A New Approach for Geographical Information System-Supported Mapping of Traffic Accident Data

Ismail Bulent GUNDOGDU, Fatih SARI and Omur ESEN, Turkey

Keywords: Mapping, Linear clusters, GIS, Traffic accidents, Probable Hot Pieces

SUMMARY

Result values that are obtained from diverse queries and cartographic products need to be sufficiently comprehensible to their users. Geographic Information System GIS is the most effective way of examining and evaluating the results of analyses which use a multitude of data and different criteria.
In the research into reasons for traffic accidents, several statistical methods can be used for determining the most critical points, called Hot Spots, where traffic accidents frequently happen.

Traffic accident causes and results are not limited to the point where the accident occurs.
Since the Hot Spot concept started to appear in accident analysis, literature has cited many methods for detecting Hot Spots. In this study, instead of classic Hot Spot detecting methods, the Hot Pieces (HPCS) method is used to determine the most critical pieces.

The Probable Hot Spots (PRHS) process aims to anticipate dangers and prevent the occurrence of accidents.

To establish the feasibility of the project, Konya city's traffic accident data for the last 10 years for highways and junctions of central roads have been collected.

A New Approach for Geographical Information System-Supported Mapping of Traffic Accident Data

Ismail Bulent GUNDOGDU, Fatih SARI and Omur ESEN, Turkey

1 INTRODUCTION

Increasing world population and corresponding vehicle count create new problems in traffic-flow. Rapidly-increasing mortality and injury from traffic accidents now have a notable priority.

A report from the World Health Organization (WHO) and the World Bank (WB) (2004) on road traffic accidents and injuries estimated 1.2 million people are killed in road crashes each year and as many as 50 million are injured worldwide (Rathinam et al., 2007; Rutter and Quine, 1996; Hilakivi et al., 1989). Estimation of the World Bank, traffic accidents will be the third reason of death in 2020.

A questionnaire done by 800 people whose relatives lost their lives in traffic accidents in European countries revealed that 37% of those had tendency to commit a suicide and 64% of those had depression in subsequent 3 years. Further, another questionnaire applied in Turkey revealed that 50% of 240 people whose relatives lost their lives in traffic accidents suffered from insomnia and 39.2% suffered from hysteria in addition to depression (Arslan,2007).

Traffic accidents on highways tend to cluster on straight roads, which can cause sleepiness and inattention, or on substantially curved roads with a limited view or in forest or sloping areas. Accidents have complicated factors owing to these and many other features.

A lot of complex factors and differing locations of accident occurrences have to be examined in information systems. An information system consists of a range of operations from observation to analysis-presentation, and the system aims to increase user-decision capability. At the end of the process the user can select the best decision by asking “what ” and “why” questions (Yomralıoğlu, 2002).

GIS makes it possible to analyse data spatially and to make models for solving problems, creating different databases at all levels, from producing particular scenarios to obtaining results. Spatial and textual data concerning actual accidents are analysed together in such studies. Many different illustrations are possible on maps by use of statistical estimations based on a great amount of data. Hence, points or locations with greater accident probability for any reason can be determined. Determination of an accurate analysis method with GIS, which provides many such methods, together with accurate estimations and essential facts about the case, is crucial for accurate results. Definition and treatment of the term Hot Spot, which is the main actor in such studies, will be reviewed shortly in this chapter.

Areas of concentrated crime are often referred to as Hot Spots. Researchers and police use the term in many different ways. Some refer to Hot Spot addresses (Eck and Weisburd, 1995; Sherman, Gartin, and Buerger, 1989), others refer to Hot Spot blocks (Taylor, Gottfredson, and Brower, 1984; Weisburd and Green, 1994), and others examine clusters of blocks (Block and Block, 1995).

This study differs from others in that it analyses "pieces" instead of accident points and increases the capability of examining roads by determining "probable pieces" before a traffic accident happens.

2 LITERATURE REVIEW

Traffic accidents are the consequence of traffic movements, which take place uniquely along a road network. The existing methods for detecting dangerous locations, such as the Kernel or the local spatial autocorrelation method (Flahaut et al, 2003; Steenberghen et al, 2004), are based on Euclidean distances, and thus disregard the specific nature of traffic movement.
Traffic accidents tend to be concentrated in clusters in geographic (e.g. Yamada & Thill, 2004); Accidents are more likely to occur at dangerous locations. Concentrations of traffic accident occurrences suggest spatial dependence between accidents and common causes. These "black zones", or zones with significantly high accident numbers, can be detected by several geostatistical techniques. This identification and analysis of locations producing more accident than the average, is hence an important step in traffic accident prevention. (Aerts et al., 2006).

Many vision algorithms depend on the estimation of a probability density function from observations. Kernel Density Estimation techniques are quite general and powerful methods for the problem, but have a significant disadvantage in that they are computationally intensive (Elgammal et al, 2003).

In the study on Eskişehir (Güvenal et al., 2006) accident intensity analysis is realised by use of accident data for junction points, defined as Hot Spots, which have seen numerous accidents,.

In the study on Isparta (Tuncuk et al., 2004) besides the definition of Hot Spots by use of accident numbers, accident day and hours data are added to the database for detecting Hot Spots related to specific times.

All these studies are aimed at detecting junction points, or Hot Spots, where accidents happen more frequently than elsewhere. But only detecting Hot Spots is not a permanent solution to the real problem. The crucial issue is detecting accident points by specifying the different reasons why certain junction points suffer traffic accidents.

3. MATERIAL AND METHODS

Konya is not only the largest province of Turkey in the context of territorial size but also has the longest road network, with a 2957 km state road in a country with a total of just 61 939 km of roads.. It lies eighth in the context of accident rates, third in the context of mortality rates and fifth in the context of numbers of injuries, according to statistics for 2006.

In this paper data relating to the duty area of the District Traffic Agency of Konya province for the last ten years were studied in order better to clarify GIS-supported studies to prevent traffic accidents. Total length of the related highways is 724 km and urban roads have 103 junction points. Figure 1 deals with satellite-image digitised roads which relate to the project.

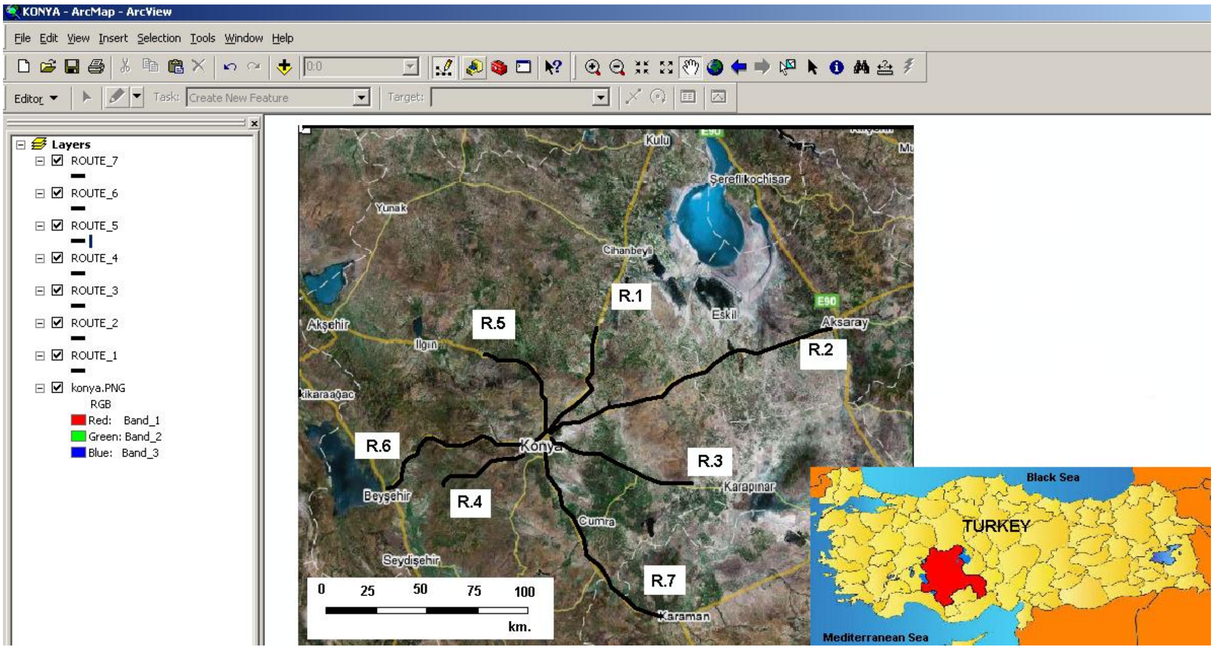

Figure 3-1: Location of the related highways (7 routes)

Since Konya is almost exactly in the middle of Turkey, it can be designated as a junction point where all territories of the country are linked to each other. The city, which was the capital of the Seljuk Empire, has significant historical importance that attracts a great number of both foreign and native tourists resulting in increased in Turkey, which commonly employs highways for transportation. On the other hand, the province of Konya has a smooth topography without any rough regions. Being so straight, the highways in Konya make it possible to drive even faster.

Data evaluation

Accident records for the last ten years have been obtained from Konya Police Department and loaded into the database on MS Access to set up the GIS-based application. 7 existing routes were digitised initially. Route numbers where the accidents occurred, together with km, date, time, weather conditions, types of vehicle, numbers of deaths and injuries and estimated damage, were all entered so that analyses from many perspectives could be carried out. For defining routes, digitisation and analyses were performed on the software ‘Arc GIS 9.2’ by use of maps with a scale of 1/25 000. Thus, the routes where the Hot Spots would be fixed were determined. Roads were digitised with maps on a scale of 1/25 000 and defined with a WGS84 coordinate system. Hence, the study was based on real coordinates and lengths. That is to say, maps were digitised completely and accurately, which made it possible to update the accident database easily and to add GPS coordinates later. Compatible input of detailed information about 3856 accidents was the most time-consuming part of such a study. 2137 accident data of 103 junction points in urban areas were used to represent Probable Hot Spots by use of Kernel Density Estimation.

3.1.HPCS Method

Scrutiny of traffic accidents shows that do not constitute just one event at just one point. For this reason, it would be wrong to consider only one point in determining Hot Spots. Assessment of HPCSs would be more a precise guideline for studying risk factors and accident regions together in linear clusters.Dufays et al.’s (2004) study is about the disadvantages and advantages of two-dimensional clusters and linear clusters. The following advantages and disadvantages were found when both techniques were applied to accident clustering in an urban environment (Figure 2).

Advantages of two-dimensional clusters

In an urban environment with a dense road network, accident locations are frequently based on proximity characteristics. In these environments, two dimensional clusters may suggest causal relationships. An example is the identification of accident concentrations near schools.

Disadvantages of two-dimensional clusters

With two-dimensional clustering techniques, the specificity of traffic flows is not taken into account. On roads with clear traffic characteristics, two-dimensional clusters may even give a false impression of a spatial distribution of accidents. The identification of spatial clusters by means of circles with a fixed radius creates a problem near the margins of the grid. In grid cells closer to the edge than the diameter of the search radius, the number of points is not computed for the totality of the circle.

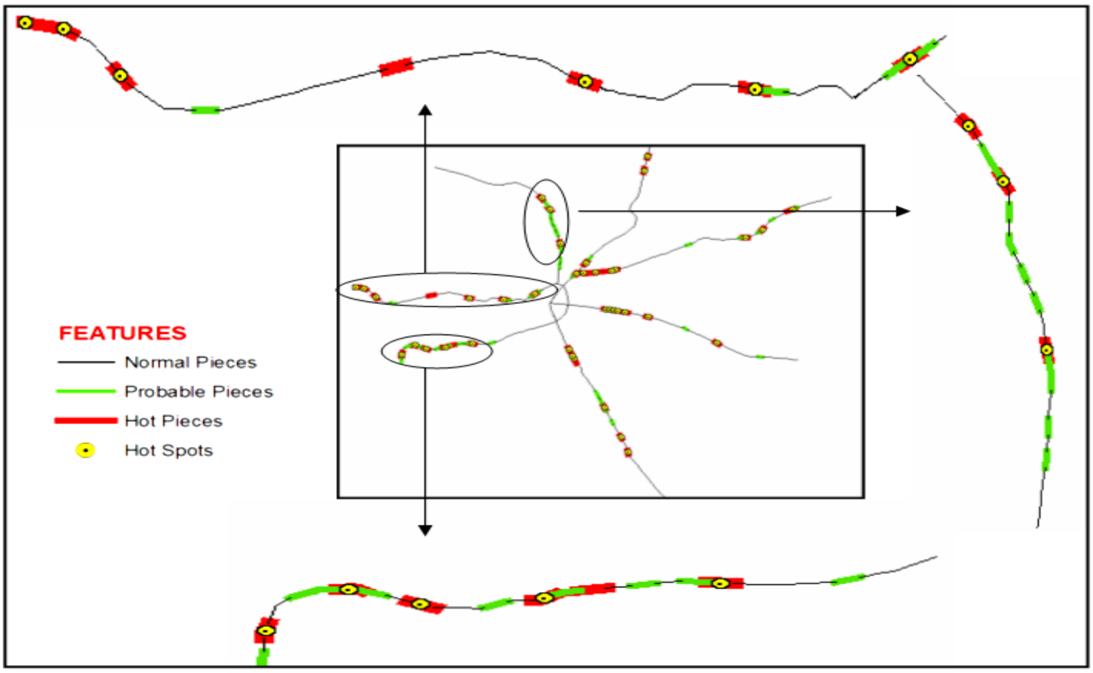

Figure 3-2: Hot pieces (HPCS) on the roads (zoomed 3 views)

Advantages of linear clusters

Linear clusters give a representation which is consistent with the linear flow of traffic. By computing the linear clusters on routes representing traffic flows, the concentrations of accidents are related to connectivity. This may suggest causal dependence. An example is the location of accident clusters on routes connecting industrial sites with major highways.

Disadvantages of linear clusters

In the linear method, edge problems occur at the end of each road. For points located closer to the end of the segment than the search radius, no accidents are counted beyond the end of the segment, thus distorting the concentrations near the edges. Without this adjustment, an artificial concentration is created due to double counts of accidents near intersections. Because the roads were combined into routes, the edge problem is reduced to the end of each route. In an urban environment with a dense road network, however, identification of meaningful routes may not be straightforward.

Considering to the advantages of Hot Pieces, in this study, the roads are divided into 1 kilometer pieces in highways and 724 pieces are examined as linear clusters for every pieces. In this study, The highway routes and statistical analysis examined separately so the disadvantages of linear clusters will not affect the results. Related results can be shown in Figure 2.

3.2. PRHS Method

One of the most vital and important innovations of this study is Probable Hot Spots (PRHS). The basic goal of PRHS is to anticipate and prevent premature accidents. Given the purpose for determining Probable Hot Spots, in this chapter two kinds of statistical analysis methods are used. These are Getis Ord Gi and Kernel Density Estimation. The details of determining Probable Hot Spots and Hot Spots are discussed below.

Getis Ord Gi:

This analysis technique is the most utilized method for detecting Hot Spots on highways. The statistics Gi and Gi*, introduced by Getis and Ord (1992) for the study of local patterns in spatial data, were extended and re-written in 1995.

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^* \bar{x}^*}{s^* \left\{ \left[\left(nS_{1i}^* \right) - W_i^{*2} \right] / (n-1) \right\}^{1/2}}$$

Where w_{ij} (d) is a spatial weight vector with values for all cells j within distance d of target cell i, W_i^* is the sum of weights, S_{1i}^* is the sum of squared weights and s^* is the standard deviation of the data in the cells.

In the Getis Ord Gi method, the probability of being a Hot Spot'value must be higher than 1,645 threshold value which is obtained from $z=(x_i-x_0)/s$ related with F normal distribution table according to 95% probability. In Figure 3, 1,645 threshold value(get from t distribution table) is shown as a red line and 1,000 Probable Hot Spot threshold value is shown as a green line. The points are called Probable Hot Spots, which have a value between 1,645 and 1,0345. The Probable Hot Spots threshold value depends on the accident numberand the scale of the project.

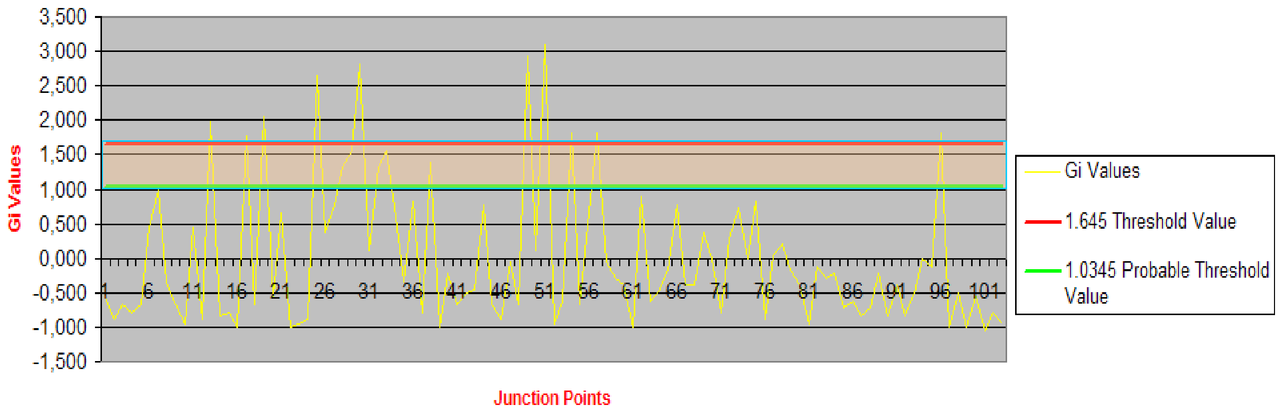

Figure 3-3: Gi values on junctions calculated by accident number.

Figure 4 represents Probable Hot Spots of junction points in Konya's urban area which are obtained by Kernel Density Estimation. This method is used for determining and representing Probable Hot Spots owing to the ease with which it can be applied to analysis of road networks like urban roads.

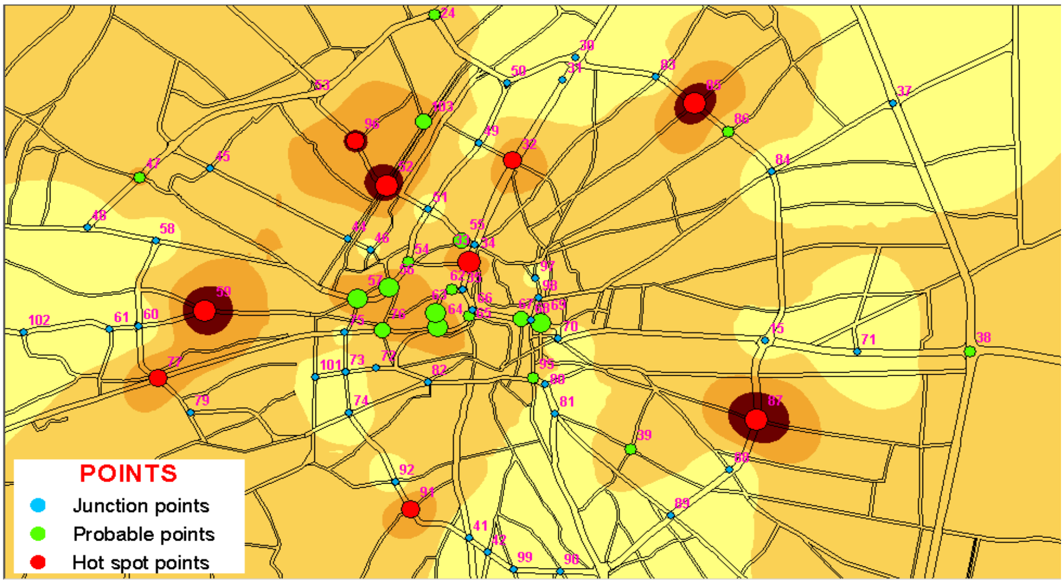

Figure 3-4: Relations of probable Hot Spots with normal Hot Spots on junction points by Getis Ord Gi.

<u>Kernel Density Estimation</u>

The Kernel method introduced by Rosenblatt (1956) has received considerable attention in nonparametric estimation of probability densities (Wu et al, 2002).
Kernel Density Estimation has been a popular technique for analysing one and two-dimensional data; see Bowman and Azzalini (1997), Scott (1992), Simono (1996), Wand and Jones (1995) for examples. Density estimates provide useful information about features in the data (Duong et al., 2007).

The existing spatial clustering techniques have disadvantages when applied to a road network instead of a single road segment. The Kernel method results in a grid over the total study area with a dangerousness measure for each grid cell, even if there is no road in the grid cell. This gives the false impression of large extents for the black zones. The local spatial autocorrelations method requires the aggregation of accidents in Basic Statistical Units (BSU’s) (Flahaut et al, 2002).

Kernel Density Estimation is able to quickly and visually identify hotspots from large datasets and therefore provide a statistical and aesthetically satisfactory outcome. The advantages of these surface representations particularly of road accidents are that they can provide a more realistic continuous model of accident hotspot patterns reflecting the changes in density which are taken difficult to represent using geographically constrained boundary basin models such as the transport network or census tracts. Over the years there have been a number of spatial tools developed which help in the understanding of the changing geographies of point patterns. The most promising of these tools is Kernel Density Estimation (Sabel et al 2005).

In Figure 5, Probable Hot Spots are represented by use Getis Ord Gi, owing to the good facilities for applying this method to the highways. The numbers represent accident numbers and kilometres.

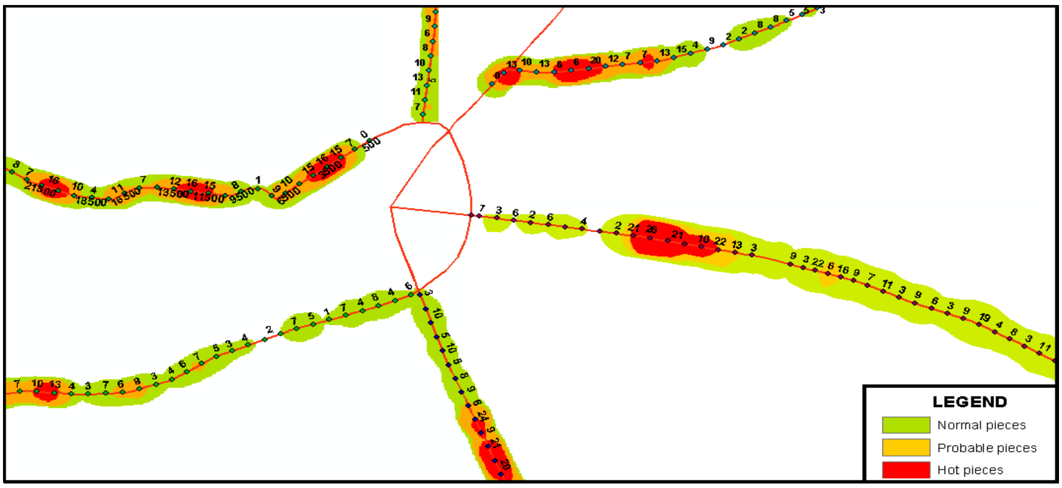

Figure 3-5: Representation of different feature pieces with accident number

There are many advantages to the use of Kernel Density Estimation as opposed to statistical hotspot and clustering techniques such as K-means. The main advantage for this particular method lies in determining the spread of risk of a accident, outlined earlier. In defining a cluster one is overlooking the buffer around it which will ultimately possess a degree of risk of accident for the people who enter it. This degree of risk would not be measured using the clustering techniques.

Kernel; Distribution with density function $p(x)$, an estimate $\hat{p}(x)$ of the density at x can be calculated using

$$\hat{p}(x) = \frac{1}{N} \sum_{i=1}^{N} K_{\sigma}(x - x_i)$$

Where K_{σ} is a “Kernel function” with a bandwidth (scale) σ .

4. RESULTS

Apart from classical illustrational techniques, Hot Pieces (HPCS) on roads divided into 1 km. segments are shown by their grading according to their numerical values. Hence, thematic illustration distinguishes them from others. The disadvantages of point representation are clear, whereas linear clusters eliminate them, thus providing us with a more sensitive analysis. Seven arteries leading into Konya, the largest province of Turkey, were studied for data-collection. Thus, we have the chance of more comprehensive involvement in problems through extensive data about the entire roads of a province and not just data for a single road-segment.

Another aspect of this study is that beside investigation of Hot Spots by means of data of previous years’ accidents, Probable Hot Spots (PRHS) were illustrated and highly-potential Hot Spots were determined. These latter are candidates for Hot Spots in the near future. So premature accident can be anticipated easily. Table 1 shows that number of probable hot spots which will be hot spot how many accidents later.

Table 1
Relation of between PRHS number and necessary accident number to be Hot Spot

	Necessary accident number to be Hot Spot															
PRHS Number	Urban area junctions		Route 1		Route 2		Route 3		Route 4		Route 5		Route 6		Route 7	
	1-3	4-6	1-3	4-6	1-3	4-6	1-3	4 - 6	1-3	4-6	1-3	4 -6	1-3	4-6	1-3	4-6
1		1	1	1	2		2	2		1	1	2	1		1	
2	1	1		1					1		1		1			4
3					1	1		1	1		1			1		
4				1					1							
5										1						
Total	1	2	1	3	3	1	2	3	3	2	3	2	2	1	1	4

In the table every "x" sign represents the necessary PRHS traffic accident number which identifies Hot Spots in the future.
In this table, accident numbers are grouped into two: (1-3), (4-6). (1-3) group shows the necessary PRHS accident number which is a first priority for attention. In the same mean (4-6) group has second priority. This table shows the points that need urgent preventative measures. These results have great importance for accident prevention.

REFERENCES

Aerts , K., Lathuy a, C., Steenberghen T., Thomas I., 2006. 19th ICTCT workshop Proceedings SESSION 4: Spatial clustering of traffic accidents using distances along the network.

Arslan, F., 2007. Emniyet Genel Mudur Yardimcisi. http://www.trafik.gov.tr/icerik/bildiriler/A6-87.doc

Biley, T. C., and Gatrell, A. C., 1996. Interactive Spatial Data Analysis (Harlow, UK:Longman).

Block, R., and Block C., 1995. "Space, Place and Crime: Hot Spot Areas and Hot Place of Liquor-Related Crime." In J.E. Eck and D. Weisburd (eds.), Crime and Place(vol. 4, pp. 145–184). Monsey, NY: Criminal Justice Press.

Bowman, A. W. and Azzalini, A. , 1997. Applied Smoothing Techniques for Data Analysis. Oxford University Press, Oxford.

Dufays, T., Steenberghen, T., Flahaut, B., Thomas, I., 2004. Intra-urban location and clustering of road accidents using GIS:a Belgian example. International Journal of Geographical Information Science. 18,169-181

Duong, V., 2007. Feature Signi¯cance for Multivariate Kernel Density Estimation Tarn Duong, Arianna Cowling, Inge Koch & M. P. Wand School of Mathematics & Statistics University of New South Wales Sydney, Australia.

Eck, J.E., and D. Weisburd, 1995. "Crime Places in Crime Theory." In J. E. Eck and D. Weisburd (eds.), Crime and Place (vol. 4, pp. 1–33). Monsey, NY: Criminal Justice Press.

Elgammal, A., Duraiswami, R and Larry S. D., 2001. "Efficient computation of kernel density estimation using fast gauss transform with applications for segmentation and tracking," in Proc. Of IEEE 2nd Int. workshop on Statistical and Computational Theories of Vision, Vancouver..

Flahaut, B., Thomas, B., 2002. Identifier les zones noires d'un reseau routier par l'autocorrelation spatiale locale – Analyses de sensibilité et aspects operationnels, Revue internationale de Géomatique, 12:245–261.

Guvenal, B., Cabuk, A., Yavuz, M.I., 2006. Trafik Kazalarinin Azaltilmasi Amaciyla Cografi Bilgi Teknolojilerinden Yararlanilmasi, 3. Cografi Bilgi Sistemleri Bilisim Gunleri, Fatih Universitesi, Istanbul.

Hilakivi, I., Veilahti, J., Asplund, P., Sinivuo, J., Laitinen, L., Koskenvuo, K., 1989. A 16-factor personality test for predicting automobile driving accidents of young drivers. Accid. Anal. Prev. 21 (October (5)), 413–418.

Ord, J.K., Getis, A., 1995. Local Spatial Autocorrelation statistics: Distribution issues and an application. Geographical analysis. 27 (4):286-306.

Rathinam, C., Nair, N., Gupta, A., Joshi, S., Bansal, S., 2007. Self-reported motorcycle riding behaviour among school children in India. Accident Analysis and Preventation. 39, 334–339

Rutter, D.R., Quine, L., 1996. Age and experience in motorcycling safety. Accid. Anal. Prev. 28 (January (1)), 15–21.

Sabel, C., Kingham, S., Nicholson, A. and Bartie, P., 2005. Road Traffic Accident Simulation Modelling - A Kernel Estimation Approach. In The 17th Annual Colloquium of the Spatial Information Research Centre University of Otago, Dunedin, New Zealand, pp. 67-75.

Scott, D. W., 1992. Multivariate Density Estimation: Theory, Practice, and Visualization. John Wiley & Sons Inc., New York.

Sherman, L.W., Gartin, P.R., and Buerger M.E., 1989. "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place." Criminology, 27(1), 27–55.

Simono, J. S., 1996. Smoothing Methods in Statistics. Springer-Verlag, New York.

Steenberghen, T., Dufays, I. Thomas, I., Flahaut, B., 2004. Intra-urban location and clustering of road accidents using GIS: a Belgian example, International Journal of Geographical Information Science, 18(2):169-181.

Taylor, R.B., Gottfredson, S.D., and Brower S., 1984. "Block Crime and Fear: Defensible Space, Local Social Ties, and Territorial Functioning." Journal of Research in Crime and Delinquency, 21, 303–31.

Wand, M. P. and Jones, M. C., 1995. Kernel Smoothing. Chapman and Hall Ltd., London.

Weisburd, D., and L. Gren, 1994. "Defining the Street Level Drug Market." In D.L. MacKenzie and C. Uchida (eds.), Drugs and Crime: Evaluating Public Policy Initiatives. Thousand Oaks, CA: Sage Publications.

Wu, W.B., Mielniczuk, J., 2002. Kernel Density Estimation For Linear Processes. The Annals of statistics. Vol 30 No.5, 1441-1459

Tuncuk, M., Karasahin. M., 2004. Cografi Bilgi Sistemleri Kullanilarak Trafik Kaza Kara Noktalarinin Tespiti: Isparta Örnegi, 3. Cografi Bilgi Sistemleri Bilisim Günleri, Fatih Universitesi, Istanbul.

Yamada, I & J.-C. Thill, 2004. Comparison of planar and network K-functions in trafficaccident analysis, Journal of Transport Geography, 12:149–158.

Yomralıoglu, T., 2002. Cografi Bilgi Sistemleri.Temel kavramlar ve uygulamalar. 2. Baski, pp:36, Trabzon. ISBN 975-97369-0-X

BIOGRAPHICAL NOTES

Dr. Ismail Bulent Gundogdu, born in 1968. Graduated in 1989 as First rank in Surveying and obtaining doctorate degree in 1997. 1991-1993 Lecturer, 1998- Assist.Prof.Dr. in Selcuk University, Engineering and Architecture Faculty, Konya/Turkey. More than 35 published paper,1 book, 2 CD about cartography, computer aided design, GIS etc. 2005-2006 Head of Cartography Department Division. Memberships of Chamber of Turkish Surveying Engineers. Interestings: Cartography, Visualisation, Multimedia Cartography, Delphi programming, Image proccessing.

Omur Esen

2005- Expert in Selcuk University Construction Offices Presidency. Konya/Turkey. PhD student in Cartography Department. Memberships of Chamber of Turkish Surveying Engineers.

Fatih Sarı

2007- Master Student in Selcuk University, Engineering and Architecture Faculty, Cartography Department, Konya/Turkey. Memberships of Chamber of Turkish Surveying Engineers.

CONTACTS

Ismail Bulent Gundogdu
Selcuk Universitesi
Muhendislik Mimarlik Fakultesi
Jeodezi ve Fotogrametri Muh.Bl.
Konya
TURKEY
Tel. +90 555 721 47 69
Fax +90 332 241 06 35
Email:bgundogdu@selcuk.edu.tr
Web site:http://www.bgundogdu.selcuk.edu.tr

Fatih Sari
Selcuk Universitesi
Fen Bilimleri Enstitusu
Konya
TURKEY
Tel. +90 555 468 86 82
Fax +90 332 241 06 35
Email:fatihsari_85@hotmail.com

Omur Esen
Selcuk Universitesi
Insaat Dairesi Baskanligi Kampus
Konya
TURKEY
Tel. +90 332 223 25 09
Fax +90 332 241 01 00
Email:omuresen@hotmail.com