Sigma J Eng & Nat Sci 9 (4), 2018, 383-394

Publications Prepared for Transist 2017,
10th International İstanbul Transport Congress

Research Article

IDENTIFICATION OF ACCIDENT BLACK SPOTS USING NETWORK SCREENING: THE CASE OF SOGUTLUCESME-15 TEMMUZ SEHITLER BRIDGE CORRIDOR

Abdullah MALTAŞ*[1], Halit ÖZEN[2], Abdulsamet SARAÇOĞLU[3]

[1] *Yildiz Technical University, Dept. of Civil Engineering, Esenler-İSTANBUL;* ORCID:0000-0002-2595-8536
[2] *Yildiz Technical University, Dept. of Civil Engineering, Esenler-İSTANBUL;* ORCID:0000-0003-4031-7283
[3] *Yildiz Technical University, Dept. of Civil Engineering, Esenler-İSTANBUL;* ORCID:0000-0002-6695-4865

Received: 01.10.2018 Revised: 19.11.2018 Accepted: 21.12.2018

ABSTRACT

In order to reduce the number and the effects of traffic accidents on a roadway, various countermeasures are taken into consideration. The first step to decide the proper countermeasures is identifying Accident Black Spots (ABS) and then improving the site regarding the different type of the countermeasures to reduce the effect of the traffic accidents. As well as many methods used to identify the ABS in literature, network screening technics, that simple ranking, sliding window and peak searching, are defined in Highway Safety Manual published by AASHTO in 2010. In these technics, there are various performance measures like average crash frequency, equivalent property damage only, etc. to rank the roadway segment. Based on the ranking, ABS are identified and prioritize to decide and implement the countermeasures.
In this study, data for fatal-injured traffic accidents that occurred at Sogutlucesme-15 Temmuz Sehitler Bridge corridor in Istanbul between 2011-2013 are provided by Istanbul Directorate of Security and the data was transferred into geographical information systems (GIS) with that way the data was related with the geographical location. The corridor was split into 10 m long segments in GIS. Three different types of performance measures are considered to rank the segments based on the above-mentioned network screening technics. K-means clustering method was used to identify the ABS in this study. As a result of the study, the K-means clustering method is accomplished to identify the ABS and sliding window technic is the most appropriate methods to identify the ABS.
Keywords: Network screening, GIS, accident black spots, K-means clustering.

1. INTRODUCTION

Traffic, an essential part of daily life, is an important issue in transportation engineering. The planning, operation, and management of transportation are directly related to traffic, and managing the traffic properly is important to decrease the congestion, fuel consumption, environmental effects, etc. One of the key issues of transportation engineering topics is traffic accidents.

The World Health Organization (WHO) emphasizes the loss of life and property caused by traffic accidents. Death numbers caused by accidents are in the 9th place among all death forms

* Corresponding Author: e-mail: maltas@yildiz.edu.tr, tel: (212) 383 51 85

(in the 1st place among deaths of young people, i.e. 15-29 years of age) nowadays and WHO predicts that these death numbers will rise to the 7th place by 2030 [1]. So, various countermeasures are taken to reduce traffic accidents. One of these countermeasures is to detect where the accidents are experienced frequently and which are named accident black spots (ABS). With the identification of black spots, the accidents occurring at these sites can be examined in more detail and improvements can be carried out.

Various methods are taken part in literature. The classical approach to identify black spots is the marking of all accident points on the map and the identification of the site that included most of the marks [2]. Geographical Information Systems (GIS) that developed in recent years are used frequently in accident analysis by mapping the coordinates of the accidents [3, 4, 5, 6].

In literature, studies on the analysis of accidents include various statistical and numerical methods, besides data mining, temporal or spatial analysis and so. In studies, traffic accidents are analyzed by using Kernel Density Estimation [5, 6, 7], Moran's I and Getis-ord GI methods [6], clustering methods including K-means [5, 7], Bayesian Networks and Empirical Bayes with performance measures [8, 9, 10].

In addition to the many methods used in the identification of ABS, the network screening technics (simple ranking, sliding window, peak searching) in the Highway Safety Manual (HSM) published by AASHTO in 2010 are also examined for alternative studies. By using these technics, it is possible to identify an accident black spot according to 13 different performance measures and datasets [11]. Also, network screening technics can be used with different methods and performance measures or can be used as the first step of accident analyses [12, 13].

Despite all precautions taken, it is a fact that accidents cannot be precisely prevented as long as the human factor is in traffic. However, it is aimed to reduce the number of accidents as much as possible and to take countermeasures for this. In this study, we purposed to detect ABS by using GIS, network screening technics identified in HSM and K-means clustering method.

2. ROADWAY SAFETY MANAGEMENT PROCESS

According to HSM, there are six steps in the road safety management process. These can be performed as a whole process or singularly as seen in

Figure 1. The six steps of the roadway safety management process are;

- Network Screening: Reviewing a roadway network to identify and rank sites on the basis of potential to reduce the crash frequency.
- Diagnosis: Evaluating accident and historical site data to identify the model of accidents.
- Select Countermeasures: Identifying the factors that may cause accidents at a site, and identifying the necessary countermeasures to reduce the average crash frequency.
- Economic Appraisal: Identifying the projects that are economically feasible by evaluating the benefits and costs of the countermeasures to be taken.
- Prioritize Projects: Evaluating economically the countermeasures that can be taken in the relevant sites in order to identify the most appropriate improvement projects according to the criteria such as cost, mobility, and environmental impact.
- Safety Effectiveness Evaluation: Evaluating the efficiency of a countermeasure at the site in reducing the crash frequency or severity [11].

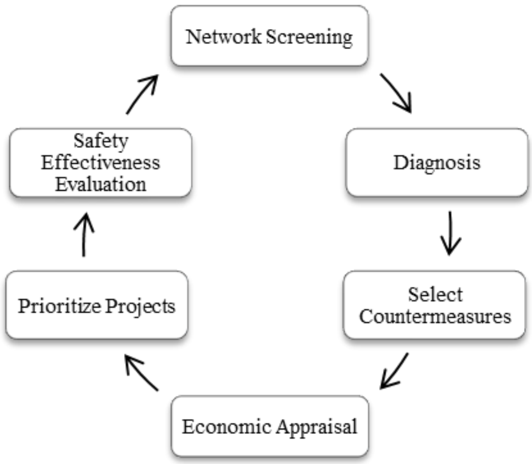

Figure 1. Roadway safety management process [11]

2.1. Network Screening

Network screening is a process to review a roadway network to identify and rank sites according to the probability of a reduction in the crash frequency with the implementation of a countermeasure. There are five main steps in the network screening process. These steps are; establishing focus, identifying the network and establishing reference populations, selecting performance measures, selecting a screening method and finally screening and evaluating results as shown in Figure 2.

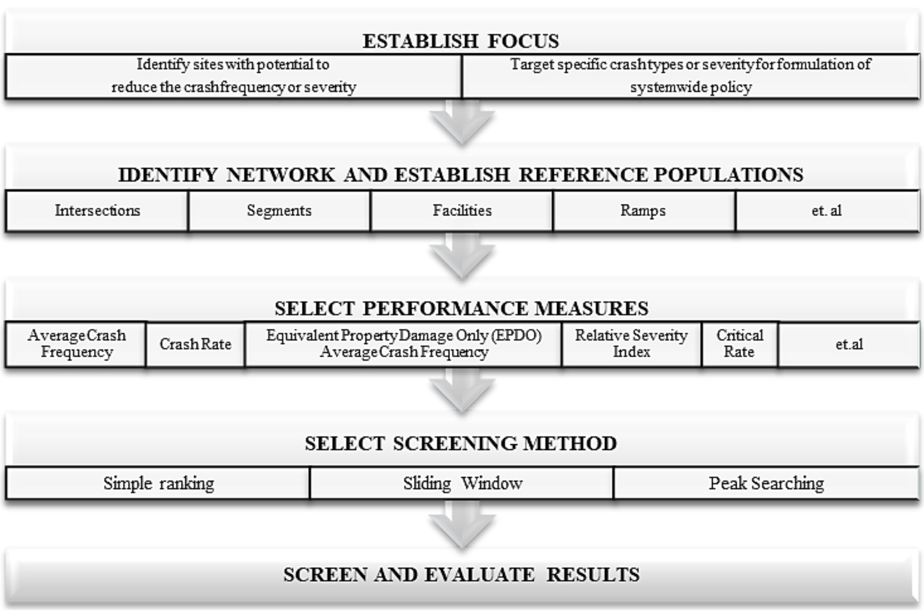

Figure 2. The network screening process [11]

2.2. Performance Measures

In HSM, there are 13 different performance measures which presented in Table 1 identified. Data availability, regression-to-the-mean bias, and performance threshold, as well as applying a systemwide policy, are directly related to the selection of the performance measures. One or more of them may be used according to targets and available datasets. Performance measures refer to numerical values identified to reduce the crash frequency or severity (property damage only, injured, fatal) considering data in a site. In this study, average crash frequency (ACF), equivalent property damage only average crash frequency (EPDO-ACF) and relative severity index (RSI) performance measures are used in the light of available data.

Table 1. Data needs for performance measures [11]

	Data and Inputs				
Performance Measure	**Crash Data**	**Roadway Information**	**Traffic Volume**	**Calibrated SPFs and Overdispersion Parameters**	**Other**
ACF	X	X			
Crash Rate	X	X	X		
EPDO-ACF	X	X			EPDO Weighting Factors
Relative Severity Index	X	X			Relative Severity Indices
Critical Rate	X	X	X		
Excess Predicted ACF Using Method of Moments	X	X	X		
Level of Service of Safety	X	X	X	X	
Excess Predicted ACF Using Safety Performance Functions (SPFs)	X	X	X	X	
Probability of Specific Crash Types Exceeding Threshold Proportion	X	X			
Excess Proportion of Specific Crash Types	X	X			
Expected ACF with Empirical Bayes (EB) Adjustment	X	X	X	X	
EPDO-ACF with EB Adjustment	X	X	X	X	EPDO Weighting Factors
Excess Expected ACF with EB Adjustment	X	X	X	X	

For ACF, it is required number and location of accidents (or crashes) in given site and period. First, accident coordinates are marked on the map and the roadway network is divided into segments with a particular length. Then, these segments are ranked according to the number of accidents, and finally, segments are identified for improvements.

EPDO-ACF performance measure requires the crashes data including severity, location and crash costs by severity. This performance measure assigns weighting factors to crashes by severity (fatal, injury, property damage only) by transforming fatal and injured accidents into property damage only (PDO). First, EPDO weights are calculated for fatal, injury, and PDO crashes using by Equation 1. Here, $f_{y(weight)}$ is corresponding to weighting factor based on crash severity, CC_y is a crash cost for crash severity and CC_{PDO} is a crash cost for PDO crash severity. Then, for each segment, the EPDO weights are multiplied by the corresponding number of fatal, injury, and PDO crashes and these are added as shown in Equation 2. Here, $EPDO_i$ is corresponding to total EPDO score for i^{th} segment, $N_{(weight),i}$ is number of crashes by severity for i^{th} segment. Finally, segments are ranked and identified for improvements.

$$f_{(weight)} = EPDO = \frac{CC_y}{CC_{PDO}} \tag{1}$$

$$EPDO_i = f_f * N_{f,i} + f_{inj} * N_{inj,i} + f_{PDO} * N_{PDO,i} \tag{2}$$

f_f : Fatal crash weight
f_{inj} : Injury crash weight
f_{PDO} : PDO crash weight

It is needed crash costs by severity to calculate weighting factors. In this study, it is used crash costs defined by Federal Highway Administration (FHWA) because there are no available local crash costs. As seen in Table 2, the adjusted crash costs by severity for the year 2016 as described in the report published by FHWA and EPDO weights [14].

Table 2. Crash costs by severity and EPDO weights [14]

Severity	Crash Cost (2016, $)	Weight (EPDO)
Fatal	5888800	561
Injured	119600	11
PDO	10500	1

RSI, another performance measure that used, requires crash data including type, location, and RSI crash costs. Firstly, for each segment, the number of crashes for each crash type is multiplied by their RSI crash cost. Then, average RSI value is calculated per segment by using Equation 3 and on reference population by using Equation 4. Segments are ranked and the average RSI cost per segment is compared to the average RSI cost for its respective population. The resulting RSI performance measure shows whether a site is experiencing higher crash costs than the average for other sites with similar characteristics.

$$\overline{RSI}_i = \frac{\sum RSI_j}{N_i} \tag{3}$$

$$\overline{RSI}_{ref\ pop} = \frac{\sum RSI_i}{\sum N_i} \tag{4}$$

RSI_i : Total RSI cost for the i^{th} site
RSI_j : RSI cost for each crash type, j
N_i : Number of observed crashes at the i^{th} site
$\overline{RSI}_{ref\ pop}$: Average RSI cost for the reference population

Table 3 shows the adjusted crash cost estimates by crash type for the year 2016 as described in the report published by FHWA [14].

Table 3. Crash cost estimates by crash type [14]

Crash Type	Crash Cost (2016, $)	Crash Type	Crash Cost (2016, $)
Struck pedestrian	421100	Rear-end	43900
Struck fixed object	138000	Sideswipe	49200
Struck parked car	28900	Opposite direction	548500
Rollover/Run-off-road	350500	Undefined	80100

2.3. Network Screening Technics

It is described three technics in HSM; simple ranking, sliding window and peak searching. The simple ranking is a simple method and can be applied to nodes and roadway segments. However, in segments, the results of this technic are not as reliable as the other technics'. In this method, performance measures are calculated for all sites and the results are ranked from high to low.

In the sliding window method, the roadway is split into equal sections and a window of a specified length (consisted of minimum three sections) is moved along the roadway from beginning to end. Sliding window process is described in Figure 3. Performance measure, requested for improvements or chosen by data availability, is calculated for each segment and then segments are ranked from high to low.

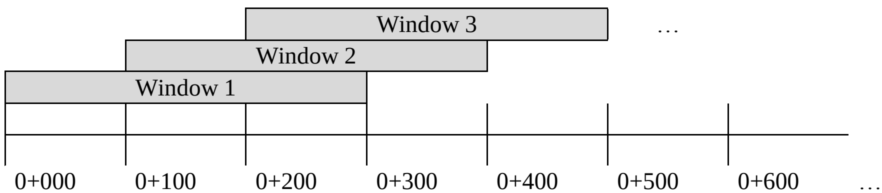

Figure 3. Sliding window process

In the peak searching method, firstly, the roadway is divided into segments defined by two specific endpoints. First, the segments are split into sub-segments by an equal length which is generally 100 m for Iteration 1. Performance measures are calculated for each sub-segment and the results are subjected to precision testing. The precision of the performance measures is assessed by calculating the coefficient of variation (CV) of the performance measure. Segments are ranked based on the maximum performance measure of the sub-segments that meet the desired precision level (should be CV≤0,5). If none of the performance measures at the end of the first iteration, sub-segments are found to have the desired precision, the length of each sub-segment is increased to 200 m for Iteration 2 and so on (Figure 4). If the calculated CV is less than or equal to the limit value, the performance measure meets the desired precision level. The performance measure for a given sub-segment can be considered to rank the segment.

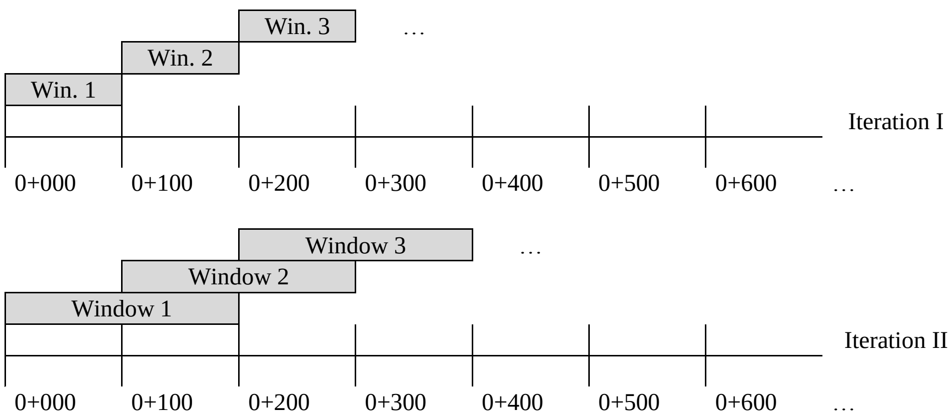

Figure 4. Peak searching process for a roadway segment

3. ACCIDENT BLACK SPOT ANALYSES: THE CASE OF SELECTED CORRIDOR

In HSM, it is not identified how ABS are identified. So, it has been used K-means clustering method to identify black spots. In the K-means clustering method which is frequently used in many studies, firstly, it is necessary to determine how many clusters the data should be divided. The number of clusters is indicated by "*k*" and the method consists of three steps.

- *k* data are selected randomly from the data and these are assigned as a center for each of *k* clusters.
- The distances between the remaining data and each cluster center are calculated. Afterward, each data is assigned to the closest cluster.
- Finally, the centers of the clusters (average of data in the cluster) are recalculated. According to the new centers, data is reassigned to the closest clusters. This process continues until the clusters are stable [15].

The results obtained by performance measures were divided into 5 clusters using SPSS software for each network screening method and hazard classification was defined in Figure 5. As a result of the analyses, the sections assigned to Cluster 5 that shown in red color in Figure 5 were identified as ABS.

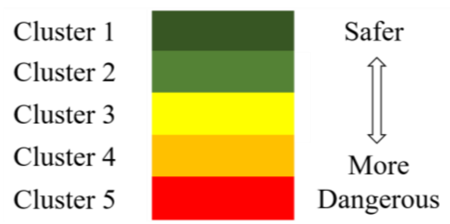

Figure 5. Hazard classification

In this study, the accidents occurred on the 7320-meter-long corridor between Sogutlucesme and 15 Temmuz Sehitler Bridge were analyzed (Figure 6). It happened 206 fatal or injured accidents, four people were passed away and 300 injured over here between 2011 and 2013. The roadway was divided into segments the length of 100 and 250 m for the simple ranking method. To perform the sliding window method, the corridor was split into 10 m segments and a window in two different lengths that were 100 and 250 m was moved alongside the corridor. For peak searching method, the roadway was split into eight segments based on merging and diverging location on the corridor as seen in Figure 6.

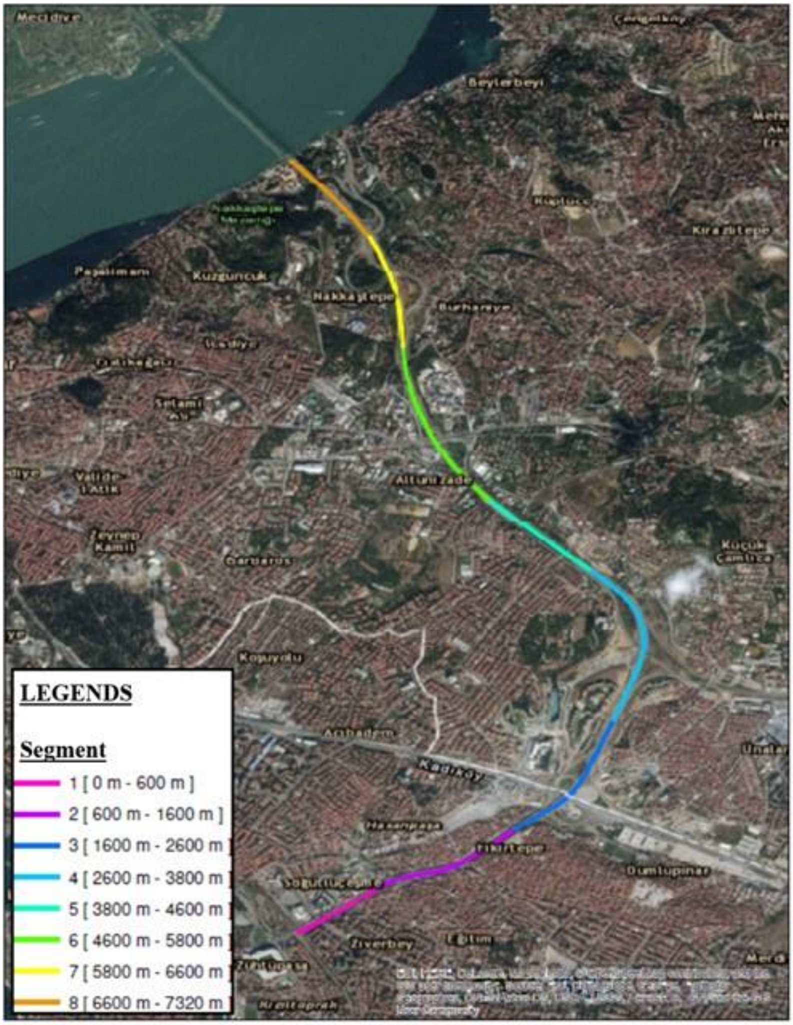

Figure 6. Study corridor and segments based on peak searching

When segments were ranked by ACF performance measure for 100-meter-long sections, it has been seen that the most dangerous section is between 6+900-7+000 km with ACF is equal to 19 and ACF is 27 between 6+910 km and 7+010 km according to simple ranking and sliding window methods, respectively. For 250-meter-long sections, the most dangerous section is between 6+750-7+000 km with ACF is equal to 47 and 6+900-7+150 km with ACF is equal to 35 according to simple ranking and sliding window methods, respectively, are presented in Figure 7.

When segments were ranked by EPDO-ACF performance measure for 100-meter-long sections, it has been seen that the most dangerous section is between 6+900-7+000 km with EPDO is equal to 858 and EPDO is 1001 between 6+910 and 7+010 km according to simple ranking and sliding window methods, respectively. For 250-meter-long sections, the most dangerous section is between 6+750-7+000 km EPDO is equal to 1078 and between 6+810 and

7+060 km EPDO is equal to 1144 according to simple ranking and sliding window methods, respectively, as shown in Figure 8.

When segments were ranked by RSI performance measure for 100-meter-long sections, it has been seen that the most dangerous sections are between 6+900 and 7+000 km with RSI is equal to 1570400$ and 7+220-7+320 km with RSI is equal to 1354900$ according to simple ranking method and RSI is 2309600$ between 6+810 and 7+000 km according to sliding window method. For 250-meter-long sections, the most dangerous section is between 6+750-7+000 km with RSI is equal to 2713900$ and RSI is 3420200$ between 6+660 and 7+150 km according to simple ranking and sliding window methods, respectively, are depicted in Figure 9.

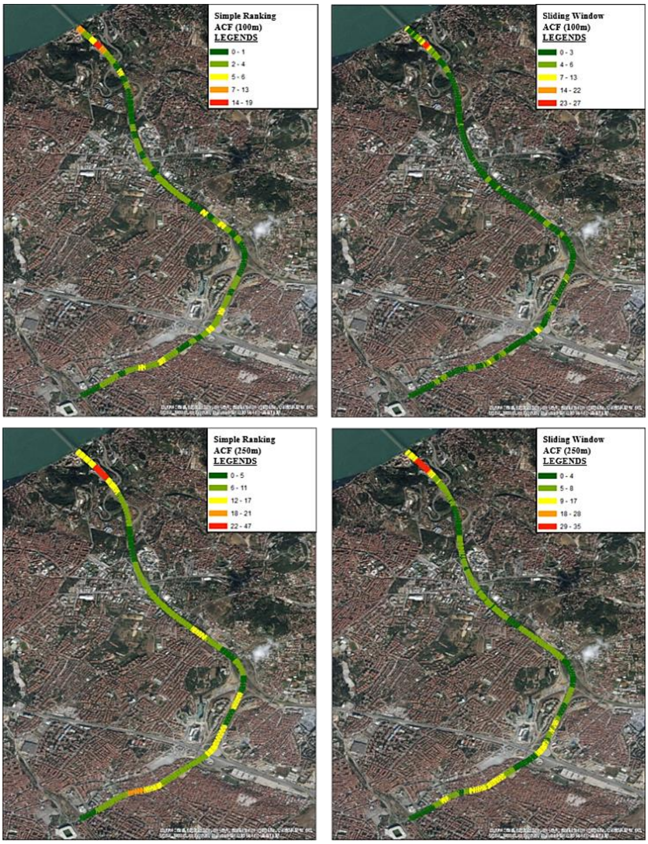

Figure 7. Application of simple ranking and sliding window methods for ACF

The roadway is divided into eight segments that have two endpoints and 100-meter-long sub-segments to perform the peak searching method. The desired precision level was achieved for this length as a result of the first iteration. The results of the analysis based upon the peak searching method (only calculated CV ≤ 0.5) are given in Table 4. It can be seen that different sub-segments are identified the most dangerous by performance measures. This is because; the costs of accidents are different according to crash severity and type.

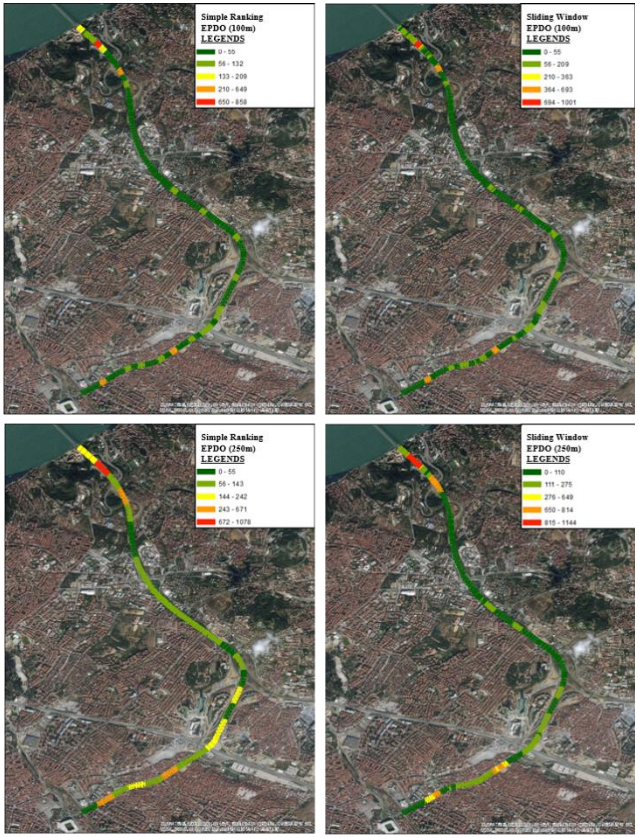

Figure 8. Application of simple ranking and sliding window methods for EPDO-ACF

Table 4. Most dangerous sections by peak searching method for a length of 100 m (Iteration 1)

Sub-Segment (Window)	Start	Finish	CV_{ACF}	CV_{EPDO}	CV_{RSI}
I_IV	0+300	0+400	0.52	0.39	1.39
I_V	0+400	0+500	0.52	5.22	0.44
I_VI	0+500	0+600	0.39	5.22	0.47
II_IV	0+900	1+000	0.28	1.31	0.90
II_VII	1+200	1+300	0.33	2.62	0.29
II_VIII	1+400	1+500	0.55	0.29	1.44
III_II	1+700	1+800	0.65	0.57	0.34
III_V	2+000	2+100	0.32	0.44	0.37
III_VIII	2+300	2+400	0.32	0.33	0.50
IV_III	2+800	2+900	0.43	0.35	0.44
IV_XII	3+700	3+800	0.34	0.56	0.33
V_III	4+000	4+100	0.27	0.30	0.31
VI_II	4+700	4+800	0.40	0.39	0.37
VI_V	5+000	5+100	0.40	0.31	0.38
VI_XI	5+600	5+700	0.30	0.39	0.62
VII_VII	6+400	6+500	0.32	0.34	0.45
VII_VIII	6+500	6+600	1.60	5.05	0.34
VIII_IV	6+900	7+000	0.30	0.32	0.22

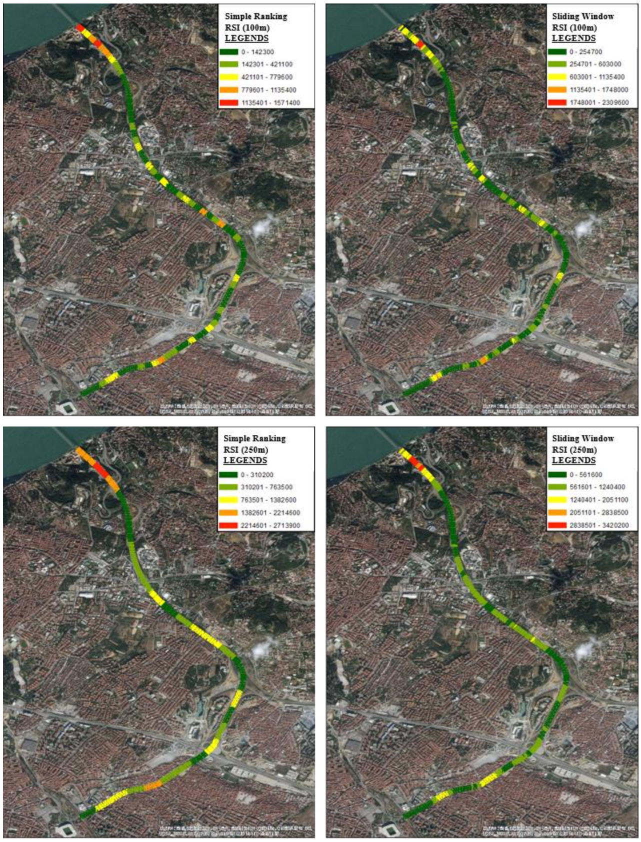

Figure 9. Application of simple ranking and sliding window methods for RSI.

According to the results of the three performance measures, segment VIII (6+900-7+000) is the location where accidents occurred frequently for peak searching method. Fatal accidents occurred in segments I, II and VII, have affected the analysis according to the EPDO-ACF performance measure. According to the RSI performance measure, the various costs of accident types have caused some differences in the ranking of dangerous segments.

4. DISCUSSION AND CONCLUSION

In this study, ABS has been identified by simple ranking, sliding window and, peak searching using by K-means clustering method in Istanbul, Sogutlucesme-15 Temmuz Sehitler Bridge corridor. In the analysis, three different performance measures such as ACF, EPDO-ACF, and RSI were used.

According to three performance measures and network screening methods that used, 15 Temmuz Sehitler Bridge tollbooths segment was identified and marked with red color as the first ABS section of the corridor in the study corridor. Also, the secondary and tertiary risky sections are colored by orange and yellow respectively in Figure 7, 8 and 9. As seen in the analysis, the tollbooths segment, where has a high traffic volume, has been the riskiest site regarding all of the performance measures. Therefore, the factors that cause accidents in this region should be identified and effective countermeasures should be taken. In consequence of replacing tollbooths with open toll system, a reduction in the number of accidents is expected.

When considering the total length of the segments classified as primary, secondary and tertiary crash zones by simple ranking, sliding window and peak searching methods; it is seen that these lengths are shorter in sliding window method. Therefore, the segments identified hazardous by the sliding window method would be the most efficient method to identify the countermeasures and the priority of the segments.

Given the fact that based on EPDO-ACF and RSI performance measures risky locations were differentiated, due to the fact that they target to achieve different results. So, it can be said that a performance measure is selected to identify sites with the potential to reduce the crash frequency, crash severity or specific crash types for a formulation of systemwide policy.

REFERENCES

[1] WHO, (2015) Global Status Report on Road Safety 2015, *World Health Organization*, Italy.

[2] Murat, Y. Ş., Şekerler, A., (2009) Trafik Kaza Verilerinin Kümelenme Analizi Yöntemi ile Modellenmesi, *İMO Teknik Dergi*, 20(3), 4759-4777.

[3] Karaşahin M, Terzi S., (2003) Coğrafi Bilgi Sistemleri ile Isparta-Antalya-Burdur Karayolunun Kara Nokta Analizi, *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi,* 9(3), 305-311.

[4] Saplıoğlu M, Karaşahin M., (2006) Coğrafi Bilgi Sistemi Yardımı ile Isparta İli Kent İçi Trafik Kaza Analizi, *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 12(3), 321-332.

[5] Anderson TK., (2009) Kernel Density Estimation and K-means Clustering to Profile Road Accident Hotspots, *Accident Analysis & Prevention*, 41(3), 359-364.

[6] Kara C, Akçit N., (2015) Traffic Accident Analysis Using GIS: a Case Study of Kyrenia City, *3rd International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2015),* March 2015, Paphos, Southern Cyprus.

[7] Kundakci E., (2014) Identification of Urban Traffic Accident Hot Spots and Their Characteristics by Using GIS, MSc Thesis, *The Graduate School of Natural and Applied Sciences, Middle East Technical University*, Ankara, Turkey.

[8] Dereli MA et al., (2015) Coğrafi Bilgi Sistemleri Destekli Trafik Kaza Kara Nokta Belirleme: Ampirik Bayes Uygulaması, *Harita Teknolojileri Elektronik Dergisi*, 7(2), 36-42.

[9] Gregoriades A., Mouskos K. C., (2013) Black Spots Identification Through a Bayesian Networks Quantification of Accident Risk Index, *Transportation Research Part C: Emerging Technologies*, 28, 28-43.

[10] de Ona, J., López, G., Mujalli, R., & Calvo, F. J., (2013) Analysis of Traffic Accidents on Rural Highways Using Latent Class Clustering and Bayesian Networks, *Accident Analysis & Prevention*, 51, 1-10.

[11] AASHTO (2010) Highway Safety Manual, 1st Edition, *American Association of State Highway and Transportation Officials,* Washington DC, USA.

[12] Manepalli URR, Bham GH., (2016) An Evaluation of Performance Measures for Hotspot Identification, *Journal of Transportation Safety & Security,* 8(4), 327-345.

[13] Young J, Park PY., (2014) Hotzone Identification with GIS-Based Post-Network Screening Analysis, *Journal of Transport Geography*, 34, 106-120.

[14] Council F, Zaloshnja E, Miller T, Persaud B., (2005) Crash Cost Estimates by Maximum Police-Reported Injury Severity within Selected Crash Geometries, *Federal Highway Administration*, Georgetown Pike, USA.

[15] Han J, Kamber M., Pei, J., (2012) Data Mining Concepts and Techniques, 3rd Edition, *Morgan Kaufmann Publishers,* San Francisco, USA.