International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 3, September 2014)

# The Validity of Geostatistical Analyst in Determining the Risky Zones of Highways and Planning New Roads.

Ismail Bulent Gundogdu
Department of Geomatics Engineering, Faculty of Engineering,
Selcuk University.

***Abstract*-In urban planning, the development of the city for the future must be controlled. Therefore, in the city planning phase, it is necessary to plan for all events.**
**By applying Geostatistical Analyst (GA), this study differs from previous studies which have determined the hot spots using linear analysis. The aim is to map and determine graduated hot or safe zones using not only traffic accident numbers but also other traffic accident criteria (AC). AC1 is the number of mortalities criterion, AC2 is the number of injured people criterion, AC3 is number of accidents with damage only criterion.**
**First, all data were examined using statistical analysis. Second, probability and prediction maps were produced by kriging interpolation methods for future planning. The last aim is to illustrate maps with different properties in order to provide a guide for new routes or settlement planning.**
***Keywords*- Geostatistical Analyst; GIS; kriging; probability maps; prediction maps; traffic accidents**

## 1. Introduction

The paper is organised as follows: after the next section on the data and methodology, I set up and interpret the accident mortality numbers, numbers of injured, and numbers of accidents with damage only of distribution models. Then I analyse the implications of traffic accident data models for urban spatial structure. The last section is the discussion and results.

The Turkish Statistical Institute (TUIK) facts reveal that 1296634 traffic accidents, 2934 of which resulted in mortality (3750 deaths), 88042 of which resulted in injuries (268079 injuries) and 1143082 of which resulted in economical damage, occurred in Turkey in 2012.

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 [1]

Several precautionary procedures are conducted to prevent or decrease the destructive problem of traffic accidents to a minimum level. Plotting, which needs to be conducted by analyzing causes of accidents through a great amount of data, is one of the major precautionary measures that might prevent accidents. The most specified analysis and plotting can be carried out using GIS nowadays.

This study differs from others in that the mapping is carried out according to the different AC. So, in this phase, GA must be explained in detail. No studies of traffic accidents that applied GA were encountered. However, some studies discussed similar problems: transportation planning, discussed in an editorial article [2], the relationship between noise pollution and traffic [3], risk reduction in urban planning [4], [5], [6], [7].

## 2. Using of Kriging Enterpolation Technique

Kriging interpolation techniques have proved to be popular in many areas such as agriculture, mining, geology, environmental science, building, cartography, risk management, and so on. For example, they have been used for modelling the spatial variability of tropical rainforest soils [8], soil mapping [9], [10], modelling the spatial distribution of human diseases [11], water table mapping [12], [13], mapping the abundance of fish in the ocean [14], rainfall mapping [15], and for detailed mathematical approaches [16]. There are detailed studies of the spatial distribution of measured variables (e.g., [17], [18], [19], [20], [21], [22], [23].

Calculating the variance is the most important and distinctive peculiarity of kriging in comparison with other techniques. The value is the criterion for the reliability of the estimated value. If the calculated variance is smaller than the variance of certain values, the estimated value is reliable for the unsampled point or area. [24], [25], [21] state that kriging is better than other interpolation techniques using obtained geostatistical results.

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 $\mu(s)$ and a random auto correlated error $\varepsilon(s)$. $s$ indicates the location of a point.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 3, September 2014)

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 $u_a$. [26].

The variogram is the function which characterises the dependence of variables between two different points in space.

## 3. Materials and Method

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.

Konya is not only the largest province of Turkey in the context 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 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 in 2012.

Accident records for last ten years have been obtained from Konya Police Department and input into the database using MS Access to set up the GIS-based application. To define accident points, digitisation and analyses were performed using maps with a scale of 1/50 000

Figure 1 shows 7 digitised routes on the satellite image. In this process, besides all AC, dates, times, and types of vehicle data were recorded in the main dataset. During this period, 3841 accidents with damage only, 4512 injured persons, and 374 fatalities were registered.

To test all datasets for mapping, ArcGIS version 9.2 was used.

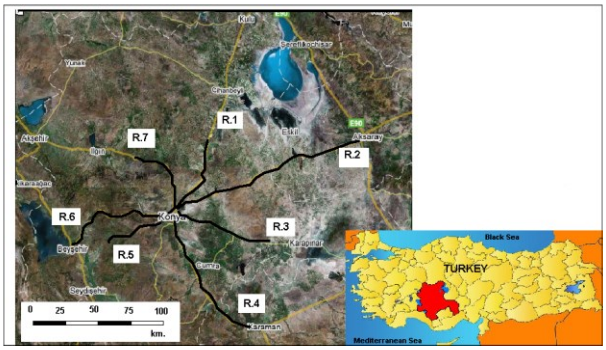



Figure 1: Location of related highways (7 routes)

## 4. Methodology

First of all, the characteristics of all data must be evaluated. As mentioned above, during this phase all analyses were carried out for all AC separately. Table 1 shows the histogram values for all AC, and Figure 2 shows a QQPlot graphic for AC3 as an example.

Table 1: Histogram values calculated by Accident Criteria.

| AC | Min | Max | Mean | Std. Dev. | Skewness | Kurtosis | 1st Quantile | Median | 3rd Quantile |
|---|---|---|---|---|---|---|---|---|---|
| **AC1** | 0 | 10 | 0.827 | 1.551 | 2.644 | 11.211 | 0 | 0 | **1** |
| **AC2** | 0 | 61 | 10.669 | 9.526 | 1.611 | 6.538 | 4 | 9 | **15** |
| **AC3** | **0** | **46** | **8.846** | **6.474** | **1.333** | **5.968** | **4** | **8** | **12** |

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 realised with a skewness value close to zero. According to the values in Table 1, our data are normally distributed. However, a normal QQPlot also allows us to compare the data with a normal distribution. Figure 2 shows that the data have a normal distribution with a few extreme points. Because of the normal distribution of the data, it is not necessary any transformation.

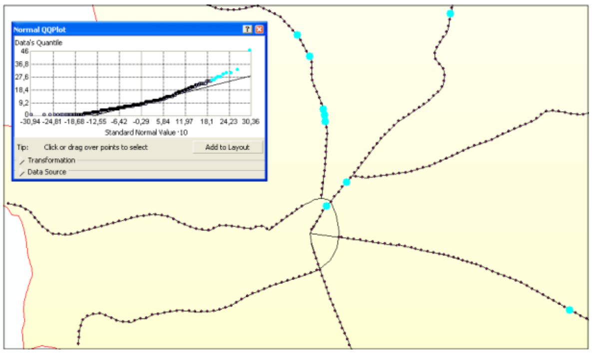
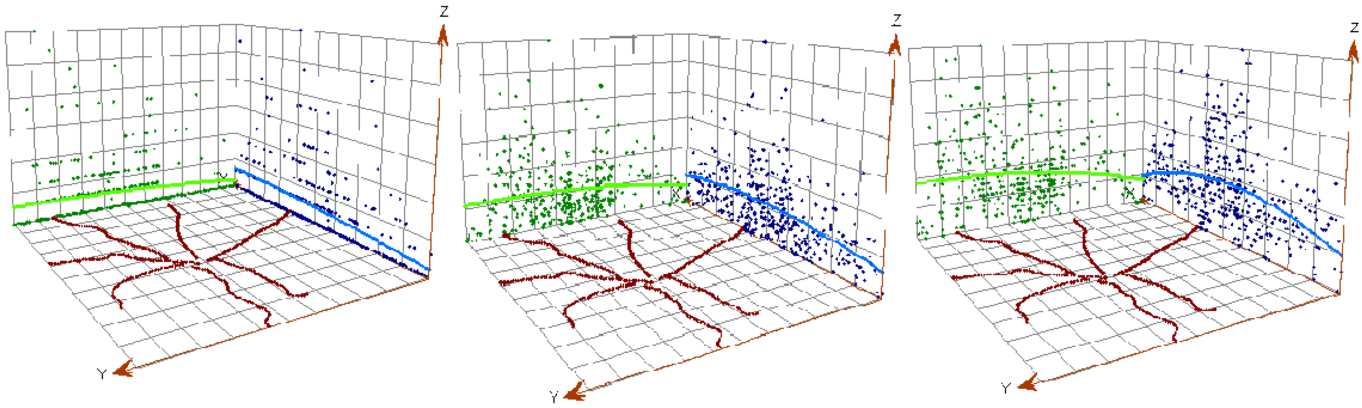


Figure 2: Example of a QQPlot graphic for AC3.

This stage investigates whether or not there is a trend. According to Figure 3, there are no trends in AC1, AC2, or AC3. If a trend has an upside-down "U" shape, this shows that there is a global trend. This trend could be removed using the second interpolation option.



(a) (b) (c)

Figure 3: Trend analysis graphics for AC1 (a), AC2 (b), and AC3 (c).

According to [27], 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 values constituted with all AC. All values which are important to show AC and to plan for the future are in shown the table. To carry out mapping with the help of the Table2, the rules must first be considered primarily as below.

The best map can be produced with the trial and error method by changing parameters.

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

Figure 4 shows prediction, probability, and indicator maps in sequence. The geometric interval method has been used to provide more detailed representation. The algorithm of this method was specifically designed to accommodate continuous data. It produces a result that is visually appealing and cartographically comprehensive. The threshold value has been accepted as the mean of point values in this study. But the value can be changed for different aims or requirements. Probability maps have been produced by exceeding the threshold value. Indicator kriging maps have been produced by not exceeding the threshold value.

This is the most important type of map to interpret for the prevention of traffic accidents. Figures 4(g), (h), (i) have been produced using covariance values.

Table 2: All calculated statistical values necessary for interpretations.

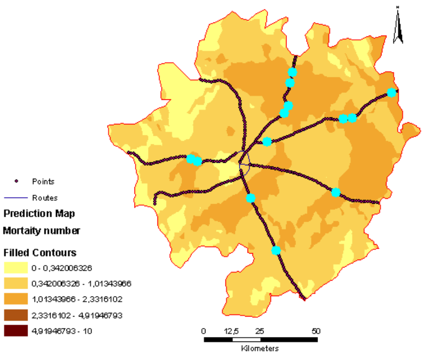
| Accident criteria / Statistical values | Mortality Number (AC1) | Injured Number (AC2) | Accident with damage only (AC3) |
|---|---|---|---|
| **Trend** | - | - | - |
| **Anisotropy** | - | - | - |
| **Major range** | 94068.6 | 25684.0 | 41408.8 |
| **Minor range** | 72870.5 | 14201.2 | 22979.9 |
| **Nugget** | 2.147 | 65.811 | 26.245 |
| **Sill** | 0.390 | 39.242 | 22.569 |
| **Ordinary kriging results** | | | |
| **Mean** | -0.005733 | 0.071280 | 0.072350 |
| **Root-Mean-Square** | 1.545 | 8.927 | 5.069 |
| **Average Standard Error** | 1.489 | 8.841 | 5.559 |
| **Mean Standardized** | -0.004072 | 0.007698 | 0.01269 |
| **Root-Mean-** | 1.037 | 1.012 | 0.914 |

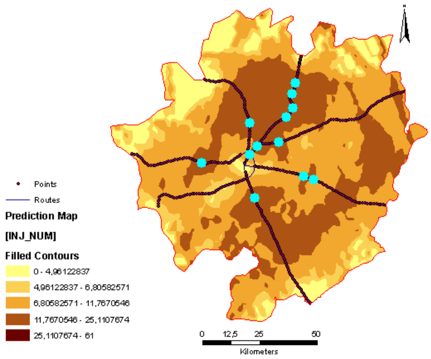
International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 3, September 2014)

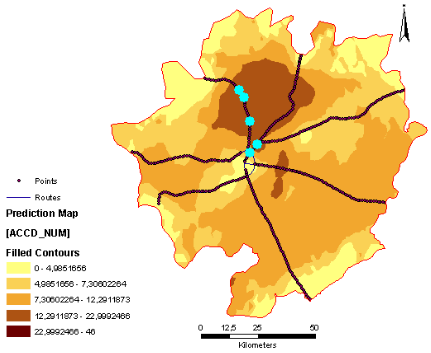
| | | | |
|---|---|---|---|
| Square Standardized | | | |
| Indicator kriging results | | | |
| Mean | -0.003741 | 0.002228 | 0.001666 |
| Root-Mean-Square | 0.4894 | 0.4686 | 0.4175 |
| Average Standard Error | 0.4923 | 0.4703 | 0.4145 |
| Mean Standardized | -0.007875 | 0.004029 | 0.003278 |
| Root-Mean-Square Standardized | 0.994 | 0.997 | 1.008 |

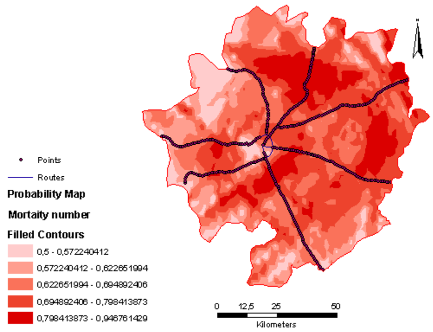
For the correct interpretation of the values shown in the table, the following rules must be taken into consideration.

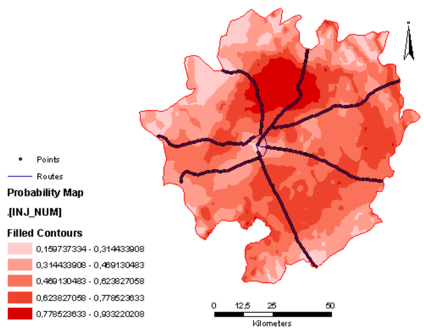
Generally, the best model is the one that has the standardized mean nearest to zero, the smallest root-mean-square prediction error, the average standard error nearest the root-mean-square prediction error, and the standardized root-mean-square prediction error nearest the one [28]. Since the models were controlled by the numerical values, it was not necessary to map the standard errors.

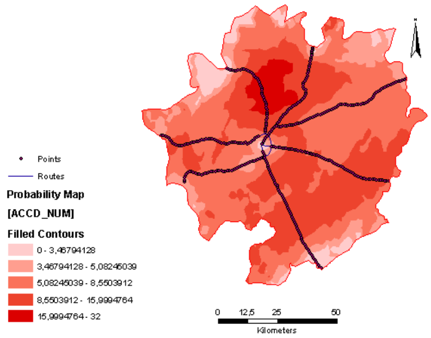


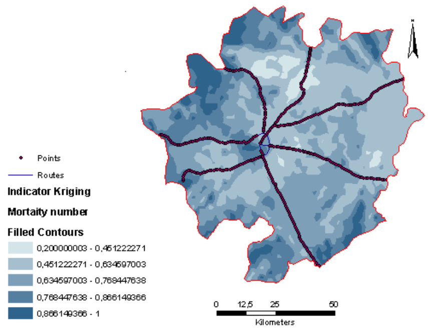



International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 3, September 2014)

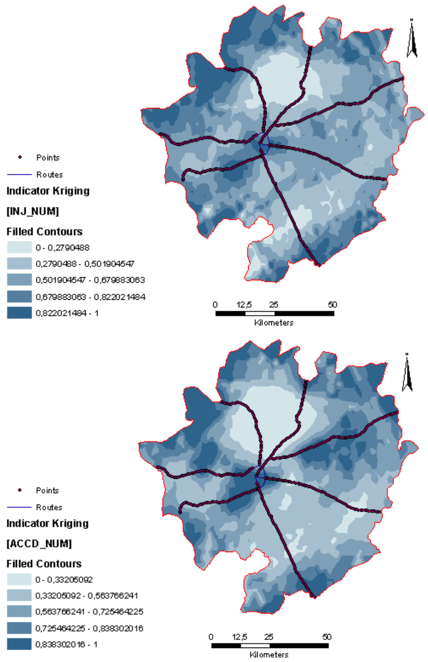



Figure 4: Result maps according to AC1, AC2, and AC3: (a), (b), (c) Prediction maps; (d), (e), (f) Probability maps; (g), (h), (i) Indicator kriging maps; for number of mortalities, number of injured persons, and number of accidents with damage only, respectively.

## 5. Discussion and Results

Since routes have been evaluated as sections of 1 km clusters, the study attempted to combine the advantages of two-dimensional and linear clusters. Each point is at the centre of its own cluster.

Deterministic and geostatistical methods are examined with two groups in GA. These methods are basically similar as 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 values. Therefore, deterministic applications were not performed in this study. This study has tried to understand the characteristics of traffic accidents.

The study gives different results from others concerning GA applications for traffic accidents in the literature. In addition, it differs from other studies in that it uses not only the number of total accidents but also different traffic accident criteria.

When the QQPlot graphic is considered with all outlier AC values, there are problematic points. These problematic points are shown in a light colour in Figures 4(a), (b), and (c).

Table 3 shows the number and distribution of problematic points.

Table 3: Number and distribution of problematic points according to all AC, and risky kilometre sections for each route.

| | Route 1 | Route 2 | Route 3 | Route 4 | Route 5 | Route 6 | Route 7 | Total |
|---|---|---|---|---|---|---|---|---|
| **AC1** | 4 | 4 | 1 | 2 | - | 2 | - | **13** |
| **AC2** | 6 | 3 | 3 | 1 | - | 1 | 3 | **17** |
| **AC3** | 2 | - | - | - | - | - | 3 | **5** |
| **Total** | 12 | 7 | 4 | 3 | - | 3 | 6 | **35** |
| **Risky km values** | | | | | | | | |
| | **22 km–80 km** | **23 km–33 km** | **11 km–19 km** | **7 km–14 km** | - | **18 km–24 km** | **7 km–43 km** | |

According to Table 3, 35 out of a total of 422 points are risky points triggering accidents. As mentioned above, not only points but also all clusters must be evaluated. The start and end points of kilometre sections of hot clusters must be identified on each road. The evaluation of hot zones, which is the main aim of similar studies, can be carried out with the help of maps.

In the light of the values shown in Table 3, the most important factor is AC2 at the point of accidents. First AC2, then AC1, must be pursued. Route 1 needs improvement most urgently, followed by Routes 2 and 7.

Furthermore, the intersection kilometres for three ACs are given at the bottom of the Table 3. These values have been obtained from the maps shown in Figure 4. These are the most important values needed to solve the problems. In fact, when Route 1 is observed it is seen that it connects the city to the capital of the country, is straighter than other routes, and has high traffic speeds. Precautions about reducing the speed must be put on the agenda urgently.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 3, Issue 3, September 2014)

***References***

[1] Arslan, F., 2007. Emniyet Genel Müdür Yardımcısı. http://www.trafik.gov.tr/icerik/bildiriler/A6-87.doc

[2] Editorial., 2008. Urban transportation planning in transition. Transport Policy. 15, 66-72

[3] Seto, E.Y.W., Holt, A., Rivard, T., Bhatia, R., 2007. Spatial distribution of traffic induced nnoise exposures in a US city: an analytical tool for assessing the health impacts of urban planning decisions. International Journal of Health Geographics. 6:24.

[4] Wamsler, C., 2006. Main streaming risk reduction in urban planning and housing: a challenge for international aid organisations. Blackwell Publishing. USA.

[5] Stathopoulos, A., Dimitriou, L., 2008. Fuzzy modelling approach for combined forecasting of urban traffic flow. Computer-Aided Civil and Infrastructure Engineering. 23, 521-535.

[6] 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

[7] 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

[8] Yemefack, M., Rossiter, D.G., Njomgang, R., 2005. Multiscale characterization of soil variability within an agricultural landscape mosaic system in Cameroon. Geoderma 125 (1-2), 117-143

[9] Leopold, U., Heuvelink, G.B., Tiktak, A., Finke, P.A., Schoumans, O., 2005. Accounting for change of support in spatial accuracy assessment of modelled soil mineral phosphorous concentration. Geoderma 130 (3-4) 368-386

[10] Lopez, G, F., Jurado, E., M., P., Barragan, C., G., Torres, L., 2005. Using geostatistical and remote sensing approaches for mapping soil properties. European Journal of Agronomy. 23 (33), 279-289

[11] Pleydell, D.R.J., Rasul, F., Tourneux, F., Danson, F.M., Graham, A.J., Craig, P.S., Giraudoux, P., 2004. Modelling the spatial distribution of Echinococcus multilocularis infection in foxes. Acta tropica 91 (3), 253-265

[12] Desbarats, A.J., Logan, C.E., Hinton, M.J., Sharpe, D.R., 2002. On the kriging of the water table elevations using collateral information from a digital elevation model. Journal of Hydrology 255 (1-4), 25-38

[13] Finke, P.A., Brus, D.J., Bierkens, M.F.P., Hoogland, J., Knotters, M., de Vries, F., 2004. Mapping groundwater dynamics using multiple sources of exhaustive high resolution data. Geoderma 123 (1-2) 23-39

[14] Rivoirard, J., 2002. On the structural link between variables in kriging with external drift. Mathematical Geology 34 (7) 797-808

[15] Lloyd, C.D., 2005. Assessing the effect of integrating elevation data info the estimation of monthly precipitation in Great Britain. Journal of Hydrology 308 (1-4), 128-150

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

[17] Shinn, N.J., Bing-Canar, J., Cailas, M., Peneff, N., & Binns, H., 2000. Determination of spatial continuity of soil lead levels in an urban residential neighborhood. Environmental Research Section A, 82, 46–52.

[18] Boruvka, L., & Kozak, J., 2001. Geostatistical investigation of a reclaimed dumpsite soil with emphasis on aluminum. Soil and Tillage Research, 59, 115–126.

[19] Biester, H., Muller, G., & Schoer, H.F., 2002. Estimating distribution and retention of mercury in three different soils contaminated by emission from chlor-alkali plants: Part 1. The Science of the Total Environment, 284, 177–189.

[20] 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.

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

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

[23] 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.

[24] Pearce, JL; Rathbun, SL; Aguilar-Villalobos, M; Naeher, LP. 2009 Characterizing the spatiotemporal variability of PM2.5 in Cusco, Peru using kriging with external drift Atmospheric Environment, 43; 12; p2060-p2069

[25] Laguardia, G., 2011 Representing the precipitation regime by means of Fourier series. International Journal of Climatology. 31-9, p1398-1407

[26] 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

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

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