Pedestrian Accident Analysis in Delhi using GIS

Shalini RANKAVAT[a], Geetam TIWARI[b]

[a]*Research Scholar, Indian Institute of Technology, New Delhi, 110016, India; E-mail: shalini.rankavat@gmail.com*
[b]*Professor, Indian Institute of Technology, New Delhi, 110016, India; E-mail: geetamt@gmail.com*

Abstract: Walking is the most sustainable and most used mode of transportation in Indian cities. From 2006-2009, an estimated 8503 fatalities in road traffic crashes occurred in Delhi in which pedestrians contribute almost 51%. The objective of this study was to investigate the potential of utilizing geographic information systems (GIS) in identifying pedestrian accident-prone locations. Crash data were geocoded in ArcGIS over the digitized road map of Delhi. Results highlighted four pedestrian fatal hotspot locations in Delhi. Fatal pedestrian accidents were significantly clustered ($p<0.05$) over the Ring Road of Delhi at intersections. The analysis showed that the pedestrian fatalities at arterial road intersections involving cars and buses were 27% and 14% of all fatalities involving cars and buses respectively.

Keywords: Pedestrian, pedestrian fatalities, GIS, Spatial analysis, Delhi traffic.

1. INTRODUCTION

According to official statistics, 105,725 people were killed in road traffic crashes in India in 2006 (NCRB, 2007), which is nearly 10% of the total road traffic deaths in the world. The fatality rate has increased from 36 fatalities per million persons in 1980 to 95 fatalities per million persons in 2006 (Mohan, 2009). The total number of people killed in road traffic crashes has increased to 1,36,834 in 2011(NCRB, 2012).

The proportion of fatal accidents in total road accidents has consistently increased since 2001 from 17.6% to 23.9% in 2010 .The severity of road accidents, measured in terms of persons killed per 100 accidents, has also increased from 19.9 in 2001 to 26.9 in 2010. (MORTH,2010). From 2006-09 the share of fatal crashes out of total crashes was 24% in India (MORTH, 2010) and 25% in Delhi (Delhi Police, 2009).

Delhi, like most Indian cities, has a mixed traffic. The road space is shared by many modes of different sizes and technologies. This is partly because large numbers of people need to walk between their places of residence and their places of work. The traffic stream include pedestrians, bicycles, animal and human drawn vehicles along with modern cars, trucks, buses, two-wheeler and three-wheelers. Nearly 32% of all commuter trips in Delhi are walking trips(Tiwari,2003; Transport Department, 2009).

From 2001-2009, an estimated 36,376 pedestrians' crashes occurred in Delhi. These crashes resulted in some 8,697 pedestrians fatalities.

Figure 1 shows the share of pedestrian fatalities in Delhi from 2006-09 (Delhi Police, 2009). Pedestrians account for nearly 51% of total fatal accidents in Delhi. It indicates that pedestrians have the largest share in total fatalities. This is because pedestrians continue to share the road space in the absence of infrastructure specifically designed for non-motorized vehicles; they are exposed to higher risks of being involved in a road traffic accident. A study of mid-block conflicts by Tiwari *et al*. (1998) presents information about the use of road space by different road users. The fourteen sites studied showed that maximum mixing of pedestrians and non-motorized vehicles with motor-vehicles occurs at the bus stops.

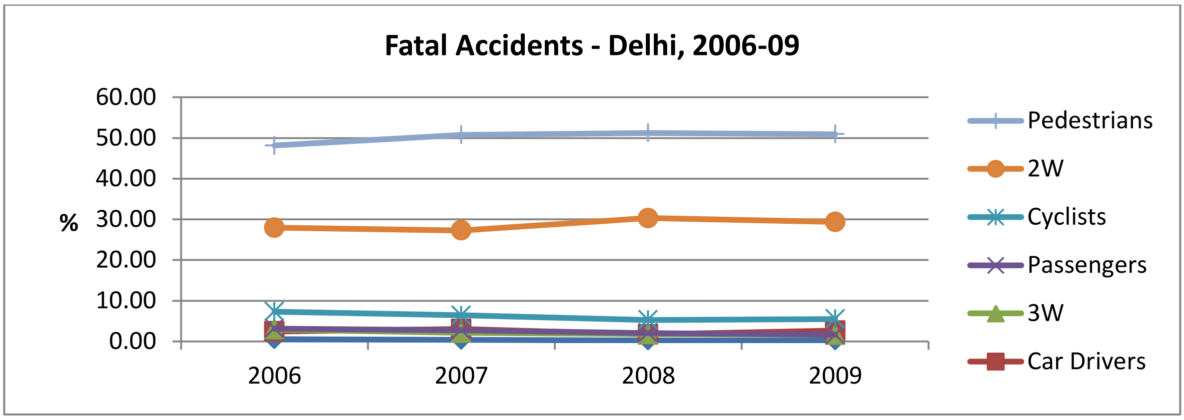

Figure 1. Road Traffic Fatalities in Delhi, 2006-09

One of the most critical questions that traffic safety engineers face is where to implement safety counter measures so that they can have the most significant impact. Therefore, there is a need to identify and analyze the locations which are hazardous, also known as black spots. Pedestrian high crash locations have to be identified to implement problem specific pedestrian safety countermeasures so as to improve pedestrian safety. Many studies have addressed the issue of safety analysis using Geographic Information Systems (GIS)(Levine *et al*., 1995; Affum and Taylor, 1995; Austin *et al*., 1997; Kim and Levine, 1996; Miller, 1999; Pulugurtha *et al*., 2007). These findings state that GIS is a useful tool for analyzing geographic context of crashes.

Since traffic crashes have a spatial dimension, statistical analysis combined with spatial attributes presents a better method of understanding traffic crashes. Geographical Information system (GIS) is being widely used to perform spatial data management and analysis operations.

The main objective of this paper was to investigate the potential of utilizing geographic information systems (GIS) in identifying pedestrian accident-prone locations in Delhi.

2. LITERATURE REVIEW

Studies in recent years have focused on the issue of safety analysis using GIS techniques. Simple crash plotting, or geocoding crash locations, is the most common GIS technique used for safety studies. GIS turns statistical data, such as number of crashes, and geographic data, such as roads and crash locations, into meaningful information for spatial analysis and mapping. Using GIS it is relatively simple to combine information received annually on crashes and determines any correlation such as type of street and adjacent land use. GIS also assists in identifying any factors that were contributing to those crashes and/or potential solutions to reduce those crashes.

In transportation safety applications, GIS has been widely used to geo-code accident locations, developing pin maps of crashes and database queries as performed by Levine *et al*.(1995); Affum and Taylor (1995); Austin *et al*. (1997); Kim and Levine (1996) ; Miller (1999). However, some researchers Levine *et al*.(1995) ; Pulugurtha *et al*.(2007) incorporated some of the powerful analytical tools available in GIS software such as buffer, nearest neighbor method, simple density and kernel density estimation method of crash cluster detection to show spatial distribution of crashes at the road network level.

Schneider *et al*. (2001) explains the importance of methods using GIS to identify where the pedestrian crash problem exists so that a greater number of pedestrian crashes can be prevented in the future.

Ziari and Khabiri (2005) presented the development and findings of crash data from police reports (Iran) and how they are being used in GIS. The authors developed a tool that generates a contour map identifying areas of high crash occurrence determined by crash density and clusters of crashes involving pedestrians and bicyclists.

Lai and Chan (2004) presented potential use of GIS technology in road accident analysis of Hong Kong. Various point-pattern techniques have been illustrated. It is by observing these distinctive distributional patterns that the hidden behavior of data is revealed. A clustered distribution is the concentration of points in one or a few segments of places. It helps to emphasize the problem or concentrated areas.

Steenberghen *et al*. (2004) have also shown the usefulness of GIS and point pattern techniques for defining road-accident black zones in one Belgian town, Mechelen. Satiennam and Tanaboriboon (2003) examined the pedestrian accidents in KhonKaen Municipality, Thailand. To identify locations where pedestrian crash problems exist in KhonKaen Municipality, the ArcView GIS spatial was employed to create the pedestrian crash density map.

Prasannakumar *et al*. (2011) evaluated road accident hot spots in a South Indian city using hot-spot analysis (Getis-OrdGi* statistics) in GIS. Anderson (2006) compared spatial methods of GIS for measuring road accident hotspots. Jang *et al*. (2013) used spatial Kernel density estimation method in GIS to measure the concentrated density of pedestrian crashes.

In this study, using GIS tools we identified pedestrian accident-prone areas, pedestrian accident-prone roads which have clustering of accidents and the factors (vehicles) involved in these accidents in Delhi, India.

3. METHODS

3.1 Usage of GIS for the study

The digital road map of Delhi was imported in the data frame of the Arc Map window and saved as "roads" layer. Before starting the editing, the imported digital map and the data frame need to have the common 'projected co-ordinate' system so as to make their scales the same. The co-ordinate system used throughout this project is WGS 1980 UCS.

The pedestrian fatal accidents' data in Delhi in four years from 2006-09 was collected from Delhi police. Excel sheet with accident details' given by police was imported in GIS by giving X and Y coordinates of each accident (considering all as points) by searching the accident locations in Google Map. For every point marked, a Feature Identity (FID) is created automatically on GIS, which is a whole number.

After importing the four years accident data as points, the file was saved as Shape file (.shp). All the details about accident points such as "Place of occurrence", "Time", "Date", etc., given in excel sheet were saved as attribute table in "4 yrs accident" layer.

Further analysis was done using SQL (structured query language) in the attribute table of four years accident points and spatial analysis using spatial analyst tool and spatial statistics tool.

4. DATA ANALYSIS

4.1 SQL

By writing simple statements in SQL, one can visualize various information like: accidents on roads; Peak time accidents; pedestrian hit by bus, car, scooter etc.; Accidents in the years 2006, 2007, 2008 and 2009 individually with road wise distinction and a lot of other combinations can be worked out! This way a clear picture is visualized with locations of occurrence of accidents.

Analysis of the four years' accident data showed that the Ring road and the outer Ring road in Delhi had higher count of accidents than other arterial roads. Ring road had 793 fatal accidents and 1818 non-fatal accidents, which comprise of about ~10% of total fatal accidents and 7% of total non-fatal accidents in Delhi. While outer ring road had about ~6% of total fatal accidents in Delhi. When we analyzed total fatal accidents over these two roads consequently in these years we did not find any trend. If we select pedestrian as victim then we find that count of pedestrian fatal accidents increased per year over Ring Road and Outer Ring Road both as shown in Table 1.

Table 1. ORR & RR of Delhi, 2007-2009

	Pedestrian Fatal Accidents	
Year	OUTER RING ROAD	RING ROAD
2007	43	93
2008	47	99
2009	59	110

We analyzed the rate of accidents per year over the arterial roads of Delhi with pedestrian as victim. Road lengths were taken from GIS map of Delhi. Table 2gives the statistics of fatal accidents' rate of pedestrians in top eight roads. We find that Ring road tops the list again.

Table 2. Top eight roads of Delhi in decreasing order of pedestrian fatal Accidents' rate per year, 2006-09

S. No.	ROADNAME	PEDESTRIAN FATAL ACCIDENTS' (per km per year)
1	RING ROAD	2.22
2	MEHRAULI BADARPUR RD	1.58
3	GRAND TRUNK(GT) ROAD	1.50
4	SHIVAJI MARG (Najafgarh road)	1.28
5	OUTER RING ROAD	1.03
6	NH-8	0.90
7	ROHTAK ROAD	0.83
8	AURBINDO MARG	0.65

We also analyzed these roads based on rate of accidents (per km) per year of pedestrians by private cars and buses whose count are 404 and 353 respectively, and found the list as shown in Table 3. Here bus includes blue line buses, DTC buses, mini/RTV buses, state buses and school buses. Blue line buses' accident count is 208 and that of DTC buses is 72. Table 3is showing that the ordering of the roads almost reversed if we compare pedestrian hits by car and buses.

Table 3. Top eight roads of Delhi in decreasing order of pedestrian fatal accidents' rate with Cars and Buses, 2006-09

	Cars		Buses	
S. No.	ROADNAME	PEDESTRIAN FATAL ACCIDENTS' (per km per year)	ROADNAME	PEDESTRIAN FATAL ACCIDENTS' (per km per year)
1	NH-8	0.25	MEHRAULI BADARPUR	0.27
2	RING ROAD	0.18	SHIVAJI MARG	0.23
3	AURBINDO MARG	0.15	G T ROAD	0.17
4	MEHRAULI	0.13	ROHTAK ROAD	0.13
5	OUTER RING ROAD	0.12	RING ROAD	0.10
6	G T ROAD	0.12	OUTER RING ROAD	0.10
7	SHIVAJIMARG	0.10	AURBINDO MARG	0.00
8	ROHTAK ROAD	0.09	NH-8	0.00

Further analysis by writing SQL statements shows that the pedestrians who were killed by buses in four years were 35% on minor roads, 22% over arterial roads, 15% over sub-arterial roads, 11% over collector roads and rest 17% was not captured in given buffer distance (based on respective ROW). The pedestrians who were killed by cars in four years were 43% on minor roads, 20% over arterial roads, 14% over sub-arterial roads, 5% over collector roads and rest 18% was not captured in given buffer distance (based on respective ROW).Major proportion of fatal pedestrian crashes were the cases of 'hit and run'. They are 50% of the records. The pedestrians fatalities "unknown" in four years were 37% on minor roads, 28% over arterial roads, 13% over sub-arterial roads,4% over collector roads and rest 17% was not captured in given buffer distance (based on respective ROW).

We analyzed pedestrian accidents involving bus, car and unknown category over intersections and midblock of arterial roads. This was done in GIS by choosing buffer radii of 150m at arterial intersections and of 15m at arterial-sub arterial intersections. Results are shown in Table 4.

Table 4. Fatal pedestrian accidents' involving bus, car and unknown over arterial roads

	Arterial to arterial road intersections	Arterial to sub-arterial road intersections	Mid-blocks of arterial roads
Bus (%)	14	10	76
Cars (%)	27	8	65
Unknown (%)	55	15	30

We find that the crashes by bus over arterial roads' mid-blocks are higher than cars. The crashes due to unknown category are higher over the intersections of arterial roads.

4.2 Spatial analysis

The purpose of this procedure was to determine if locations exist in which “accident clusters” or specific accident types occur. A cluster is defined as a group of crashes that is in relatively close proximity to a single location (point) or corridor (line).

4.2.1 Hot spot analysis

Hot spot analysis is a spatial statistics tool. Hot-spot analysis was done for major roads to see the sections of road where there is clustering of accidents. This tool identifies statistically significant spatial clusters of high values (hot spots) and low values (cold spots).The Hot Spot Analysis tool calculates the Getis-OrdGi* statistic (pronounced G-i-star) for each feature in a dataset. The resultant z-scores and p-values tells where features with either high or low values cluster spatially. This tool works by looking at each feature within the context of neighboring features. A feature with a high value is interesting but may not be a statistically significant hot spot. To be a statistically significant hot spot, a feature will have a high value and be surrounded by other features with high values as well. The local sum for a feature and its neighbors is compared proportionally to the sum of all features; when the local sum is very different from the expected local sum, and that difference is too large to be the result of random chance, statistically significant z-scores results. It creates a new Output Feature Class with a z-score and p-value for each feature in the Input Feature Class. A high z-score and small p-value for a feature indicates a spatial clustering of high values. A low negative z-score and small p-value indicates a spatial clustering of low values. The higher (or lower) the z-score shows the more intense the clustering. A z-score near zero indicates no apparent spatial clustering.

The G-statistic is calculated using a neighborhood based on a distance that we specify. If the “neighboring” feature is within the specified distance of the target feature, then that pair will be assigned a weight of 1; else, the pair will be assigned a weight of 0. Then, the statistic is calculated using Eq.(1).

$$G(d) = \frac{\sum_i \sum_j w_{ij}(x_i)(x_j)}{\sum_i \sum_j (x_i)(x_j)} \quad (1)$$

where x_i = the measured attribute of interest at location i

x_j = the measured attribute of interest at location j

w_{ij} = a weight indexing the location of i relative to j (this is 1 if locations i and j are within the distance you specified and 0 if not)

Polygon features around the road were created by giving a buffer of 20 m on both sides of centerline of the road. Spatial join tool was used to count the number of accidents in each polygon. The resultant field containing the number of accidents in each polygon became the input field for analysis. Fixed distance band option was chosen for spatial relationship. Threshold distance or distance band chosen was 1m for analyzing the features within this distance. The output of the Gi function is a z score for each feature. The z score represents the statistical significance of clustering for a specified distance

Map in Figure 2shows the critical sections of the road based on clustering of accidents. We can see the majority of clustering of accidents is over the intersections of the road. E.g. if we analyze Ring road only we find that out of 18 intersections, clustering was found in 12 intersections (i.e. ~67%) where we are 99% sure of clustering of accidents. Figure 2shows the locations which are highlighted are having clustering of accidents with statistical significance of > 2.58 standard deviation. Road types considered are Arterial, sub-arterial and collector roads. Minor roads are not included in this hot-spot analysis.

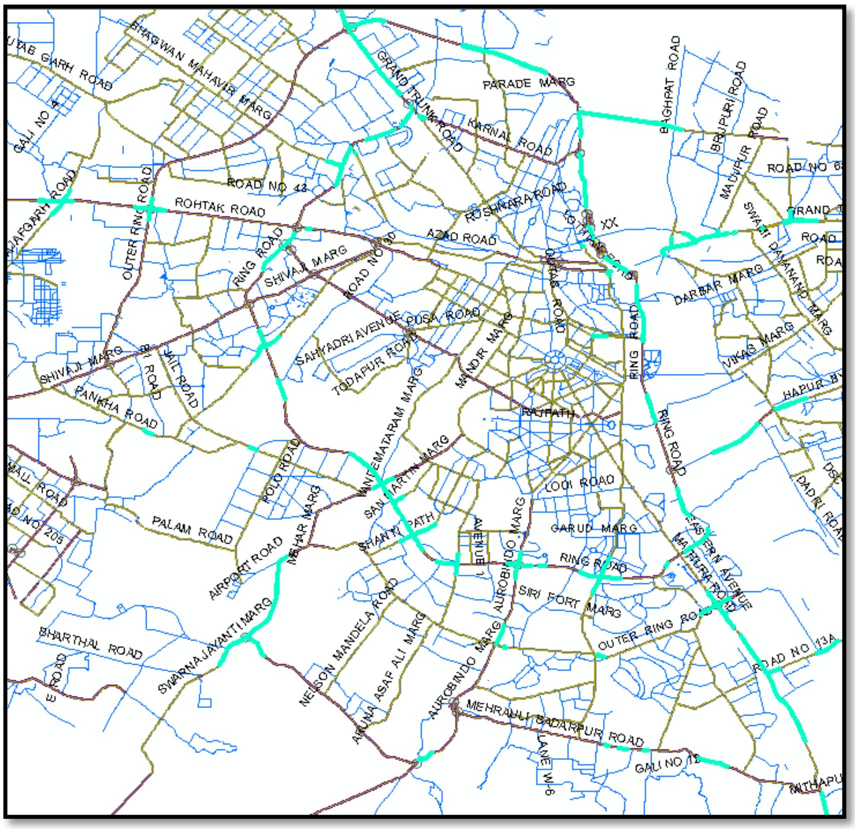

Figure 2. Critical Road sections for pedestrian accidents in Delhi, 2006-09

4.2.2 Kernel density

Kernel Density is a spatial analyst tool. Kernel Density analysis method was used to calculate the density of accidents with a search radius of 50m.

The plotted pedestrian crashes may show clustering and dispersion throughout the study area. Clusters of crashes are generally observed in densely populated metropolitan areas. Locations having clusters of crashes can be identified using the density feature available in widely used commercial GIS software. Density is a measure of the quantity of an item per unit of area. In a GIS environment; density can be calculated using Kernel Method (ESRI, 2002). The Principle of Kernel function is

$$f(x,y) = \frac{1}{nh^2}\sum_{i=1} K\left(\frac{d_i}{h}\right) \qquad (2)$$

where,

f (x, y) : density estimate at the location (x, y),
n : number of observations,
h : bandwidth or kernel size,
K : kernel function, and
d_i : distance between the location (x, y) and the location of the i_{th} observation

The Kernel Method divides the entire study area to a pre-determined number of cells and applies a circular neighborhood around each crash. Cell size chosen was 1m x1m. It then applies a quadratic kernel function (Silverman, 1986) that goes from being the highest at the point of crash to zero at the neighboring boundary. Individual cell values are assigned to each of the cell, and these are the values of the kernel function at the cell center. The final density of each cell is calculated by adding its individual cell values. The radius of the circular neighborhood affects the resulting density map. The greater is the radius; the flatter would be the resulting kernel.

Density map of pedestrian fatal accidents in Delhi of four years (2006-09) was created and is shown in Figure 3.

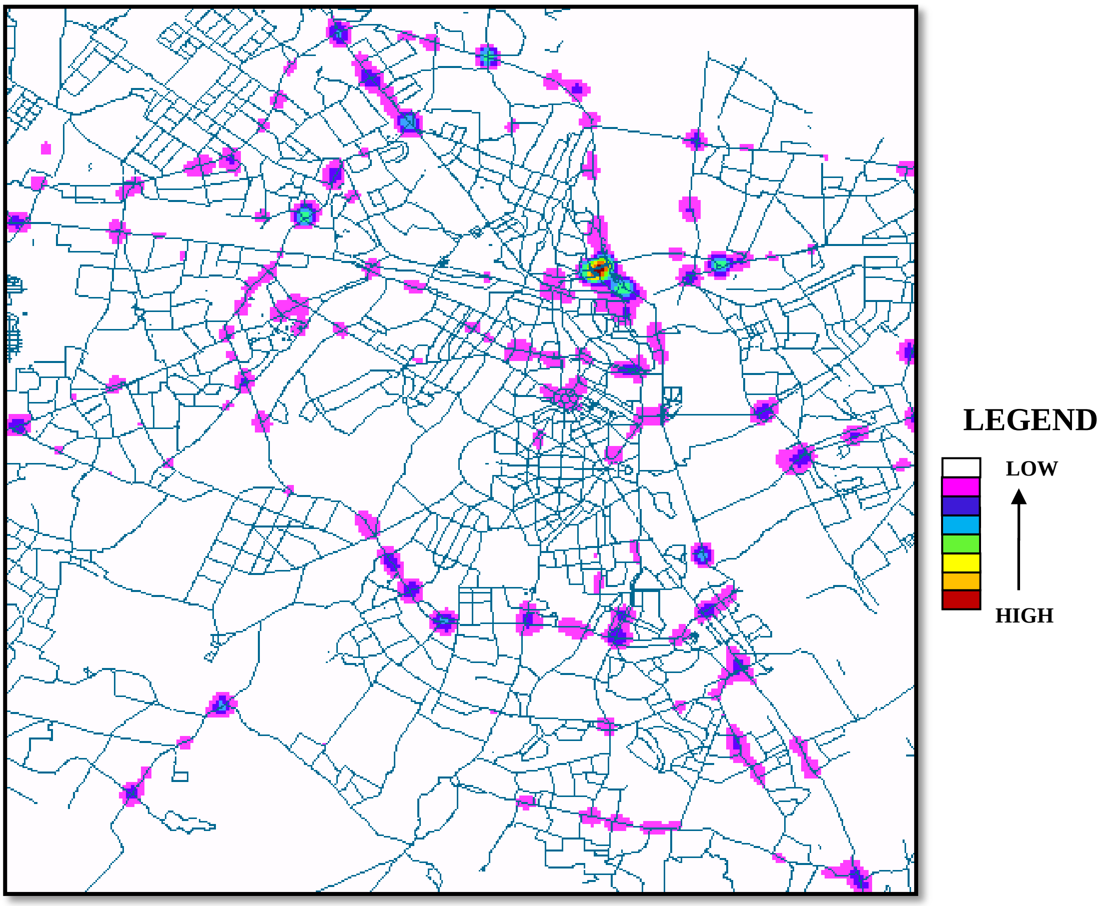

Figure 3. Density map for pedestrian accidents in Delhi, 2006-09

Once high density zones are defined, pedestrian safety programs can be focused in them with greatly increased efficiency (NHTSA, 1998).

It is clear from the map that density of pedestrian accidents is higher in North Delhi, near ISBT. If we see the density maps of pedestrian accidents' locations (Figure 4) according to their hierarchy of severity of risk we find the area surrounds the Ring Road again, shown in Figure 4 by marking circles.

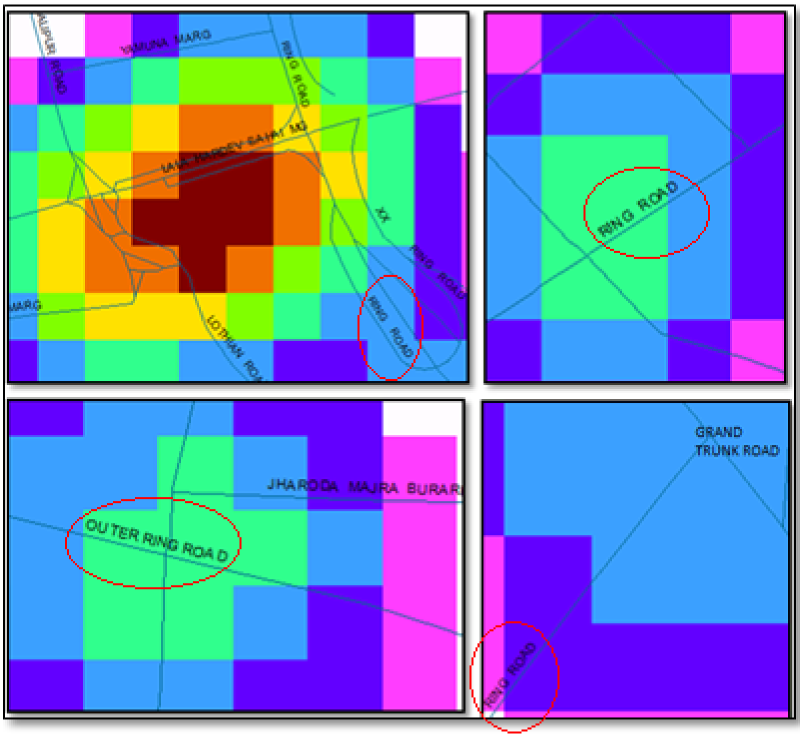

Figure 4. Density map of four major hotspots for pedestrian accidents in Delhi, 2006-09

After determining the hotspots we explored the percentage of pedestrian fatal accidents over the days of week and different time of day at the hot-spot location i.e. near to ISBT and Ring road.

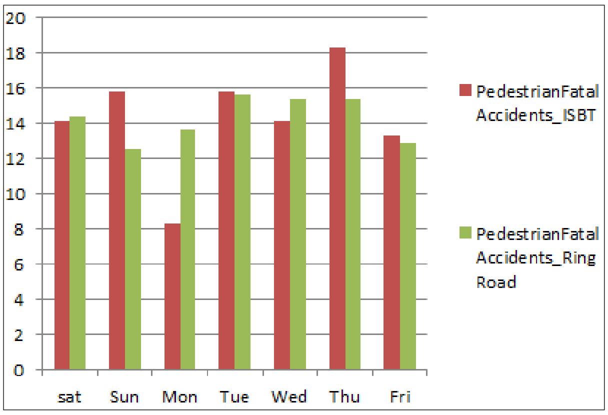

Figure 5. Pedestrians' fatalities based on days of the week

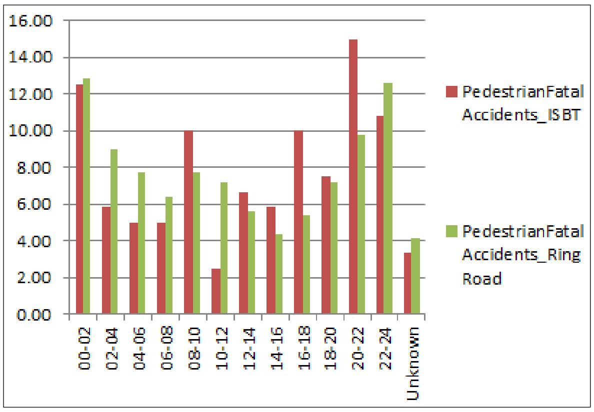

Figure 6. Pedestrians' fatalities based ontime of the day

We find that pedestrian fatal accidents' percentage is almost the same in all days of a week as shown in Figure 5. It is slightly lesser on Mondays near to ISBT due to most of the market being closed, so the exposure is less as compare to other working days.

If we compare the percentage of pedestrian fatal accidents at these two locations based on time of the day as shown in Figure 6then we find that high percentage of fatalities occur during night time between 8pm and 10pm near to ISBT. Even during late night when the pedestrians are very sparse then also count of the pedestrian road traffic fatalities are significantnear to ISBT (12am to 2am) and ring road (10pm to 2am).

5. CONCLUSION & DISCUSSION

It has been illustrated how the use of GIS can effectively help in the processing of accident data, and for performing complex spatial analysis. GIS helps tremendously in the visualization of the problem of road accidents.

The geographic distribution of pedestrian victims highlights the widespread insecurity people experience in Delhi when walking. Fatal crash density is higher near ISBT in Delhi, where population density is also high. To reduce the risk some resources should be dedicated specifically to pedestrian safety. On superimposing GIS map kml files over google earth map of Delhi it was found that roundabouts are having less number of accidents and clustering of accidents is found over the junctions and near foot of flyovers. The results display statistically significant and consistent patterns of clusters for pedestrian crashes over a four-year period and justify the use of spatial statistical techniques.

Findings can be used to understand the correlation between built environment and pedestrian safety, to prioritize the high-density zones for intervention efforts, and to formulate research hypotheses for investigating pedestrian crashes.

Traffic accident analysis would have been more accurate, easier and descriptive if the traffic accident reports were more detailed and formatted properly. If the X and Y coordinate of location using GPS were saved in police database then plotting of accident points would have been easier and results of analysis more reliable.

REFERENCES

Affum, J. K., Taylor, M. A. P. (1995) Technology tools for transportation professionals

– moving into the 21st century: Resource papers for the 1995International conference, Integrated GIS database for road safety management. Institute of Transportation Engineers, Washington, D.C., 189-193.

Anderson, T. (2006) Comparison of spatial methods for measuring road accident 'hotspots': a case study of London, *Journal of Maps, 55-63*

Austin, K., Tight, M.,, Kirby, H.(1997). The use of geographical information systems to enhance road safety analysis. *Transportation Planning and Technology*, 20 (4), 249-266.

Delhi Police, India, 2009.

ESRI.(2002). *Arc/Info documentation* – Version 8.2. Environmental Scientific Research Institute, Inc., Redlands, CA.

Flahaut, B., Mouchart, M., San Martin, E. , Thomas, I. (2003) The local spatial autocorrelation and the kernel method for identifying black zones, A comparative approach. *Accident Analysis and Prevention*, 35(6), 991-1004.

Jang, K., Park, S. H., Kang, S., Song, K.H., Kang, S., Chung, S. (2013) Evaluation of Pedestrian Safety: Geographical Identification of Pedestrian Crash Hotspots and Evaluating Risk Factors for Injury Severity, *TRB Annual Meeting*

Kim, K., Levine, N. (1996). Using GIS to improve highway safety. *Computers, Environment and Urban Systems,* 20(4), 289-302.

Lai, P., Chan, W. (2004). GIS for Road Accident Analysis in Hong Kong, *The International association of Chinese professionals in Geographic Information Science* 10(1), 58-67.

Levine, N., Kim, K., Nitz, L. (1995) Spatial analysis of Honolulu motor vehiclecrashes: part I: spatial patterns, *Accident Analysis and Prevention*, 27(5), 663-674.

Levine, N., Kim, K. , Nitz, L. (1995) Spatial analysis of Honolulu motor vehicle crashes: part II: zonal predictions, *Accident Analysis and Prevention*, 27(5), 675-685.

Mohan, D., Tsimhoni, O., Sivak, M.,Flannagan, M. (2009) Road Safety in India: Challenges and Opportunities, *UMTRI*-1

MORTH [Ministry of Road Transport and Highways]. (2010). Road Accidents in India

Miller, J. S. (1999) What value may geographic information systems add to the art of identifying crash countermeasures?, Report VTRC 99-R13.

NCRB [National Crime Records Bureau]. (2012). Accidental deaths & suicides in India - 2011.New Delhi: Ministry of Home Affairs, National Crime Records Bureau

NHTSA (1998). Zone Guide for Pedestrian Safety Shows How To Make Systematic Improvements. Traffic Tech, Issue 181, HS-042 731, National Highway Traffic Safety Administration, Washington, D.C. (available athttp://safety.fhwa.dot.gov/).

Prasannakumar, V., Vijith, H., Charutha, R.,Geetha, N. (2011) Spatio-Temporal Clustering of Road Accidents: GIS based Analysis and Assessment, *Procedia Social and Behavioral Sciences*, 21, 317-325.

Pulugurtha, S. S., Krishnakumar, V. K. , Nambisan, S. S. (2007) New methods to identify and rank high pedestrian crash zones: An illustration, *Accident Analysis and Prevention*, 39(4), 800-811.

Schneider, R. J., Ryznar, R. M., Khattak, A. J. (2004)An accident waiting to happen: a spatial approach to proactive pedestrian planning, *Accident Analysis & Prevention*, 36., 193-211.

Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*.Chapman and Hall Ltd., New York, NY

Steenberghen, T., Dufays, T., Thomas, I., Flahaut, B. (2004) Intra-urban location and clustering of road accidents using GIS: a Belgian example. *International Journal of*

Geographical Information Science, 18[2], 169-81.

Satiennam, T.,Tanaboriboon, Y.(2005) COMPREHENSIVE TRANSPORTATION POLICY DEVELOPMENT USING TRAFFIC SIMULATION TECHNIQUE FOR KHON KAEN CITY, THAILAND. *Proceedings of the Eastern Asia Society for Transportation Studies* 5, 2352-64.

Tiwari, G. (2003) Transport and land-use policies in delhi, *Bulletin of the World Health Organization*, 81

Tiwari, G. (1998)TOWARDS A SUSTAINABLE URBAN TRANSPORT SYSTEM: PLANNING FOR NON-MOTORIZED VEHICLES IN CITIES, Transportation Research and Injury Prevention Programme, Indian Institute of Technology ,Delhi

Ziari, H.,Khabiri, M. (2005) Applied GIS software for improving pedestrian and bicycle safety, *Transport*, XX(4), 160-164