Accident Black-Spot Validation using GIS

Rajiv Gupta, Senior Professor
[1]Manpreet Singh, Master student
Civil Engineering Department
Birla Institute of Technology and Science, Pilani- 333031

Abstract:

Recent reports by World Health Organization (WHO) indicate India has highest number of fatal accidents. Main concern, regarding road safety in a developing country, further amplifies in absence of proper framework to analyze the accident related data and produce solutions towards the same, using modern tools & techniques, in semi-urban or rural setting. This paper is an attempt to develop data analysis tools that can help prioritize the study of the most accident prone intersections amongst all the intersections in the study area, in order to produce feasible solution to prevent accident growth rate from surging.

Since the location of accidents is not very accurately known, therefore after mapping of the points, hot spot analysis tool is used to identify the significant hot spots based on spatial autocorrelation. This study shows that the roadway conditions play a major role in occurrence of accidents and therefore has to be sufficiently studied and improved upon and hence can be a prime indicator in improvement prioritization. The analytical process validates the results of hot spot analysis by comparing them with areas having significant accident rates; termed as black spots which require urgent attention corresponding to the relevant parameters like roadway conditions, design characteristics, surface conditions etc. In this study, the accident data from Chandigarh has been used. Geographical Information Software (Arc GIS 10.1) is used for accident mapping of the entire city.

About the Authors:

Rajiv Gupta

Rajiv Gupta is Senior Professor of Civil Engineering at BITS Pilani. He has completed his B.E., M.E and Ph.D. from BITS, Pilani. In last 27 years of teaching and research he has authored number of books and course development material, published more than 125 research papers in renowned journals and conferences. He is involved in a World Bank and DST sponsored projects to implement rainwater-harvesting schemes in 6 villages of Rajasthan. He has been working in area of Remote sensing and GIS from last 15 years.

Email: rajiv@pilani.bits-pilani.ac.in

Contact: 8741959971

Manpreet Singh (first author)

He is a Master student of BITS, Pilani in Civel Engineering department.

Email: manpreet_singh04@yahoo.com

Introduction

Road accidents are a major cause of loss all over the world. They are the cause of over 50 million injuries every year and sum up to over US$ 65 billion in the developing countries. These deaths are more in number than many other accidents due to different modes put together and the cost to the society is much more than the official aid received by the developing countries all over the world. In India, more than 100,000 lives are lost in road accidents every year while 400,000 people are injured. Sweden has set itself a target of Vision Zero by 2020; it might be difficult to achieve something similar in India in the near future on such a grand scale.Primary aim of any transportation infrastructure is to provide mobility. Accidents pose as threats to the improvement of the system, which need to be controlled in order to achieve the objective. Primarily in India, we need to do a lot more to reduce accident-prone locations and to help us understand the ways to improve the road safety scenario.

Mapping the accident locations on a map and visually identifying the clusters is very tedious and not an accurate solution. This is mainly because of the following three reasons:

- Inaccuracies: Mistakes in plotting the data points manually may lead to incorrect interpretation of the pattern
- Difficult to identify clusters in a location where accidents are very high
- Disregard for statistical and spatial analysis and a possibility of a personal bias

One of the continuing goals of conducting this study is to convey the feasibility of road safety analysis study, among the safety engineers and others in the system. However, due to the variety of implementations of GIS that exist within various organizations, developing models in highway safety analysis requires an understanding of the requirements of GIS, Data collection & digitization systems, and GIS-based highway safety analysis applications. Using GIS, one can merge accident related data with other factors such as traffic related data and represent the results in a user friendly way.

Many Studies have been conducted for predicting the crash locations and also providing solutions to reduce the number of accidents. Considering the grave situation it is very important to come up with some models for prediction of accident-prone location taking in to account the road surface and design characteristics accounting for the accident occurrence. In many developed countries GIS has been widely used for mapping the accident locations. Many academics and government bodies are working on producing new tools and improving the existing ones for road safety analysis.

A traffic accident has various different characteristics associated with it. For proper traffic accident analysis use of GIS technology has become an inevitable tool (Ghosh, Parida, & Uraon, 2004). One such example can is given in one of the reports prepared by Federal Highway Administration of US Department of Transportation (Smith). The document explains the methodologies and application of the safety analysis tools developed by them. The primary goal of this report is to discuss the ways of combining Geographic Information System (GIS) and Road Safety analysis. The report tries to reduce the gap between acceptance of the concept of safety analysis tools and their actual implementation by an organization. Several concepts, their background and ways of implementation have been discussed in the report.

As the differing situations, at different locations, imply varying level of importance of each data point in accident analysis, we need to prioritize the locations depending on various parameters. It is demonstrated in detail, in a work by (Mandloi & Gupta, 2003). Weights for each point have been determined by giving weights to several parameters affecting accident analysis. Based on the total value of all the parameters at any particular situation, weights for the data point at the location can be determined. The analysis includes the location of the areas where the accidents are found to repeat and then finding out the reasons and classifying them accordingly.

Sarkar & Malleshwari, 1995 have suggested a method for identifying accident-prone locations and rates them according to accident density (AD), and proportion of fatal accidents (PFA) but due to the non-availability of proper accident data in our country the analysis becomes a major problem.

According to (Mohan, 2009) road safety policies in India must concentrate on the following issues to reduce the incidence of road traffic injuries: pedestrians and other non-motorist in urban areas; pedestrians, other non-motorists, and slow moving vehicles on highways; motorcycles and small cars in urban areas; over-involvement of trucks and buses; night-time driving; and wrong way drivers on divided highways. There is an urgent need to upgrade police data collecting procedures so that necessary information is available for scientific analysis.

There may be various reasons for an accident, one such paper from (Mondal, 2011)] identifies rainfall as one of the reasons, It states Road crashes are a complex interaction of different parameters like road, vehicle, environment, human etc. Skidding of road vehicles is considered as one of the major causes of road accidents occurring all over the world. Skidding, caused by lack of tire to- road friction, is one of the most important single causes of traffic accidents. This paper aims to critically analyze the weather and wet road related crashes.

Various models maybe built to predict future accidents based on the present traffic data, accident data, and population or growth rate of vehicles. This helps in evolving a major policy to reduce the growth of personalized vehicles and simultaneously increase the use of public transport systems (Valli, 2005). In his paper he develops models by analyzing the road accident data at an all India level as well as for major metropolitan cities. The data for the 25-year period from 1977 to 2001 were analyzed to build models to understand the nature and extent of the causes of accidents using the concept of Smeed's formula and Andressen's equations. Desai, 2011 tried to develop another accident model. Ahmedabad city is taken as case study and Accident data related to hourly-classified traffic volume per lane extracted from classified traffic volume count survey of Ahmedabad city. A liner regression model is developed in this study exhibits satisfactory goodness-of fit and prediction success rate. Singh & Suman, 2012 selected a stretch of NH-77 from Hajipur to Muzaffarpur. The accidental data was collected for last eleven years, 2000-2010 from the Police Stations where FIR was lodged. The data collected were analyzed to evaluate the effect of parameters influencing the accident rate and a model was developed. And according to the model, accident prediction represents that the number of accidents per-km-year increases with AADT and decreases with improvement in road condition.

While these methods, as used in previous works, have proven to be efficient, they need to be improved in a way so that those can suit our requirements. Therefore an attempt has been made in this paper to provide an improved methodology trying to bridge the gaps found.

Study Methodology

In the places where the information related to accidents is not digitized in any way, implementing methods to use the data remains a problem thus resulting in lack of proper analysis. Such is the case in India and similarly in many other developing countries. In order to suggest a methodology for the safety analysis at such places, possible methods of implementation are used. Feasibility of their implementation needs to be checked and improvisation and customization be made accordingly. The methodology for accident mapping and analysis used can be briefed through the following points:

- Data acquisition: undertaken through police records and statistics available online
- Digitization: To digitize map available from net, so that it can be used in computer based analysis.
- Mapping: includes processes such as choosing the appropriate map as per the requirement in accident analysis.
- Reclassification: Preprocessing of the available data to use it for analysis
- Analysis: Hot Spot Analysis done one the Mapped Data.
- Inferences.

Hot spot

It is based on statistical values and most statistical tests are based on null hypothesis. The null hypothesis for the pattern analysis tools (Clustering done in hot spot analysis) is Complete Spatial Randomness (CSR), either of the features themselves or of the values associated with those features. The z-scores and p-values returned by the pattern analysis tools tell you whether you can reject that null hypothesis or not. For pattern analysis tools it is the probability that the observed pattern was created by some random process. Z score and P values are Statistical values. Z score is standard deviation and p value is the probability. Main concept in it is that the values in the middle of the normal distribution (z-scores for example), represent the expected outcome. When the value of the z-score is large and the probability is small (in the tails of the normal distribution), however, the results are somewhat unusual and generally very interesting.

For the Hot Spot Analysis tool, for example, "unusual" means either a statistically significant hot spot or a statistically significant cold spot. This tool calculates the resultant Z score and presents features with either high or low values cluster spatially. The Gi* statistic value given for each feature in the dataset is a Z score. For statistically significant positive Z scores, the larger the Z score is, the more intense is the clustering of high values (hot spot). For statistically significant negative Z scores, the smaller the Z score is, the more intense the clustering of low values (cold spot).

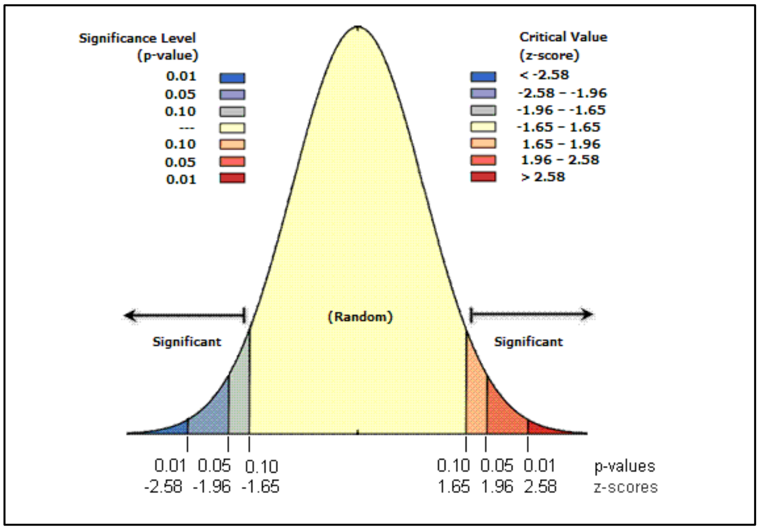

Figure 1 Distribution*

***Source (ESRI Website, 2013)**

Table 1 shows the confidence level for various standard deviation and probability values.

Table 1 Standard Deviations and Probability Confidence Range*

z-score (Standard Deviations)	p-value (Probability)	Confidence level
< -1.65 or > +1.65	< 0.10	90%
< -1.96 or > +1.96	< 0.05	95%
< -2.58 or > +2.58	< 0.01	99%

Black spot

So far, there is no assured definition of black spots. There are different approaches to data collation and variations in areas, investigating bodies differ in defining what constitutes a Black Spot. The biggest hurdle in identifying an accident black spot is to determine the locations with highest accident rate and the causes for these accidents. At certain sites, the level of risk will be higher than the general level of risk in surrounding areas. Crashes tend to concentrate at these relatively high-risk locations. Locations having an abnormally high number of crashes are described as crash hazardous, hot spot or black spot sites. In the study, black spots have been assumed to be the intersections having a high rank (attributed to roadway conditions).

Identification of Hot Spots in the data

This method considers spatial attributes (such as population density variation across the area, location of workplaces etc.) and calculates their interdependence. Local method of spatial autocorrelation is more useful in its application for finding statistically significant cluster of accident locations. This tool was used to calculate Geary's ratio (local) for all the data points. Inverse distance method and zone of indifference are used to get the spatial relationships. Such points can be easily displayed as shown in figure 4, where red colored locations are statistically significant cluster formations and light colored ones represent statistically insignificant cluster locations.After the georeferencing, digitization and preparation of attribute tables the following are the steps in Analysis –

1. Mark the accident location points.
2. Project the data points to the required coordinate system using Projections Tool.

3. Aggregate the incident data using Integrate and Collect Event Tool. On running this tool features within some specified default distance of each other snap to the same location thus creating a 'stack of the coincident features with same X Y coordinates.
4. Finally after collecting the events Hot Spot Analysis (Getis- Ord Gi*) (Spatial Statistics tool) is used with I Count as input field for calculations.
5. The final results obtained can be interpreted with basics of hot spots mentioned above.

Prioritization of Stretches using Black spots

While conducting this analysis the model only takes into account the road related factors such as road geometries and surface characteristics, which may lead to accidents. The factors considered for evaluating accident prone locations on road are as follows:

- Road width.
- Number of lanes in each direction.
- Approximate number of vehicles per day.
- Drainage facilities.
- Surface condition of the pavement.
- Presence of shoulders, edge obstructions and median barriers.
- Controlled, Uncontrolled Intersection.
- No of Conflict Points.
- Length of Road before the intersection

In order to model the above-mentioned factors and achieve the desired result, a step-by-step procedure as given below is adopted.

- Scan the map containing the desired road network and input this image to the Arc GIS software for digitizing.
- Digitize the road network making sure that every link is separated link and assign id number to every link.
- Specify the attributes for every road link using the questionnaire provided.
- Export the road attribute table generated in dbase format so that it can be imported by Arc view.
- Join the road attribute table to the digitized road map and prioritize the road network for accident occurrence using total weights assigned to every link.
- Accident black spots on a given road network are ranked by result obtained from prioritization.

Prioritization

In Prioritization suitable weights are assigned to different factors so as to achieve a desired result. In this model, the various factors, which may influence the occurrence of accidents on roads, are assigned weights on a scale of 0-10 in a way that the factors, which tend to increase the probability of the accidents have lower weights. In order to prioritize roads for occurrence of accidents, the various factors considered and the weights assigned to them are given in table 2:

Table 2: Factors Used with relative weights

	Factors Affecting Occurrence of Accidents	Possible Variations	Rank
1	No of Lanes in Each Direction	4	10
		3	8
		2	6
		1	4
2	No of Vehicles per day	Less than 1000	10
		Less than 2500	8
		Less than 5000	6
		Greater than 5000	4
3	Width of Road	More than 15m	10
		10.1-15m	8

		7.5-10.5	6
		6.1-7.5m	4
		Less than 6m	10
4	Presence of Shoulders	Yes	10
		No	4
5	Surface Condition of Road	Flexible	10
		Rigid	8
6	Drainage Condition	Good	8
		Satisfactory	6
		Poor	4
		No Drainage	2
7	Presence of Traffic Lights	Yes	10
		No	4
8	Provision of Median	Yes	10
		No	4
9	Roundabout	Yes	10
		No	4
10	Length of Road before junction	100m	10
		300m	8
		500m	6
		700m	4
		1000m	2
11	Conflict Points	24	10
		17	9
		16	8
		13	7
		12	6
		11	5
		7	4

The final weight assigned to each road link is obtained by adding all the individual weights and normalizing the value using maximum weight (in this case 110) that can be assigned. Hence,

$$\text{Total weight} = (\Sigma \text{ Individual Weights}) \times 100/110$$

Thus road links with high final weight are less prone to accidents than the road link with low final weight. The classification of roads for occurrence of accidents based on final weights is shown in table 3:

Table 3 Total Weights

Final Weight	Accident Prone Level
80-100	Very Low
60-80	Low
40-60	Medium
0-40	High

Weighted Average –

The various factors considered for affecting the occurrence of accidents may not have similar effect. Every factor will have a different level of involvement for an accident to take place. For example Presence of Shoulder and AADT cannot be given same weight age because more traffic may be a greater factor in occurrence of accident as compared to whether a shoulder is present on the side of a road or not. Similarly Drainage may combine the road surface condition as reason for an accident but length of road before the junction will always be a different parameter. Thus providing different weights to each of the factors and then using the weighted average sum to get the final ranks would provide a better scenario in deciding the more sensitive factors. The formula used for weighted sum is as follows –

Total score = ∑ (Normalized weight age assigned X Individual weight X100)/110

Weights may be given based on expert comments. Different weights assigned to the factors are given in Table 4 –

Table 4- Weights Assigned

	Factors Affecting Occurrence of Accidents	**Weight**
1	No of Lanes in Each Direction	.4
2	No of Vehicles per day	.5
3	Width of Road	.2
4	Presence of Shoulders	.1
5	Surface Condition of Road	.4
6	Drainage Condition	.6
7	Presence of Traffic Lights	.4
8	Provision of Median	.3
9	Roundabout	.5
10	Length of Road before junction	.6
11	Conflict Points	.3

Accident Analysis of the city of Chandigarh: A Case Study

To measure and use the advanced technological tools a case study of Chandigarh city has been taken. The accident data for the year 2012 was collected from Chandigarh traffic police and hot spot analysis was performed after identifying the accident points on the digitized map of the city following the steps mentioned earlier. Some Intersections from the manual analysis of accident data were chosen for the study. After visiting the chosen intersections, approximate ranks were assigned to each of them related to the factors observed. The following places were selected for investigation of Black Spots –

- Airport Light Chowk
- Kali Bari Light Point
- Housing Board Chowk
- Railway Light Point Intersection
- Sec 46,47/48,49 Light Point
- Aroma Light Point
- Kala Gram Chowk
- Tribune Chowk
- Transport Chowk

Table 5 gives the final weights obtained after the calculations.

Table 5 Final Scores Obtained

Intersection	Weights Sum	Weighted Weights
Airport Light Chowk	82	7.15
Kali Bari Light Point	80	6.95
Housing Board Chowk	89	7.91
Aroma Light Point	84	7.36
Kala Gram Chowk	88	7.7
Tribune Chowk	91	8.4
Transport Chowk	94	7.95
Railway Light Point	93	8.3
Sec 46,47/48,69 Light Point	85	7.53

Results

After the completion of analysis on Arc GIS following results were identified. Figure 3 shows the hot spots identified across the city. As from the Table 5 it is clearly visible that Kali Bari Light Point has minimum weight i.e. 80 and also least score according to the weighted average method. Thus according to the concept of prioritization this has to be the most accident prone point. This is also cross checked by the results from hot spot analysis rank 1 intersection roads lie the hot spot area of the map(refer Figure 4). Also manual analysis of accident data from the police shows that there have been maximum repeated numbers of accidents on this intersection. Some parameters normally unaccounted for like length of the road before the intersection, sight distance, control of entry and number of conflict points were taken into consideration in the study. These parameters directly or indirectly affect the speed of the vehicle which is one of the major reasons for an accident to occur. Different weights were assigned to each of the factors and they show better results as compared to the normal sum method. Some differences in results were obtained showing better accountability of factors in the method. Thus, weighted mean would provide more ease and refinement in the prioritization of stretches. Therefore speed calming measures at such intersections with length of road connecting the point need to be taken.

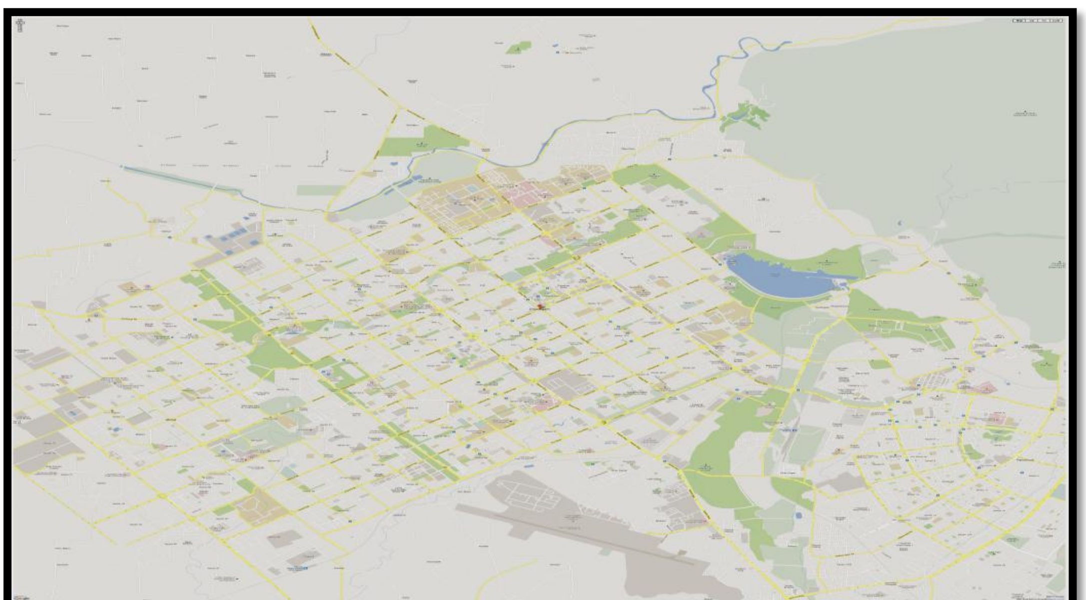

Figure 2: Reference Map of Chandigarh City

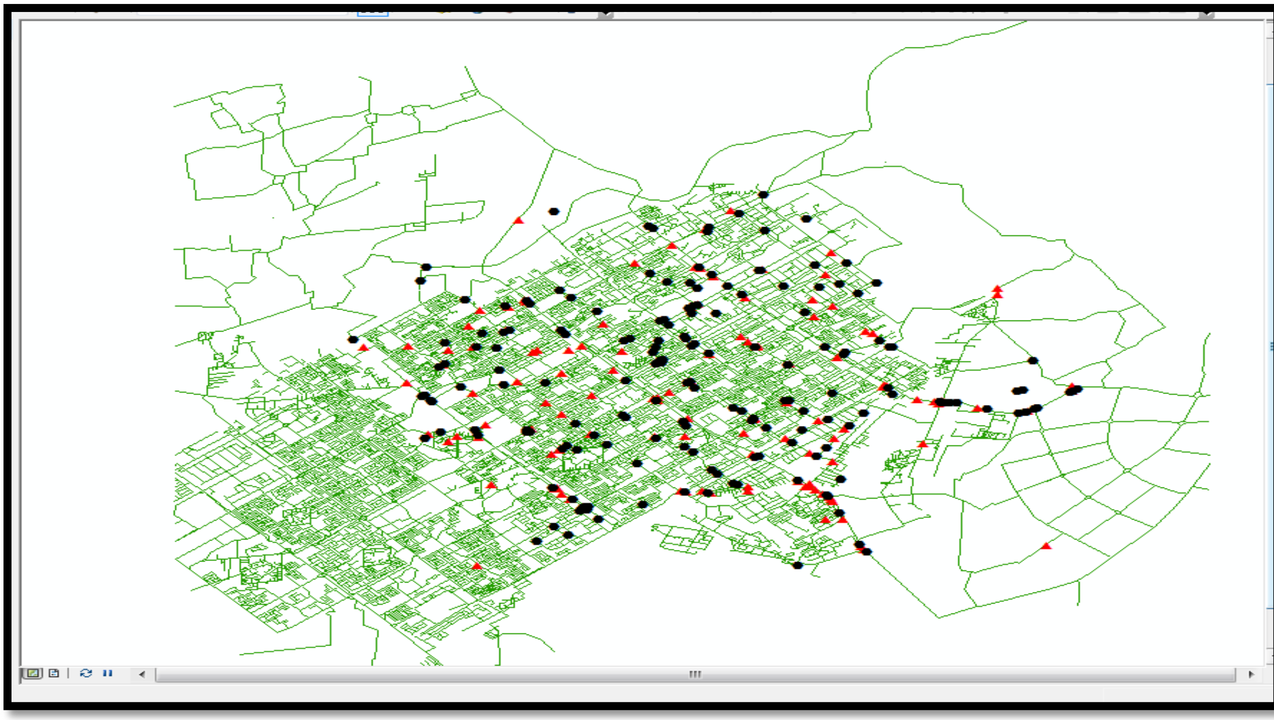

Figure 3: Marked Accident Locations

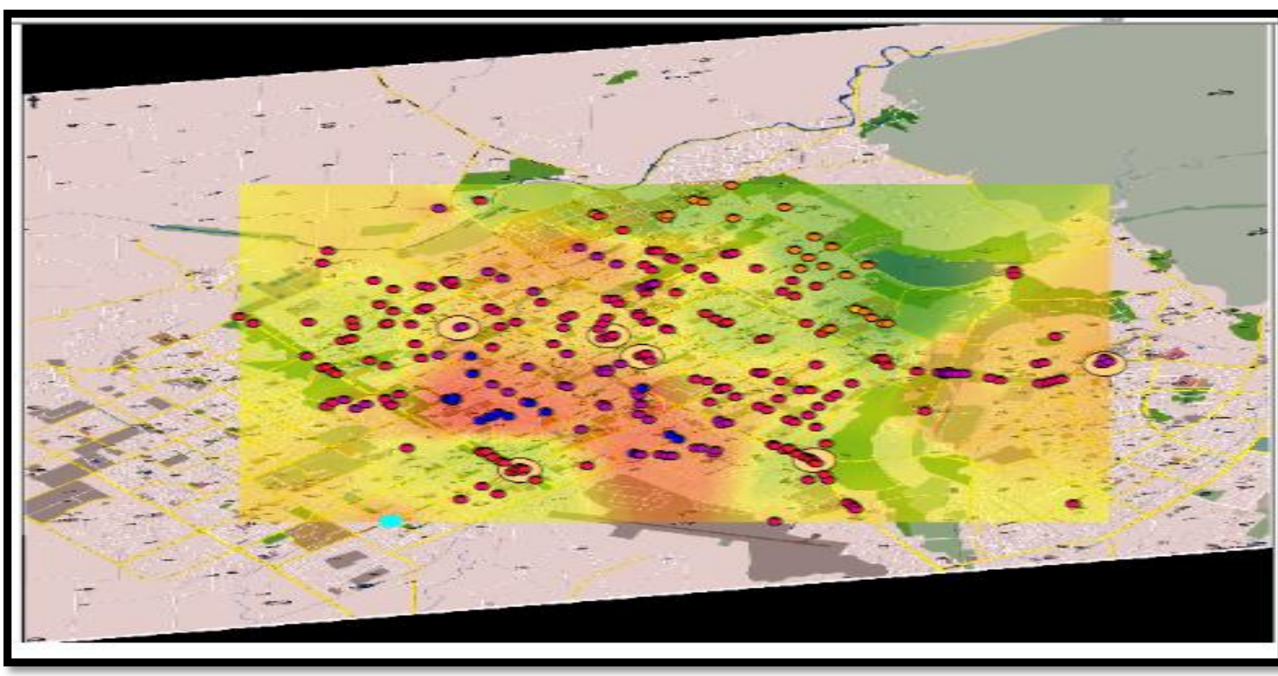

Figure 4: Hot Spot Analysis Results

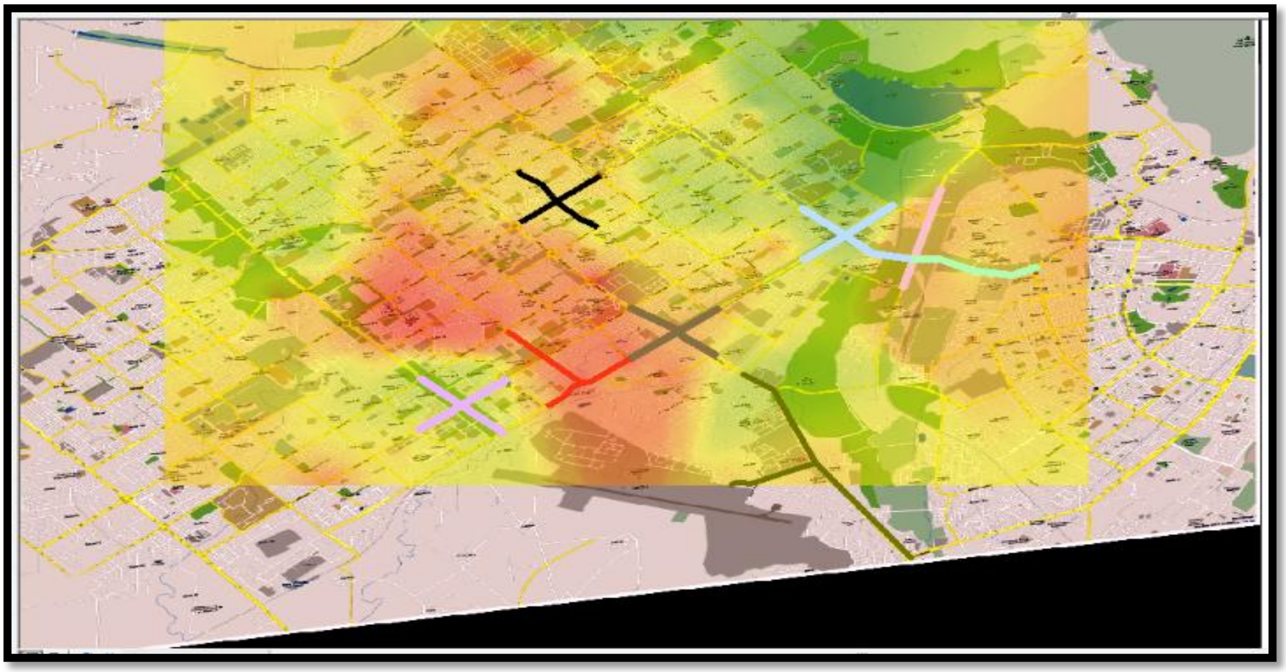

Figure 2: Comparison of Black Spots and Hot Spots

Conclusions and Recommendations

The main advantage of using this approach is that the results can be validated easily by taking into account some essential parameters and is not time consuming. Therefore the results obtained from the entire process can be used for planning road safety measures. These results may act as a quick reference for decision making of prioritization of roads in the city.

A. Overview – Feasibility of use of GIS for accident analysis for the Chandigarh city was checked.
B. Data Collection – The FIR records available with the police stations are considered to be authentic for implementation in the analysis as there is no other viable source for the data availability. The accuracy of point location is not very high because of non-availability of exact coordinates and hence approximation needs to be done.
C. Analysis – Due to some issues such as lack of digitized data, improper data collection, vagueness in the point's identification a need for use GIS for the same is established.
D. Recommendations –
   - Proper data collection methods need to be taught to the personal reporting the accident.
   - Proper Performa needs to be made to incorporate all the details relating to the accident occurrence.
   - Use of the model on a larger scale for more accuracy.

Acknowledgement

Finally we would like to thank Dr. A.K. Sarkar for his guidance throughout the work and Chandigarh Traffic Police for their support and cooperation in providing the accident data required for the analysis.

References

[1] Al-Ramdan, B. (2002). Introduction to Geographic Information Systems and its Application. Review 2002: A Chapter In *College of Environmental Design Book,* KFUPM, Dhahra, 113-120.

[2] Anderson, T. (2007). Comparison of spatial methods for measuring road accident 'hot-spots'- A case study of London. *Journal of Maps*, v2007, 55-63.

[3] Aylo, A. (2010). *Developing a GIS Safety Analysis System- A case study.* MS Thesis at University of Dayton, USA.

[4] Desai, M. M. (2011). Roads Accident study Based on regression Model- A case Study of Ahmedabad City. *National Conference on recent Trends in Engineering and Technology.*

[5] Deshpande, N., Chanda, I., & Arkatkar, S. (2011). Accident Mapping and Analysis Using Geographical Information Systems. *International Journal of Earth Sciences and Engineering*, 342-342.

[6] Ghosh, S. K., Parida, M., & Uraon, J. K. (2004). Traffic accident analysis for Dehradun city using GIS. *ITPJ*, 40-54.

[7] Mandloi, D., & Gupta, R. (2003). Evaluation of Accident Black Spots on roads using GIS. *6th Annual International Conference, Map India, Paper No. 38.*

[8] Mitra, S. (2008). *Enhancing Road Safety: A GIS based methodology to identify potential areas of improvement.* Project Report Submitted to William nad Barbara Leonard University Transportation Center, CSU- San Bernardino.

[9] Mohan, D. (2009). Road Accidents in India. *Transport Research and Injury Prevention Programme, IIT Delhi.*

[10] Mondal. (2011). Are road accidents affected by Rainfall- A case study from a large Indian Metropolitan City. British *Journal of Applied Science and Technology*, 16-26.

[11] Sarkar, A. K., & Malleshwari, N. (1995). Identification of accident prone location on city streets- a case study. *New horizons in roads and road transport, Volume 2*.

[12] Singh, R. K., & Suman, S. K. (2012). Accident Analysis and prediction of model on National Highways. *International Journal of Advanced Technology in Civil Engineering*. Volume 1 Issue2.

[13] Singh, S. K., & Mishra, A. Road Accident Analysis-A case Study of Patna city. *Urban Transport Journal 2*. 60-75

[14] Smith, R. (2001). Implimentation of GIS based highway safety analysis- Bridging the gap. Report No. *FHWA-RD-01-039, U.S.* Department of Transportation

[15] Valli, P. P. (2005). Road accident model for large metropolitan cities of India. *IATSS Research Vol.29 No 1*.

[16] Accident FIR Records – Chandigarh Traffic Police