs are significantly less than one, implying that the point patterns of pedestrian crashes can be described as statistically clustered, as other studies have shown [41, 61].

Global Moran's I analysis revealed the differences between the observed values of the global Moran's I of the investigated crash counts for two street lighting conditions. The observed values for the crash counts under the SL condition were all negative (ranges between -0.072 to -0.029) and statistically non-significant with the pseudo-p-values above the threshold of 0.05 for all considered spatial weights matrices. In contrast, for the crashes under the NSL condition noticeably higher values of global Moran's I were observed with one statistically significant global Moran's I value of 0.182. This observation indicates the presence of a high degree of positive spatial autocorrelation for crashes under the NSL condition. The authors suggested that the greatest difference in the observed values of global Moran's I, which can be seen in the case of distance-based spatial weights, could be attributed to the number of neighbours incorporated into the calculation of global Moran's I, as stated by Lee and Khattak [62].

The LISA analysis provides interesting results on localised clustering, where the spatially segregated H-H clusters are located in the coastal regions of Croatia, while the concentrated L-L clusters are mainly situated in Slavonia, close to the centre of Croatia. One possible explanation could be attributed to the factors associated with the roadways and the behaviour of drivers and pedestrians that differ between the identified locations of the H-H and L-L clusters. Higher design speed and speed limits, long-distance travel, fewer unprotected pedestrian access and pedestrian crossing facilities accompanied by impaired surrounding lighting conditions could contribute to the higher concentration of night-time pedestrian crashes at the locations of the H-H indicators.

The H-L and L-H clusters are spatially segregated by the L-L clusters and the H-H clusters, respectively, situated in Croatia proper, Dalmatia and Slavonia. While the statistically significant H-H clusters, evincing a high number of fatally or severely injured pedestrians, are located in the outskirts of Croatia, the L-L clusters with a lower number of severely injured pedestrians are more likely to occur closer to the centre of Croatia. This finding is in agreement with what was found by Gademu et al. [54] and Lee and Khattak [62] and could be attributed to the different characteristics of the road, pedestrian and driver, as reported by Siddiqui et al. [63]. Previous research by Khatun et al. [64] reported that areas with a higher concentration of serious crashes located on highways are often characterised by a lack of streetlights. Furthermore, Hossain et al. [14] studied the patterns of pedestrian crashes at night on roads and intersections and found that most pedestrian crashes at night are associated with roads without physical separation and streetlights.

Additionally, driver and pedestrian conditions such as drug or alcohol involvement were reported as factors that contribute to the higher risk of pedestrian involvement in crashes at night [14, 65]. The results showed that the used approach effectively reveals spatial dependence between night-time pedestrian crashes without street lighting. Furthermore, the method identified counties where the higher counts of crashes were concentrated during the study period. The integration of the distance-based spatial statistics with the global and local spatial autocorrelation overcame the drawbacks of the KDE method.

However, this study is limited only to one spatial unit, which cannot reflect the influence of the scale and zoning effects on the observed results from a broader perspective; thus, the different units may produce different cluster patterns. Additionally, the coarser level of spatial aggregation cannot yield more details about additional potential clusters around the identified significant clusters at the county spatial scale. Therefore, it is recommended to further investigate the robustness of the presented findings across different spatial units. Secondly, errors may arise from the input crash database, as some records are not clear about their locations. Moreover, some crashes can be reported with incorrect information on street lighting and visibility conditions. Thirdly, the quality of the present street lighting (for instance, the luminance level) is not considered in this paper, since the data on the current quality of lighting are not collected by the police. Furthermore, the information about road lighting was determined only as the presence or absence of street lighting without considering the distance from a street lighting point to the accident location.

Future work could be improved by considering different spatial units to assess the replicability of the presented findings regarding the crash clustering, as well as incorporating other spatial data providing more detailed information regarding the street lighting, such as the night-time satellite images. In addition to that, the present paper is aimed only at investigating the presence of spatial autocorrelation with respect to different street lighting conditions. Therefore, another analysis of factors associated with the spatial clusters of crashes and their variations in space is needed.

5. CONCLUSION

This paper aims to investigate the spatial autocorrelation of night-time pedestrian crashes in the Republic of Croatia under two different street lighting conditions using global and local Moran's I statistics. The results showed that the spatial statistics employed can provide further information on the clustering of crashes in the study area with respect to different street lighting conditions.

ANND analysis confirmed that the spatial patterns of the crashes for both street lighting conditions exhibit statistically significant spatial clustering, with a more pronounced clustering for crashes with the presence of street lighting. In addition to that, the G function was applied to measure the distribution of the nearest neighbour distances for both patterns, which confirmed the clustering of the crashes. The function also revealed that the clustering of the crashes under the condition SL occurs at small distances compared to crashes under the NSL condition where the clustering is somewhat scattered across the larger distances. Two-sided Monte Carlo tests confirmed that both crash patterns deviate from the CRS. Global Moran's I analysis revealed high and statistically significant positive spatial autocorrelation of the crash counts without street lighting for the distance-based spatial weights matrix. Moreover, the crash counts with street lighting evince weak and statistically non-significant negative spatial autocorrelation for both the contiguity-based and distance-based spatial weights. According to the results from the global Moran's I analysis, the analysis of LISA was applied only to crashes without the presence of street lighting, since values of global Moran's I for crashes with street lighting indicated no or sporadic spatial autocorrelation. The local Moran's I analysis revealed that Istra County, Primorje-Gorski Kotar County and Zadar County stand out as the counties with the highest number of pedestrian crashes. Furthermore, these locations were identified as the most serious locations with respect to the levels of severity of the injury. On the other hand, the locations with the lowest number of pedestrian crashes were identified in the counties in the Slavonia region.

Despite the limitations of this paper, it can be stated that night-time pedestrian crashes without street lighting evince some spatial dependence on both the global and local scales compared to crashes with street lighting. At the global scale, the dependence is significantly positive for the distance-based weight matrix, which reflects the wider interactions between neighbouring counties as people's mobility can cross the borders of neighbouring counties. At the local level, spatial dependence was found in several distinct clusters of counties. This clustering appears in counties located in the coastal regions of Croatia, where the lack of street lighting could be considered an important factor contributing to a higher concentration of night-time pedestrian crashes. These findings can attract more attention from road authorities to the area of road lighting design, which can improve road safety at certain locations. Concrete steps leading to a reduction in crashes caused by insufficient street lighting vary from case to case and each case must be assessed individually. Generally speaking, the lighting design process is carried out with the aid of lighting-calculation software, which calculates lighting parameters for considered lighting installation, road surface and road geometry. However, the authors suggest that further inspections that account for lighting conditions and pedestrian activity are conducted.

REFERENCES

[1] Bougna T, Hundal G, Taniform P. Quantitative analysis of the social costs of road traffic crashes literature. *Accident Analysis & Prevention*. 2022;165:106282. DOI: 10.1016/j.aap.2021.106282.

[2] Atasayar H, Fleischer M, Donabauer M. Annual statistical report on road safety in the EU 2024, 2024. Last visited 2024-04-09. URL: https://road-safety.transport.ec.europa.eu/european-road-safety-observatory/data-and-analysis/annual-statistical-report_en.

[3] Piatkowski D, Krizek K, Handy S. Accounting for the short term substitution effects of walking and cycling in sustainable transportation. *Travel Behaviour and Society*. 2015;2(1):32–41. DOI: 10.1016/j.tbs.2014.07.004.

[4] Johansson Ö, Wanvik PO, Elvik R. A new method for assessing the risk of accident associated with darkness. *Accident Analysis & Prevention*. 2009;41(4):809–815. DOI: 10.1016/j.aap.2009.04.003.

[5] Shanshal D, Babaoglu C, Basar A. Prediction of fatal and major injury of drivers, cyclists, and pedestrians in collisions. *Promet - Traffic & Transportation*. 2020;32(1):39–53. DOI: 10.7307/ptt.v32i1.3134.

[6] Wood J. Nighttime driving: visual, lighting and visibility challenges. *Ophthalmic and Physiological Optics*. 2020;40(2):187–201. DOI: 10.1111/opo.12659.

[7] Mohamed MG, Saunier N, Miranda-Moreno L, Ukkusuri S. A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. *Safety Science*. 2013;54:27–37. DOI: 10.1016/j.ssci.2012.11.001.

[8] Haleem K, Alluri P, Gan A. Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis & Prevention*. 2015;81:14–23. DOI: 10.1016/j.aap.2015.04.025.

[9] Scarano A, Rella Riccardi M, Mauriello F, Montella A. Analysis of contributory factors of fatal pedestrian crashes by mixed logit model and association rules. *International Journal of Injury Control and Safety Promotion*. 2023;30(2):195–209. DOI: 10.1080/17457300.2022.2116647.

[10] Fan W, Song L, Li Y, Liu P. Mixed logit approach to analyzing pedestrian injury severity in pedestrian-vehicle crashes in North Carolina: Considering time-of-day and day-of-week. *Traffic Injury Prevention*. 2021;22(7):524–529. DOI: 10.1080/15389588.2021.1940983.

[11] Ferenchak N, Gutierrez R, Singleton P. Shedding light on the pedestrian safety crisis: An analysis across the injury severity spectrum by lighting condition. *Traffic Injury Prevention*. 2022;23(7):434–439. DOI: 10.1080/15389588.2022.2100362.

[12] Kemnitzer C, et al. An investigation of driver, pedestrian, and environmental characteristics and resulting pedestrian injury. *Traffic Injury Prevention*. 2019;20(5):510–514. DOI: 10.1080/15389588.2019.1612886.

[13] Kim J, Ulfarsson G, Shankar V, Mannering F. A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis & Prevention*. 2010;42(6):1751–1758. DOI: 10.1016/j.aap.2010.04.016.

[14] Hossain A, et al. Exploring nighttime pedestrian crash patterns at intersection and segments: Findings from the machine learning algorithm. *Journal of Safety Research*. 2023;87:382–394. DOI: 10.1016/j.jsr.2023.08.010.

[15] Montella A, Aria M, D'Ambrosio A, Mauriello F. Data-mining techniques for exploratory analysis of pedestrian crashes. *Transportation Research Record*. 2011;2237(1):107–116. DOI:10.3141/2237-12.

[16] Ferenchak N, Ghodrat Abadi M. Nighttime pedestrian fatalities: A comprehensive examination of infrastructure, user, vehicle, and situational factors. *Journal of Safety Research*. 2021;79:14–25. DOI: 10.1016/j.jsr.2021.07.002.

[17] Crabb GI, Crinson LF. The impact of street lighting on night-time road casualties. *Engineering, Environmental Science*. 2008. URL: https://api.semanticscholar.org/CorpusID:106828872.

[18] Aprimah Apuseyine B, Ackaah W, Afukaar F. Road traffic crashes at night-time: characteristics and risk factors. *International Journal of Injury Control and Safety Promotion*. 2020;27(3):392–399. DOI: 10.1080/17457300.2020.1785508.

[19] Hossain A, Sun X, Thapa R, Codjoe J. Applying association rules mining to investigate pedestrian fatal and injury crash patterns under different lighting conditions. *Transportation Research Record*. 2022;2676(6):659–672. DOI: 10.1177/03611981221076120.

[20] Gu Z, Peng B. Investigation into the built environment impacts on pedestrian crash frequencies during morning, noon/afternoon, night, and during peak hours: a case study in Miami county, Florida. *Journal of Transportation Safety & Security*. 2021;13(9):915–935. DOI: 10.1080/19439962.2019.1701164.

[21] Ewing R, Dumbaugh E. The built environment and traffic safety: A review of empirical evidence. *Journal of Planning Literature*. 2009;23(4):347–367. DOI: 10.1177/0885412209335553.

[22] Ćosić M, Šimunović L, Jakovljević M. Relationships between external factors and pedestrian accident blackspots – a case study of the city of Zagreb. *Promet – Traffic & Transportation*. 2019;31:329–340. DOI: 10.7307/ptt.v31i3.3119.

[23] Kučinić T, Rajsman M, Kunštek A, Velić Ž. Analysis of pedestrian road safety in the republic of Croatia. *International Scientific Journal "Trans & Motauto World"*. 2023;8(2):53–56. URL: https://stumejournals.com/journals/tm/2023/2/53.

[24] QGIS.org. QGIS Geographic Information System. Open Source Geospatial Foundation Project, 2024. URL: https://qgis.org/en/site/.

[25] Gehlke CE, Biehl K. Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*. 1934;29(185A):169–170. DOI: 10.1080/01621459.1934.10506247.

[26] Taylor PJ, Openshaw S. A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. *Statistical Applications in the Spatial Sciences*. Ed. Wrigley N (Pion, London) 1979;127–144.

[27] Baumann JH, Fischer MM, Schubert U. A multiregional labour supply model for Austria: The effects of different regionalisations in multiregional labour market modelling. *Papers of the Regional Science Association*. 1983;52(1):53–83. DOI: 10.1007/BF01944095.

[28] Briant A, Combes PP, Lafourcade M. Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*. 2010;67(3):287–302. DOI: 10.1016/j.jue.2009.09.014.

[29] Openshaw S. The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*. 1984. URL: https://cir.nii.ac.jp/crid/1570291225725496704.

[30] Fotheringham AS, Wong DWS. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A: Economy and Space*. 1991;23(7):1025–1044. DOI: 10.1068/a231025.

[31] Chen S, Kuhn M, Prettner K, Bloom D. The global macroeconomic burden of road injuries: estimates and projections for 166 countries. *The Lancet Planetary Health*. 2019;3(9):e390–e398. DOI: 10.1016/S2542-5196(19)30170-6.

[32] Daniels S, et al. A systematic cost-benefit analysis of 29 road safety measures. *Accident Analysis & Prevention*. 2019;133:105292. DOI: 10.1016/j.aap.2019.105292.

[33] Moeinaddini M, Asadi-Shekari Z, Zaly Shah M. The relationship between urban street networks and the number of transport fatalities at the city level. *Safety Science*. 2014;62:114–120. DOI: doi:10.1016/j.ssci.2013.08.015.

[34] Quddus, M. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. *Accident Analysis & Prevention*. 2008;40(4):1486–1497. DOI: doi:10.1016/j.aap.2008.03.009.

[35] Lee J, Abdel-Aty M, Cai Q. Intersection crash prediction modeling with macro-level data from various geographic units. *Accident Analysis & Prevention*. 2017;102:213–226. DOI: 10.1016/j.aap.2017.03.009.

[36] Ukkusuri S, Hasan S, Abdul Aziz HM. Random parameter model used to explain effects of built-environment characteristics on pedestrian crash frequency. *Transportation Research Record*. 2011;2237(1):98–106. DOI: 10.3141/2237-11.

[37] Openshaw S, Taylor PJ. The modifiable areal unit problem. In: *Quantitative Geography: A British View*. Edited by N. Wrigley and R. Bennett, 1981;60–69.

[38] Longley P, Goodchild M, Maguire D, Rhind D. Geographic Information Science and System. Wiley, Chichester, 2001.

[39] Amiri AM, Nadimi N, Khalifeh V Shams M. GIS-based crash hotspot identification: a comparison among mapping clusters and spatial analysis techniques. *International Journal of Injury Control and Safety Promotion*. 2021;28(3):325–338. DOI: 10.1080/17457300.2021.1925924.

[40] Anderson T. Kernel density estimation and k-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*. 2009;41(3):359–364. DOI: doi:10.1016/j.aap.2008.12.014.

[41] Nicholson A. Analysis of spatial distributions of accidents. *Safety Science*. 1998;31(1):71–91. DOI: 10.1016/S0925-7535(98)00056-3.

[42] Keskin F, et al. Analysis of traffic incidents in METU campus. *Procedia - Social and Behavioral Sciences*. The 2nd International Geography Symposium-Mediterranean Environment 2010. 2011;19:61–70. DOI: 10.1016/j.sbspro.2011.05.108.

[43] Saba J, Sesugh A, Danladi G. The spatial patterns of road traffic crash black spots and emergency facilities in federal capital city. *Journal of Geoscience and Environment Protection*. 2023;11:121–134. DOI: 10.4236/gep.2023.112008.

[44] Rangam H, Sivasankaran SK, Balasubramanian V. Visual hazardous models: A hybrid approach to investigate road hazardous events. *Accident Analysis & Prevention*. 2024;200:107556. DOI: 10.1016/j.aap.2024.107556.

[45] Yang J, et al. Spatial distribution patterns and intra-specific competition of pine (pinus yunnanensis) in abandoned farmland under the sloping land conservation program. *Ecological Engineering*. 2019;135:17–27. DOI: 10.1016/j.ecoleng.2019.04.026.

[46] Clark P, Evans F. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 1954;35(4):445–453. URL: http://www.jstor.org/stable/1931034

[47] Bivand R, Pebesma E, Gómez-Rubio V. Applied Spatial Data Analysis with R. 2013.

[48] Baddeley A, et al. On tests of spatial pattern based on simulation envelopes. *Ecological Monographs*. 2014;84:477-489. DOI: 10.1890/13-2042.1.

[49] Baddeley A, Rubak E, Turner R. Spatial Point Patterns: Methodology and Applications with R. *Chapman and Hall/CRC Press*, 2015.

[50] Anselin L. Local indicators of spatial association–LISA. *Geographical Analysis*. 1995;27(2):93–115. DOI: 10.1111/j.1538-4632.1995.tb00338.x.

[51] Cliff AD, Ord JK. Spatial Processes: Models and Applications. Pion, London, 1981.

[52] Anselin L. An Introduction to Spatial Data Science with GeoDa: Volume 1: Exploring Spatial Data (1st ed.). Chapman and Hall/CRC, 2024. DOI: 10.1201/9781003274919.

[53] Dubé J, Legros D. Spatial Autocorrelation, chapter 3, pages 59–91. John Wiley & Sons, Ltd, 2014. DOI: 10.1002/9781119008651.ch3.

[54] Gedamu WT, Plank-Wiedenbeck U, Wodajo BT. A spatial autocorrelation analysis of road traffic crash by severity using moran's i spatial statistics: A comparative study of Addis Ababa and Berlin cities. *Accident Analysis & Prevention*. 2024;200:107535. DOI: 10.1016/j.aap.2024.107535.

[55] Saeed TU, Nateghi R, Hall T, Waldorf B. Statistical analysis of area-wide alcohol-related driving crashes: A spatial econometric approach. *Geographical Analysis*. 2020;52(3):394–417. DOI: 10.1111/gean.12216.

[56] Maingi J, Mukeka J, Kyale D, Muasya R. Spatiotemporal patterns of elephant poaching in south-eastern Kenya. *Wildlife Research*. 2012;39(3):234–249. DOI: 10.1071/WR11017.

[57] Liu P, Le KG, Lin LT. Traffic accident hotspot identification by integrating kernel density estimation and spatial autocorrelation analysis: a case study. *International Journal of Crashworthiness*. 2022;27(2):543–553. DOI: 10.1080/13588265.2020.1826800.

[58] Anselin L. A local indicator of multivariate spatial association: Extending Geary's c. *Geographical Analysis.* 2019;51(2):133–150. DOI: 10.1111/gean.12164.

[59] Fischer M, Wang J. Spatial data analysis: Models, methods and techniques. *Springer Nature*, 2011. DOI: 10.1007/978-3-642-21720-3.

[60] Wang Z, et al. Spatially varying relationships between land subsidence and urbanization: A case study in Wuhan, China. *Remote Sensing*. 2022;14(2):291. DOI: 10.3390/rs14020291.

[61] Apardian R, Alam BM. Pedestrian fatal crash location analysis in Ohio using exploratory spatial data analysis techniques. *Transportation Research Record*. 2020;2674(11):888–900. DOI: 10.1177/0361198120950717.

[62] Lee M, Khattak A. Case study of crash severity spatial pattern identification in hot spot analysis. *Transportation Research Record*. 2019;2673(9):684–695. DOI: 10.1177/0361198119845367.

[63] Siddiqui N, Chu X, Guttenplan M. Crossing locations, light conditions, and pedestrian injury severity. *Transportation Research Record*. 2006;1982(1):141–149. DOI: 10.1177/0361198106198200011.

[64] Khatun MS, Hossain MA, Kabir MA, Rahman MA. Identification and analysis of accident black spots using Geographic Information System (GIS): A study on Kushtia-Jhenaidah national highway (N704), Bangladesh. *Heliyon*. 2024;10(3):e25952. DOI: 10.1016/j.heliyon.2024.e25952.

[65] Rahman M, Kockelman K, Perrine K. Investigating risk factors associated with pedestrian crash occurrence and injury severity in Texas. *Traffic Injury Prevention*. 2022;23(5):283–289. DOI: 10.1080/15389588.2022.2059474.

František Kekula, Bernard Kosovec, Darko Babić, Pavel Hrubeš

Zkoumání a porovnávání prostorových shluků nočních dopravních nehod s chodci na základě různých stavů veřejného osvětlení na úrovni prostorové jednotky okresu

Abstrakt

Článek se snaží stanovit roli veřejného osvětlení v prostorovém shlukování nočních dopravních nehod s účastí chodců v Chorvatské republice. Zahrnuta byla pětiletá (2018-2022) data o noční dopravní nehodovosti s chodci za uvažovaných podmínek bez přítomnosti veřejného osvětlení a s přítomností veřejného osvětlení. Nejprve byly použity statistické metody založené na vzdálenostech pro posouzení míry prostorového shlukování vzorů nehod a jejich odchylek od úplné prostorové náhodnosti. Za druhé byla provedena globální analýza pomocí Moranova I za účelem vyšetření míry prostorové autokorelace ročních počtů nehod agregovaných do 21 správních jednotek Chorvatské republiky. Nakonec byly použity Lokální indikátory prostorové asociace (LISA) k identifikaci hotspotů. Výsledky ANND analýzy potvrzují významné shlukování nehod pro obě uvažované podmínky veřejného osvětlení. Byly získány odlišné hodnoty globálního Moranova I pro obě podmínky s vysokou a statisticky významnou pozitivní hodnotou pro počty nehod bez přítomnosti osvětlení. Analýza lokálního Moranova I odhaluje, že hodnoty Vysoká – Vysoká jsou lokalizovány ve správních jednotkách na pobřeží Chorvatska, zatímco hodnoty Nízká – Nízká se objevují ve východní části Chorvatska, vedle Slavonie. Výsledky naznačují, že neadekvátní světelné podmínky mají vliv na shlukování nočních dopravních nehod s účastí chodců.

Klíčová slova
noční dopravní nehody s chodci; veřejné osvětlení; statistické metody založené na vzdálenostech; prostorová autokorelace; globální Moranovo I; lokální Moranovo I.