CRASH PREDICTION: EVALUATION OF EMPIRICAL BAYES AND KRIGING METHODS

Uday R. R. Manepalli
Civil, Architectural and Environmental Engineering,
Missouri University of Science & Technology
Rolla, MO, USA, E-mail: umr75@mail.mst.edu

Ghulam H. Bham
Corresponding Author
Civil, Architectural and Environmental Engineering,
Missouri University of Science & Technology
Rolla, MO, USA, E-mail: ghbahm@mst.edu

Submitted to the 3rd International Conference on Road Safety and Simulation,

September 14-16, 2011, Indianapolis, USA

No. of words in abstract =228

No. of words in text = 3880 + 1250 (4 tables and 1 Figure) = 5130

ABSTRACT

Crash frequency prediction plays an important role in traffic safety for providing precautionary measures to reduce severity of the crashes and investment decisions. The Highway Safety Manual negative binomial regression to estimate safety performance functions and when crash history is available uses the Empirical Bayes (EB) method to predict crash frequency. Recent studies have used Kriging methods to predict AADT. This paper explores the use of Kriging method to predict crash frequency. Crash severity is derived from crash frequency and literature review indicated use of different weights for calculation of crash severity index. The Kriging and EB methods are compared in predicting crash frequency and crash severity index subject to sensitivity of weights over time and space. Crash data for I-630 in Arkansas were chosen for the same. The best method for prediction of crash frequency and crash severity index is recommended for use based on crash history, during the three years analysis period. The Kriging methods performed better than the EB method for medium term (3 years) prediction crashes. However, both methods over estimated the crash frequency when medium term crash data were used. When crash frequency was considered both Kriging and EB methods performed similarly for short term crash prediction (1 year). When crash severity was considered, Kriging performed better than the EB method when crash severity weights according to the Highway Safety Manual were used.

INTRODUCTION

The estimated highway crash cost to the society in 2000 was a staggering $230 billion a year according to NHTSA (2006). This cost may increase further every year. It is important to identify and predict the crashes accurately. Once predicted, other factors like roadway factors, weather conditions, engineering factors, and driver behavior can be studied at these locations for current conditions and necessary measures can be implemented.

Several methods like Box-Jenkins, neural networks, nonparametric regression, Gaussian maximum likelihood, time series analysis, etc. are used extensively for prediction (forecasting) in various fields. When prediction of crashes is considered empirical Bayes (EB), crash prediction models like negative binomial models, etc. have been used. Crash prediction is based on existing data at several locations. There are many factors which influence the crash frequency, which makes crash prediction stochastic. Incorporating these factors leads to accurate crash prediction, however, it becomes more complicated and the analyst should require special training and skills to use these techniques. With the availability of geographic information systems (GIS) and evolution of spatial analysis techniques, researchers have started to explore methods that exploit the spatial content of the data. Kriging has been used in the past as a technique for spatial interpolation. Kriging presumes autocorrelation in error terms of unobserved factors, as a function of distance. Kriging eclipsed traditional methods for predicting annual average daily traffic (AADT). Recent research (Eom et al., 2006, Wang and Kockelman, 2009, and Selby and Kockelman, 2011) identified that the error rate in predicting AADT values were very low when Kriging methods were applied. Additionally they suggested that the implications can be extended in predicting crash rates. It would be beneficial if the crashes were predicted with lower error rates. Kriging methods can be applied when the location of the crash is known. Therefore, the main objective of this paper is to predict crashes using Kriging as an extrapolation technique and compare the results with the EB method.

In this paper, three years (2000-2002) of crash data from I-630 highway of Arkansas were used. The following sections describe the existing literature on related topics, the data and the method used, and the results of analysis are presented last.

LITERATURE REVIEW

The Empirical Bayes (EB) method was first introduced by Hauer et. al. (2002). A detailed and systematic approach of the EB procedure was presented. It also included the prediction of crashes in the future based on available data from recent years. Traditionally, crash prediction models have used negative binomial regression to assess highway safety based on crash counts and crash rates (Shankar et. al., 1995; Poch and Mannering, 1996; Abdel-Aty and Radwan, 2000; Savolainen and Tarko, 2005). Negative binomial regression models were preferred over linear regression and Poisson regression models as they are generalizations of Poisson models and address complex data that is over-dispersed and correlated. Similarly, other predictive models have been developed based on the limitations of the previous models. Some other studies recently predicted AADT using time series, neural networks, nonparametric regression, and Gaussian maximum likelihood (GML) techniques (Tang et. al., 2003, Lam et. al., 2006).

However, after the availability of GIS, researchers have started geo-coding the crash locations and consider the effects of unmeasured confounding variables such as spatial autocorrelation. Hotspot identification (Truong and Somenahalli, 2011), cluster analysis (McMohan, 2000), high spatial crash concentrations (Hakkert and Mahalel, 1978; Songchitruksa and Zeng, 2010; Depue, 2003; Norden et. al., 1956, McGuigan, 1991; Peled et. al. 1996), database queries (Levine et. al., 1995a; Levine et. al., 1995b; Affum and Taylor, 1995; Austin et. al., 1997; Kin and Levine, 1996; Miller, 1999) and other methods like kernel density estimation (Flahaut et. al., 2003; Pulugurtha et. al., 2007) and spatial autocorrelation (Pulugurtha et. al., 2007) have been extensively used in the area of highway safety. Eom et al., 2006, Wang and Kockelman, 2009, and Selby and Kockelman, 2011, used Kriging models to predict AADT for roadways. Results indicate that kriging predicted AADT values with less percentage of error and can also be implemented for other studies like crash rate, traffic speeds, etc. However, they used spatial interpolation methods for predicting the AADT values. Selby and Kockelman, 2011 due to the limitations of ArcGIS, developed Kriging models and coded them in MATLAB.

Based on the existing literature it is evident that there are several accurate methods that can be used for prediction. However, EB and Kriging (spatial extrapolation) are used in this paper for prediction of crashes on I-630 using three years of crash data. EB was chosen as it is one of the most popular techniques used in highway safety. Kriging extrapolation was chosen as it was not explored previously.

DATA

Three years (2000-2002) of crash data on I-630 interstate highway in Arkansas was used for predicting crashes for 2003 in this paper. I-630 is 7.4 miles in length and is located in one county i.e., Pulaski County. The average crash rate of this highway for the analysis period is around 198 crashes per mile. The crash data was projected in ArcGIS using linear referencing. Table 1 presents the summary statistics of the crash frequency by year for I-630 from 2000 to 2003.Some of the major contributing factors are presented in terms of the percentage of crashes. From Table 1, for column 4, the remaining percentage of crashes represents straight roadway alignment i.e., for year 2000 among 500 crashes, 15% occurred on curve and 85% occurred on straight roadway profile. Similarly for column 5, the other factors include level roadway profile and unknown.

Table 1. Summary of crash statistics

Year (1)	CF (2)	AADT (vehs/yr) (3)	Curve* (4)	Grade* (5)	Types of Collisions* (6)			Weekends* (7)	DUI-Yes* (8)
					Rear-end (a)	Sideswipe Same Direction(b)	Single Vehicle Crashes (c)		
2000	500	90563	15	19	50	10	23	19	4
2001	523	93073	13	31	55	14	18	15	9
2002	443	93528	12	27	49	16	23	12	5
2003	537	97250	10	20	60	14	19	15	5

METHODOLOGY

Kriging was chosen to spatially extrapolate the crashes based on the existing data. The analysis was carried out in two stages: a) each year's crash data were used to predict the next year i.e., 2000 year crash data was used to predict 2001 crashes. Similarly, 2001 crash data was used to predict 2002; and so on, and b) aggregated three year crash data (2000-2002) was used to predict crashes for 2003. The error rates were compared in both cases. Both cases were chosen to determine the accuracy of the prediction using short term and medium term data. This process was also implemented by EB prediction methods and the results were compared. To perform a detailed study the prediction was carried out for crash frequency and crash severity. A crash severity index was generated using crash severity weights proposed by the Highway Safety Manual (HSM) (2010), Blincoe et. al. (2002), and Geurts et. al. (2004). The details of kriging and EB methods, crash severity index, and the goodness of fit are described next in detail.

Kriging

The method of kriging was first developed by Matheron (1963) based on the work of Krige (1951) to predict ore reserves. After several decades kriging has been applied in air quality analysis (Bayraktar and Turalioglu, 2005), natural resource analysis (Emerson, 2005), etc. The major application of this technique is to predict values at unmeasured locations while assessing the errors of these predictions (Wang and Kockelman, 2009). They rely on the notion that unobserved factors are autocorrelated over space, and the levels of autocorrelation deceases with distance. A trend estimate, μ(s), is determined which can be defined as (Wang and Kockelman, 2009):

$$Z_i\,(s) = \mu_i\,(s) + \varepsilon_i\,(s) \qquad (1)$$

where: $Z_i\,(s)$ is the variable of interest (crash frequency or crash severity index) and s gives the location of the site 'i'. $Z_i\,(s)$ is composed of a deterministic trend $\mu_i\,(s)$ and a random error term $\varepsilon_i\,(s)$. These random errors are autocorrelated over space. Features of trend or the expected value of *Z(s)* results in three types of kriging namely: Ordinary, Simple and Universal Kriging. However, in this paper universal kriging was preferred to other kriging methods as the trends depend on explanatory variables and unknown regression coefficients. ArcGIS "Geostatistical Analyst" (ESRI 1996) was used to fit and then apply the universal kriging. The correlation between Z(s) and Z(s + h) does not depend on actual locations, but only distance 'h' between two sites. This is possible by assuming weak stationary in all the three cases. This indicates a constant variance of 2γ(h) for any s and h, where γ(h) can be expressed as:.

$$\gamma(h) = \frac{1}{2}\, var\,[Z(s+h) - Z\,(s)\,] \tag{2}$$

where: var [Z (s + h) – Z(s)] is the variance between counts s and s + h. One of the major steps is to select an appropriate curve or semivariograms model that best fits the relationship between γ and h. When 2γ(h) is plotted (along y-axis) versus distance, is called semivariograms. There are three models that best explains the relationship i.e., exponential, spherical and Gaussian. In this paper only spherical model was chosen and the specifications are as follows:

$$\gamma(h)= \begin{cases} c_0 + c_1[1.5\frac{h}{a} - 0.5(\frac{h}{a})^3] & \text{if } 0 < h < a \\ c_{0} + c_1 & \text{if } h>a \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The models spherical, exponential and Gaussian rely on parameters that describe their shape while quantifying the level of spatial autocorrelation in the data. c_0 is called the nugget effect, and reflects discontinuity in the variograms origin, as caused by factors such as sampling error and short scale variability; a is called range, and this determines the threshold distance at which γ(h) stabilizes (Wang and Kockelman, 2009). $c_0 + c_1$ is the maximum γ(h) value, called sill, and c_1 is referred as partial sill (Cressie, 1993). Figure 1 illustrates the semivariogram.

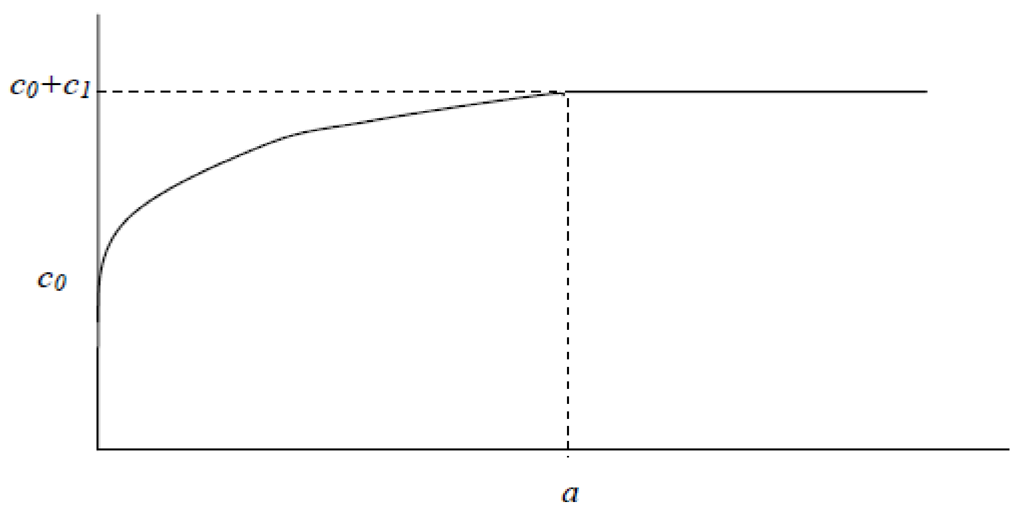

Figure 1 Illustration of semivariogram (Wang and Kockelman, 2009)

The EB method is described in detail by Hauer et. al., 2002 and is not presented in this section. However, the safety performance function, AADT values and the crashes (crash frequency or crash severity index) are required for crash prediction using the EB method.

Crash Severity Index (CSI)

A high CSI indicates a large number of fatal crashes or crash frequency of various levels of severity. The CSI was computed as:

$$CSI = S1*W1 + S2*W2 + S3*W3 + S4*W4 + S5*W5 \tag{4}$$

where:

- $S1$ = frequency of crashes involving fatalities,
- $S2$ = frequency of crashes involving incapacitating injuries,
- $S3$ = frequency of crashes involving moderate injury,
- $S4$ = frequency of crashes involving complaint of pain,
- $S5$ = frequency of crashes involving property-damage-only (PDO), and
- WI = weights assigned to a given crash severity level.

The weights used are based on the comprehensive crash costs per person for each type of crash. The costs are $4,008,900 for a fatality; $216,000 for major injury; $79,000 for a minor injury; $44,900 for complain of pain, and $7,400 for property damage only (HSM, 2010). The weights represent the ratios of comprehensive crash costs to the cost of PDO crashes. For example, the weight of a fatal crash is calculated as the cost of such a crash ($4,008,900) divided by the cost of a property-damage-only crash ($7,900) and is thus equal to 542. Other weights were computed similarly and rounded to the nearest zero. They are 29, 11, 6 and 1 for major injury, minor injury; complain of pain and property damage only, respectively. Similarly, Blincoe et. al., 2002 suggested the weights of 1330:1 for severe vs non-severe crashes. Geurts et. al., 2003 suggested the weights of 5, 4, 3, 2, and 1 for fatal to PDO crashes. Table 2 presents the crash severity weights used by various studies.

Table 2 Crash Severity Weights

Severity Type	Severity Weights		
	HSM, 2010	Blincoe et. al., 2002	Geurts et. al., 2003
Fatal	542	1330	5
Major Injury	29	1330	4
Minor Injury	11	1	3
Complain of Pain	6	1	2
Property Damage Only	1	1	1
Variance	56345	52987	3
Average	118	533	3

Goodness of fit

In order to compare and validate the two methods, kriging and EB, the following error ratios were determined:

$$Error_i = \frac{C_{est,i} - C_i}{C_i} \tag{5}$$

where:

$C_{est,\,i}$ =estimated crash frequency or crash severity index at the i^{th} location

C_i =the actual (true) crash frequency or crash severity index at the i^{th} location

The prediction method with the minimum the rate of error is the best method. For better presentation of results, the median and the average error rates for the entire highway are presented in this paper.

The error rates were computed based on Equation (5) i.e., for instance if there are three section length of 0.1 miles each which had 48, 34 and 89 actual or true crashes for an year; the predicted crashes using some method (Kriging or EB) estimates it to be 67, 28 and 120 respectively. The error rates for each section are using equation (5) would be 0.3958, -0.1765, and 0.3483. The median and the average of the error rates are computed based on these three sections. The method which has lower median and average error rates is considered to be better than the other, which is used in this paper.

RESULTS

The analytical results for kriging and the EB method are presented in this section. Four cases in each method are compared and presented. The comparison is carried out based on the overall median and mean error rates for each method.

Table 3 presents the parameter estimates for the spherical function of the semivariogram for the year 2000. Different semivariogram specifications were estimated and compared for each crash frequency and crash severity models.

Similarly, for each year i.e., 2001 and 2002 the semivariogram parameter estimates were determined. This was carried out for the aggregated crash data from 2000-2002. From Table 3, the results indicate that the nugget values increased from CF, to CSI (5), CSI (542) and CSI (1330). This is due to the variance in the weights which can be observed from Table 2. As variance increases, the nugget values and sill values also increase. This was also noticed, when the aggregated data from 2000-2002 was used. Nugget values were ‘zero’ for the CF and CSI (5) models in these cases, indicating a stabilized function.

Table 3 Semivariogram parameter estimates for Spherical function

	Nugget (c_0)	Sill (c_0+c_1)	Range (a)
2000			
Crash Frequency	80.1	177.5	5497.5
CSI (542)[@]	2483.9	2767.7	5822.6
CSI (5)[#]	221.72	317.99	5278.1
CSI (1330)[*]	2493800	2898200	10241

@CSI (542) indicates weights suggested by HSM, 2010

#CSI (5) indicates weights suggested by Geurts et. al., 2003

*CSI (1330) indicates weights suggested by Blincoe et. al., 2002

It can be observed that as the crashes are assigned to severity weights the nugget and sill values increases which indicates that it is more distance dependent, as the crashes now are not equal i.e., when kriging is performed the effect of the crash near to another is not same as they are unequal in magnitude. This can be seen directly from the results presented in Table 3.However, it should also be noted that difference in magnitude of the weights had a significant impact i.e., the difference between S1 (fatal) when weights suggested by HSM, 2010 are used is 542 times property damage whereas its 1330 times property damage when weights proposed by Blincoe are used. Therefore, it can be stated that when crashes are treated with severities and weights are assigned to them spatial autocorrelation is more distance dependent. Also, the variance increases with distance. From Table 3 it can be inferred that in most cases the semivariogram flattens in the range within 5200 (ft) and 5900 (ft), which indicates that the effect of variance between two crashes stabilizes with the specified range value is nearly 1 mile. This might be higher or lower depending on the data set used in terms of highway length, distance between crashes and the crash rates. In this paper, the highway was limited to 7.4 miles hence these distances are acceptable.

Table 4 presents the summary of results of the entire highway of I-630. The error rates for median and average were compared to one another i.e., relative to other method. It can be noticed that for most cases except for 2003(b) the prediction error was less than the actual values (true values) for both kriging and EB methods. The model results for 2003(b) indicated higher prediction error than the actual values. This indicates that short term data for prediction have resulted in lower values than the actual values and mid-term data for prediction have resulted in higher values than the actual values. Based on the four cases, Kriging method performed better than EB for CSI (542) and CF. EB method predicted than kriging for CSI (5). Both performed similar when CSI (1330) was used. In three of the four cases presented, kriging predicted better than EB, for mid-term data i.e., prediction of 2003 crashes based on 2000-2002 aggregated crash data. The effects of the different weights used for different crash severity levels can be observed in the model form of the kriging method. Overall, results indicate that kriging and EB methods can be used for prediction of crashes.

Table 4 Summary of Results: Error Rates

	CF		CSI (542)@	
	Median	Average	Median	Average
Kr 2001	0.10	0.15	0.08	0.06
EB 2001	0.01	-0.02	0.20	0.24
Kr 2002	0.11	0.23	0.13	0.39
EB 2002	0.18	0.31	0.20	0.45
Kr 2003 (a)	-0.02	0.01	-0.06	0.15
EB 2003 (a)	0.05	0.01	-0.36	0.02
Kr 2003(b)	2.17	2.04	2.26	2.43
EB 2003 (b)	2.11	2.21	1.96	2.51
	CSI (5)#		CSI (1330)*	
	Median	Average	Median	Average
Kr 2001	-0.09	-0.13	0.51	17.78
EB 2001	0.01	0.08	34.32	45.05
Kr 2002	0.15	1.23	12.73	31.87
EB 2002	0.14	0.37	3.90	31.42
Kr 2003 (a)	-0.10	-0.04	-0.45	6.40
EB 2003(a)	-0.06	-0.04	0.00	5.99
Kr 2003(b)	2.19	2.51	9.13	42.09
EB 2003 (b)	2.03	2.11	72.97	84.37

Kr represents Kriging

@CSI (542) indicates weights suggested by HSM, 2010

#CSI (5) indicates weights suggested by Geurts et. al., 2003

*CSI (1330) indicates weights suggested by Blincoe et. al., 2002

2003 (a) represents prediction for 2003 based on 2002

2003 (b) represents prediction for 2003 based on 2000-2002

Shading represents the best among the two methods i.e., Kr (orange) and EB (grey)

CONCLUSIONS

This paper proposes a straightforward process that can be used to predict crashes using ArcGIS's geospatial analyst tools i.e., kriging. The results were compared to the EB method using crash data from I-630 in Arkansas. Three years of crash data were used for the same. Results from kriging and the EB methods indicate over estimation of crashes when mid-term data i.e., aggregated crash data were used. However, this over estimation to some extent can be considered as it is a known fact that under-reporting of crash data often occurs (Elvik and Myssen, 1999, Blincoe et. al., 2002; Hauer and Hakkert, 1989) especially for non-severe crashes. Results from this paper indicate that kriging methods can be used for crash predictions, however, if the crashes are geo-coded and projected on ArcGIS; the computation becomes easy and straightforward. In this paper, only spherical function was used, however, the use of exponential and Gaussian functions and there effects in crash predictions should be explored. The predicted values can be used for further analysis in identifying factors that affect crashes, and investment decisions can be made for providing counter measures.

ACKNOWLEDGEMENTS

The authors are grateful for the data provided by the Arkansas Highway and Transportation Department (AHTD) and three anonymous reviewers for providing useful comments.

REFERENCES

Abdel-Aty, M., and Radwan, A. E. (2000). Modeling traffic accident occurrence and involvement. Accident Analysis and Prevention, Vol. 32, No. 5, pp. 633-42.

Affum, J. K., and Taylor, A. P. (1995). Technology Tools for Transportation Professionals—Moving Into the 21st Century. Resource Papers for the 1995 International Conference, Integrated GIS Database for Road Safety Management. Institute of Transportation Engineers, Washington, D.C., pp. 189–193.

Austin, K., Tight, M., and Kirby, H. (1997). The Use of Geographical Information Systems to Enhance Road Safety Analysis. Transportation Planning and Technology, Vol. 20, No. 4, pp. 249–266.

Bayraktar, H. And Turalioglu, F .S. (2005). A Kriging based Approach for Locating A Sampling Site- In the assessment of Air Quality. Stochastic Environment Research and Risk Assessment, Vol. 19, No. 4, pp. 301-305.

Blincoe, L., Seay, A., Zaloshnja, E., Miller, T., Romano, E., Luchter, S., and Spicer, R. (2002). The economic impact of motor vehicle crashes, NHTSA Technical Report, Washington, DC.

Cressie, N. (1993). Statistics for Spatial Data, Wiley Interscience: New York.

Depue, L. (2003). Safety Management Systems: A Synthesis of Highway Practice. NCHRP Synthesis 322, Transportation Research Board of the National Academies, Washington, D.C.

Elvik, R., and Myssen, A. (1999). Incomplete accident reporting: meta-analysis of studies made in 13 countries. In Transportation Research Record: Journal of the Transportation Research Board, No. 1665, TRB, Washington, D.C., pp. 133-140.

Emerson, X. (2005). Simple and ordinary Kriging Multigaussian kriging for estimating recoverable reserves. Mathematical Geology, Vol. 37, No. 3, pp. 295-319.

Eom, J. K., Park, M. S., Heo, T. Y., and Huntsinger, L. F. (2006). Improving the Prediction of Annual Average Daily Traffic for Nonfreeway Facilities by Applying a Spatial Statistical Method. Artificial Intelligence and Advanced Computing Applications, No. 1968, pp. 20-29.

ESRI, 1996. ArcView Geostatistical Analyst. Environmental Systems Research Institute.

Flahaut, B., Mouchart, M., Martin, E, S., and Thomas, I. (2003). The local spatial autocorrelation and the kernel method for identifying black zones A comparative approach. Accident Analysis and Prevention, Vol. 35, pp. 991-1004.

Geurts, K., Wets, G., Brijs, T., and Vanhoof, K. (2004). Identification and Ranking of black spots: Sensitivity Analysis. In Transportation Research Record: Journal of the Transportation Research Board, No. 1897, TRB, Washington, D.C., pp. 34-42.

Hakkert, A. S., and Mahalel. D. (1978). Estimating the Number of Accidents at Intersections from a Knowledge of the Traffic flow on the Approaches. In Journal of Accident Analysis and Prevention, Vol. 10, pp. 69-79.

Hauer, E., and Hakkert, S. (1989). The extent and some implications of incomplete crash reporting. In Transportation Research Record: Journal of the Transportation Research Board, No. 1185, TRB, Washington, D.C., pp. 1-10.

Hauer, E., Harwood, D. W., Council, F. M., and Griffith, M.S. (2002). Estimating safety by the Empirical Bayes Method: A Tutorial. In Transportation Research Record: Journal of the Transportation Research Board, No. 1784, TRB, Washington, D.C., pp. 126-131.

HSM, Highway Safety Manual, (2010). Section 4-30 and Table 4A-1. First Edition.

http://www.esri.com/library/userconf/proc96/TO50/PAP005/P5.HTM, accessed November, 2009

Kim, K., and Levine, N. (1996). Using GIS to Improve Highway Safety. Computers, Environment, and Urban Systems, Vol. 20, No. 4, pp. 289–302.

Krige, D. G. (1951). A statistical approach to some mine valuations and allied problems at the Witwatersrand, Master's Thesis of the University of Witwatersrand.

Lam, W. H. K., Tang, Y. F., Chan, K. S., and Tam, M. L. (2006). Comparison of Four Modeling Techniques for Short-term AADT Forecasting in Hong Kong. Journal of Transportation Engineering, Vol. 129, No. 3, pp. 271-277.

Levine, N., Kim, K., and Nitz, L. (1995a). Spatial Analysis of Honolulu Motor Vehicle Crashes—Part I: Spatial Patterns. Accident Analysis and Prevention,Vol. 27, No. 5, pp. 663–674.

Levine, N., Kim, K., and Nitz, L. (1995b). Spatial Analysis of Honolulu Motor Vehicle Crashes—Part II: Zonal Predictions. Accident Analysis and Prevention, Vol. 27, No. 5, pp. 675–685.

Matheron, G. (1963). Principles of Geostatistics. Economic Geology, Vol. 58, pp. 1246-1266.

McGuigan, D. R. D. (1981). The Use of Relationships between Road Accidents and Traffic Flow in "Black-Spot" Identification. Traffic Engineering and Control, pp. 448–453.

McMahon, P. A Quantitative and Qualitative Analysis of the Factors Contributing to Collisions between Pedestrians and Vehicles along Roadway Segments. Masters Project, Department of City and Regional Planning, University of North Carolina at Chapel Hill, NC. 1999.

Miller, J. S. (1999). What Value May Geographic Information Systems Add to the Art of Identifying Crash Countermeasures? Report VTRC 99–R13.

NHTSA (2006). http://www-nrd.nhtsa.dot.gov/Pubs/810623.PDF, Accessed March. 10, 2010.

Norden, M., J. Orlansky and H. Jacobs. (1956). Application of Statistical Quality-Control Techniques to Analysis of Highway-Accident Data. Bulletin 117, HRB, National Research Council, Washington, D.C., 1956, pp. 17-31.

Peled, A., Haj-Yehia, B., Hakkert, A.S. (1996). ArcInfo Based Geographical Information System for Road Safety Analysis and Improvement.

Poch, M., and Mannering, F. L. (1996). Negative binomial analysis of intersection accident frequencies. Journal of Transportation Engineering, Vol. 122, No. 2, pp. 105-113.

Pulugurtha, S. S., Krishnakumar, V.K., and Nambisan, S.S. (2007). New methods to identify and rank high pedestrian crash zones: An illustration. In Journal of Accident Analysis and Prevention, Vol. 39, No. 4, pp. 800-811.

Savolainen, P. T., and Tarko, A. P. (2005). Safety Impacts at Intersections on Curved Segments. In Journal of Transportation Research Record, No. 1908, Transportation Research Board, Washington, D.C., pp. 130-140.

Selby, B., and Kockelman, K. M. (2011). Spatial prediction of AADT in unmeasured locations by Universal Kriging. Proceedings of Transportation Research Board 90th Annual Meeting, Washington D. C., 2011.

Shankar, V. N., Mannering, F. L., and Barfield, W. (1995). Effect of roadway geometrics and environmental conditions on rural accident frequencies. Accident Analysis and Prevention, Vol. 27, No. 3, pp. 371-390.

Songchitruksa, P., and Zeng, X. (2010). Getis-Ord Spatial Statistics for Identifying Hot Spots Using Incident Management Data. Proceedings of Transportation Research Board 89th Annual Meeting, Washington D. C.

Truong, L. T., and Somenahalli, S. V. C. (2011). Using GIS to Identify Pedestrian-Vehicle crash hot spots and unsafe bus stops. Journal of Public Transportation, Vol. 14, No. 1, pp. 99-114.

Wang, X., and Kockelman, K. M. (2009). Forecasting Network Data: Spatial Interpolation of Traffic Counts using Texas data. In Transportation Research Record: Journal of the Transportation Research Board, No. 2105, TRB, Washington, D.C., pp. 100-108.