Current Trends in Technology and Sciences
ISSN: 2279- 0535. Volume: 3, Issue: 5 (Aug-Sept. 2014)

Geostatistical Analyst using the Junction Risk Factor to Analyse and Prevent Urban Traffic Accidents

Ismail Bulent Gundogdu
Department of Geomatics, Faculty of Engineering, Selcuk University. Turkey, bgundogdu@selcuk.edu.tr

Abstract - Geostatistical Analyst (GA) gives information about the reliability of models in addition to continuous surface mapping of the whole area for which data have been obtained. Maps produced by GA allow interpretation from different views in order to improve existing problematic routes or junctions and new routes in future urban planning.
The Junction Risk Factor (JRF), which it has been developed and proved recently Gundogdu (2010), is different from previous methods aiming to prevent traffic accidents using more points of view. The most important difference is that the JRF produces the value with present junction data, but does not consider the statistical data of previous years. The calculated values can be helpful not only for existing junctions or routes but also for future applications.
It is possible to decide which criteria must be considered for new routes and junctions with the help of continuous surface maps including hot zones.
Desired probability and prediction maps based on different cartographic rules can be produced in GA by the kriging technique. In this study, in order to provide recommendations, 121 junction points in an urban centre of Konya, Turkey, are discussed. First the digitisation of points collecting datasets and calculations of the JRF values were performed, and then the reliability of the maps produced were evaluated by the rules of GA.

Keywords - Junction Risk Factor, Prediction maps, Probability maps, Traffic accident.

1. INTRODUCTION

Inner city or urban traffic can doubtlessly be considered one of the biggest problems in developing cities. Urban traffic has its importance in supplying transitions between the urban citizens and their vital functions. Especially in cities which are developing and have the population density in their centres, the present traffic sometimes can not respond to the need. Since damaging or fatal accidents occur in urban areas due to heavy traffic, new searches for solutions to the trouble have been brought about. The financial damage and the casualties have been increasing swiftly and traffic jams and their effects on both citizens and the environment have become a major problem [1].

Although many methods are compiled in the following chapter and many studies have been carried out in order to stop traffic accidents, none of them have been sufficient to resolve that problem entirely. There are many varying factors which contribute to this problem.

The main target of this study is not to state the hot spots by taking the former years' data; the target here is to determine how, using the newly developed algorithm, variables of traffic affecting the result linearly or counter-linearly to reach the minimum JRF value ought to be found. Variables of the JRF are depend on existing factors of junction and routes as numerical and time period. For new roads, the same variables ought to be taken up and they ought to be built and signalization also ought to be adjusted by means of this aspect. In this way, some regulations at present junctions and at prospective junctions, taking care of the variables in the pattern, are going to unroll the sources of accidents and are going to ensure the reduction of damages.

The variables of JRF are independent of each other, and changing one never affects the other. This structure shows the superiority of JRF. over others.

Because of the irregular distribution of junction points, mapping by traditional interpolation techniques is insufficient to evaluate the accident dataset. It is not the main purpose of this study to determine the hot spots. The main aim is to produce continual surface risk maps for the new roads and settlements for future planning. Therefore, risk maps are the fundamental products which consider all JRFs.

As a result, the aim is to demonstrate the appropriateness of using GA for vehicle traffic movements.

A review of the previous literature found that accident prediction models for urban roads were studied by [2] used Getis-Ord Gi statistics to identify spatial patterns of different types of weather-related crashes. In addition to the analysis of the impact of adverse weather conditions on road safety, they explained the Getis-Ord Gi method in detail. [3] used spatial statistics in their article to investigate changes since the early 1980s in offence patterns for residential burglary, theft of and from cars, and vandalism in Stockholm city. [4] also studied spatial relationships between local spaces and global space. [5] examined the use of process models to improve spatial analysis and used Getis-Ord Gi statistics. [6] also used different statistical methods to draw more kinds of maps.

In GIS studies, new approaches were used by ([7], [8], [9]) in applying GIS to environmental simulation models and by [10] for geographic representation. ([[1], 11], [12], [13], [14], [15], [16], [17], [18]) studied risk maps and [19] studied geostatistical models for solving cartographic problems.

Current Trends in Technology and Sciences
ISSN: 2279- 0535. Volume: 3, Issue: 5 (Aug-Sept. 2014)

2. Study Area

Konya is not only the largest province of Turkey in terms of territorial size but also has the longest road network, with a 3651 km state road in a country with a total of just 385748 km of roads. It is eighth worst in terms of accident rates, third in terms of accident fatality rates, and fifth terms of numbers of injuries, according to statistics for 2012. In a seven year period, a total of 2240 accidents occurred; there were 18 accidents with fatalities, 322 accidents with injuries, and 2023 accidents with only damage.

To define 121 junctions, digitisation and analyses were performed using maps with a scale of 1/25 000. In this process, besides all accident numbers, junction criteria, dates, times, and types of vehicle data were recorded in the main dataset.

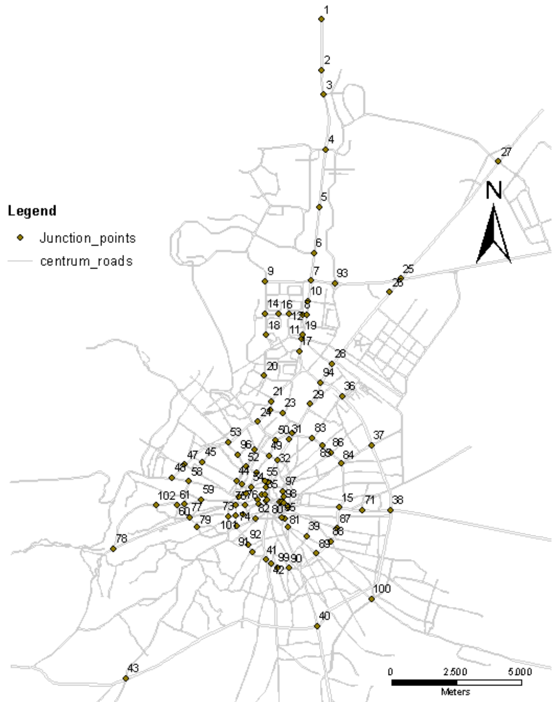

Fig.1: 121 junction points in the city centre of Konya.

3. Generated Algorithm

The main target here is to calculate JRF values for each junction. The JRF at a junction is a value computed at the point of each traffic light and is gained after these calculations are summed. Furthermore, it is accepted that all other risk factors (ORF) will be added into that coefficient. However, because of the ORF which have the same effect at all other junctions, like drunk drivers, weather conditions or the faults of drivers or vehicles, these were not included in the formulization. The main intent here is to investigate the effects of the certain factors that cause accidents. In this manner, these factors can also provide alterations to present roads and can inform the authorities to take them into consideration. An analysis which will be performed according to the accidents can provide us with the result on urban roads. However, only by making an analysis and presentations considering the accidents can help to find the accident ratio of the junction instead of revealing the troubles. As a result of finding real and valid solutions, a model which can provide the most effective way to reach the target must be used, instead of finding out which junction is the most dangerous or at which junction the most accidents occurred (Gundogdu 2010).

Factors that affect the formulization:

Numerical factors: The number of routes that cross the junction is named a numerical factor. The number of roads of the junction affects the total JRF.

Interval factors: These factors must be examined in two groups: first, the average number of vehicles passing through the road in a day and, second, the waiting duration at the red light.

Data related to length: These can also be examined in two groups as the width of the road and the distance from the previous traffic or warning light.

Data related to slope: Since the city of Konya was settled on flat land, the slope was accepted as ineffective and it has not been included in the calculation.

In accordance with that,

$$JRF_i = \sum_{i=1}^{\mu} \frac{l_i \cdot n_i \cdot \lambda_i}{S_i \cdot \eta_i} \qquad (1)$$

If a value is to be evaluated: here, μ : the number of traffic lights at the junction, l_i : the distance from the previous traffic or warning light, n_i : roads crossed at the junction, λ_i : average number of cars that will pass in a day, S_i : the width of the road and η_i : the coefficient of traffic flow.

For η_i

$$\eta = \frac{\delta}{t_R} \qquad (2)$$

can be written. δ is the number of cars waiting at the red light and δ effects η_i directly proportionally; t_R is the duration of the red light which effects η_i inversely proportionally.

Here, the values of l_i, n_i, λ_i increase to JRF.

Therefore, these values expose traffic jams or risks of traffic accidents. As far as possible, these values must be decreased. If it is not possible to lessen some values of the variables, the S_i or η_i value must be increased. In this situation, the most practical solution is to increase η_i or achieve a minimum t_R.

4. Using of Kriging Interpolator for JRF

According to [20] kriging is a moderately quick interpolator that can be exact or smoothed depending on the measurement error model. It is very flexible and allows the user to investigate graphs of spatial

Current Trends in Technology and Sciences
ISSN: 2279- 0535. Volume: 3, Issue: 5 (Aug-Sept. 2014)

autocorrelation. Kriging uses statistical models that allow a variety of map outputs including predictions, prediction standard error, standard error of indicators, and probability. The flexibility of Kriging can require a lot of decision making. Kriging assumes that the data comes from a stationary stochastic process. A stochastic process is a collection of random variables that are ordered in space and/or time such as elevation measurements. The selection of a Kriging method is based on the autocorrelation of a variable between two points that is formulated as follows:

$$Z(s) = \mu + \varepsilon(s)$$

Z(s) Consists of two parts: a deterministic trend $\mu(s)$ (i.e. flow direction) and a random auto correlated error $\varepsilon(s)$. The symbol s simply indicates the location of a junction point. Because $\mu(s)$ is a deterministic trend, the selection of a Kriging method is based on whether a directional trend exists or not. If a directional trend is unknown, then an ordinary Kriging method is appropriate. If a directional trend is known, then a simple Kriging method should be selected.

The ordinary Kriging formula is generally given by:

$$Z^*(u) = \sum_{a=1}^{n(u)} \lambda_a(u) Z(u_a) + \left[1 - \sum_{a=1}^{n(u)} \lambda_a(u)\right] m$$

where, $Z^*(u)$ is the ordinary Kriging estimate at spatial location u, n(u) is the number of the data used at the known locations given a neighbourhood, $Z(u_a)$ are the n measured data at locations u_a located close to u, m is mean of distribution, $\lambda_a(u) =$ weights for location u_a computed from the spatial covariance matrix based on the spatial continuity (variogram) model, which is given by:

$$\lambda(h) = \frac{1}{2n} \sum_{i=1}^{n} \left(z(u_i) - z(u_i + b) \right)^2$$

Where, n is the number of data pairs separated by distance b, $z(u_i)$ and $z(u_i + b)$ are the data values at locations separated by distance b.

Ordinary Kriging assumes a constant but unknown mean and estimates the mean value as a constant in the searching neighbourhood [20].

The variogram is the function which characterises the dependence of variables between two different points in space. According to [21], in contrast to classical statistical methods, semivariogram models are interested in all data points whether regularly distributed or not with respect to both time and location.

5. Methodology and Result

First of all, the characteristics of the data must be evaluated. Table 1 shows the histogram values

Table 1: Histogram values calculated by JRF.

	Min	Max	Mean	Std. Dev.	Skewness	Kurtosis	1st Quantile	Median	3rd Quantile
JRF	**0.0023**	**2.1092**	**0.4645**	**0.4969**	**1.214**	**3.6048**	**0.0489**	**0.1856**	**0.7936**

If the mean and median values are close, it can be said that the data have a normal distribution. A histogram shows whether distributions of data are symmetrical or not. Symmetrical data can be realized with a Skewness value close to zero. According to the values in Table 1, our data are not normally distributed or exactly symmetrical.

Two effects may influence the results. One of them is the global trend and the other is anisotropy. When a study is carried out on a two dimensional surface, sometimes the semivariogram and covariance functions must be investigated not only according to distance but also according to direction. This is called anisotropy. Anisotropy is determined by random errors, and it is different from a global trend.

In addition, unlike the circular isotropic model, an ellipse must be chosen in the searching neighbourhood to determine the neighbourhood number and location. In interpolation, the major axis must have the same direction as the anisotropic direction. This is taken into consideration in the calculation of the values in Table 2

According to [22], the ratio of nugget to sill can indicate statistical autocorrelation. If the value is smaller than 0.25, there is strong autocorrelation; if the value is between 0.25 and 0.75 there is moderate spatial autocorrelation; otherwise there is no spatial autocorrelation. Table 2 shows all geostatistical checking.

Table 2: All calculated statistical values required for interpretation.

Statistical values	JRF	
	With nugget effect	Without nugget effect
Trend	Log.	Log.
Anisotropy	✓	✓
Major range	23548.7	3807.99
Minor range	12554.5	3807.99
Nugget	0.20779	0
Sill	0.25086	0.25086
Nugget sill ratio	0.82	-
Mean	-0.02352	0.01149
Root-Mean-Square	0.4576	0.444
Average Standard Error	0.4699	0.2749
Mean Standardized	-0.049	-0.01313
Root-Mean-Square Standardized	0.9718	1.703

Finally, before the mapping, the performance of values can be controlled by cross validation. Cross validation investigates which model has the best performance.

Current Trends in Technology and Sciences
ISSN: 2279- 0535. Volume: 3, Issue: 5 (Aug-Sept. 2014)

Figure 2 shows prediction maps produced by JRF values with the most problematic eleven junctions (a), and a prediction map without the nugget effect (b). Furthermore, probability maps can be produced in respect of threshold values that are exceeded or not exceeded. A related example is shown in Figure 2(c). The indicator prediction value shows the probability of exceeding a threshold value. The most important maps to interpret to prevent traffic accidents are of this type. Figure 2(d) was produced using covariance values.

Representations of maps illustrated using Equal Interval techniques have generally been used for probability maps to clarify extreme values so far. Therefore, all maps are produced using this method.

For the correct interpretation of the values shown in the table, the following rules must be taken into consideration.

Generally, the best model is the one that has the standardised mean nearest to zero, the smallest root-mean-square prediction error, the average standard error nearest to the root-mean-square prediction error, and the standardised root-mean-square prediction error nearest to 1 [23].

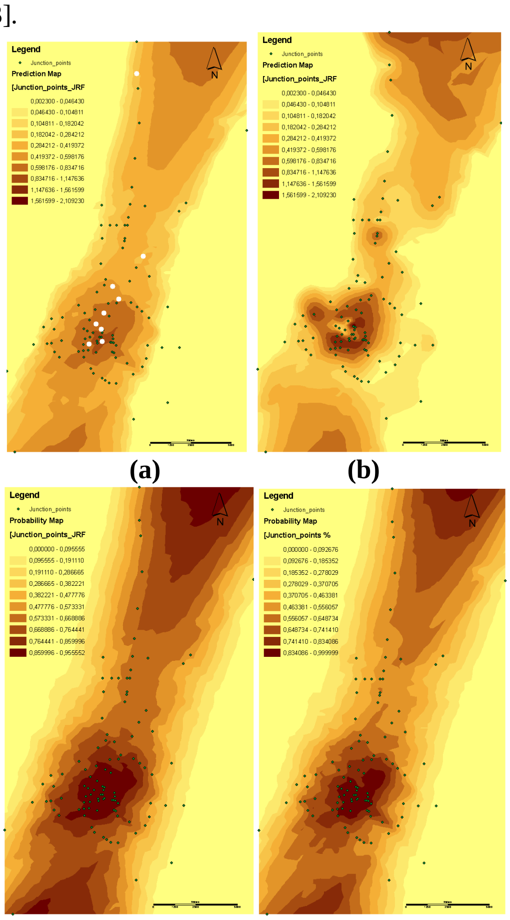

Fig. 2: Result maps: (a) prediction map, (b) prediction map without nugget effect, (c) probability map, (d) probability map for JRF values greater than the mean

6. Summary and Discussion

Up to now, the Kernel and Getis-Ord Gi methods have been used to prevent traffic accidents or to determine critical points for statistical study in the literature. It is accepted that these methods are reliable for these kinds of studies. The methods use all the data related to the area together. The variables of these methods' mathematical models consider the total number of criteria and statistical coefficients for whole area.

The risk of accidents can be reduced and traffic flow can become untroubled. In the light of these suggestions, each of the 121 junction points in Konya has been evaluated and the 27 most dangerous junctions' results are given in Table 3. Suggestions for all the characteristic junctions are shown in the table. For instance, five suggestions for eight junctions, three suggestions for three junctions, and two different suggestions for sixteen junctions are given in the table. In reality, the results could be observed in the city. The variables do not alter from one country to another. So, the JRF model can easily be used in other areas by GA [13].

If the method is evaluated alongside GA, it can be said that deterministic and geostatistical methods are examined with two groups in GA. These methods are basically similar as the data for nearby points are close. On the other hand, when the mathematical and statistical methods are used together, geostatistical methods are not only used as interpolation methods but also give an opportunity for interpretation. The method also gives information about the reliability of predicted areas. Therefore, deterministic applications were not performed in this study. This study has tried to understand the reliable and risky zones.

As a result, in the light of the calculated values, result maps were produced with the aim of helping to create new city road models. Maps which have statistical value and can be visualised will of greater help to city planners [13].

Table 3. Suggestions for the junctions which have high JRF values are mentioned according to the variables of the formula.

Suggestions for the junctions which have big value more than mean	Junction numbers (total 25)			
	6	3	4	12
l_i must be shortened	√	√		
n_i must be controlled	√			
λ_i must be controlled	√		√	√
S_i must be optimized	√	√	√	
t_R time must be set for η_i	√	√		√

References

[1] Gundogdu, I.B., 2010. Traffic management in Konya by Junction Risk Factor. Proceedings of the Instution of Civil Engeneers-Municipal Engineer. 163 (2010), pp:87-96

[2] Ghazan K. G., Qin, X., Noyce, D. A., 2006. 85th Annual meeting of the Transportation Research Board, Washington

[3] Ceccato, V., Haining, R., Signoretta, P., 2002. Exploring Offence Statistics in Stockholm City Using Spatial Analysis Tools. Annals of the Association of American Geographers, 92:1, 29— 51

[4] Boots, B. and Okabe, A., 2007. Local statistical spatial analysis: Inventory and prospect. International Journal of Geographical Information Science, 21:4, 355 — 375

[5] Laffan, S. W., 2002. Using process models to improve spatial analysis. International Journal of Geographical Information Science, 16:3, 245 — 257

[6] Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: From equations to case studies. Computers Geosciences, 33 (2007) 1301-1315

[7] Viger, R.J., 2008. The GIS Weasel: An interface for the development of geographic information used in environmental simulation modeling. 34 (2008) 891–901

[8] Vienneau, D.; de Hoogh, K.; Briggs, D. 2009. A GIS-based method for modelling air pollution exposures across Europe. Science of the Total Environment. 408-2, p255-266.

[9] Bajat, B., Pejović, M., Luković, J., Manojlović, P., Ducić, V., Mustafić, S. 2013. Mapping average annual precipitation in Serbia (1961-1990) by using regression kriging. *Theoretical & Applied Climatology*. 112-2, pp1-13.

[10] Goodchild, M. F., Yuan, M., Cova, T. J., 2007. 'Towards a general theory of geographic representation in GIS. International Journal of Geographical Information Science, 21:3, 239 -260

[11] Soares, A., Pereira, M.J., 2007. Space–time modelling of air quality for environmental-risk maps: A case study in South Portugal. Computers Geosciences 33 (2007) 1327–1336.

[12] Hansen, D., Lauritsen, J M., 2010. Identification of black spots for traffic injury in road intersections dependence of injury definition. Injury Prevention Vol. 16, pp:261-267

[13] Gundogdu, I.B., "Geostatistical analyst usung the junction risk factor to analyse and prevent urban traffic accidents." *Latest Trends in Energy Environment and Development*. June 3-5, 2014. Salerno, Italy.

[14] Prasannakumar ,V., Vijith, H., Charutha, R., Geetha, N. 2011. Spatio-Temporal Clustering of Road Accidents: GIS Based Analysis and Assessment In International Conference: Spatial Thinking and Geographic Information Sciences 2011, Procedia - Social and Behavioral Sciences. 21:317-325

[15] Hill, L., Rybar, J., Baird, S., Concha-Garcia, S., Coimbra, R., Patrick, K. 2011. Road safe seniors: Screening for age-related driving disorders in inpatient and outpatient settings. Journal of Safety Research.42-3, pp165-169.

[16] Pan, F., Zhu, P., 2011. Design optimisation of vehicle roof structures: benefits of using multiple surrogates. International Journal of Crashworthiness, 16-1, pp85-95.

[17] Yu, R., Abdel-Aty, M. 2013. Investigating the different characteristics of weekday and weekend crashes. Journal of Safety Research. Vol. 46, pp91-97

[18] Gu, X., Sun, G., Li, G., Huang, X., Li, Y., Li, Q. 2013. Multiobjective optimization design for vehicle occupant restraint system under frontal impact. Structural & Multidisciplinary Optimization. 47-3, p465-477.

[19] Liu, X. H., Kyriakidis, P. C. and Goodchild, M. F. 2008. Population-density estimation using regression and area-to-point residual kriging. International Journal of Geographical Information Science, 22:4, 431 — 447

[20] Kumar, A., Maroju, S., and Bhat, A., 2007., Application of ArcGis geostatistical analyst for interpolating environmental data from observations. Wiley Interscience DOI 10.1002/eg. 10223

[21] Cetin, M., Tulucu, K., 1998. Dogu Akdeniz bolgesinde aylik yagislarin yersel degisimlerinin jeoistatistik yontemlerle incelenmesi. Tr. Journal of Engineering and Environmental Science 22, pp 279-288

[22] Ahmadi, S.H., Sedghamiz, A., 2007. Geostatistical analysis of spatial and temporal variations of groundwater level, Environ Monit Assess 129, pp:277-294

[23] Johnston, K., Ver Hoef, J.M., Krivoruchko, K., and Lukas, N., 2003. Using ArcGis Geostatistical Analyst, ESRI, 300 pp.

AUTHOR'S PROFILE

Dr. Ismail Bulent Gundogdu, 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 60 published paper,1 book, 2 CD about cartography, computer aided design and GIS. 2005-2006 Head of Cartography Department Division. Now, Head of Geodesy Department Division and Deputy Dean of Engineering Faculty. Interestings: Cartography, GIS, Geostatistical analyst. (http://www.bgundogdu.selcuk.edu.tr)