SPEED-RELATED TRAFFIC ACCIDENT ANALYSIS USING GIS-BASED DBSCAN AND NNH CLUSTERING

B. Topcuoglu [1,2 *], T. Memisoglu Baykal [2,3], H. Tuydes Yaman[2]

[1] Parabol Yazılım, METU Technopolis, Ankara, Turkey – baran.topcuoglu@paraboly.com
[2] Dept. of Civil Engineering, Middle East Technical University, Ankara, Turkey – (baran.topcuoglu, tugbam, htuydes)@metu.edu.tr
[3] Dept. of Geomatic Engineering, Artvin Coruh University, Artvin, Turkey - tugbamemisoglu@artvin.edu.tr

Commission IV, WG IV/4

KEY WORDS: Speed-related traffic accident, spatial clustering, GIS, DBSCAN, NNH Clustering, Turkey.

ABSTRACT:

To ensure road safety and reduce traffic accidents, it is essential to determine geographical locations where traffic accidents occur the most. Spatial clustering methods of hot spots are used very effectively to detect such risky areas and take precautions to minimize or even avoid fatal or injury accidents. This study aims to determine speed-related hot spots in the pilot Mersin Region, which includes seven cities in the central-southern part of Turkey. Two different hot spot clustering methods, the Nearest Neighbourhood Hierarchical Clustering Method (NNH) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) Method, were employed to analyse traffic accident data between 2014-2021, obtained from the General Directorate of Highways. CrimeStat III program, which is free software, was used to detect NNH clusters, while the DBSCAN clusters were obtained using the open-source GIS program QGIS, which was also used to visualize and evaluate the results comparatively. As a result of the analysis, it was determined which method gave more effective results in determining the traffic accident risk clusters. These clusters were analysed based on road geometries (intersection or corridors). In addition, by considering the areas where repeated accidents have occurred over the years, future predictions of traffic accidents have been estimated.

1. INTRODUCTION

Resulting in many deaths and injuries every year, traffic accidents are a significant problem globally. It is even more critical in developing countries, such as Turkey, where the motorization rates increase rapidly (WHO, 2021). When analysed, it is seen that the probability of occurrence of traffic accidents is not random. Factors such as road condition and geometry, traffic flow and speed characteristics as well as topography of the land play an essential role in the event of an accident at one point (WHO, 2021). For this reason, traffic accidents tend to cluster in some areas; and it is critical to identify such areas and take the necessary measures to ensure road safety and reduce traffic accidents (The Worldbank, 2009). Identifying accident prone locations can help taking necessary precautions and preventing them and resulting losses including fatalities.

Geographical Information Systems (GIS), provides a noticeable impact on identifying and visualizing traffic accident hot spots. In particular, GIS-based spatial hot spot analysis is one of the fundamental analyses used to locate traffic accident hot spots. Traffic safety literature shows that analysing the locations by considering the hot spots with spatial clustering methods plays a very active role in examining the tendency of traffic accidents; this enables detection of common features of hot spot locations (road type, function, traffic conditions, etc.). There are many GIS-based hot spot analysis techniques in the literature: some of these methods include density analysis, interpolation analysis, pattern analysis, mapping clusters analysis, and density-based analysis (Levine et al. 1995a, 1995b; Prasannakumar et al., 2011; Harirforoush and Bellalite 2016; Colak et al., 2018; Choudhary et al., 2015; Erdogan et al., 2008; Erdogan, 2019). Among these, Nearest Neighbourhood Hierarchical (NNH) and (DBSCAN) are two commonly used methods for traffic accident clustering detection. NNH algorithm, coded in the free CrimeStat software, performs point-based clusters directly and presents convex hulls of cluster zones (Levine, 2013; Levine, 2014; Kundakci, 2014; Kundakci and Tuydes-Yaman, 2014; Türe-Kibar and Tuydes-Yaman, 2020; Keskin et al., 2011; Drawve, 2016). DBSCAN method also focused on grouping geocoded traffic accidents (Mohammed and Baiee, 2020; Agrawal et al., 2018; Zhang et al., 2018; Wang et al., 2019; Szénási et al., 2018; Szénási et al., 2020; Islam et al., 2021), with slightly different processing of the point data and produces list of nodes with appointed cluster numbers or noise ones, which are not clustered.

Speed is one of the five main factors (i.e. seatbelt use, drinking and driving, etc.) contributing heavily to the losses in the traffic accidents. However, this factor has a two-sided effect mechanism: while it can be a reason causing accident, it can also increase the severity of the accidents. In Turkey, speed related run-off road (ROR) accidents are the third largest type, which requires further attention to understand and prevent. When speed related accident hot spots were researched, it is also necessary to include other accident types that may have association with speeding such as hitting fixed objects, etc.

This study aimed to determine the risky areas where speed-related traffic accidents. Due to the popular use and easy integration with GIS programs, NNH and DBSCAN are selected as the two methods to be employed in this study to detect speed related traffic accident hot spots. NNH is also currently employed by the General Directorate of Highways (GDH) in Turkey. The results of the two methods were compared to detect speed-related problematic road locations/elements. Evaluations were made at the regional scale,

* Corresponding author

then the determined clusters depending on the road geometry, were evaluated at the local scale. Moreover, it has been tried to make future predictions by giving examples from the points where the accidents are repeatedly clustered depending on the years.

The traffic accident data is obtained from the GDH for the Mersin Region (including 7 cities) for the period of 2014-2021. The analysis and visualization stages in this study were carried out through the QGIS program, which is one of the open-source GIS software.

MATERIALS AND METHODS

2.1. Study area

In this study, the Mersin Region of the GDH was chosen as the study area for the analyses. The region extends to an area of 61.683 km^2 and includes the provinces of Mersin, Adana, Osmaniye, Hatay, Kahramanmaras, Gaziantep and Kilis; the total population of the region is 9.8 million residents and there are 2.977 million vehicles registered (URL-1). More importantly Mersin has an international port, acting as a major truck trip attraction/production point, connected by motorways extending to the eastern part of the region. The study area and the average traffic volumes on the state highways/motorways in the region was shown in Figure 1 (URL-2, 2022).

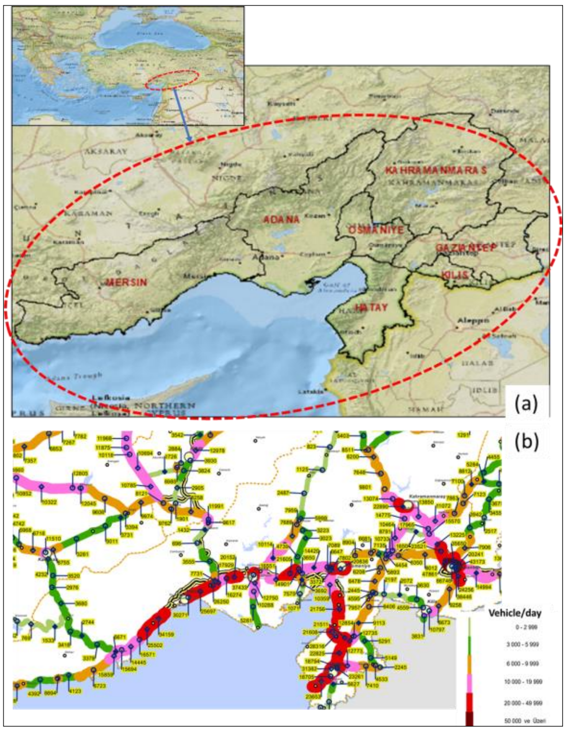

Figure 1. Mersin Region (a) cities and (b) traffic volumes on the state highways in 2019

2.2. Study Framework

This study was planned in five phases:
i) Spatial and non-spatial traffic accident data for the pilot region were obtained from the GDH.
ii) Data were organized and then transferred to the geographic database for GIS-based analyses.
iii) The hot spots were detected using the NNH and the DBSCAN methods: The CrimeStat III program was used for the former, and the open-source GIS program QGIS was used for the latter. All results were analysed, visualized, and evaluated through the QGIS program.
iv) Obtained results are evaluated according to years. With the analysis results, the differences in the clustering of traffic accidents between the two selected methods. Road geometry also examined these evaluations, and different cluster areas were examined more locally.
v) Future predictions have been made for points where accidents have repeatedly clustered over years.

2.3. Data and the GIS database

The GDH provided the traffic accident data for the study area as a part of a going funded project for 2014-2021 as obtained from the General Directorate of Security (GDS). The data in excel format has been arranged and associated with spatial data to be used in the analysis. All data were defined WGS_1984_Web_Mercator_Auxiliary_Sphere as the projection system, integrated into the geographic database, and prepared for analysis. Since speed-related traffic accidents and hot spot analyses were performed in the study, the database was suited to include speed-related accidents. Among 15 entries in the "accidentType" attribute, four of them were joined to create the subset describing speed-related accidents: i) Rollover/drift/overturned, ii) run-off-road (ROR), iii) crashing into a stationary vehicle, and iv) collision with obstacle/object. According to the accident type year by year, several samples are listed in Table 1.

Year	*Total # of accidents*	*Total number of speed-related accidents*			
		Rollover/ drift/ overturned	Run-off-road	Crashing into a stationary vehicle	Collision with obstacle/ object
2014	4554	1702	1335	479	1038
2015	5302	2041	1470	462	1329
2016	4996	1825	1338	482	1351
2017	4691	1737	1381	428	1145
2018	4829	1861	1437	434	1097
2019	4659	1787	1500	337	1035
*2020**	4307	1896	1284	226	861
*2021**	4967	2226	1547	228	906

* The years with mobility restrictions due to pandemic conditions

Table 1. Speed-related accident totals in Mersin Province

Due to the restrictions in the intercity mobility during the pandemic years of 2020-2021, there were changes in the pattern and level of traffic accident numbers. Thus, there were evaluated separately for two years. To be compatible, the previous years were also grouped into 2-year analysis periods of 2014-2015, 2016-2017, and 2018-2019. Note: the data included the accidents from the police regions, while a small part of the accidents falling into the rural parts within the jurisdiction of the gendarmery was omitted due to a lack of data format compatibility problems between the two datasets.

2.4. NNH Clustering Method

NNH Clustering is a hot spot spatial clustering method that detects accident hot spots. This method considers two types of criteria for spatial mapping clustering of spatial point data: the threshold distance (**d**), which is the Euclidean distance between

https://doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-487-2022 | © Author(s) 2022. CC BY 4.0 License.

each pair of data points, and used as a search radius value in the algorithm. The second parameter is the minimum number of points that must be present in a cluster (**n_{min}**). If a threshold distance is chosen for operation, pairs of points with smaller distances are clustered together. Although both the threshold distance and the number of points in a cluster are specified, clusters are created depending on the given parameters. In order to create a new cluster, the number of defined observations must be closer than the threshold value. Second and higher-order clusters are then created in the same way until only one cluster remains or the grouping criteria fail (Kundakci E, 2014; Kundakci and Tuydes-Yaman, 2014; Levine, 1996; Ture Kibar and Tuydes-Yaman, 2020).

At the point of realizing this method, the CrimeStat III program, which was developed especially for hot spot clustering analysis of crimes, is widely used. The CrimeStat III is a crime mapping software program developed by Ned Levine, was used in this study to obtain NNH clustering results (Levine, 1996). This program is a free downloadable Windows-based crime mapping software program that performs spatial and statistical analysis and is designed to interface with GIS. Apart from crime mapping, it also enables hot spot analyses in different areas. Detection of traffic accident hot spots is one of these studies. Analysis outputs can visualize in two different formats: Convex hull and ellipse (Figure 2). Convex hulls are polygons that fully cover all clustered points. These are pretty sensitive and can identify the actual area where the hot spot happens. Ellipse is more like a symbolic representation of the cluster (Levine, 1996). As an option, there is a fixed search distance radius in the menus.

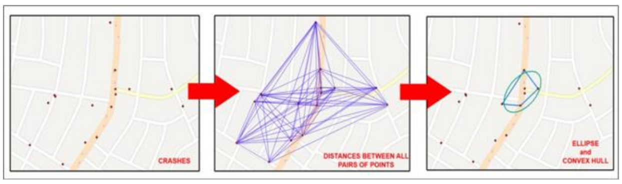

Figure 2. NNH clustering algorithm (Kundakci and Tuydes-Yaman, 2014)

2.5. DBSCAN Clustering Method

DBSCAN Clustering is a method for finding specific predefined events and hot spots. This method is a data clustering algorithm proposed by Martin Ester, Hans-Peter Kriegel, Jörg Sander and Xiaowei Xu in 1996 (Ester et al., 1996). The algorithm, moreover, is open source and recommended for noisy data in large spatial databases (Ester et al., 1996). This method identifies a cluster as the most densely connected set of points possible. There are two main criteria addressed in this method: Epsilon (**Eps**) and minimum points (**n_{min}**). The maximal radius of the neighbourhood is epsilon, and the minimal number of points in the epsilon-neighbourhood to describe a cluster is minimum points. This clustering algorithm separates the point data into three different forms (Schubert et al., 2017).

The DBSCAN algorithm takes two parameters as input and checks all objects, starting with any object in the database. If the controlled object has been included in a cluster before, it passes to the other object without processing. If the object is not clustered before, a region query by doing this, it finds the **Eps** neighbors of the object. If the number of neighbors is more than **n_{min}**, it names this object and its neighbors as a new cluster. Then, new neighbors are created by querying the new region for each previously unclustered neighbor finds. If the number of neighbors of the region query is more than **n_{min}**, it is included in the cluster (URL-2, 2022; URL-3, 2022).

2.6. Comparison of NNH and DBSCAN

The two algorithms used in this study are fundamentally different from each other. NNH is an example of agglomerative clustering and uses average linkage when creating higher order clusters. However, after creating the first order sets, the algorithm eliminates the clusters that do not meet the **n_{min}** criterion and this procedure works iteratively. Hence, NNH does not include all points in the data set into clusters. NNH differs from traditional clustering algorithms with this feature. On the other hand, DBSCAN proceeds iteratively on a point-based basis over the given criteria, and firstly determines the data points to be clustered and marked as noise. This makes DBSCAN an example of density-based clustering. In this respect, it can be mentioned that DBSCAN has a chaining feature and thus the clusters it creates are random in shape, while cluster outputs of NNH tend to be convex or ellipsoidal in shape (Levine, 1996; Ester et al., 1996).

Within the scope of this study, the minimum data point requirements (**n_{min}**) of both methods will be kept the same; for the sake of comparisons, the distance threshold values, **d_{max}** and **Eps**, will be kept the same even though they have fundamentally different scopes.

3. RESULTS AND DISCUSSIONS

3.1. Speed-Related Traffic Accident Analysis Using GIS-Based NNH and DBSCAN Clustering Methods

The traffic safety literature generally takes a 3-year data period while searching for accident hot spots, as they are generally rare events in terms of temporal distribution. A rule of thumb suggests searching for minimum of 5 accidents/3 years as a practical limit among traffic safety analysts. In Turkey, the GDH uses a **d_{max}**= 100 meter and **n_{min}**= 5 accidents to form a cluster within a one-year period, when using NNH. By using these parameters, separate analyses are made for each year in the accident data of the last three years. In the last three years, places that overlap spatially have been accepted as hazardous (also called blackspot) locations.

In this study, a two-year period (2014-2015, 2016-2017, 2018-2019, and 2020-2021) were used as the analysis period. Total number of accidents in Table 2 dropped from 9856 (2014-2015) to 9174 accidents during 2020-2021 which may be affected by the pandemic related mobility restrictions; thus, a normalized analysis based on the vehicle-km-travelled (VKT) values of the region should be performed before concluding anything regarding the clustering and accident patterns.

In the NNH analysis, varying distance threshold values of **d_{max} = {**50, 100 and 200 meters} and a minimum number of 5, 10 and 15 accidents (**n_{min}**) were used. While locating traffic clusters using the DBSCAN method. QGIS software was used for this method of analysis and visualizing maps. The **n_{min}** = {5, 10 and 15 accidents} and **Eps = {**50, 100 and 200 meters} were used. While similar results were obtained during the 2014-2015 and 2016-2017 periods, smaller number of clusters was obtained for the remaining two periods. While the pandemic period may be due to lower traffic volumes and mobility restrictions, it is necessary to further evaluate the drop in 2018-2019.

n_{min}		2014-2015			2016-2017			2018-2019			2020-2021		
	# Speed-related accidents	9856			9687			9488			9174		
	d_{max} /Eps (m)	50	100	200	50	100	200	50	100	200	50	100	200
5	NNH	19	60	174	14	52	163	16	50	142	13	40	142
	DBSCAN	34	114	275	33	95	240	32	85	216	20	79	237
10	NNH	3	**6**	17	0	**7**	18	3	**6**	18	1	**8**	12
	DBSCAN	3	9	28	1	8	24	4	9	30	3	8	19
15	NNH	0	0	3	0	1	4	1	1	3	0	1	3
	DBSCAN	0	1	4	0	1	4	1	1	4	0	1	3

Table 2. Clustering results using NNH and DBSCAN with various n_{min} *(5/10/15 accidents) and* ***d*** *threshold (50m, 100m, 200m)*

As expected, keeping a small $\mathbf{n_{min}}$ = 5 accidents/ 2 years value detected many hot spots by both methods. In 2014-2015 period, there were 60 clusters detected by NNH, when $\mathbf{d_{max}}$= 100 m was assumed, which was raised to 174 clusters, when distance threshold was doubled. For the same period, DBSCAN method found 1.6-1.9 times more clusters. When $\mathbf{n_{min}}$ = 10 accidents/ 2 years was assumed; cluster numbers were reduced significantly, as 10 accidents were also required to be closed to each other, producing only 6 clusters by NNH and 9 clusters by the DBSCAN with $\mathbf{d_{max}}$= 100 m. However, searching for $\mathbf{n_{min}}$ = 15 accidents/ 2 years with a distance threshold value of even 200 m did not produce many clusters.

3.3. Comparison of the NNH and DBSCAN Clustering Methods

The visual depiction of the maps of the years 2014-2015, 2016-2017, 2018-2019, and 2020-2021 were produced and visualized (Figure 3). The black points represent all speed related data points, where numbers and letters represent the locations of the clusters found by DBSCAN and NNH, respectively. It can be seen that majority of the accidents were not clustered as they were spread out in a very large geography.

When looking at the maps, the clusters shown with numbers represent the DBSCAN results, and the clusters shown with the letters represent the NNH results and it is observed that these clusters spread to the provinces in the selected region. At some points, the cluster outputs of the two methods overlap; at some points, only the DBSCAN method detects clusters, and the NNH method cannot detect clusters at the specified point. These results prove that the DBSCAN method is more advantageous than the NNH method in the clustering of accidents. These produced maps are sufficient for a holistic evaluation of the region. However, it is not significant based on road geometry. For this reason, the clusters identified in the study were analyzed and mapped on a more regional basis, and the cluster outputs obtained by the two methods were tried to be revealed more clearly.

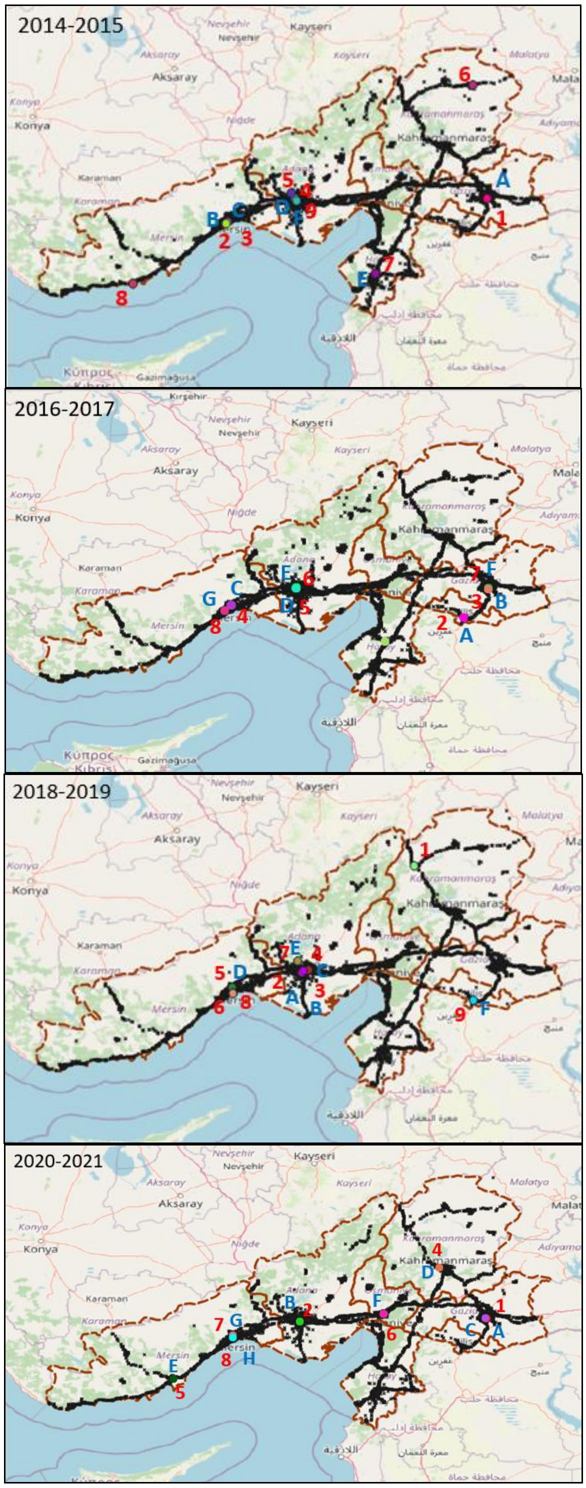

Figure 3. Mersin Region Speed-related Traffic Accident Maps for the study periods of (2014-2015), (2016-2017), (2018-2019), (2020-2021)

<u>**For 2014-2015 period:**</u> the number of clusters in which both methods show joint clustering was determined as 6. These clusters from the provinces within the specified region are 1 in Hatay, 1 in Gaziantep, 2 in Adana, and 2 in Mersin. When evaluated based on roads, it is seen that the regions where the clusters are located on five intersections and one corridor. When the differences in the methods are examined, it is observed that there are three different regions that the DBSCAN method

detects, and the NNH method does not. Therefore, it has been determined that the results obtained, the DBSCAN method, give more effective results in the clustering of accidents. In addition, the results obtained showed that the clusters detected by the DBSCAN method and not detected by the NNH method are 1 cluster in Adana and this cluster is in the corridor region, 1 cluster in Kahramanmaras and this cluster is in the intersection area, and finally, 1 cluster in Mersin and this cluster is in the corridor region.

For 2016-2017 period: the number of clusters in which both methods show joint clustering was determined as 7. These clusters from the provinces within the specified region are 1 in Kilis, 2 in Gaziantep, 2 in Adana, and 2 in Mersin. When evaluated based on roads, it is seen that the regions where the clusters are located on seven intersections. When the differences in the methods are examined, it is observed that there is one different region that the DBSCAN method detects, and the NNH method does not. In addition, the results obtained showed that the clusters detected by the DBSCAN method and not detected by the NNH method are 1 cluster in Hatay, and this cluster is in the corridor region.

For 2018-2019 period: the number of clusters in which both methods show joint clustering was determined as 6. These clusters from the provinces within the specified region are 1 in Kilis, 4 in Adana, 1 in Mersin. When evaluated based on roads, it is seen that the regions where the clusters are located on four intersections and two corridors. When the differences in the methods are examined, it is observed that there are three different regions that the DBSCAN method detects, and the NNH method does not. In addition, the results obtained showed that the clusters detected by the DBSCAN method and not detected by the NNH method are 1 cluster in Kahramanmaras and this cluster is in the intersection region, 1 cluster in Mersin and this cluster is in the intersection region and 1 cluster in Mersin and this cluster is in the intersection regions.

For 2020-2021 period: the number of clusters in which both methods show joint clustering was determined as 8. These clusters from the provinces within the specified region are 1 in Kilis, 3 in Mersin, 2 in Gaziantep, 1 in Kahramanmaras, and 1 in Osmaniye. When evaluated based on roads, it is seen that the regions where the clusters are located on six intersections and two corridors. When the differences in the methods are examined, it is observed that there is a no different region that the DBSCAN method detects, and the NNH method does not. When the all results of both methods are examined, it can be determined that the DBSCAN method catches a cluster with a noticeable difference compared to the NNH method.

In order to reveal the differences between the evaluations and methods for the four periods more clearly, the results were examined on a road scale at this stage. Evaluations were made by overlapping the outputs of both methods. The maps were shown in Figure 4, Figure 5, Figure 6, and Figure 7, respectively.

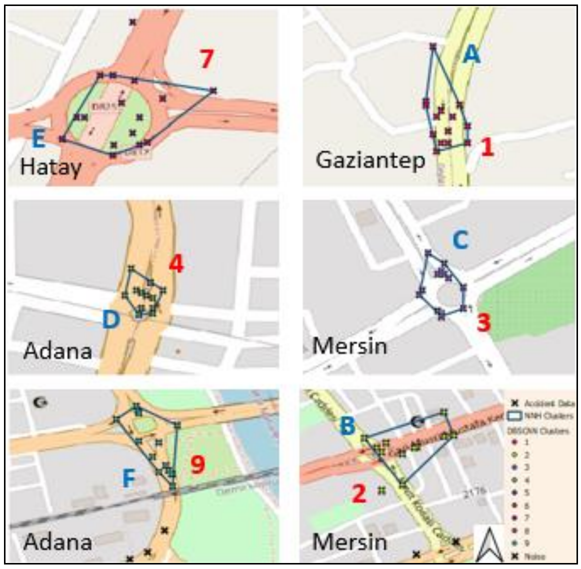

Figure 4. Mersin Region Speed-related Accident Clusters Map of 2014-2015 (NNH and DBSCAN are overlaid)

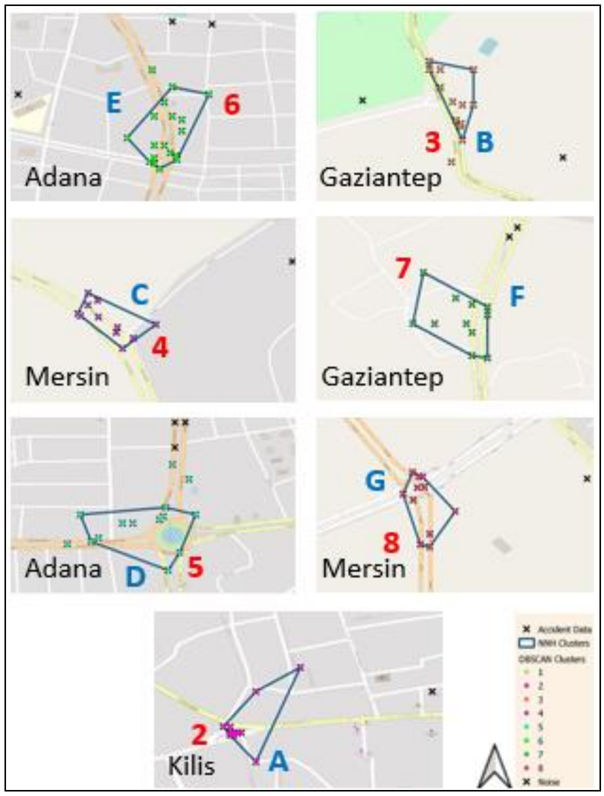

Figure 5. Mersin Region Speed-related Accident Clusters Map of 2016-2017 (NNH and DBSCAN are overlaid)

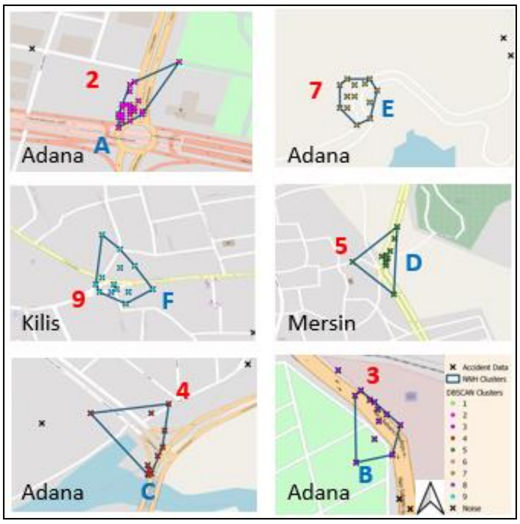

Figure 6. Mersin Region Speed-related Accident Clusters Map of 2018-2019 (NNH and DBSCAN are overlaid)

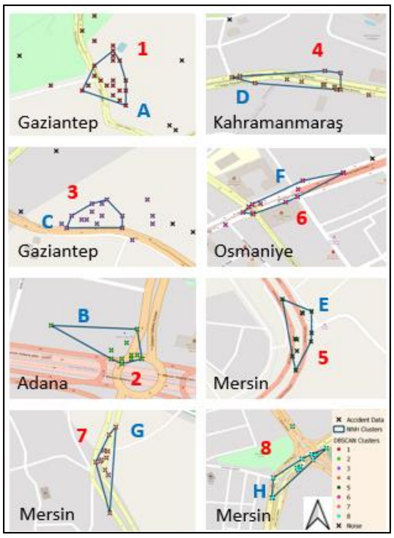

Figure 7. Mersin Region Speed-related Accident Clusters Map of 2020-2021 (NNH and DBSCAN are overlaid)

In the last stage of the study, the detection of clusters found in the same location in different 2-year periods using the 100 meters and 10 accidents parameters was examined. At this point, maps were produced only for the detected regions and considerations were made on the examples. As in Figures 3-7, blue letters represent NNH and red numbers represent clusters detected with DBSCAN (Figure 8, Figure 9, Figure 10, Figure 11).

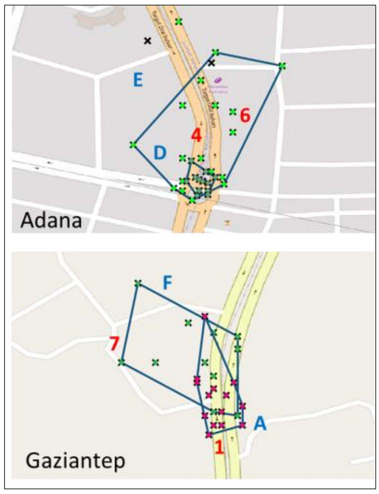

Figure 8. Mersin Region Repeatedly-detected Clusters in Different Periods (2014-2015, 2016-2017)

The clusters labelled as D-4 and 1-A were produced from the accidents of the period 2014-2015, while the remaining from 2016-2017. A total of two regions, one each, were detected in Adana and Gaziantep for the successive time periods 2014-2015 and 2016-2017, whereas no cluster was detected in the same region for the next four years. This may imply that taken countermeasures in these regions were effective and result in a decrease in the density of speed-related traffic accidents.

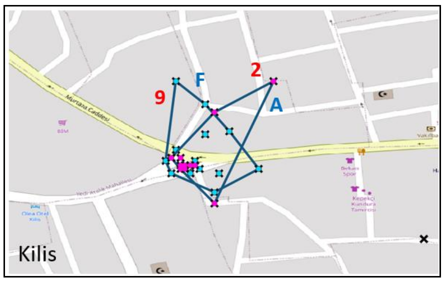

Figure 9. Mersin Region Repeatedly-detected Clusters in Different Periods (2016-2017, 2018-2019)

Another region having clusters in successive time periods is detected in Kilis for the periods of 2016-2017 and 2018-2019. Similarly, it was observed that the traffic accidents that occurred in the following period did not form a cluster. This can be explained by the effectiveness of the precautions taken, or this might be interpreted as the alteration in transportation habits due to pandemic conditions.

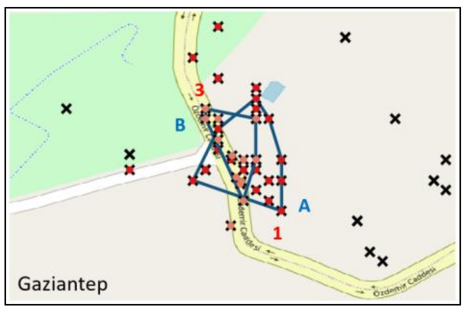

Figure 10. Mersin Region Repeatedly-detected Clusters in Different Periods (2016-2017, 2020-2021)

https://doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-487-2022 | © Author(s) 2022. CC BY 4.0 License.

The region detected in Gaziantep for the periods of 2016-2017 and 2020-2021 is different than beforementioned regions in terms of their successiveness. Although there did not appear any clusters in 2018-2019 after the detected one in 2016-2017, the accidents in 2020-2021 form a cluster again. This can reveal that the countermeasures were inadequate to prevent speed-related accidents, and may be revised to reduce the number of crashes for upcoming years.

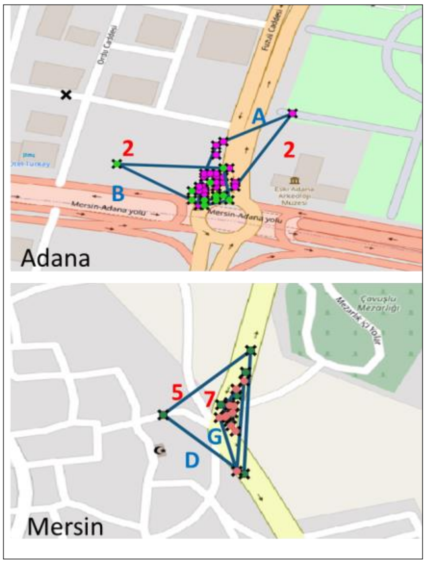

Figure 11. Mersin Region Repeatedly-detected Clusters in Different Periods (2018-2019, 2020-2021)

For the periods 2018-2019 and 2020-2021, two regions having overlapping clusters were detected and both are in the intersections. In Adana example, it is seen that the accidents were clustered at the same part of the rotary intersection. Also, in Mersin example, cluster area gets smaller with the same number of data points as the time progresses. Therefore, these clusters point out specific spots in both examples, and it can be predicted that these intersections will remain as areas with high accident density in coming years unless necessary conditions are met.

4. CONCLUSION

Traffic safety is a major problem, especially in developing countries facing rapid motorization. There are different factors causing and increasing the severity of traffic accidents; and speed is a major factor affecting both. In Turkey, speed related accidents rank third, which shows the urgent need to understand their nature and mechanisms. To reduce speed-related traffic accidents, it is necessary to locate hazardous parts of the road network prone to them, which are also called hot spots. It is generally addressed using hot spot clustering analysis methods, but, different parameters of available methods yield to varying locations.

In this study, the speed related traffic accident hot spots were detected using two different methods: NNH and DBSCAN. While both works with geocoded traffic accident point locations, the clustering techniques employed results in different hot spot location and sizes, which are compared to evaluate the persistent ones in a region. The study area was selected as the Mersin Region in Turkey, spanning over 7 cities in the southern part of Turkey. The data from 2014-2021 were used as sliced in four 2-year periods.

The study was first evaluated based on the selected region in terms of both methods, and then it was handled in a more local region based on the road geometry. The favourable aspects of the methods used in determining traffic accident clusters compared to each other were evaluated. The results obtained from the study show that the DBSCAN method achieves visibly effective clusters compared to the NNH method. As a result of the analyses, it was seen that both methods detected traffic accident clusters in common, but the DBSCAN method captured other clusters that the NNH method could not detect. Therefore, it is thought that GDH commonly uses the NNH method in clustering traffic accidents, but it would be more meaningful to consider the DBSCAN method based on the obtained results. Thus, it is predicted that GDH will take more effective planning and precautions in traffic accident cluster regions that the NNH method cannot detect.

In this study, the points where the accidents are repeatedly clustered depending on the years were examined and results were evaluated on selected sample maps. It is envisaged that these outputs can be used effectively in future predictions. In particular, a more detailed examination of the area can contribute to the studies to reduce traffic accidents.

As a result, it is thought that this study will contribute to the following objectives: (1) this study will be a base for studies of the same type, (2) it will help in the selection of the method to be made in the determination of the points where traffic accidents are clustered on a regional and local scale, and (3) it will guide in the determination of the places where the accidents occur repeatedly depending on the years and will be a basis for taking the necessary measures.

ACKNOWLEDGEMENTS

The study was carried out within the scope of TUBITAK Industry Projects Support Program called the Project of Design and Development of the Spatial and Multidimensional Assessment and Risk Rating Analytical Platform (KoDeRDAP) for Traffic Safety. The project continues under the management of Parabol company. We thank them for providing the spatial data necessary for the study.

REFERENCES

Agrawal, K., Ruth, V.M., Nandini, Y., Sravani, K., 2018. Analysis of road accident locations using DBSCAN algorithm. *IJSRST*, 4(8), 462–467.

Choudhary, J., Ohri, A., & Kumar, B., 2015, December. Spatial and statistical analysis of road accidents hot spots using GIS. In 3rd Conference of Transportation Research Group of India (3rd CTRG).

Colak, H. E., Memisoglu, T., Erbas, Y. S., & Bediroglu, S., 2018. Hot spot analysis based on network spatial weights to determine spatial statistics of traffic accidents in Rize, Turkey. *Arabian Journal of Geosciences*, 11(7). https://doi.org/10.1007/s12517-018-3492-8.

Drawve, G., 2016. A Metric Comparison of Predictive Hot Spot Techniques and RTM. *Justice Quarterly*, 33(3). https://doi.org/10.1080/07418825.2014.904393.

Erdogan, S., 2009. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *Journal of Safety Research*, 40(5).

https://doi.org/10.1016/j.jsr.2009.07.006.

Erdogan, S., Yilmaz, I., Baybura, T., & Gullu, M., 2008. Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar. *Accident Analysis and Prevention*, 40(1). https://doi.org/10.1016/j.aap.2007.05.004

Ester, M., Kriegel, H. P.., Sander, J., & Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *In Proceedings of the 2nd ACM International Conference on Knowledge Discovery and Data Mining (KDD)*.

Harirforoush, H., 2019. Spatial and Temporal Analysis of Seasonal Traffic Accidents. *American Journal of Traffic and Transportation Engineering*, 4(1). https://doi.org/10.11648/j.ajtte.20190401.12.

Islam, M. R., Jenny, I. J., Nayon, M., Islam, M. R., Amiruzzaman, M., & Abdullah-Al-Wadud, M., 2021, August. Clustering algorithms to analyze the road traffic crashes. In 2021 International Conference on Science & Contemporary Technologies (ICSCT) (pp. 1-6). IEEE.

Keskin F, Yenilmez F, Colak M, Yavuzer I & Düzgün HS, 2011. Analysis of traffic incidents in METU campus. *Proc Soc Behav Sci*, 19:61–70.

Kundakci E., 2014. Identification of urban traffic accident hot spots and their characteristics by using GIS, The Graduate School of Natural and Applied Sciences of Middle East Technical University, Master Thesis, Ankara, Turkey.

Kundakci, E. and Tuydes-Yaman, H., 2014. Understanding the Distribution of Traffic Accident Hotspots in Ankara, Turkey. TRB 2014 Annual Meeting.

Levine, N., 1996. Spatial Statistics and GIS: Software Tools to Quantify Spatial Patterns, *Journal of the American Planning Association*, 62(3). ttps://doi.org/10.1080/01944369608975702

Levine, N., Kim, K. E., & Nitz, L. H., 1995a. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. Accident Analysis & Prevention, 27(5), 663-674.

Levine, N., Kim, K. E., & Nitz, L. H., 1995b. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. *Accident Analysis & Prevention*, *27*(5), 675-685.

Levine, N. 2013. Hot spot analysis of zones. CrimeStat: Spatial Statistics Program for the Analysis of Crime Incident Locations, Version 4: 242960–242995.

Levine N, 2014. CrimeStat IV: Part I: Program Overview. Ned Levine &Associates, Houston, TX, and the National Institute of Justice, Washington DC.

Mohammed, A. F. and Baiee, W. R., 2020. The GIS based Criminal Hotspot Analysis using DBSCAN Technique, 2nd International Scientific Conference of Al-Ayen University (ISCAU-2020), IOP Conf. Series: Materials Science and Engineering 928 (2020) 032081, doi:10.1088/1757-899X/928/3/032081.

Prasannakumar, V., Vijith, H., Charutha, R., & Geetha, N., 2011. Spatio-temporal clustering of road accidents: GIS based analysis and assessment. *Procedia-social and behavioral sciences*, 21, 317-325.

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., Xu, X., 2017. DBSCAN Revisited, revisited. *ACM Transactions on Database Systems*, 42(3). https://doi.org/10.1145/3068335

Szénási, S., Felde, I., Kertész, G., & Nádai, L., 2018, November. Comparison of Road Accident Black Spot Searching Methods. In 2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI) (pp. 000247-000250). IEEE.

Szénási, S., Felde, I., & Kertész, G., 2020, November. Comparison of Accident Black Spot Identification Methods based on GPS Coordinates. In 2020 IEEE 20th International Symposium on Computational Intelligence and Informatics (CINTI) (pp. 000061-000064). IEEE.

The World Bank, 2009. http://web.worldbank.org/archive/website01213/WEB/0__CO-72.HTM#Why

Ture Kibar, F. and Tuydes-Yaman, H., 2020. GIS-based evaluation of the speed limit increase on urban arterial traffic safety in Ankara. *Arabian Journal of Geosciences*, 13(12). https://doi.org/10.1007/s12517-020-05486-5

URL-1, 2022. https://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Bolgeler/5Bolge/Bolge5.aspx

https://www.kgm.gov.tr/Sayfalar/KGM/SiteTr/Bolgeler/5Bolge/Harita.aspx

URL-2, 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-statistics/how-density-based-clustering-works.htm.

URL-3, 2022. https://en.wikipedia.org/wiki/DBSCAN.

Wang, T., Ren, C., Luo, Y., & Tian, J., 2019. NS-DBSCAN: A density-based clustering algorithm in network space. *ISPRS International Journal of Geo-Information*, 8(5), 218.

WHO, 2021. Road traffic injuries, https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Zhang, Y., Han, L. D., & Kim, H., 2018. Dijkstra's-DBSCAN: Fast, Accurate, and Routable Density Based Clustering of Traffic Incidents on Large Road Network. *Transportation Research Record*, 2672(45), 265-273.