JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

JIMMA INSTITUTE OF TECHNOLOGY
FUCULTY OF CIVIL AND ENVIRONMENTAL ENGINEERING
HIGH WAY ENGINEERING STREAM

ANALYSIS OF TRAFFIC ACCIDENTS HOT SPOT AREAS USING SPATIAL STATISTICS AND GEOSTATISTICAL METHODS OF GEOGRAPHIC INFORMATION SYSTEM:

A CASE STUDY OF GIBE BRIDGE TO SEKORU TOWN ROAD SEGMENT

A Thesis Submitted to the School of Graduate Studies of Jimma University in Partial Fulfillment of the Requirements for the Degree of Master of Science in Highway Engineering

By: Lukas Embaye

2018
Jimma, Ethiopia

JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

JIMMA INSTITUTE OF TECHNOLOGY

FACULTY OF CIVIL AND ENVIRONMENTAL ENGINEERING

ANALYSIS OF TRAFFIC ACCIDENTS HOT SPOT AREAS USING SPATIAL STATISTICS AND GEOSTATISTICAL METHODS OF GEOGRAPHIC INFORMATION SYSTEM:

A CASE STUDY OF GIBE BRIDGE TO SEKORU TOWN ROAD SEGMENT

A Thesis Submitted to The School of Graduate Studies of Jimma University in Partial Fulfillment of the Requirements for The Degree of Master of Science in Highway Engineering

By: Lukas Embaye

Advisors:

Main Advisor: Dr.-Ing Fekadu Fufa **(PhD)**

Co-Advisor: Eng. Abel Tesfaye **(MSc.)**

2018

Jimma, Ethiopia

DECLARATION

I, hereby declare that the work which is being presented in this research study entitles "Analysis of traffic accidents hot spot areas using spatial statistics and geostatistical methods of geographic information system: a case study of Gibe bridge to Sekoru town road segment" is original work of my own, and it has not been presented for degree in any other university.

Submitted by:

Lukas Embaye	_________________	____/____/_____
Name	signature	date

JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

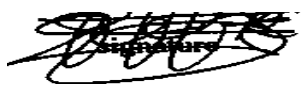
JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING

As members of the Examining Board of the Final M.Sc. Open Defense, we certify that we have read and evaluated the thesis prepared by: Mr. Lukas Embaye Entitled: Analysis of traffic accidents hot spot areas using spatial statistics and geostatistical methods of geographic information system: a case study of Gibe bridge to Sekoru town road segment. And recommended that it be accepted as fulfilling the thesis requirement for the degree of Master of Science in Civil Engineering under highway engineering stream.

Eng'r. _______________ ____/____/______

Chairman Signature date



Dr.-Ing Fekadu Fufa (PhD ______/____/_____

Advisor date

Eng'r. Abel Tesfaye (MSc.) ________________ _____ /____/_____

Co-Advisor signature date

Eng'r. Markos Tsegaye _________________ ____/_____/_____

Internal Examiner Signature date

Dr. Bikila Teklu (PhD) ________________ _____/_____/_____

External Examiner Signature date

ACKNOWLEDGEMENT

Above all I would like to thank the Almighty God, who gave me the opportunity and strength to achieve whatever I have achieved so far. Next, I would like to express my deep-felt gratitude to my advisors **Dr.-Eng Fekadu Fufa (PhD)** and **Engr**. **Abel Tesfaye** for their continuous guidance, encouragement, and enduring support shaped this work considerably and made the process of doing this thesis to learn invaluable experience.

I would also like to express my gratitude to Dr.-Ing Tamene Adugna, Dr. Asfaw Kebede, Eng'r Getachew Kebede, Eng'r Tareqegn Kumera, Eng'r Markos Tsegaye who supported and accompanied me during all the progress of this study.

Finally, to all my parents and friends deserve special thanks in every moment of my life to Mr. Henok Addissie, Mr. Melese Haile, Mr. Milashu Sisay Mr. Dagmawi Kedir, Mr. Tesfaye Teklu, Mr. Gebremedhin Teklay, Ms. Lemlem Tekle, Mr. Kaleab Embaye, Ms. Kidist Gebremedhin and Ms. Tsionmariam Abrha.

ABSTRACT

Geographic Information System (GIS) technology has been a popular tool for identifying hotspots in highways and analyzing traffic accident data. Many institutions and researchers are using GIS for accident analysis. The geometry of Gibe bridge to Sekoru town highway segment (total of 64.4 km) has many curves. 50.71% of the road segment is dominated by curves with design speed 35-40 km/hr and an average grade of 6% from the whole segment. Thus, this worst road geometric condition leads to high amount of traffic accident problems in the study area. The traffic accidents occurred along this segment need to be analyzed with GIS tools.

Therefore, the main objective of this study is to analyze traffic accident hotspots along the road segment from Gibe bridge to Sekoru town using GIS tools. The study identified sites of accident hotspots and major causes of traffic accidents along the road segment depending on 5 years' (2013–2017) of property damage only (PDO), slight injuries, serious injuries and fatal accident data obtained from Jimma Zone, Sekoru and Yem Districts Police Offices. The hot spot areas are analyzed using spatial statistics and geostatistical methods. The spatial statistical analysis includes Getis-Ord Gi, Anselin Local Moran Index, Moran Index, and hotspot optimization. The geostatistical approaches also contains inverse distance weighting, empirical Bayesian kriging and kernel smoothing density method with supportive methods such as the geographical weighting matrix, exploratory regression and ordinary least square.*

As observed from the GIS spatial auto correlation analysis results for accident data in the study area, the GiPValue of Shen Debitu curves (around Abelti) and Kumbi comes $P < 0.05$ and $P < 0.1$ which is in between 90 and 95% confidence level (Gi_Bin) with GiZScore values > 1.96 and > 1.65, respectively. The GiPValue of Natri, western Saja, Simini and Birilea river (around eastern part of Sekoru town also comes $P < 0.1$ which is 90% confidence level (Gi_Bin) with GiZScore values > 1.65. in addition, the Kernel Smoothing Density estimation, Inverse Distance Weighting and Empirical Bayesian Kriging method (geostatistical method) analysis method also have shown that shen debitu curves (around Abelti), Kumbi, Natri, western Saja, Simini and Birilea river (around eastern part of Sekoru town) were hotspot areas.

High accident zones were concentrated in Abelti (Shen Debitu), Kumbi, Natri, Saja and Sekoru cluster due to high speed, many number of curves, teenage drivers, night driving, design defects, and improper sight distance. This research recommends redesigning, reconstructing curves across the segment, limiting the maximum speed and developing road infrastructure along the Gibe bridge to Sekoru town road segment would decrease the occurrence of traffic accidents in the identified hotspot areas. It shall be also appropriate to use GIS tools in identifying traffic accident hotspot areas, if applied in the road network of Ethiopia.

Keywords*: Geographic Information System; Hot spot areas; Traffic accident; GiPvalue; GiZscore, Gi_Bin.*

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYM

AADT	Annual Average Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
ASD	Available Sight Distance
CSR	Conceptualization of Spatial Relationship
DBMS	Data Base Management System
DEM	Digital Elevation Model
EBK	Empirical Bayesian Kriging
ERA	Ethiopian Road Authority
GCP	Geographic Coordinate Points
GCS	Geographic Coordinate System
Gi*	Getis-Ord statistics method
GIS	Geographic Information System
IDW	Inverse Distance Weighting
JIT	Jimma Institute of Technology
KM	Kilometer
KSDE	Kernel Smoothing Density Estimation
LMiIndex	Anselin Local Moran Index
LOS	Line of Sight
PCS	Projected Coordinate System
PDO	Property Damage Only accident
PSD	Passing Sight Distance
QQ plot	Quantile-Quantile plot
RDS	Road Design Software
RTA	Road Traffic Accident
SA	Spatial Autocorrelation
SNNP	Southern Nation and Nationality of People

SSD	Stopping Sight Distance
UTM	Universal Transvers Mercator
VKT	Vehicle Kilometer Travelling
WGS_84	World Geodetic System
WHO	World Health Organization

CHAPTER ONE

INTRODUCTION

1.1. Background

Nowadays, road traffic accidents (RTAs) are unquestionably one of the many scourges distressing the world with a greater impact on developing countries (WHO, 2013). The most negative results of developing modern transportation systems are road accidents with injuries and loss of survives. So, traffic safety is the most critical matter in agencies' transportation strategy. The identification of safety deficient areas on the highway network is aimed at a comprehensive safety program by traffic officials. One of the most important problems that traffic officials face is where and how to implement precautionary measures and provisions so that they can have the most significant impact for traffic safety (Khan et al., 2006). GIS is a very important and comprehensive management tool for traffic safety. Since 1990's, GIS technologies have been used more frequently for such studies due to the availability of low cost GIS with user-friendly interfaces (Khan et al., 2006).

Accident data, collected for many years serve as the ground base for programs designed to reduce the number of traffic accidents but none of them identifies the geographical location of the accident areas simply serve numerical analysis using traditional statistical methods. To solve these and other problems using ArcGIS tools which identifies hot spot locations is advantageous. Identifying hot spots, ranking and determining the potential safety improvement for each location among a set of sites are the main purposes of studies in traffic safety. "Hotspots," "black spots," or high crash locations are sites on a section of a highway that have an accident frequency significantly higher than expected at some threshold level of significance (Hakkert and Mahalel, 1978). Identification of these hotspots is a systematic process of detecting road sections that suffer from an objectionable high risk of crashes. It is a low-cost strategy in road safety management where a small group of road network locations is selected from a large population for further analysis of specific problems, selection of cost-effective countermeasures, and prioritization of treatment sites. These identified sites are often known by various terms in literature, as hazardous locations, hotspots, black spots, priority investment locations, collision-prone locations, or dangerous sites (Lalita et al., 2015). According to Geurts, and Wets., 2003; Hauer, et al., 2002 and Miranda-Moreno, 2005 noted that using crash data to identify high risk areas can help agencies allocate limited resources more efficiently. Hence, numerous studies have utilized statistical indexes, and procedures to define crash hot-spots. As commercial GIS software became more commonly used, traffic safety analysts have now started using spatial methods in the analysis of traffic accident, primarily for convenience and efficacy. Several papers have applied spatial analysis methods, such as the Moran I, Getis-Ord Gi* (Gi*) and Geostatistical method such as Kernel Smoothing Density Estimation (KSDE) to identify hot-spots, and they have compared their results to those obtained from more traditional methods (e.g., repeatability analysis). In a recent study, Gundogdu stated that using traditional numerical methods and Gi* analysis together could improve the accuracy of hot-spot identification. The above studies did not discuss the differences between the various methods or provide a sound strategy for choosing an appropriate method that fits their studies. According to the Highway Safety Manual, several hot-spot identification methods focus on defining areas as high risk spots only based on the frequency of crashes, instead of clustered pattern, especially when examining aggregated crash data. It is precarious for traffic engineers to only use the Gi* method for defining crash hot-spots, simply because software programs such as ArcGIS have named the method as a "hot-spot" discovery tool. Conversely, another spatial analysis method, the kernel smoothing density method, is a popular site screening tool for identifying crash hotspots, which works similarly to the sliding window method. The purpose of this method is to find out which locations of the segment are the most crashes and, consequently, require the most improvement. As traffic accidents are the major cause of death and injuries worldwide. The location in a road where the traffic accidents often occur is a hot spot. In these hot spots, accidents are not a random event, but common due to varying factors like the distance that the driver can see along the vehicle path. That is of major importance when hazardous maneuvers must be done, such as an emergency stopping or an overtaking. Several studies have shown the existence of a direct link between accident rate and available sight distance (Silyanov et al., 1973). To reduce crashes and their worst consequences in terms of injuries and fatalities, road engineers are committed in protecting road users limiting vehicles' operating speeds and increasing the sight distances from the conflict points and/or potential obstacles along the driving path. Therefore; the research would be the analysis of traffic accident hotspot areas using spatial statistical and geostatistical methods of GIS.

1.2. Statement of the Problem

Ethiopia is one of the developing countries with low level of income coupled with high rate of population growth. As part of the developing world, Ethiopia is predominantly an agricultural country with low level of urbanization. In Ethiopia, the rate of road traffic accidents (RTAs) is very high; because road transport is the major transportation mechanism along with poor road infrastructure, poor traffic laws enforcement and other factors. The Ethiopian traffic control system archives data on various aspects of the traffic system, such as traffic volume, concentration, and vehicle accidents. (WHO, 2009).

Road traffic accident has been increasing from time to time in southwestern part of the country region of Oromia and SNNP region in alarming rate. According to the Jimma Zone Police Office the high number of road traffic accident was occurred in Sekoru district. The Gibe valley- Sekoru town road segment of the Jimma- Addis Ababa highway is located in Sekoru district, Jimma Zone, Oromia Regional State and some part of the Yem special district, SNNP, southwest Ethiopia. This section of the highway is the main gate of the whole country transport comes from Addis Ababa with the southwestern part of the country. This road segment is expressed by steep and cavernous topography causing sever road traffic accidents in the region. The geometry and road condition is not suitable for smoothing driving because of so many curves (from the whole distance of Gibe valley to Sekoru town of 64.4 km; the 32.658 km which is 50.71% of the road segment is dominated by curves and the remaining 31.742 km which is 49.29% of the road segment is dominated by straights) with an average grade of 6% and in some parts like the Abelti segment shen debitu, around kumbi and north eastern part of Sekoru towns have a grade of 7 up to 9% ; and in sufficient sight distance conditions with small radius and small curve lengths and with design speed restricts between 35-40 kilometer per hour (but so many driver doesn’t follow this rule) (studio Pietrangeli Rome). The 73% of (this percent indicates the general velocity from gibe valley up to Jimma town) permits a velocity of below 80 km/hr. due to these worst road geometric conditions leads to huge amount of traffic accident problems in the study area.

Taking in to consideration the fact that traffic safety is a commonly observed problem on the subject road segment, identification of hot spot areas with related to the geometric design parameters aspects is vital, and it can create an understanding on the basic root causes of traffic accidents. “Before commercial GIS software programs were available, traffic safety analysts tended to use traditional statistical tests to define hotspots that had significantly higher crash rates. Using the traditional statistical method is inconvenient and inefficient, because traffic engineers must separate road networks into multiple segments with equal lengths (if possible), record crashes for each segment length, use older statistical methods (such as Chi-square test) to define hot spots, and show results via tabulated data. In addition, using traditional statistical methods will not show a geographical relationship between crashes and other environmental variables." So, to reduce those problems, analyzing of these hotspot areas with impressive analysis of spatial statistics and geostatistical methods of traffic accidents on the subject area from Gibe bridge- Sekoru town road segment using advanced ArcGIS tools is the best solution.

1.3. Research Questions

The main research questions to be answered through the research process are:

1. How can identify hot spot areas?
2. How to identify the major causes of the occurrence of hot spot areas?
3. How ArcGIS applications are better to compare Spatial and Geostatistical analysis?

1.4. Objective

1.4.1. General Objective

- to analyze traffic accident hotspots along the Gibe bridge- Sekoru town road segment of the Jimma- Addis Ababa Highway using GIS tools.

1.4.2. Specific Objectives

The specific objectives are:

- to identify areas of accident hotspots along the Gibe bridge- Sekoru town road segment using Geographic Information System (GIS) tools (Spatial autocorrelation and Geostatistical analyst method)
- to identify the major accidents of the occurrence of hotspots with high number of crashes
- to compare hot spot analysis using Spatial statistics and Geostatistical methods;

1.5. Significance of the Research

At the end of this research, it is expected that all the research questions are answered and possible solutions or recommended values for the selected road segment will be indicated. And, the researcher believed to provide inputs for decision making, policy change or new regulations on the way to minimize traffic accidents and way of implementation using ArcGIS tools; the study is solve by showing the appropriate areas and help to stimulate further investigation by different researchers for further actions; and this study be an important resource /reference material for those who need further investigation about traffic accidents hot spot areas using ArcGIS tools on different road segments, accidents due to geometric design elements, traffic safety, across the country and for those established government and public organizations to reduce road geometry defects and increase traffic safety at different road corridors and it can be used for academic purpose as well, and contribute to review the effectiveness of existing / current control mechanisms that were taken by the officials.

1.6. Scope of the Study

Among the factors that affect the precision of any study, the availability and reliability of the information it employs is very important. This research was discussed issues involving analyzing of traffic accident hotspot areas through this case study and the information collected from Jimma zone, Sekoru and Yem district traffic police stations of Traffic accident control and inspection office of available five-year data (2013-2017) and other geometric parameter date are collected from Ethiopian Road Authority Jimma District and head office of Addis Ababa. Other division in Ethiopia with similar experience discussed briefly. As this study is specific to Gibe bridge to Sekoru town road segment, a similar study on other areas of the country is further suggested. This study is geographically limited to Gibe bridge to Sekoru town road segment. Generally, it addresses issues related to analysis of traffic accident hotspot areas.

1.7. Limitation of the Study

One of the main objective of this study is to analyze traffic accident hotspots along the Gibe bridge-Sekoru town road segment of the Jimma- Addis Ababa Highway using GIS tools. To do such kind of research, so many problems was happened; some limitations of the study are the organizations are not willingly to cooperate and provide useful documents which are susceptible to personal judgment. incompatibility of data sources, data management, data storage, data quality, varying data definitions and formats, data standardization in crash and inventory data, compliance with these guidelines varies by agency, and it was a time-consuming process to convert data into a layer in the GIS.

CHAPTER TWO

2. LITERATURE REVIEW

2.1 Review of Previous Studies

Different scholars, international journals and studies have attempted to come up with the definition of road traffic accidents and way of analysis using different methods. However, the term continued to be ambiguous and difficult to define. This chapter explores conceptual frameworks and literature review on road traffic accidents, sight distance and different analysis methods using GIS tools and control mechanisms. As a result, the definition causes and control mechanisms of RTAs, methods of analysis using GIS tools will be discussed in detail also it aims to censoriously assess and identify main accident hot spot areas and geometric elements relevant to the traffic safety of roads including geometric requirement of roads such as lane width, nature of curves, radius of curves and gradients, length of roadway alignment and approach sight distances through the evaluation of existing mathematical and statistical models for the existence of casualties in road accident using spatial statistical and geostatistical analysis method of GIS. Before commercial GIS software programs were available, traffic safety analysts tended to use traditional statistical tests to define hotspots that had significantly higher crash rates. Using the traditional statistical method is inconvenient and inefficient, because traffic engineers must separate road networks into multiple segments with equal lengths (if possible), record crashes for each segment length, use older statistical methods (such as Chi-square test) to define hot spots, and show results via tabulated data. In addition, using traditional statistical methods will not show a geographical relationship between crashes and other environmental variables (Pei-Fen, et al., 2011). GIS software programs simplify this procedure and solve problems by providing graphical data points that can be used for mapping. They have remained one of the most popular tools for visualization of crash data and hot-spot analysis. Schneider et al. (2004) provided an excellent review of the methods, findings, and problems related to using GIS for traffic safety. Previously, some crash datasets were recorded in textual or tabular formats. These data sets were required to be transformed into geographic data before using GIS software programs.

One key element of modern transportation systems is safety. The goal of safety is to minimize the number of accidents and to reduce the severity of injuries for all users, including motorists, passengers of particular vehicles, public transport commuters, cyclist and pedestrians (Haque, et al., 2013; and Machado, et al., 2014). Several studies have been conducted to establish spatial patterns in vehicle or pedestrian crashes to identify the critical locations (Flahaut et al., 2003; Jones et al., 1996). Kim andYamashita analyzed spatial patterns of pedestrian crashes in Honolulu, Hawaii using K-means clustering techniques (Kim and Yamashita, 2007). Baratian- horghi et al. (2015) ranked the high-crash locations by developing a new methodology to predict the first and second possible entry points based on the crash locations and distance from upstream interchanges. Erdogan et al. (2008) performed a research in Turkey to evaluate accidents distribution in a highway, located in the entrance of Afyonkarahisar city. This research used two different methods of kernel density analysis and after identification of hotspots, accidents' conditions were considered hourly, daily and seasonal basis. Moreover, introducing three important hotspots, the researchers offered a number of strategies to the traffic departments in order to solve the problem (Erdogan et al., 2008).

Traffic accidents result in the second highest cost of transportation. These costs result from personal damage (injuries and wounds), fatalities, property damage (to vehicles and other public or private property), degradation of quality of life and decreases in available time for conducting activities and maintaining social relationships (Santos, et al., 2010). Due to the spatial nature of the data used in transport-related studies, geospatial technologies provide a powerful analytical method for studying traffic safety frameworks through the use of spatial analysis. Several studies describe how GIS help the integration of many transportation elements. Meyer and Sarasua (1996) envisioned a common and coordinated database system that will serve all aspects of transportation management such as congestion, pavement, bridges, safety, inter-modal activities, and public transportation. Martin (1993) did a similar study, in which he proves that incorporating GIS in a pavement management program improves the reporting and analysis of data through the production of maps and graphic displays. These problems are solved using hot spot analysis but the main problem in identifying the accident hot spot is to determine the highest rate accident locations and the cause of these accidents. So far, there is no assured definition of hot spots. Given a range of approaches to data collation and variations in areas and locations under consideration, investigating bodies differ in defining what constitutes a hot spot. At certain sites, the level of risk will be higher than the general level of risk in surrounding areas. Crashes will tend to be concentrated at these relatively high-risk locations. Locations that have an abnormally high number of crashes are described as crash concentrated, high hazard, hazardous, hot spot or black spot sites (Levine, et al., 1995).

2.1.1 Traffic Accident Hot Spot Using GIS Tools

” Hotspots”, “blackspots” or high crash locations are sites on a section of a highway that have an accident frequency significantly higher than expected at some threshold level of significance (Hakkert and Mahalel, 1978). Hot spot 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 Zscore 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). Hot spot analysis is a commonly used tool in traffic accident analysis and the goal of this technique is to identify clusters of high and low accident places on a map. This would allow authorities to better allocate resources and protect potential victims of accidents. Several studies have been conducted to establish spatial patterns in vehicle or pedestrian accidents for the identification of critical locations (Jones et al., 1996). Thomas (1996) carried out a study for hot zones using spatial autocorrelation and kernel methods on road segments. Kim and Yamashita (2004), Levine et al. (1995) analyzed spatial patterns of pedestrian crashes in Honolulu, Hawaii using K-means (kernel smoothing density) clustering techniques. These spatial patterns show areas of high pedestrian crashes which have been explained in light of various demographic characteristics. Bello (2005) explored a stratified accident analysis in the city of Richardson. This research focuses on identifying the spatial patterns of traffic accidents to school age kids compared to other traffic accidents in the city, using kernel densities. Sabel et al. (2005) developed a method using kernel estimation cluster analysis techniques to automatically identify road traffic accident hot spots in Christchurch in New Zealand. Today to explore the spatial nature of patterns, a set of spatial analysis and hotspot detection tools included in GIS software's that extract information from the records based on their spatial location.

Identification of hotspots is a systematic process of detecting road sections that suffer from an unacceptable high risk of crashes. It is a low-cost strategy in road safety management where a small group of road network locations is selected from a large population for further diagnosis of specific problems, selection of cost-effective countermeasures, and prioritization of treatment sites. These identified sites are often known by various terms in literature, such as hazardous locations, hotspots, black spots, priority investment locations, collision-prone locations, or dangerous sites (Thakali et al., 2015).

There are various approaches that are aimed at identifying hotspots. One of the well-known approach is using statistical crash models. This approach focuses on relating crashes as a function of potential variables such as road characteristics, geometric design, sight distance, traffic level, and weather factors using historical records (Hauer, 1992, Saccomanno et al., 2001, and Cheng 2008) and subsequently uses these models to identify relatively high-risk sections. The other alternative approach is a geostatistical technique. This technique differs from the previous approach by considering the effects of unmeasured confounding variables through the concept of spatial autocorrelation between the crashes event over a geographical space. Here, some distinctive spatial and geostatistical methods are evaluated and compared: one is the most widely used is the Getis-Ord Gi*, Moran index, Anselin local Moran index, Optimization hotspot analysis, and kernel smoothing density estimation, Empirical Bayesian, Inverse Distance Weighting method and kriging method. The expected crash frequency could also be estimated using a geostatistical technique by considering the effects of unmeasured confounding variables through the concept of spatial autocorrelation between the crash events over a geographical space (Anderson et al., 2009, Keshin et al., 2001, and Blazquez and Clein, 2013).

Anderson et al. applied KSDE method in the City of Afyonkarahisar, Turkey. In this study, the authors were able to detect highly crash risk sections which were highly concentrated in road intersections. Similarly, Keshin et al. and Blazquez and Ceils used KSDE and Moran's Index method to observe temporal variation of hotspots across the road network. Khan et al. used Getis-Ord Gi* statistics to explore the spatial pattern of weather-related crashes, specifically crashes related to rain, fog, and snow conditions.

A noticeable difference in aforementioned geostatistical methods is how spatial correlations are considered. For example, in the KDE method, a symmetrical kernel function, which is a function of bandwidth, is placed on each crash point generating a smooth intensity surface. Then, for a given point of interest, the crash intensity is a summation of the entire overlapping surface due to the crashes. In contrary, in the clustering technique such as nearest neighborhood a threshold value, which determines the extent of clustering in the neighborhood, is pre-specified. If the distance between crash data point pairs is smaller than the threshold value, then these crashes are grouped into the same cluster. Additional criteria such as minimum number of points to be in a cluster can also be specified. This variation in allocating different weights to the crashes occurring in its neighborhood (e.g., KSDE method) or simply grouping crashes into certain clusters clearly indicates that these techniques are likely to have different results in terms of size, shape, and location of hotspots. One of the attractive parts of the KSDE method as compared to other variants of clustering methods is that it takes into consideration of spatial autocorrelation of crashes.

Khan et al. used the Gi* to study the relationship between weather-related crashes and undesirable weather conditions (e.g., snow, rain and fog). Their results showed that weather-related crashes tended to cluster together at different locations, depending on the weather conditions. In this case, the Gi* method allowed each crash to be characterized by a weather condition; thus, identifying spatial clusters of bad weather conditions is equivalent to finding clusters of crashes that are attributed to the same bad weather conditions. This tool is convenient when the analysis prefers the attributes of a crash over the occurrence of a crash. Moreover, this method is simple and easy to implement. This could be one of the reasons that Gi* and KSDE methods are being widely used in road safety.

2.1.1.1 Spatial Hot Spot Statistical Analysis (Getis-Ord Gi*)

The Getis-Ord Gi* test, named for its creators, Professor Arthur Getis and Keith Ord, and pronounced Gi* (Getis and Ord, 1992; Ord and Getis, 1995). finds the locations and extents of point clusters in two-dimensional space. They represent a global SA index. The G_i* statistic, on the other hand, is a local spatial autocorrelation index. It is more suitable for discerning cluster structures of high or low concentration. A simple form of the G_i* statistic is (Songchitruksa, 2010) density can tell you where clusters in your data exist, but not if your clusters are statistically significant. Hotspot analysis uses vectors (not raster's) to identify the locations of statistically significant hot spots and cold spots in data Produces Z scores and P values. A high Z score and small P value for a feature indicates a significant hot spot. A low negative Z score and small P value indicates a significant cold spot. The higher (or lower) the Z score, the more intense the clustering. A Z score near zero means no spatial clustering.

The Getis-Ord statistical analysis, also known as hot-spot analysis, is a method for analyzing the related location tendency (clustering) in the attributes of spatial data (points or areas), and includes the Getis-Ord General G statistic and the Getis-Ord Gi* local statistic (Getis and Ord, 1996, Mitchel et al., 2005). The Getis-Ord General G statistic computes a single statistic for the entire study area, while the Getis-Ord Gi* statistic serves as an indicator for local autocorrelation, i.e., it measures how spatial autocorrelation varies locally over the study area and computes a statistic for each data point (Peters et al., 2015). Compared with the Moran Index, another important spatial autocorrelation analysis method, the Getis-Ord statistic gives more intuitive results and a better visual exploration, and has the advantage of distinguishing high value clusters or low value clusters. Getis and Ord (1992) develop a method for hot and cold spot analysis in a geographical space. The Getis-Ord Gi^* (d) statistic is calculated as follows (Getis and Ord, 1992):

The Getis-Ord local statistics is given as

$$Gi^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j}\right)^2\right]}{n-1}}} \quad \textbf{(1)}$$

Where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between n feature i and j, equal to total number of features and

$$\bar{X} = \frac{\sum_{j-1}^{n} x_j}{n} \quad \textbf{(2)}$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \quad \textbf{(3)}$$

The Gi* may have different definitions of the term “hot-spot” from those used by traffic safety professionals. Hot-spot in Gi* is defined as a point or area where incidents are clustered together by high values, but this is not necessarily a condition affecting traffic safety. It is understandable why Gi* would show an area as normal that does not include a clustered pattern. However, traffic engineers could underestimate the risk in these areas, if they only use a Gi* map. A more cautious method of identifying hot-spots is to combine Gi* maps with other tools, such as the KSDE maps. The Gi* might have an unreliable Z-value for small or spread out data sets because of the limitations that come with a fixed distance band. A distance band should be large enough to ensure that all features have at least one neighbor, but the band should also not be so large that it lets in too many features. Also, for skewed data, the distance band should be large enough to ensure that several neighbors (for the best results, approximately eight) are included for each feature. In rural areas, crash rates tend to be very low (skewed right) and the locations are usually widely spread out. Hence, if one uses the GI* under these circumstances, the user is likely to receive a biased Z-value. The hotspots were categorized based on the Z-value of the Gi* statistic. Categorization can based be on six classification schemes: equal interval, defined interval, quartile, natural breaks, geometric interval, and standard deviation. In the natural breaks scheme, the classes are based on natural categorizing inherent in the data. The break points are identified by the class breaks that best group similar values and maximize the differences between the classes. The features are divided into classes of the boundaries which correspond to relatively big jumps in the data values. This classification scheme was best suited to the present study (Krygier, 2010). In this paper, hotspots were categorized using natural breaks and based on the Jenks ‘algorithm. This algorithm is a common method of classifying data in a choropleth map, a type of thematic map that uses shading to represent classes of a feature associated with specific areas (e.g., a population density map). This algorithm generates a series of values that best represent the actual breaks observed in the data, as opposed to some arbitrary classificatory scheme; thus, it preserves 6 true clustering of data values. Further, the algorithm creates *k* classes so that the variance within categories is minimized (Lewis, 1996). In this study, the categorization was carried out in two categories., high and low.

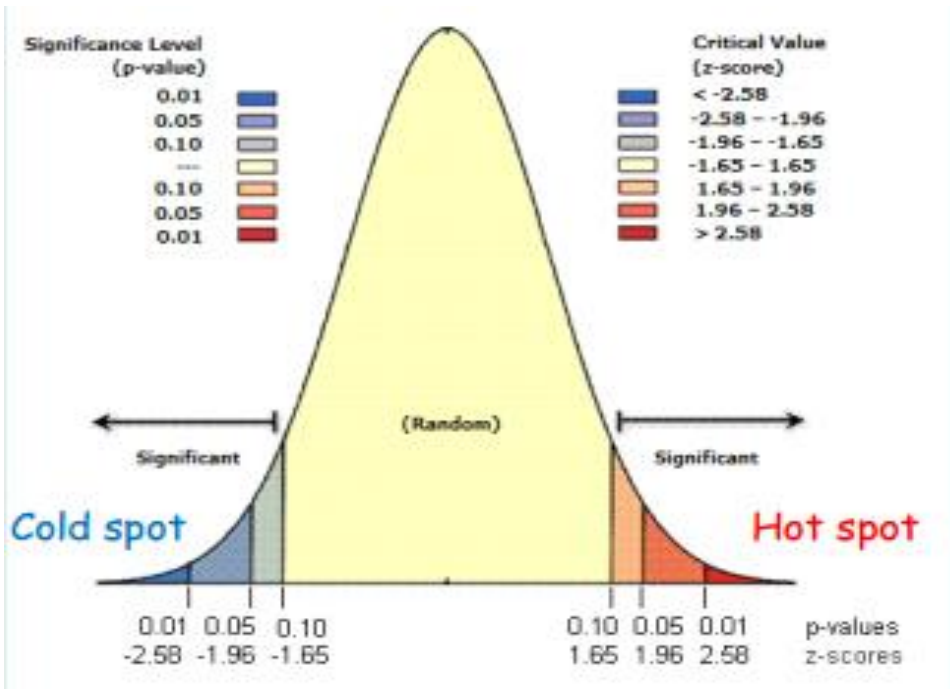


Figure 2.1: Distribution of values (ArcGIS help tool)

Table 2.1: Standard deviation and probability confidence range

Z-Score (Standard Deviation)	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%

2.1.1.2 Moran's "I" Statistic

Moran's "I" statistic (Moran, 1950) is one of the oldest indicators of spatial autocorrelation. It is applied to zones or points that have attribute variables associated with them (intensities). For any continuous variable, Xi, a mean, X bar, can be calculated and the deviation of any one observation from that mean, Sx, can also be calculated. The statistic then compares the value of the variable at any one location with the value at all other locations (Ebdon, 1988; Griffith, 1987; Anselin, 1992). Formally, it is defined as:

$$I = \frac{N * \sum_i \sum_j wij\left(X_i - \overline{X}\right)\left(X_j - \overline{X}\right)}{\sum_i * \sum_j w_{ij} * \sum_i \left(X_i - \overline{X}\right)^2} \qquad \textbf{(4)}$$

Where: N is the number of cases, Xi is the value of a variable at a particular location, i, Xj is the value of the same variable at another location (where i=j), X is the mean of the variable and wij is a weight applied to the comparison between location i and j.

2.1.1.3 Kernel Density Estimation Methods

Based on the statement of Utoyo, B., et al., 2012, Kernel Density Estimation (KDE) is a spatial data analysis capabilities of ArcGIS program. KDE analysis principle is distribute the value of a point in the distribution of events into a smooth curve (smooth) around the point of being simulated by a certain distance. The highest values are at the point under review. Value will be reduced by the farther distance of the point under review until it reaches zero value at a specified distance. In the event of an overlay of the distribution curve of an event with other events in the distribution curve, the value at the point of the overlay is a cumulative value of two or more curves. Analysis using Kernel Density tool produces output in the form of raster data. The main advantage using Kernel Density Estimation (KDE) and K-mean lies in determining the spread of the risk of accidents. The spread of risk can be defined as the area around the cluster in which there may be increased due to an accident. In addition, using this density, arbitrary units of spatial analysis that can be defined and homogeneous for all the areas that can be used as a comparison or even further as a taxonomy (classification). Manepalli et. al. (2011) to identify locations prone to several statistical methods, namely Kernel density estimation (K) and the Getis-Ord Gi* coupled with the method of conceptualization of spatial relationships (CSR). From Brunsdon, 1877; Flahaut, et al., 2005; and Xie, 2008; states that the KDE technique is used to calculate the probability density function of a distribution from which a sample has been observed by centering a probability density function on each of the observed events. The kernel estimator is a non-parametric algorithm that uses a density estimation method. This technique allows one to evaluate the local probability accident occurrence and, consequently, the probable dangerousness of a spatial unit. KDE is the most widely-used nonparametric method in recent decades. The KDE describes the distribution of the location of an event and ignores its association with values. This distribution is characterized by the density of events that occur around a centroid and represents the behavioral patterns of points or lines. In this study, the events are the locations (geocoded) of the accidents (represented by points), and the KDE is used to calculate the probability density function of each accident location. Kernel density analysis is performed by passing a moving window over the data, usually on a regular grid. The densities of the observations within a set radius are calculated for each event located on the grid, and the contributions of each observation are weighted by its proximity to the center of the moving window. Thus, the result of applying KDE is a density map (raster format). The values of each pixel represent the relationships between the concentrations of the events per unit area. In addition, KDE can be used to calculate the density of punctual events (i.e., the density of traffic accidents in a region) or linear events (i.e., the density of a road network in a zone). It is important to highlight the simplicity, satisfactory properties and good results of the KDE method. The areas where events are concentrated are identified by KDE analysis and have the highest accident rates involving accidents. These areas are called black spots (zones that reveal concentrations of accidents). The existence of hot spot zones results from the awareness of the spatial interactions between contiguous traffic accident locations. The most straightforward use of GIS for accident analysis is the examination of spatial characteristics and attributes of traffic accident locations. In fact, the use of GIS has several advantages of the use of non-spatial methods for accident analysis.

2.1.2 Traffic accident

Road traffic accidents have been recognized as one of those adverse elements which contribute to the suffocation of economic growth in the developing countries, due to the heavy loss related to them, hence causing social and economic concern (Anitha et al., 2016). Ethiopia is one of the developing countries with low level of income coupled with high rate of population growth. As part of the developing world, Ethiopia Is predominantly an agrarian country with low level of urbanization. The Economic performance of different sectors of the nation is very low. This Low performance is due to a number of constraints in relation to investment in Different sectors. The existing transport system could be mentioned as one. Transport is an important sector for facilitating different economic activities in the national economy. In Ethiopia, the rate of road traffic accidents (RTAs) is very high; because road transport is the major transportation mechanism along with poor road infrastructure, poor traffic laws enforcement and other factors. The Ethiopian traffic control system archives data on various aspects of the traffic system, such as traffic volume, concentration, and vehicle accidents (Fekadu, 2010).

Road traffic accidents (RTAs), here defined as "An accident that occurred on a way or street open to public traffic; resulted in one or more persons being killed or injured, and at least one moving vehicle was involved. Thus, RTA is collisions between vehicles; between vehicles and pedestrians; between vehicles and animals; or between vehicles and fixed obstacles" (Erdogan et al., 2008), are a major global public health problem, most of it occurs in low- and middle-income countries including Ethiopia. Pedestrians and passengers of commercial vehicles are the most vulnerable in Ethiopia, whereas in high-income countries crashes involve primarily privately owned vehicles with the driver being the main car occupant injured or killed. In the United States of America, for instance, 60% of the fatalities account to car drivers, while in Ethiopia, 5% account to drivers. This implies that in one crash the number of people killed or injured in Ethiopia is about 30 times higher than in the US. Road traffic accidents are a huge public health and development problem in Ethiopia. Its current situation requires a high level political commitment, immediate decisions and actions in order to curb the growing problem (Persson, 2008).

Road traffic crashes pose a significant burden in Ethiopia, as is the case for other developing countries. Currently, developing countries contribute to over 90% of the world's road traffic fatalities (WHO, 2009) and overall road injury disability-adjusted life year (DALYs) increased by 2.5% between 1990 and 2010, with pedestrian injury DALYs increasing by 12.9%, more than any other category (Murray, Lozano, Naghavi, & et al, 2012).

Recently, Ethiopia has become one of the fastest growing non-oil producing economies in the world (AfDB, et al., 2012). Car ownership has grown rapidly at about 7.0% per annum on average (Ethiopian Roads Authority, 2001). The construction of roads is one of the major focal areas of the government to fast-track economic growth. Although the vehicle population growth rate per annum is increasing, the number of total vehicles remains low compared to other developing countries. Currently road density and number of vehicles per 1,000 populations in Ethiopia are low compared with other African countries (Table 2.2).

Table 2.2: Paved road density ratio (source world bank 2012)

	Ethiopia	Uganda	Kenya	Ghana	Nigeria	Botswana	South Africa
Road density by area km/1000 km2 in 2001	4	8	13	58	65	10	60
Motor vehicle (per 1000 people) in 2009	4	8	23	30	31	133	162

Traffic accidents are an unfortunate legacy of the 20th century transport revolution. Although various accident countermeasures are being taken, none have proven decisive. In addition, growing budget deficits are necessitating more effective and efficient infrastructure development (Masayuki Hirasawa, 2005). The rapid and extensive increase in the number of motor vehicles since the 1950's caused certain negative results world-wide. One major problem is traffic accidents. It is possible to categories the factors that cause traffic accidents as related to humans, road and vehicles. At the design stage of a highway, it is important from the highway safety point of view to establish a harmony between the human factor and the other two. It seems that the human factor, comprising driver, pedestrian and passenger, is more dominant than road or vehicle factors in the happening of accidents. However, the control of the road factor is much easier than the human factor. Moreover, by making a geometrically good design, it is even possible to compensate for the other factors and thus decrease the number of traffic accidents. This makes it apparent that in order to establish highway safety, it is important to make a good geometric design. The main criterion for road safety is traffic accidents. The three basic aspects of transport humans, roads and vehicles, are the primary factors in accidents. Human factor seems to be the dominant cause of accidents compared to the others. However, the number of accidents can be seriously reduced if the road factor is evaluated better and highway design is made correctly. (Kumaras et al., 1985). The main problem in identifying the accident hot spot is to determine the highest rate accident locations and the causes of these accidents. This is because the recorded data were not precise enough to carry out the accident study. Some diagnosis system will only be done if the accident data were recorded systematically. The criterion which will be taking into account in identifying the accident blackspot are the accident location areas, types of road where the accidents happened, the road geometry and the light factor (Lim Yu Liang, 2005).

2.1.3 Major causes of car accidents

1. Speeding

Many drivers who caused the accidents ignored the speed limit and drove the design speed of the road segment. Speed kills, and travelling above the speed limit is easy way to cause a car accident. The faster the driver drives, the slowest the reaction time will be if the driver wants to prevent an auto accident.

2. Distracted driving

Distracted drivers diverted their attention by the influence of improper resting or depression caused accidents. Also, distracted drivers caused accidents by diverting attention from the road, to talk on a cell phone, send a text message or chewing chat.

3. Reckless driving

Careless driving ended up in a needless car accident. That is what often happens to reckless drivers who speed, change lanes too quickly or tailgate before causing a car accident. Reckless drivers are often impatient in traffic so pedestrians must be sure to take extra care around aggressive drivers.

4. Wet weather

More accidents occur during the weather got bad so do the roads interims of pavement properties. Car accidents happen very often in the rain and excessive fog because water creates slick and dangerous surface for cars, trucks, and motorcycles and often causes vehicles to spin out of control or skid while braking. To avoid a car accident, drive extra careful when it rains.

5. Running stop signs

Running stop signs caused a number of accidents. Stop signs should never be ignored, but when they are serious car accidents are often the result.

6. Teenage drivers

Youth is wasted on the young, but careful driving is never wasted on young drivers. Unfortunately, teenagers are not often known for their carefulness. When teen drivers hit the roads they don’t always know what to do and that lack of experience ended up causing a number of car accidents.

7. Night driving

Driving in the day light can be hazardous, but driving at night nearly doubles the risk of a car accident occurring. When the driver can't see what is up ahead he/she don't know what to anticipate as he/she drivers towards it. As the sun goes down, awareness of the road and cars around must go up.

8. Design defect

No product is ever made perfectly, and vehicles are no different. Vehicles have hundreds of parts, and any of those defective parts can cause a serious car accident. Many vehicles travelling the segment have had problems with design and maintenance defects.

9. Unsafe overtaking

When drivers don't make safe overtaking, it often leads to a car accident. To prevent a needles overtaking accidents, check signals and estimate the overtaking sight distance.

10. Driving under the influence of drugs

It's not only alcohol that is dangerous when mixed with drivers on the road. Drugs like chat can impair the ability of the driver to fully function as a driver. A driver without clear mind and without control over a body, can lead to serious car accidents.

11. Potholes

Drivers run the risk of losing control of their vehicles or blowing out a tire when they drive over these potholes.

12. Fatigue driving

Most of the car accidents caused by fatigue driving occurred at night.

13. Deadly curves

Many drivers have lost control of their control of their vehicles a long dangerous and sharp curve and caused car accidents. So when approaching these curves, take head of the posted speed limit and drive continually to avoid a car accident.

14. Animal crossing

While drivers are requiring to know the roles of the roadway, animals do not. Caution must be taken when animals cross roads.

15. Careless crossing of pedestrians

Pedestrians crossing roads carelessly encountered car accidents. Caution must be taken while crossing the roads.

16. Defects on roadway construction

The improper design of roadways resulted in a number of traffic accidents. Defects include improper instalment and absence of traffic control systems.

2.1.4 Traffic safety

Traffic safety is an important key and plays an integral role in sustainable transportation development areas. Now days, the main negative impact of modern road transportation systems are loss of property injuries and deaths in road accidents. The success of traffic safety and highway improvement programs hinges on the analysis of accurate and reliable traffic accident data (Anitha et al., 2016). Today, the most negative results of developing modern transportation systems are road accidents with injuries and loss of lives. So, traffic safety is the most critical matter in agencies' transportation strategy. The identification of safety deficient areas on the highway network is aimed at a comprehensive safety program by traffic officials. One of the most important problems that traffic officials face is where and how to implement precautionary measures and provisions so that they can have the most significant impact for traffic safety (Khan et al., 2006). the accuracy and comprehensiveness of traffic accident reports is very important for inputting data and spatial analysis for improving traffic safety analysis. Today there is no consensus among countries which necessary items should be included in traffic accidents reports (Demirel and Akgungor, 2002). Any highway should be able to offer driving safety. In this sense, it is especially valued that roadways satisfy drivers' needs of riding with safety and render an adequate workload. Road safety involves social and economic implications in terms of loss of human lives and of damages; it depends on different factors as road, human behavior and environment. Road geometric design carries out an important role in assuring drivers' safety. Alignment should send right information about its total length so that to guarantee users' expectancies of traveling safely with a low mental workload, since driving is a complex process of acquiring and processing information and making decision, (Altamira, 2010).

The social and economic impacts of road crashes in developing countries are not well understood. It is believed that the implications are immense and that road safety issues require more immediate attention of researchers, professionals, and politicians. Developing countries have embarked on achieving the United Nations Millennium Development Goals as a primary objective; however, the Goals do not explicitly include road safety. Despite the lack of a specific mention of road safety within economic targets, road crashes and economic productivity are linked because primary income earners within families are disproportionately represented among fatalities. At least one study has demonstrated that road crashes have a negative impact on the achievement of the Millennium Development Goals (Ericson & Kim, 2011). Therefore, the road crash problem in Ethiopia merits investigation both in its own right and because of its linkages with other development objectives. Geometric elements should assure sight distances which are one of the most important factors of road safety because they influence the driver’s behavior in choosing the right speed and in making braking and overtaking manoeuvres in safety conditions (Maltinti, 2012).

2.1.5 Traffic Safety and GIS

As the world moves more towards an automobile-based system of transportation, traffic safety has become a global concern (Trinca et al., 1988). Traffic collisions remain a principal cause of death, injury, and economic loss. while there have been, some recent developments linking traffic safety and GIS (Peled & Hakkert, 1993; Pisano, 1993; Meyer & Sarasua, 1992), given how ubiquitous traffic safety problems are, it is surprising to find just how few functioning systems actually exist. This is a pity, since a GIS-based traffic accident decision support system would be invaluable to traffic engineers and transport planners. There are several reasons why traffic safety GIS systems have not lived up to the promise of this particular technology. The principal reason is that much of traffic safety research is oriented towards driver "behavior" or "performance". Evans (1991) defines driver behavior as what drivers, actually do on the road, while performance refers to what drivers are capable of doing--their perceptual and motor skills. Human error is recorded on crash reports as a contributing factor in 70 to 90 percent of all accidents (Wilson & Burtch, 1992). For this reason, much of the research, analysis and modeling of traffic safety focuses on the characteristics, abilities, behaviors, and actions of individual drivers affected by various vehicle, roadway, and environmental factors. Indeed, one of the most basic problems in traffic safety research involves bridging the gap between individual or micro level driver characteristics which produce collisions and more aggregate-level analysis of why, when, and where collisions occur, and how they can be prevented. Certainly, GIS is a valuable research tool for traffic safety, but the challenge is to go beyond accident reconstruction or the representation of a particular event to a more aggregate level of analysis and understanding (Karl Kim and Ned Levine, 1996).

In spite of the fact that all states collect it, the quality of collision data is poor (O'Day 1993; Shinar, Treat, & MacDonald, 1983). Typically, collision data are collected by state or local police officers dispatched to the scene of a crash. Different states have different procedures and policies, standards and thresholds, and they deploy varying levels of resources for the collection, management, and distribution of crash data. Some states have numerous overlapping jurisdictions, state police agencies and local authorities, and the latter collect and manage the data in a variety of different systems (Parke & Kim, 1994).

While there are some national databases (e.g., the Fatal Accident Reporting System, National Accident Sampling System, etc.) these have not been developed with a GIS component. Moreover, the quality of locational information within police crash reports is poor. Locations are generally referenced in terms of distance to or from intersection, mile markers, or other reference points. The lack of explicit geographical standards for police crash reporting, as well as the larger problems associated with non-uniform mapping across different spatial datasets in terms of scale, level of detail, projection, and features, has also made it difficult to combine spatial datasets.

There is another more serious constraint which relates to GIS technology. It pertains to the difficulties associated with spatial analysis, and, in a phrase, it is the problem of "going beyond pin maps". In spite of the developments made in spatial analysis (Anselin, 1988; Anselin & Getis, 1992), applications have not yet filtered down to the level of traffic safety analysts. Except for geographers and some planners who are accustomed to thinking and reasoning about problems across space and time, many people are not used to conducting spatial analyses. This is unfortunate because, beyond generating maps, there is also a need to utilize GIS for problem identification as well as determination of appropriate interventions.

GIS is currently used to aid in the identification of "high accident locations", there is need for more statistical sophistication in determining what constitutes a "high accident location". That is, there is a need to find out what happens when, in addition to vehicle miles traveled, locations are also controlled for roadway characteristics, weather, socio-economic characteristics of drivers, and other factors influencing the density of collisions. In other words, more sophisticated spatial analyses would permit the design, implementation, and evaluation of new interventions that are meant to increase traffic safety. Put succinctly, the viability of a GIS system rests on an assumption of "spatial literacy", among not just the programmers and system operators, but also the planners and policy makers and others who interact with the system.

For a traffic safety GIS, a vector system may be more practical for agencies just starting out with this technology. Motor vehicle crashes are easily linked to a vector system. The crash locations are defined by points while the streets are defined by lines, and neighborhoods or areas are represented by polygons in "desktop" mapping packages. With this data syntax, each geographical object (points, segments, polygons) that exist can be manipulated independently of other objects. Most larger systems, however, such as Arc/Info and Intergraph use a hierarchical data structure in which polygons refer to previously defined lines. The latter, in turn, refer to previously defined points. Any change to an object requires a re-indexing of all related files in the system, since internal data integrity must be retained. This requires more data processing steps, more memory requirements and more processing time. Hence while the prototype system was principally developed using desktop mapping packages, certain applications were carried out utilizing Arc info.

2.1.6 Geographic Information System(GIS)

A geographic information system (GIS) is a computer system for capturing, storing, querying, analyzing and displaying geographic data. GIS represents a new paradigm for the organization of the information and the design of information system, the essential aspect of which is the use of concept of location as the basis of structuring of information systems. GIS technology can be viewed as an offshoot from two major software technologies i.e., database management system (DBMS) and computer aided design (CAD), with the addition of specialized functions for managing and analyzing spatial data i.e., data that can be referenced to a geographical location. The objective of any GIS system is to integrate hardware, software, and data for capturing, managing, analyzing, and displaying all forms of geographically referenced information. GIS allows us to view, understand, question, interpret, and visualize data in many ways that reveal relationships, patterns, and trends in the form of maps, globes, reports, and charts. GIS is a powerful tool for managing large amounts of heterogeneous data. A GIS can be effectively used to prioritize black spots on roads. The capability of GIS to link attributes data with spatial data facilitates prioritization of accident occurrence on roads and the results can be displayed graphically which can be used for planning and decision making. In the present study, GIS analysis is performed using

ARCGIS 10.4.1package. GIS benefits organizations of all sizes and in almost every industry. There is a growing awareness of the economic and strategic value of GIS. The benefits of GIS generally fall into five basic categories: Cost saving and increased efficiency; Better decision making; Improved communication; Better record keeping; Managing geographically and better analyzing methods.

2.1.7 GIS Uses and Basics in Transportation

A Geographic Information System (GIS) is a computer system for capturing, storing, querying, analyzing and displaying geographic data (41). GIS represents a new paradigm for the organization of the information and the design of information system, the essential aspect of which is the use of concept of location as the basis of structuring of information systems. Traffic accident analysis means to investigate the causes of accidents, to determine hazardous locations (Accident Prone Locations) and to determine to enhance road features, to evaluate traffic safety and enhancement (Samaila Saleh). GIS can be classified as either vector or raster, and for spatial phenomena that are continuous and varying (e.g., topography, land uses, soil types), a raster representation is more useful. Also, operations between layers are more easily conducted using a raster, since operations are applied only to each individual cell. However, raster systems can become very large because the geographical characteristics have to be stored for every data layer. That is, unlike a vector system, where one geographical file can be linked to multiple variables, each data layer in a raster system represents a separate geographical and attribute file. Thus, storage and processing speed requirements can be very demanding. The major raster GIS programs utilize workstation or mainframe computer technology, for example Erdas, GRASS, and Vicar (Karl Kim and Ned Levine, 1996). GIS permits users to display database information geographically. It can also provide a common link between two or more previously unrelated databases. The most useful aspect of GIS as a management tool is its ability to associate spatial objects (street names, milepost, route number, etc.) with attribute information (accidents, cause, etc.). Most of the documents reviewed consider the use of GIS in transportation under either for general data maintenance (primarily inventory of transportation-related incidents) or for simple data analysis (Lim Yu LIANG, 2005).

Several studies describe how GIS help the integration of many transportation elements. Meyer and Sarasua (1996) envisioned a common and coordinated database system that will serve all aspects of transportation management such as congestion, pavement, bridges, safety, inter-modal activities, and public transportation. Martin (1993) did a similar study, in which he proves that incorporating GIS in a pavement management program improves the reporting and analysis of data through the production of maps and graphic displays. GIS has been proven to work well in addressing transportation problems related to safety. Affum and Taylor (1996) described the development of a Safety Evaluation Method for a Local Area Traffic Management (SELATM), which is a GIS-based program for analyzing accident patterns over time and the evaluation of the safety benefits. GIS can also be implemented in determining roadway and surface conditions. This was proven by Gharaibeh et. al. (1994) when they proposed to use GIS to obtain statistical and spatial analyses of roadway characteristics such as safety, congestion level and pavement conditions. In recent years, there has been much discussion about GIS technology and applications across a wide variety of settings. Moreover, there have been many GIS-related developments in transportation planning and engineering (FHWA, 1993; Lewis, 1990; Kim and Levine, 1996). The power of them is rooted in the fact that GIS allows inferences to be drawn about the spatial nature of the data. Examples of GIS applications in transportation include pavement management systems that work with road segments, optimal vehicle routing, Automated Mapping Facilities Management (AMFM) used for infrastructure management, drainage design, traffic modeling and accident analysis, demographic analysis for funding justification, and the option of displaying any form of tabular data that has a spatial component.

2.2 Possible Cause of High number of Accidents

The reasons for the relatively high number of road traffic accidents includes the following (47). Some of them are: Lack of driving skills; Poor knowledge of traffic rules and regulations; Violation of speed Limit; Insufficient enforcement; Lack of vehicle maintenance; Animal drawn carts and animals; frequently using in main highways; Lack of safety conscious design and planning of road network; Disrespect of traffic rules and regulations; Lack of general safety awareness by pedestrians and; Lack of medical facility in general; which increase the severity of accidents. A review of the major assessments regarding the effect of road environment, drivers, and vehicle factors on traffic safety will be discussed in the following sub section

2.2.1 Users Environment

In Ethiopia, the posted speed is 30 km/hr in towns since the roadways are often used by mixed traffic, especially pedestrians and animal drawn carts. Observations at some sites indicate that there are many inconsistencies in design speed as well as posted speed (Mekonnen, 2007), and passing lanes are rarely provided. Although only 16.1% of the population live in urban areas, these areas account for the majority of road traffic crashes due to the clustering of vehicles in cities, for instance vehicles in Addis Ababa constitute 77% of country's vehicle population (Akloweg, et al., 2011; Population Census Commission, 2008). This implies that road traffic crashes typically occurred in busy areas. About 74.4% of fatal crashes occurred on paved roads and only 25.2% on gravel and earth roads, whereas paved roads make up only 18.62% of the road network, while 28.31% are gravel, and 53.07% earth (Ethiopian Roads Authority, 2011). Paved roads are heavily trafficked in terms of loads and volumes, mainly because the provision of asphalt paved surfaces depends on the volume of traffic. Unfortunately, actual traffic volumes are not known and so rates cannot be calculated to enable comparison between road surface types (Getu et al., 2011).

2.2.2 Road Environment

Road environment has impacts on occurrences of road traffic accidents. In developed countries, there have been continuous efforts to meet the safety standards of roads through safety audit during the planning, designing, and operation stages. In Africa road network is expanding fast, maintenance standards have started improving lately, and there is potential for improving the safety standards of the roads (Terje, A., 1998). In Ethiopia, the police have limited road and traffic engineering skill in general and thus they underestimate the contribution of roads and road environments to traffic accidents and especially they lack trainings on the subject areas (26).

2.2.3 Road Geometric Design Elements

"The geometric design of a road is said to be consistent when its alignment is in agreement with the drivers´ expectations, so that no unexpected driving maneuvers are needed. Drivers' expectations related to the actual geometry of the road sections to come influence their level of driving attention and behavior. The amount of information drivers has to process in relation to the time available for such analysis in known as mental workload. Naturally, mental workload is heavier as the geometry of the road is more complex or less predictable" (Altamira et al., 2010).

The geometric form of a road is a two dimensional alignment which is presented in two projections, the horizontal and the vertical alignment. The horizontal alignment consists of three elements, the straight, the circular curve and the transition curve. The vertical alignment consists of two elements, the straight and the vertical curve (crest or sag). Other elements of the alignment are sight distance and super elevation. From Demetrios E. Tonias point of the geometry of a typical highway comprises three basic components: cross-sectional geometry, horizontal geometry, and vertical geometry. The type, size and number of elements used in a highway are directly related to its class and the corresponding function of the highway.

2.2.4 Horizontal Alignment

Berhanu, G. 2000 stated that horizontal alignment of a roadway is defined graphically using a series of straight-lines, circular curves (with a constant radius), and spiral curves, whose radius changes regularly to allow a gradual transfer between adjacent road segments with different curve radii. Many factors, including terrain conditions, physical features and right-of way considerations, affect the design of tangent and curve sections. Inconsistent horizontal alignments of highways with sharp curves are main for their significant and confrontational safety impacts. Several researchers have shown that traffic accident in curves are so much as high. Zegeer, et al., 1992, stated that the accident rate in curves are 1.5 to 4 times higher than in tangents. Another research by Lamm, R., et al., 1999 said that from 25% to 30% of fatal accidents occur in curves. Approximately 60% of all accidents happened in horizontal curves are single-vehicle off road accidents and the proportion of accidents on wet surface is high in horizontal curves. Similar study on safety proportional by Council, 1998, stated that 62% of fatalities and 48% of other accidents occurred in curves, the first maneuver that led to the accidents was made at the beginning or at the end of the curves. Spacek, 2000, on his behavior and accidents occurrence on curves of two lane studies stated that there is a relative relationship between the speed reduction in the curve and the probability of accidents when the required speed reduction in the curve is high, the probability of error and accident also increase. Moreover, the risk is even higher when the speed reduction is unexpected or unusual. An important factor, which affects the occurrence of road traffic accidents in terms of frequency and severity, is the road alignment.

Recent studies such as the Federal Highway Administration, 2006., shows that accidents on horizontal curves are causes for concern in all countries. It has been shown in past researches that horizontal curves experience crash rates of up to 4 times the rates on tangent sections, all else being equal. The following traffic, roadway, and geometric features that influence safety at horizontal curve sections have been identified (Hummer, et al., 1987): Traffic volume on the curve and traffic mix (such as the percentage of trucks); Curve features (such as degree of curve, curve length, super elevation, presence of transition curves); Cross sectional curve elements (such as lane-width, shoulder width, shoulder type, shoulder slope); Curve section roadside hazard features (such as clear slope, rigidity, and types of obstacles); Stopping sight distance on curve (or at curve approach); Vertical alignment on horizontal curve; Distance to adjacent curves; Distance of curve to nearest intersection, driveway; Pavement friction; Presence and type of traffic control devices (signs and delineation).

Based on the research of effects of alignment in highway by Glennon, (1985), indicated that hazardous roadside design are the primary cause of crashes at horizontal curve sections. The effect of degree of curvature, tangent length, and sight distance on accident rates at horizontal curve sections was significantly influenced by the degree of curvature. The relationship between accident rate and horizontal curvature was suggested for the main that the critical radius, i.e. the radius below which accident risk increase sharply, is approximately 400m as shown in figure below (McLean, 1986).

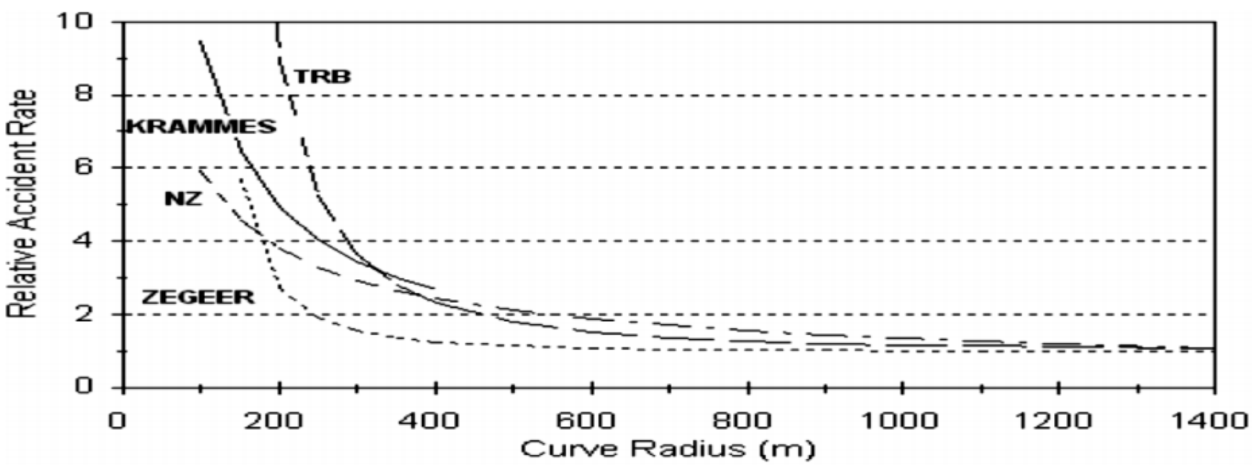


Figure 2.2: Relationship between Accident risk and Horizontal Radius (McLean. J. 1996)

A horizontal curve with R > 1300m can be regarded as the standard against which other curves can be assessed, and is therefore afforded a ‘relative risk ratio’ of 1.0. In all studies of horizontal curves, it is apparent that an improved understanding of the safety risk associated with horizontal curves can be found by considering the nature of the tangent prior to the curve (McLean, 1996).

Hauer, 2001 stated that sections with curvature of between 5 and 10 degrees have at least twice the crash rate of sections with curvature of 1 and 5 degrees, and sections with curvature of between 10 and 15 degrees have crash rates four times as great. In terms of curve radius, 200 m seems to be the point below which crash rate greatly increases. The evidence suggests that curve flattening is highly effective in reducing crashes. There are several elements of horizontal alignments which are associated with horizontal curve safety. It states that safety of a horizontal curve, its accident frequency and severity is determined by: Internal features: Radius or degree of curve; Super elevation, etc. External features: Density of curves upstream; Length of the connecting tangent sections; Sight distance, etc. that influence driver expectation and curve approach speed. Road geometric improvements or remedial measures used to ensure safety at deficient horizontal curves include the following points: Lengthening the radius of a curve to reduce accidents in sharp horizontal curves. Improving warning and guidance provided for drivers: Better sight distance, curve conspicuity; signing and marking, delineation. minor geometric improvements, modifications to the shoulder and roadside conditions; curve straightening; roadway widening at curve sections; super elevation improvements.

$$R_{min} = \frac{V_d^2}{127(e+f)} \quad \textbf{(5)}$$

Where: V_d is design speed (km/h); e is Maximum superelevation (%/100); f is Side friction coefficient (given in Tables 8-1 and 8-2, and Figure 8-1)

Table 2.3: Minimum Radii for Horizontal curves, 8% superelevation (ERA, 2001)

Design speed Vd km/hr.	20	30	40	50	60	70	85	100	120
Min. Horiz. Radius R(m)	15	30	50	85	125	175	270	395	630
Side Friction Factor (f)	0.18	0.17	0.17	0.16	0.15	0.14	0.1	0.12	0.1

Table 2.4: Minimum Radii for Horizontal curves, 4% urban Superelevation (ERA, 2001)

Design speed Vd km/hr.	20	30	40	50	60	70	85	100	120
Min. Horiz. Radius R(m)	15	35	60	100	150	215	320	490	810
Side Friction Factor (f)	0.4	0.32	0.25	0.22	0.19	0.17	0.1	0.12	0.1

2.2.4.1 Sight Distance

Sight distance is an important design element of curved highway sections and is the downstream length of a highway section that drivers can see. On long straight sections, sufficient sight distance is provided naturally. On horizontal curves, roadside objects on the insides of the curves may block sight lines and therefore reduce sight distance. Roadside objects likely to block sight lines include side slopes in cut areas, buildings, trees, retaining walls, and longitudinal safety barriers. Such reductions in sight distance may lower both safety and mobility (1). The existence of an ASD greater than the distance necessary to perform specific maneuvers is a pre-requisite to create safe driving condition in a harmonious road environment. Sight distance is one of the fundamental elements in achieving safe and efficient operation of highways. A driver’s ability to see ahead is of the utmost importance in the safe and efficient operation of a vehicle on a highway. Observing sight distances influences the functionality of a two-lane rural road because it depends on the possibility for faster vehicles to overtake slower vehicles indeed when overtaking is not possible, platoons arise and the level of service of the road decays. the distance of unobstructed view must be compared with three distances: the stopping sight distance, the passing sight distance and the changing lane sight distance. The distance of unobstructed view is the length of the road stretch that the user can see in front of him regardless of traffic, weather and lighting conditions. The three distances are defined as follow: Stopping sight distance is the minimum distance necessary for a driver to safely stop a vehicle in front of an unexpected obstacles; “Passing sight distance is the length of road stretch necessary to safely complete a passing manoeuvre when the presence of another vehicle coming from the opposite direction cannot be excluded; Changing lane sight distance is the length of road stretch necessary to change lanes in deviation manoeuvre in correspondence of particular point like intersections, exits and so on" (Cafiso et al., 2010). So, the stopping sight distance must be guarantee along the full length of the road, the passing sight distance must be guaranteed along a suitable percentage length of two lane roads according to the level of service and the changing lane sight distance must be guaranteed along roads that have more than one lane for each traffic direction or in correspondence of a special point (intersections, turnings, etc.) (F. Maltinti, 2012).

"Driving a vehicle is a continuous process of acquiring and processing information, and making decisions in accordance with what the driver is exposed to: alignment of the road he is presently seeing and his anticipation of geometric conditions to come; environment; traffic control systems present and the purpose of his trip, among other factors. The results of these analyses are the actual driving manoeuvers he performs along the road. These driving decisions can be evaluated in terms of operating speed, clearance to lane edge, vehicle trajectory, etc. Of utmost importance in highway design is the sequence of geometric elements so that there is adequate sight distance for safe and efficient traffic operation, assuming adequate light conditions, and drivers' visual acuity" (Gomez et al., 2010).

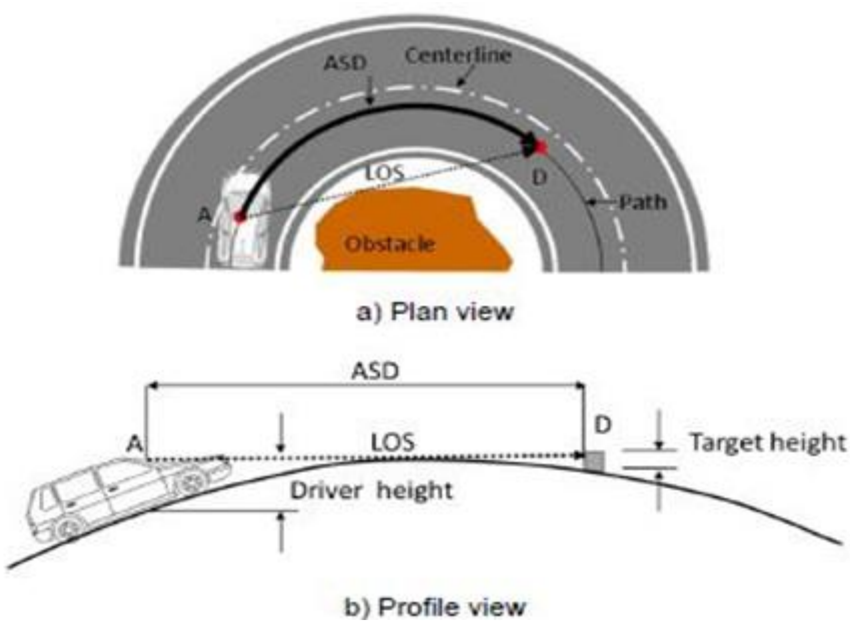


Figure 2.3: Available Sight Distance (Castro et al., 2008)

2.2.4.2 Stopping Sight Distance

The stopping sight distance is the distance ahead that a driver should be able to see so that he or she can bring his or her vehicle to a safe halt short of an obstruction or object on the road. It is crucial to consider both the vertical as well as the horizontal stopping sight distance. In curves, the horizontal stopping sight distance might be restricted by crash barriers, bushes etc. In vertical curves, sight distance can be restricted due to inclines (Brockenbrough, 2009).

$$d = (0.278) * (t) * (v) + \frac{v^2}{254f} \tag{6}$$

2.2.4.3 Passing Sight Distance

PSD is the minimum distance visible ahead which is to be provided to enable the driver of vehicle moving at design speed overtakes the slow-moving vehicles. For efficient movement of traffic in high priority road networks like Expressways, National Highways and State Highways, provision of this facility for overtaking is mandatory for attaining high Level of Service (LOS) and thus, it is equally important to design the highway for PSD (Deepjyoti et al., 2015).

2.2.5 Sight distance and traffic safety

An important element in ensuring driver safety and maintaining a roadway's operational efficiency is providing adequate sight distance the length of roadway ahead visible to the driver. Sight distance applies to four conditions that may arise when setting a project's horizontal alignment. Considering the fact that drivers receive 95% of all information from the environment by sense of sight and that the lack of visibility is direct or indirect cause of almost 40% of all traffic accidents on suburban roads (LJ. Simunovic, 2011), it can be stated with certainty that a significant role in road design belongs to sight distance testing. In this paper stopping sight distance on horizontal curves was observed. Many road traffic accidents (RTAs) occur due to information errors (e.g. the view was obstructed by a tree or the driver didn't perceive the oncoming vehicle due to an incline). In order to minimize the likelihood that such errors occur "the sight distance must be sufficient for a driver to perceive potential hazards which may or may not be conspicuous (Basacik et al., 2007)." Additionally, it must be ensured that the driver has enough time to process the perceived information and to make appropriate decisions. As a result, sight distance is a crucial aspect in highway engineering. In highway engineering, the sight distance is determined by various factors such as the driver eye height 1,07 m, brake-reaction time assumed value 2.5 sec, and object height 0,61 m (AASTHO, 2001). According to many road design standards, the existence of appropriate available sight distance along roads is of primary condition for traffic safety. “The available sight distance (ASD) in front of the driver to detect possible conflicts with unexpected obstacles is fundamental for traffic safety”. Maltinti, (2012). To avoid problems related to the limitation associated with the use of digital terrain models typically employed in RDS, the Geographic Information Systems (GIS) software can use DSM which are more flexible in the modelling of sight obstruction due to vegetation, landscapes, and vertical surfaces largely diffused in rural areas. The paper deals with the evaluation of GIS in the estimation of ASD in a typical rural road where the density of sight obstruction along the roadside is relatively high (Sanchez et al., 2014).

2.2.6 Vertical Alignment

“Vertical alignment is layout of the road on a vertical plane and consists of grades and curves. The vertical alignment of a highway consists of straight sections of the highway known as grades, or tangents, connected by vertical curves. The design of the vertical alignment therefore involves the selection of suitable grades for the tangent sections and the design of the vertical curves. The topography of the area through which the road traverses have a significant impact on the design of the vertical alignment” (Brockenbrough, 2009). The minimum lengths of crest vertical curves are determined mainly by sight distance requirements. These lengths generally are satisfactory from the standpoint of safety, comfort and convenience. An exception maybe at decision areas, such as intersections and approaches to ramp exit gores, where adequate sight distance requires longer lengths. Passing sight distance seldom can be attained on a crest vertical curve simply by lengthening the curve. Excessively long vertical curves often reduce the length of passing opportunities on the adjacent tangent sections on either side of the crest. They also can adversely impact roadway and roadside ditch drainage systems. Sag vertical curves use four different criteria for determining their lengths: headlight sight distance; passenger comfort; drainage control; and general appearance the primary control used in design is headlight sight distance. Properly designed vertical curves should provide adequate sight distance, safety, comfortable driving, good drainage, and pleasing appearance.

Poor condition of the horizontal and vertical alignments of a road can result in visual effects, which contribute to accidents and are detrimental to the appearance of the road. It may be difficult for a driver to appreciate the sight distance available on crust curve and he/she may overtake when it is insufficient for him to do so safely (Ross, 1994). Effects of vertical curve have been summarized in such a way that steep grades have higher accident rates than mild ones and grades of less than 6 percent have little effect, but grades steeper than this are associated with higher accident rates (Berhanu, 2000).

Grades

Iyinam, et al., 2000; Christo, et al., 1999, and Sarbaz, et al., 2000, stated that maximum grade is not a complete design control. The length of an uphill grade is important as well, because it affects capacity, level of service, and delay when slow moving trucks, buses, and recreational vehicles are present. The effect of grade on the performance of heavy vehicles is more pronounced than that for passenger cars. The speed of a heavy vehicle can be significantly reduced if the grade is steep and/or long. In order to limit the effect of grades on vehicular operation, the maximum grade on any highway should be selected with care. The selection of maximum grades for a highway depends on the design speed and the design vehicle. It is generally accepted that grades of 4 to 5 percent have little or no effect on passenger cars, except for those with high weigh/horsepower ratios, such as those found in compact and subcompact cars. As the grade increases above 5 percent, however, speeds of passenger cars decrease on upgrades and increase on downgrades. Accident rate is highly increasing with increasing of road grades at the point of high slope section. For the low radius horizontal curves, the accident number is high. Steeper grades above 6 per cent are generally associated with higher accident rates.

Table 2.5: Minimum values for crest Vertical curve (ERA, 2001)

Design speed (km/hr)	Rate of Vertical Curvature, K for stopping sight distance	K, for Passing Sight Distance
20	2	10
30	3	50
40	5	90
50	10	130
60	18	180
70	31	250
85	60	350
100	105	480
120	210	680

Table 2.6: Minimum values for Sag Vertical curves (ERA, 2001)

Design speed (km/hr)	Rate of Vertical Curvature, K for stopping sight distance	K, for Passing Sight Distance
20	2	10
30	4	50
40	8	90
50	12	130
60	18	180
70	25	250
85	36	350
100	51	480
120	74	680

2.2.6.1 Road cross-section Elements

A cross-section is normally consisting of the carriageway, shoulders or curbs, drainage features, and earthwork profiles. These terms are defined in the Definition portion of the manual text; major elements are repeated here for clarity (Ethiopian Road Authority, 2002). In various studies revealed that these elements are the most important road related features which affect road safety.

Carriageway- the part of the road constructed for use by moving traffic, including traffic lanes, auxiliary lanes such as acceleration and deceleration lanes, climbing lanes, and passing lanes, and bus bays and lay-byes.

Roadway- consists of the carriageway and the shoulders, parking lanes and viewing areas

Earthwork profiles- includes side slopes and back slope

Lane Widths A feature of a highway having great influence on safety and comfort is the width of the carriageway. Lane widths of 3.65m are used for Design Classes DS1 and DS2. The extra cost of 3.65 m above that for 3.0 m is offset to some extent by a reduction in cost of shoulder maintenance and a reduction in surface maintenance due to lessened wheel concentrations at the pavement edges. The wider 3.65m lane also provides desired clearances between large commercial vehicles on two-way rural highways.

Shoulders A shoulder is the portion of the roadway contiguous to the carriageway for the accommodation of stopped vehicles; traditional and intermediate non-motorized traffic, animals, and pedestrians; emergency use; the recovery of errant vehicles; and lateral support of the pavement courses.

Normal Cross fall Normal crossfall (or camber, crown) should be sufficient to provide adequate surface drainage whilst not being so great as to make steering difficult. The ability of a surface to shed water varies with its smoothness and integrity. On unpaved roads, the minimum acceptable value of crossfall should be related to the need to carry surface water away from the pavement structure effectively, with a maximum value above which erosion of material starts to become a problem. The normal crossfall should be 2.5 percent on paved roads and 4 percent on unpaved roads. Shoulders having the same surface as the roadway should have the same normal crossfall. Unpaved shoulders on a paved road should be 1.5 percent steeper than the crossfall of the roadway. The precise choice of normal crossfall on unpaved roads will vary with construction type and material rather than any geometric design requirement. In most circumstances, crossfall of 4 percent should be used, although the value will change throughout the maintenance cycle.

2.2.7 Speed

Lamm, et al., 1999 and Oxley, et al., 2004 stated that operating speed is influenced by several factors related to the driver, road and road side conditions, vehicle characteristics, traffic conditions and weather conditions and road alignment is absolutely the most important factor among road characteristics that influence driver’s speed. A speed variation along a road is a major factor in accident causation which reduces driver’s ability to control the vehicle, negotiate curves or maneuver around obstacles on the roadway. Therefore, increases the chance of running off the road or into an oncoming vehicle with speed has the greatest effect on traffic safety and the accident rate decreases as design speed increases from 60 to 80 km/h. so many researchers developed rules to help designers choose horizontal alignment sequences that would reduce operating speed variations along a route.

2.3 Traffic volume and vehicle classification

As Milton, and Mannering, 1998; believes that traffic volume is positively correlated with incidences of traffic accidents. As the number of vehicles increases through a section, the exposure to potential accidents and number of conflicts increases. Traffic volumes are considered during horizontal curve widening. Vehicle classification is also an essential aspect of traffic volume evaluation. The types of vehicles are defined according to the breakdown adopted by ERA for traffic counts: cars; pick-ups and 4-wheel drive vehicles such as Land Rovers and Land Cruisers; small buses; medium and large size buses; small trucks; medium trucks; heavy trucks; and trucks and trailers. This breakdown is further simplified, for reporting purposes, and expressed in the five classes of vehicles (with vehicle codes 1 to 5) listed in Table 2.7.

Table 2.7: Vehicle classification by class (Road Asset Management, 2001)

Vehicle code	Type of Vehicle	Description
1	Small Car	Passenger cars, minibus(up to 24-passenger seats), Taxis, Pick-ups, and Land Cruisers, Land Rovers, etc.
2	Bus	Medium and Large size buses above 24 passenger seats
3	Medium Truck	Small and medium sized trucks including tankers up to 7 tons load
4	Heavy Truck	Trucks above 7 tons load
5	Articulated Truck	Large trucks or Semi-trailer and Tanker trailers

ERA has also classified vehicles into 8 categories

Table 2.8: Vehicle classification (ERA, 2001)

Classification	Description
A. Passenger Vehicle	
Car	Cars and Taxi
Land Rover	Land rovers, Jeep, Station Wagon, Land Cruisers
Small Bus	Small Buses up to 27 passenger seats
Large Bus	Large Buses above 27 passenger seats
B. Fright Vehicles	
Small Truck	Small and Light Trucks of 3.5 tons load
Medium Truck	Medium sized Trucks of 3.6 to 7.5 tons load
Heavy Truck	Trucks and Tankers of 7.6 to 12 tons load
Truck- Trailer	Truck Trailers and Tanker Trailer above 12 tons load

(Source: Road Asset Management and Contract Implementation Coordination Directorate)

2.4 Driver characteristics

Garber N. J., Hoel L. A., 2001 stated that the driver's behavioral characteristics, such as inattention, fatigue, inexperience, and risk-taking behavior (speeding, drunk-driving, and failure to wear a seat-belt), have all been identified as factors that significantly contribute to increased crash and injury risk on roads. The driver's age also plays an important role in crash causation. The key components involved in run-off crashes is the driver's ability to control both speed and direction. It occurs when a driver is faced with a piece of unexpected or unusual information, which leads him or her to over-correct at a large steering angle and the causes of this include: Distractions such as talking, eating; Aggressive and conservative in nature; Physical problems like colorblindness, night blindness, drugs or medication; Drowsiness, fatigued, illness, or blackout; Speeding; and failure to obey signs, signals or traffic officers.

2.5 Vehicle Characteristics

Garber, and Hoel, 2001 on their traffic and highway engineering book and Saravade, 2005, on his study of fatal rollover crashes in the state of Florida and Haworth, et al., 1991, all believes that mechanical problems in vehicles are additional important factor that contributes to traffic accidents. Faulty brakes, worn tires and other vehicle defects affect the controlling of a vehicle, especially at high speeds. It has been observed that at high speeds the tires may blow out leading to loss of control. Vehicle and roadway interactions, such as skid resistance, play a major role in stopping the vehicle from encroaching the off-road features, like the shoulder, median and other traffic signage. There are problems associated with heavy vehicles on rural roads. First, there is the problem of the mix of vehicles on rural roads, and the risk of injury to vehicle occupants involved in a crash with a heavy vehicle. Secondly, there is the problem of single-vehicle crashes involving heavy vehicles, and the contribution of road features to these crashes. Although crashes involving trucks are less common than those involving smaller vehicles, they tend to be more severe than those involving smaller vehicles.

2.6 Road Traffic accident recording system in Ethiopia

Regional departments of the Traffic Police are responsible for the recording of all traffic accidents under their jurisdictions. The Federal Police Commission is responsible for national accident data compilation and processing. In each Region's District Police Station, accident data are reported manually. The traffic police accident data form contains accident classification, date, time, day of the week, year, age of the driver, sex, education of the driver, ownership of the vehicle, service year of the vehicle, defects of the vehicle, location of accident, road traffic condition, road surface condition, road junction type, weather and illumination condition, collision type, and property damage and parties injured (age, sex, physical fitness and the like) etc. Monthly reports are submitted to pertinent the Regional Police Commissions. A yearly report from the Regional Police Commission will then be submitted to the Federal Police Commission to generate national accident statistics.

CHAPTER-THREE

3. RESEARCH METHODOLOGY

3.1. Historical Background of Study Area

The Gibe valley – Sekoru road segment of the Jimma -Addis Ababa Highway is located is located in Sekoru District, Jimma Zone, Oromia, and some parts on Yem Special District, SNNP, Southwest Ethiopia with geographical location of 37º26’00’’ Longitude, 7º55’28’’ Latitude at Sekoru and 37º35’45’’ Longitude 8º13’48’’ Latitude at Gibe bridge and an altitude average of 1090 to 1845 m above mean sea level. An average area over 177.3 km^2 and is a dense, grassland biome, agricultural lands. From the design of the road segment by Studio Pietrangeli-Rome consultants the road segment lies at mountainous and escarpment land scape. This section of highway is existed trunk road of two-lane rural highway. This road segment begins (0+00 km) at the Gibe bridge which is 185.7 km from Addis Ababa at an elevation of 1090 m and after a mountainous, uphill course of 12.5 km comes out of the Gibe gorge at Abelti, elevation 1669, with an average grade of 7%. The maximum. grade is between 5 and 9%. From Abelti the road route runs on the crest line between the Gibe and the Gilgel Gibe basin for about 13.6 km on hilly route until Kumbi which is 212 km from Addis Ababa, where the route enters the Gilgel Gibe basin crossing several tributaries of the main river on a predominantly mountainous alignment to reach, after 40 km, the town of Sekoru, 250 km from Addis Ababa. The segment of the road under study passes through the towns of Natri, 226 km from Addis Ababa and, Saja that is 238 km from Addis Ababa. Table 3.1 shows that 50.71% of the road segment Sekoru Town-Gibe valley is curve while straight sections cover 49.29% of the road segment.

Table 3.1: Curves and tangent in Sekoru town-Gibe valley (Studio Pietrangeli-Rome)

Street (km)	Length (km)			% Curves
	Curves	Straight	Total	
0+00 up to12.8	6.743	6.057	12.8	52.67969
12.8 up to 26.4	6.723	6.877	13.6	49.43382
26.4 up to 40.2	6.971	6.829	13.8	50.51449
40.2 up to 54.2	5.863	7.337	13.2	44.41667
54.2 up to 64.4	6.358	4.642	10.2	57.8
Total	32.658	31.742	64.4	

Percentage	50.71118	49.28882		

It can be clearly seen that from 0+00 to 64+400 km design speed must be limited to 40 km/h due to; high density of curves, mainly with radii less than 120 m and longitudinal slopes mostly between 5 and 7%. And only in a few isolated streets of widening alignment with somewhat reduced longitudinal slope, the design speed raised to 50 km/hr. There are two streets of road where the concentration of narrow bends (radii less than 40 m) and steep longitudinal slope (6-8%) reduce the speed limit to 25-35 km/hr., i.e. the last portion of the Gibe gorge near Abelti (0+00-5+00 km) and between the towns of Saja and Sekoru (52+900 km and 64+4002 km) where the road crosses the valleys of the kosho and simini rivers. The geometric characteristic of the road permits generally low to very low vehicle speeds: 73% (this percent indicates the general velocity from Gibe valley up to Jimma town) of the bends will permit a velocity of below 80 km/h but the route morphology restricts the design speed to 40-60 km/h in large portion of the road but as stated earlier the design speed of Gibe valley to Sekoru town is limited to 35-40 km/h. The speed limit in the villages and town areas has been uniformly fixed at 35 km/h as per the design consultants. The total length of the study road segment is 64.4 km. location of study area is presented below by creating map using shape files from ArcMap 10.4.1. the map consists the basic element such as title, legend, scale bar, coordinate and grids. As shown in the figure below the gap between the graticule grids (graticule grids are a type of grids which consists the latitude and longitude of places) is 10 minute. The coordinate system is World Geodetic System 1984 (WGS_1984); its projection is universal transverse Mercator (UTM); datum is WGS_1984; false easting 500,000.00; false northing 0.00; central meridian 39; latitude of origin is 0.00 and its unit is meter.

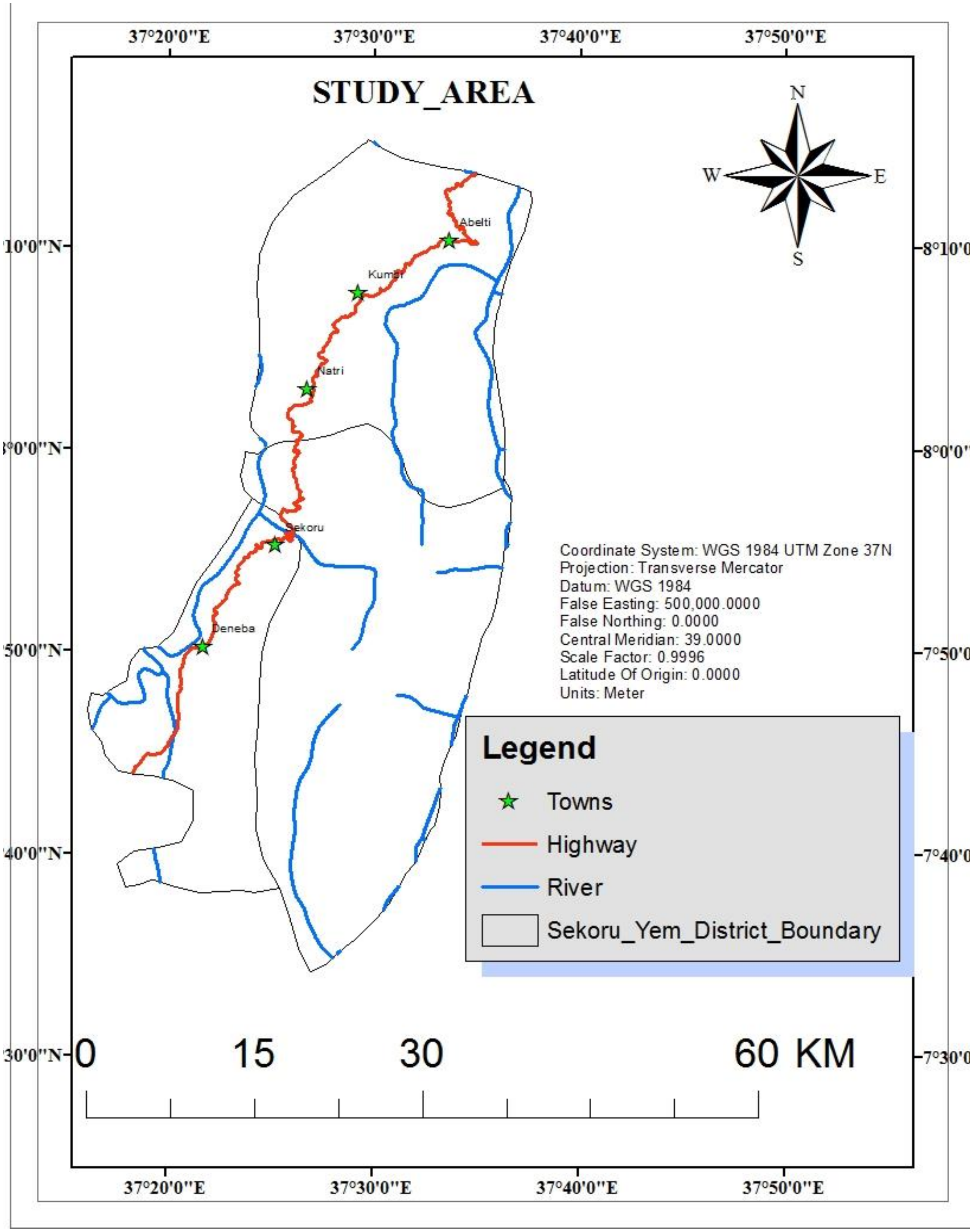


Figure 3.1: Location of Study area

3.1.1. Geology

The geology of the area consists of the volcanic products of the Magdala group and the Ashangi group. The Magdala group consists mainly in acid lavas, volcanic products ranging from rhyolites to trachyte, accompanied by tuffs and ignimbrites and interbedded with lavas and agglomerates of basaltic composition. This group is probably of the Pliocene era or older. The Ashangi group consists predominantly in alkaline basalts with interbedded pyroclastic associated with the rift valley faulting. This group is from the Paleocene to Miocene era. The Ashangi group is found in the Gibe valley and the Magdala group appears in the section of Abelti to Sekoru road segment (source Studio Pietrangeli-Rome consultants).

3.1.2. Weather Condition and Demographics

The area is classified as temperate and generally humid with the exception of the Gibe valley which is classified as warm. The monthly average minimum and maximum temperature (°C) for 45-year data is annually maximum of 25°C and minimum of 17°C and The area receives maximum precipitation occurs during the three months' period, June to September and moderate precipitation from March up to May. The rainfall peaks in July or August at about 300 to 350 mm per month, with minimum rainfall in December and January Existing rainfall observation records in the Sekoru district of 36 years' data, generally the area receives an average of the annual rainfall ranges from 989.7 mm to 1612.6 mm. (Source Jimma district Metrological agency).

3.2. Study Period and Study Design

The research was carried out from April 2017 to February 2018. The research was conducted by using descriptive, analytical and exploratory methods and also designed in the way that important and exact information is getting to identify traffic accident hot spot analysis with related factors like traffic accident, traffic safety, and geometric design of road and available sight distance. The design of the study is as follows. Using the GIS tools,

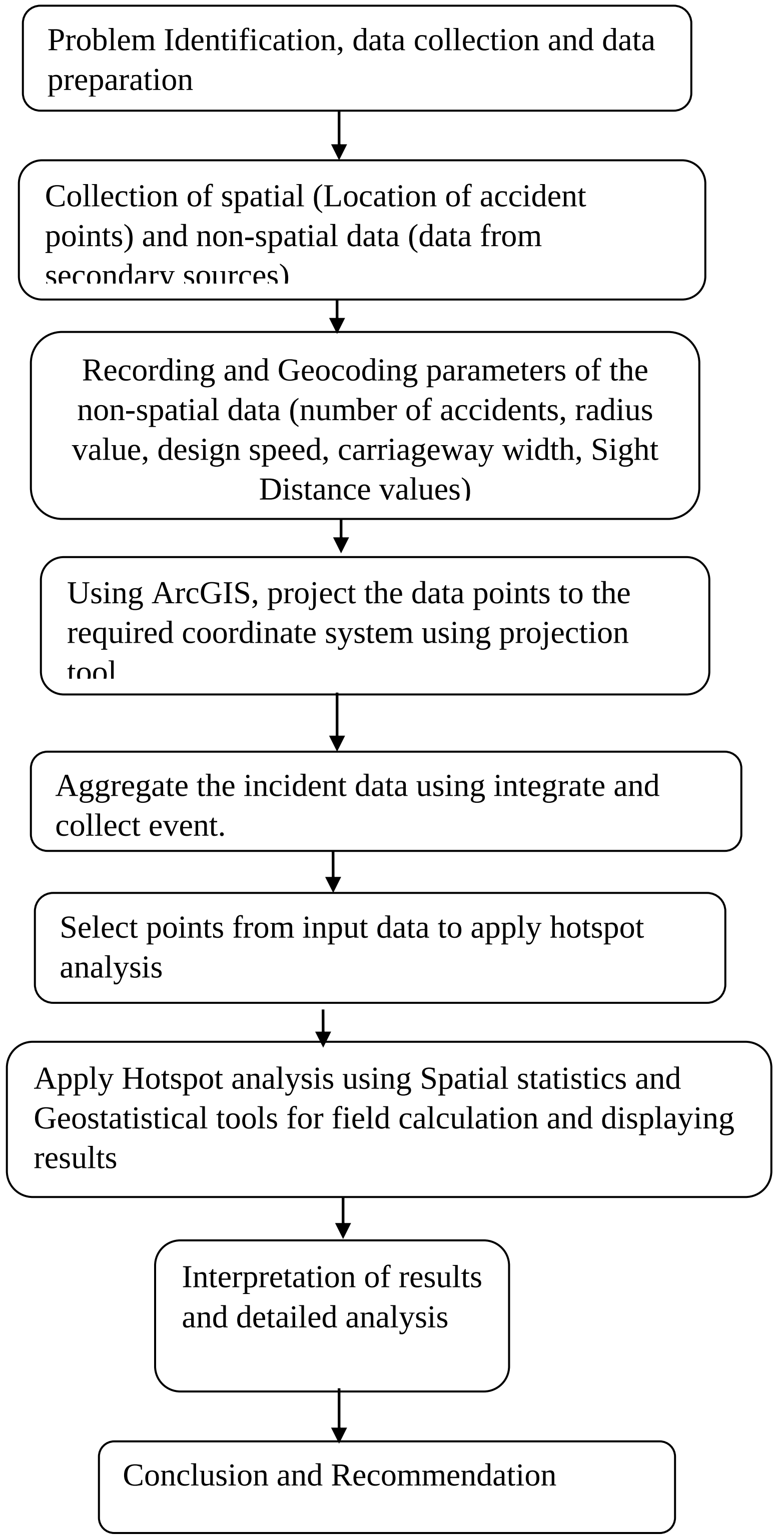


Figure 3.2: Study Design

3.3. Sampling Technique

Since Non-probability sampling represents a group of sampling techniques (e.g. purposive sampling) and has free distribution that help researchers to select a unit sample from a population, this sampling method is adopted.

3.4. Materials

3.4.1. Software

The following software are capable for sight distance and hotspot analysis; these are ArcGIS 10.4.1; Micro-soft Excel, AutoCAD. ArcGIS tools used in this study are tools like line of sight, observer point, target features, viewshed, visibility analysis, Spatial Statistical analysis (such as Getis-Ord hot spot analysis, Anselin Local Moran Index analysis, Optimization Hotspot analysis, Moran Index analysis), Geostatistical Analyst tools such as Inverse Distance Weighting, Empirical Bayesian, Kernel Smoothing density and related tools.

3.4.2. Input data

The DEM which enables more accurate results, Traffic accident data was collected from Jimma zone police station, Sekoru district police station and Yem district police station. Next, collecting road geometric parameters such as radius of curve, design speed, coefficient of friction, sight distance and surveying data which consists the latitude, longitude and elevation from ERA and other respective bodies; downloading some elevation points from Google earth (points first originated as kml file then converting these files into csv comma delimited excel files) collecting surveying data using GPS at field gathering. The method is used to identify the hotspot areas, the resulting points can be evenly spaced or not and are saved as a shape file or in a feature class of Geodatabase.

3.5. Variables

The independent variables are: radius of curve; length of curve; design speed; gradient; curve widening; carriageway width; stopping sight distance; super elevation; vehicle and driver eye height; traffic accident and the dependent variable is the analysis of traffic accident hotspot areas.

3.6. Data Collection and Processing Methods

Primary and secondary data were collected for the study. Secondary data collection included the collection of required accident data for the past five years, 2013 – 2017, from Sekoru District Police Department, Yem District Police Department and Jimma Zone Police Station. Exact locations of accidents, types of injury, type of vehicles were collected. The traffic volume count gives the measure of how many vehicles pass through a particular location during a period of time expressed by Annual Average Daily Traffic (AADT). Road geometric design data were obtained from Ethiopian Road Authority. (DEM_90) was obtained from Ethiopian Mapping Agency. Primary data such as road inventory data, photo capturing, traffic sign, and coordinates were collected from the field survey of the road segment identified for the study. actual spot.

3.7. Data Processing and Analysis

This study pursues to find answers to the problems through the analysis of traffic accident hot spot areas. The general procedure followed for the analysis of the data, primarily consists of performing spatial statistical and geostatistical analysis coupled with subjective evaluation. spatial statistical and geostatistical analysis was adopted to determine whether the dependent variable (rate of traffic accidents) was significantly related to the independent variables (Gradient, Radius, Carriageway Width, and age of drivers) using the relative importance of these factors. Exploratory data analyses were conducted using ArcGIS 10.4.1 software to investigate interactions and checking of assumptions underlying the use of geostatistical analysis techniques, specifically using map result, histogram, quantile-quantile relationship, trend analysis and semivariogram plots of the data. Accident data extracted from the case study sites were used to validate the spatial statistical and geostatistical analysis. It is important to state that for all analysis P-value < 0.01, P-value < 0.05 and P-value < 0.1 was considered to be statistically significant. The steps involved in data processing are shown in Figure 3.3. below

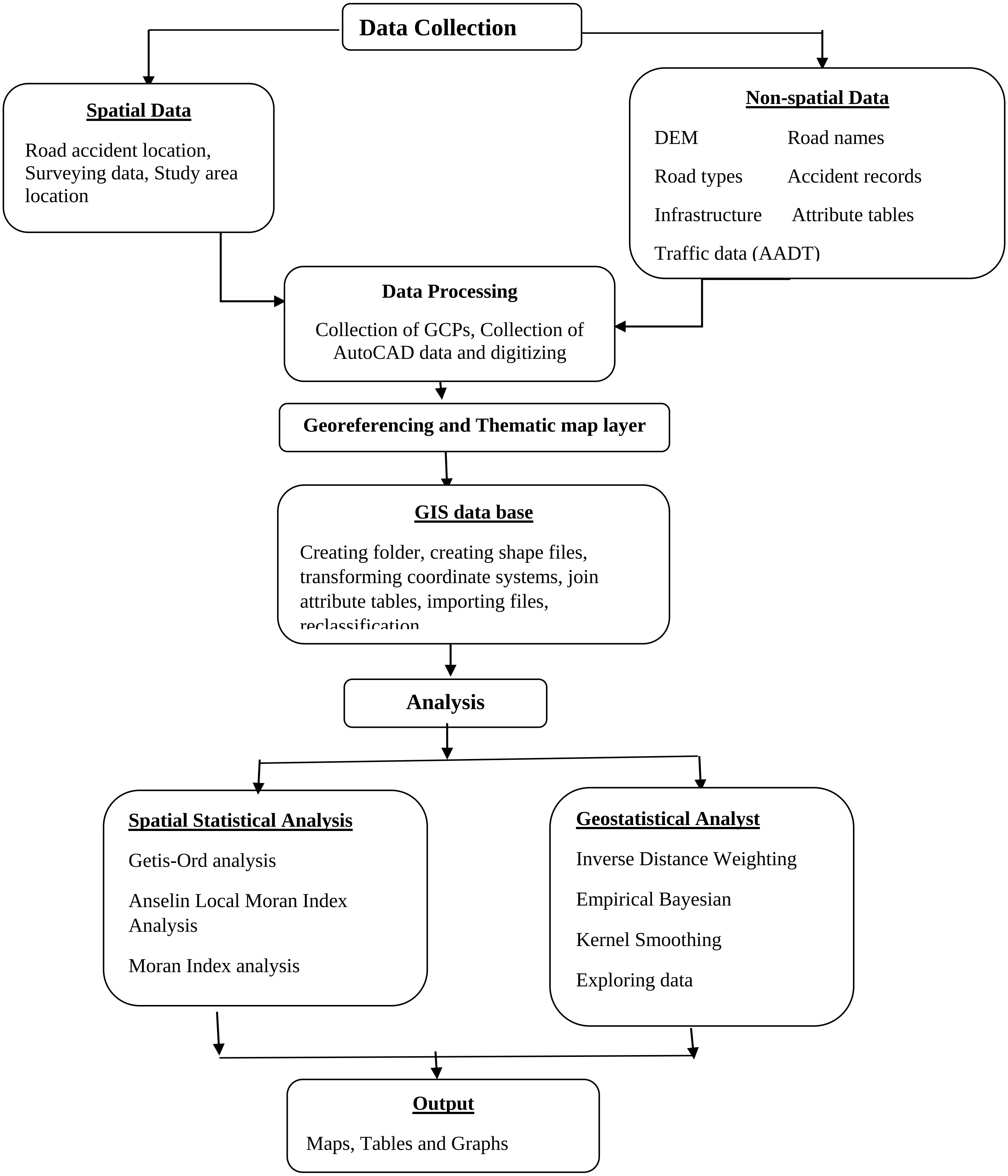


Figure 3.3: Flow chart of Data Processing

3.7.1. Define projection

Before starting all over work procedure adjusting the unit and setting of workspace of GIS must be accomplished. Adjusting ArcGIS 10.4.1 to Projected Coordinate System(PCS)- Universal Traverse Mercator (UTM)- World Geodetic System 1984(WGS 1984)- Northern Hemisphere- WGS 1984 UTM Zone 37N is obligatory. and the SI unit is adjusted to meter by following required steps and the detailed steps are described on Appendix A.

3.7.2. Geo referencing and Digitizing

All of collected data do not have usually contain same coordinate information as to where the area represented on the map fits on the surface of the earth. To establish the relationship between an image coordinate system and a map (x, y) coordinate system geo referencing the raster data has to be done. Digitizing is the process of encoding the geographic features in digital form as x, y coordinates. It was carried out to create spatial data from existing hard copy maps and documents. In the Present work, the geo referenced raster image of horizontal alignment of Sekoru district is digitized using ArcGIS10.4.1. This type of digitization is called onscreen digitization. Road network of the study area was digitized as line, point and polygon features. Accident locations are digitized as point and polygon features. The above spatial data were organized in a feature class. The exact location of accidents was identified by using measure tool in ArcGIS10.4.1. By using the measure tool, the spatial location of a particular accident can be marked by knowing its distance from a particular station.

3.7.3. Assigning attributes

All vector data (line, polygon, point features) will contain separate attribute tables. Here each point of road is labeled with its corresponding name or location with the help of the Google map and information of design files obtained from respective designers and consultants. Similarly, the accident location attribute tables were created contains the following data type of injury, exact area of occurrence obtained respective district police station officials information.

For available sight distance on specified road, the advanced geographic information system(ArcGIS) tools were used. Thus, from the findings correlations have developed and interpreted using a purpose built statistical modeling programs to solve sight distance problems as well as traffic safety problems using Spatial statistical and Geospatial hot spot analysis tools. Final conclusions have been drawn after the interpretation of each explanatory variable.

3.7.4. Traffic accident data

Five years (from 2013 to 2017) of traffic accident data were collected from Jimma zone police station, Sekoru and Yem district Police stations. The data is classified based on accident severity for fatal, severe, slight and property damage only (PDO) accident severity categories. The geographic coordinate details for all accident data types were collected and spatially joined with the district boundary shape file. Figure 3.4 illustrates the accident distribution related to their geographical coordinates overall the district. The data also includes the total PDO accident cost in ETB for each year but it doesn't include in the analysis.

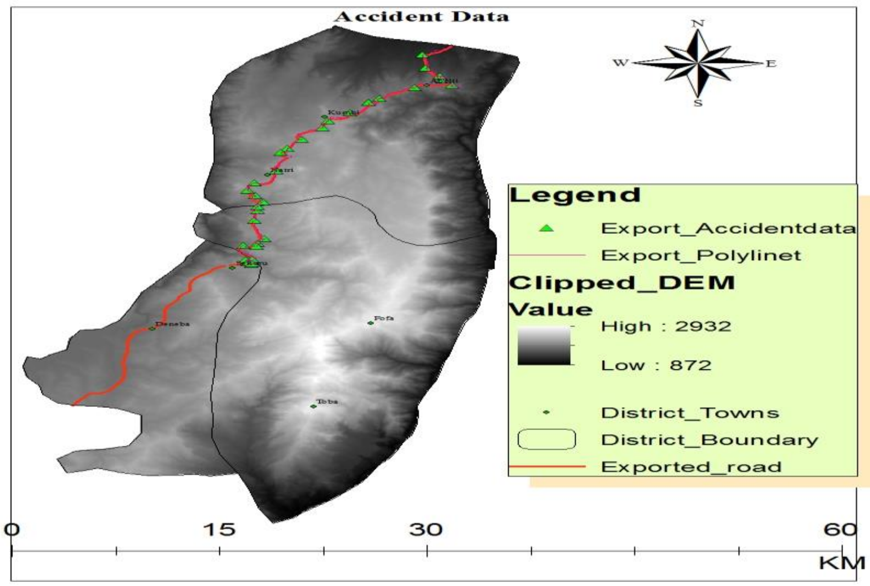


Figure 3.4: Accident distribution overall the highway(from 2003-2017)

Table 3.2: Annual Variation of Accidents by their degree of severity (Jimma zone and Sekoru district police officials)

No	Traffic Accident data of Sekoru wereda (Gibe bridge-Sekoru town)						
	Year	Fatal	Serious accident	Slight accident	PDO	TOTAL	Accident cost/Birr
1	2013	10	1	4	29	44	2209100.0
2	2014	10	3	2	27	42	977900.0
3	2015	7	10	3	21	41	2960850.0
4	2016	2	2	4	30	38	2136685.0
5	2017	13	7	12	25	57	3178900.0
	TOTAL	42	23	25	132	222	9268688

Table 3.3: AADT versus Accident Rate of Study area

YR	AADT	KM	VKT	ACCIDENT	AR
2005	1125	65	73125	44	0.329704
2006	1453	65	94445	42	0.243673
2007	1536	65	99840	41	0.225018
2008	1631	65	106015	38	0.196405
2009	1759	65	114335	57	0.27317

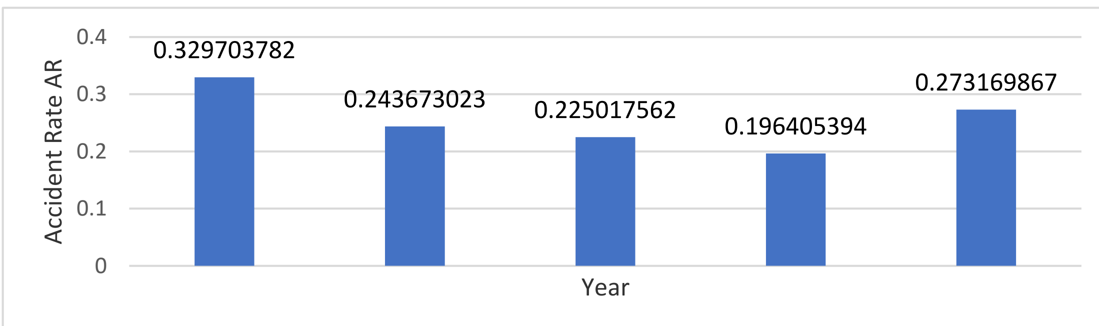


Figure 3.5: Accident rate versus Year

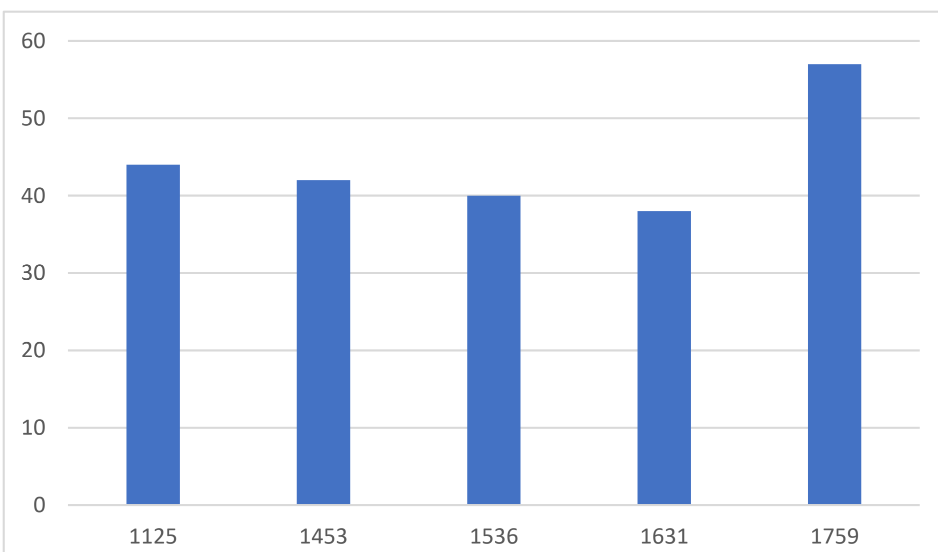


Figure 3.6: AADT versus Total Accident

3.7.5. Traffic data (AADT)

A 14 year (from 2002-2015) of average annual daily traffic AADT data have been collected from Ethiopian road authority ERA Jimma District Office for different vehicle types. These AADT data are used to predict the traffic trend and in the calculation of traffic accident rates.

Table 3.4 : AADT of Welkitea to Jimma road section (ERA)

YEAR	LENGTH	CARS	BUSES	TRUCK	TRUCK TRAILER	TOTAL
2002	188	116	145	275	39	575
2003	188	121	145	284	34	584
2004	188	126	164	315	55	660
2005	188	115	164	358	48	685
2006	188	166	233	456	77	932
2007	188	154	262	460	45	921
2008	188	138	255	423	72	888
2009	188	141	296	462	78	977
2010	188	121	272	434	82	909
2011	188	142	328	429	106	1005
2012	188	178	376	638	103	1295
2013	188	159	344	517	105	1125
2014	188	198	439	690	126	1453
2015	188	226	501	683	126	1536

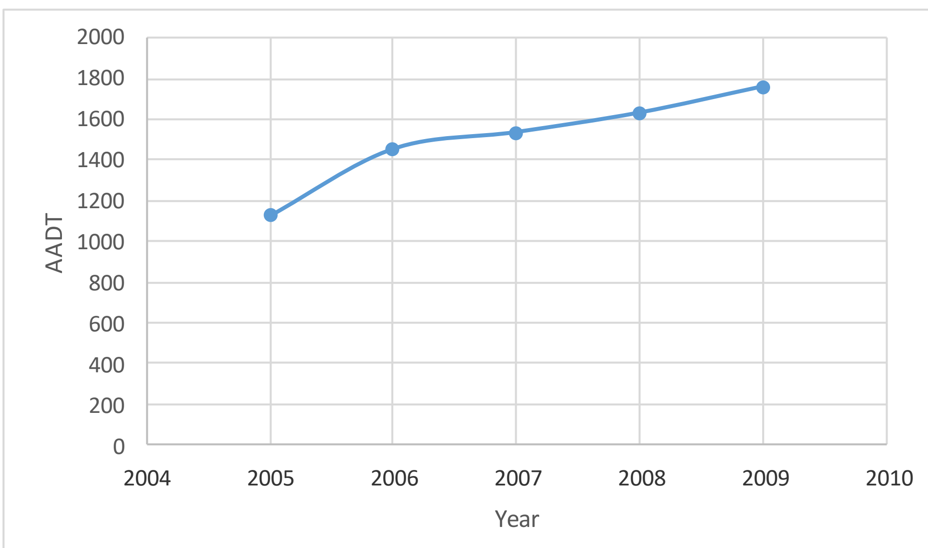


Figure 3.7: AADT versus Year

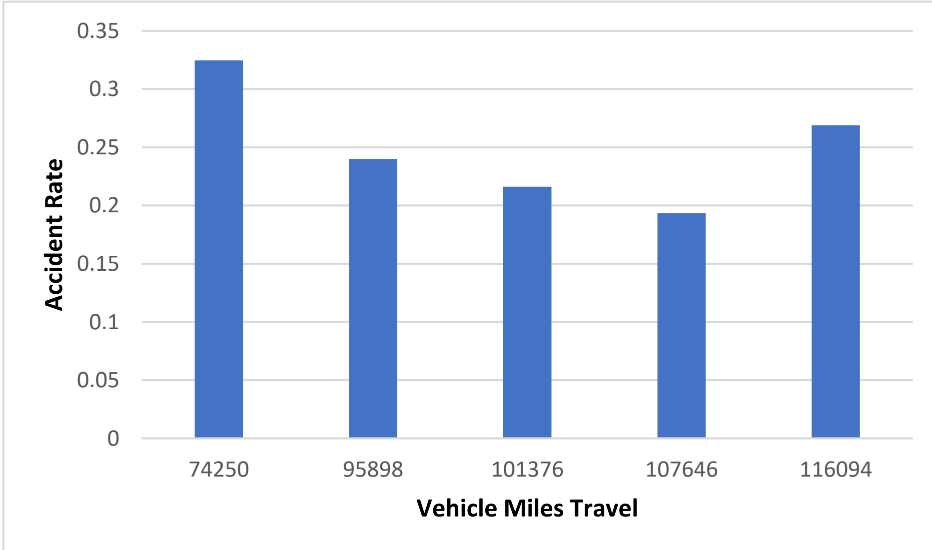


Figure 3.8: Accident rate versus Vehicle miles travel

3.8. Methodology Used

This study mainly uses a data of hotspot areas collected by GPSmap 62s (Garmin model) of the respective areas. Then after importing the surveying data and with related to the secondary data processing using ArcGIS tools with their application methods. These are Spatial Statistical analysis and Geostatistical analyst for traffic accident hot spot analysis. Each methods consists sub methods such as under the Spatial statistical analysis like Getis-Ord Gi*(Gi_Bin, GiZscore and GiPvalue), Optimization hotspot analysis, Anselin Local Moran Index (Clusster outlier relationship), Moran Index, Exploratory regression, Weighted regression and Least square methods are included. Under the Geostatistical analyst methods like Inverse distance weighting, Kernel smoothing density, Empirical Bayesian and exploratory data (Histogram, Semivariogram or covariance, quantile-quantile relation ship, and trend analysis) are included. The following paragraphs discuss the theoretical background and application of each method.

3.8.1. Spatial Autocorrelation

The first law of geography states that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). The basic principle of spatial autocorrelation (SA) is similar and is defined as the correlation of a variable with itself through space whether the points spatially autocorrelated or not are based on the distance of their neighborhood. SA of results positive means they are high clustered and negatives are randomly autocorrelated. As Mitra, 2009 stated that Spatial Autocorrelation are not expound why locations that have clusters of statistically significant crashes have higher incidence of crashes than other locations; therefore, spatial autocorrelation methods cannot pinpoint factors that cause crashes. Except they already tell us the hotspot and coldspot location with the included attribute table values.

Based on the official website of ESRI spatial autocorrelation is defined as sample values taken close to one another are more alike than samples taken far away from each other. Some interpolation methods require an explicit model of spatial autocorrelation (for example, kriging), others rely on assumed degrees of spatial autocorrelation without providing a means to measure it (for example, Inverse Distance Weighting), and others do not require any notion of the spatial autocorrelation in the dataset. Note that when spatial autocorrelation exists, traditional statistical methods (which rely on the independence among observations) cannot be used reliably.

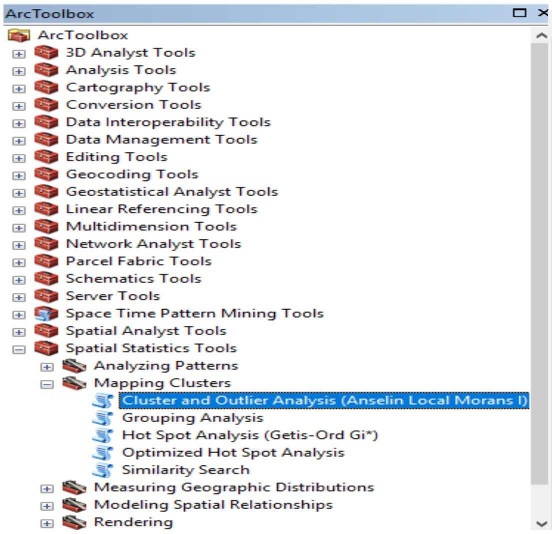
3.8.1.1. Spatial Hot Spot Statistical Analysis (Getis-Ord Gi*)

G-statistics, developed by Professor Arthur Getis and J.K Ord, analyze evidence of spatial pattern (Getis and Ord 1992; Ord and Getis, 1995). They represent a global spatial autocorrelation index. The Gi* statistic, in other means, is a local spatial autocorrelation index. Compare local averages to global averages. Gi* statistic includes the value of the point in its calculation Gi excludes this value and only considers the value of its nearest neighbors (within distance) against the global average (which also does not include the value at site i) The standardized Gi* is essentially a Z-value and can be associated with statistical significance. This method calculates Gi* value of each point/event in a particular area, where the values are analyzed and influenced by the number of events around them. The results of this method of analysis of Z Score and P Value. Point/hotspot is the point of incident/event that has a value of Z-Score is high (positive) and P-value zero (Utoyo et al., 2012). Positive and negative Gi* statistic with high absolute values correspond to clusters of crashes with high and low values events, respectively. A Gi* close to zero means a random distribution of events. the location of Getis-Ord Gi hot spot analysis in ArcGIS on the drop down choices of Arc Toolbox located on spatial statistical analysis then under the mapping cluster.

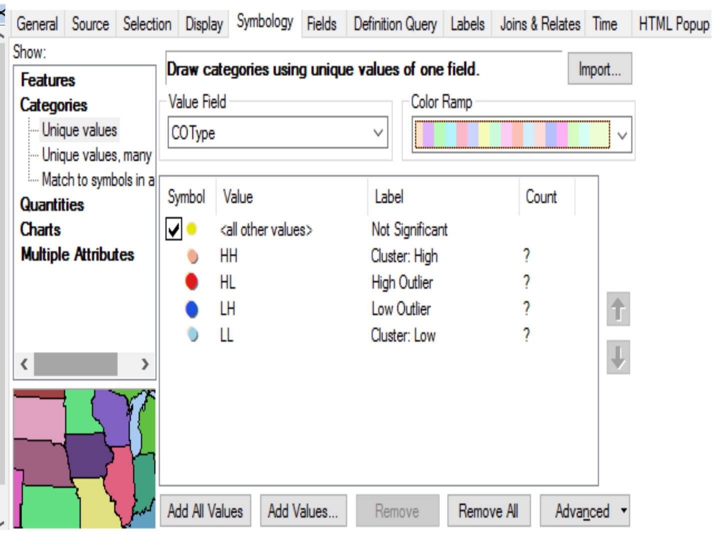
In order to analyze using the Getis-Ord Gi* tool the following requirements are needed; these are Input features at least 30 objects, projected with value field. In order to use the Getis-Ord statistics method data must be projected. The conceptualization of spatial autocorrelation relationships; distance method; standardization; distance band or threshold distance(optional); self-potential field (optional) and weight matrix file (optional) must be choose one of them from these parameters: Inverse Distance –Closer features are weighed more heavily than features that are further away. Inverse Distance Squared –Same as inverse distance, but weight decreases more dramatically over distance. Fixed Distance Band –Every feature within a fixed distance is included, every feature outside that distance is excluded. Zone of Indifference –Combination of Inverse Distance and Fixed Distance Band. Polygon Contiguity –Only features that share a border are included. In inputs conceptualization of spatial relationship Esri recommends Fixed Distance Band because of consistent scale of analysis across each/every layer and must provide adequate neighbors and appropriate scale. For each analysis, Fixed Distance Band and a minimum number of neighbors (expands distance band where needed). Appendix C1 shows an example of thematic map layer of fatal accidents with its Gi* ZScore value which represents in graphical form such as features value, categories representation, quantities representation (either in graduated color or in graduated and proportional symbol representations), charts and multiple attributes form. The values represented most of the time in quantities graduated color form because of easily to represented the result in symbol, range and label form.

3.8.1.2. Anselin Local Moran Index (Cluster and Outlier Analysis)

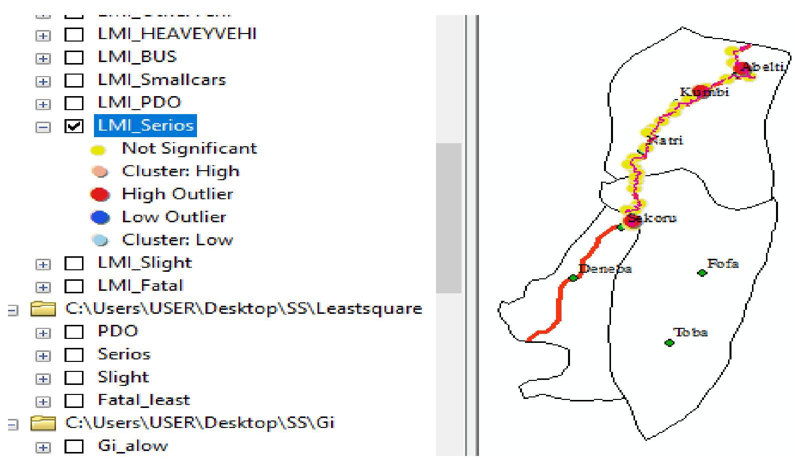
Anselin Local Moran I method or Clusters and Outliers Analysis is a statistical method that takes into account equation (value) events/feature similarity. This method aims to identify the clusters with high and low values and points/events that surrounded the incident with different values. Analysis using this method will produce output in the form of Local I index, Z score, P Values and Cluster Type outliers. Analysis using Cluster and Outliers Analysis (Anselin Local Moran I) defines accident traffic, and accident black spot as a point of having a CO Type "HH" (clusters of high values). Figure 3.9 shows how the cluster and outlier analysis method works.



(a) Location of ALMI



(b). sample result using thematic map layer



(c) Table of content

Figure 3.9: Work flow of Anselin Local Moran Index

3.8.1.3. Spatial Autocorrelation (Moran I) hot spot analysis

Moran Index is a global statistic that tells whether there is clustering or dispersion, but it does not inform about the location of a cluster. Compare if the value for each observation is similar to those that neighbor it. appendix C2 shows the Moran index value of slight accidents happened in the Highway of Sekoru town to Gibe valley. As shown in appendix C2 the spatial autocorrelation Moran index values are given in a normal distribution graph form. This method requires for analyzing accidents uses all accidents data as input feature class, specific injury types as input field, fixed distance as conceptual spatial relationship and Euclidian as distance method. Global Moran index analysis tool calculates the input data and presents the result in the form of normal distribution curve, z-values on x-axis and p-values on y-axis. The clustering of the result as cold spot and hot spot classification is given as dispersed, random and clustered.

3.8.2. Geostatistical Analyst Toolbar

Geostatistical Analyst contains tools for exploratory spatial data analysis and a geostatistical wizard to lead you through the process of creating a statistically valid surface. These tools are not part of the Geoprocessing tools; however, the results obtained from them can be used in the Geoprocessing tools. Geostatistics is a class of statistics used to analyze and predict the values associated with spatial or spatiotemporal phenomena. It incorporates the spatial (and in some cases temporal) coordinates of the data within the analyses. Many geostatistical tools were originally developed as a practical means to describe spatial patterns and interpolate values for locations where samples were not taken. Those tools and methods have since evolved to not only provide interpolated values, but also measures of uncertainty for those values. The measurement of uncertainty is critical to informed decision making, as it provides information on the possible values (outcomes) for each location rather than just one interpolated value. Geostatistical analysis has also evolved from uni- to multivariate and offers mechanisms to incorporate secondary datasets that complement a (possibly sparse) primary variable of interest, thus allowing the construction of more accurate interpolation and uncertainty models. (ESRI, 2010) The geostatistical wizard is accessed through the geostatistical analyst toolbar as shown in Appendix A2. To explore data on the geostatistical analyst, click on the geostatistical analyst, next click explore data and finally click on the required form among the given options of exploring data such as histogram, normal quantile-quantile plot, trend analysis, etc. if histogram is selected layer and attribute are selected for the data source and the results represented by count, minimum value, maximum value, mean, standard deviation, skewness, kurtosis, 1st quartile, median and 3rd quartile.

The Geostatistical Wizard is a dynamic set of pages that is designed to guide you through the process of constructing and evaluating the performance of an interpolation model. Choices made on one page determine which options will be available on the following pages and how you interact with the data to develop a suitable model. The wizard guides you from the point when you choose an interpolation method all the way to viewing summary measures of the model's expected performance. A simple version of this workflow (for inverse distance weighted interpolation) is represented graphically on Appendix A2.

3.8.2.1. Examining the distribution of data

Primarily, the distribution of the data was examined in order to fit the best model to use for prediction. It is clear from the histogram presented in Appendix B1 that the total serious accidents of the district skewed right. The quantile-quantile (QQ) plot in Appendix B1, approaches a straight line, thus, the data are normally distributed.

3.8.3. Symbolization and Categorization of the Accident Hotspots

The accident hot spots are symbolized and categorized based on the Z-value, Index value and P-value of the Spatial statistics. symbolization can be based on features (single symbol); categories which include unique value, unique value many fields and match to symbols in a style; quantities which include graduated colors, graduated symbols and proportional symbols; charts which include pie charts, bar/column and staked; multiple attribute which includes quantity by category. Whereas categorization of fields can be on seven classification schemes: manual classification, equal interval, defined interval, quantile, natural breaks (jenks), geometric interval and standard deviation. in the equal interval and defined interval scheme values are represented similarly, in the quantile scheme lower values densely represented or closely to the y-axis, in the geometric inter more similarity with quantile because minimum values more compressed in the same side whereas in the standard deviation values displayed by more generalized, in the natural breaks schemes values are classified by the class breaks that best group similar values and maximize the difference between the class and this method is best suited to the present study

3.8.3.1. Inverse Distance Weighting

Inverse distance weighted (IDW) interpolation explicitly makes the assumption that things that are close to one another are more alike than those that are farther apart. To predict a value for any unmeasured location, IDW uses the measured values surrounding the prediction location. The measured values closest to the prediction location have more influence on the predicted value than those farther away. IDW assumes that each measured point has a local influence that diminishes with distance. It gives greater weights to points closest to the prediction location, and the weights diminish as a function of distance, hence the name inverse distance weighted. Weights assigned to data points are illustrated in Appendix A2 for fatal accidents that occurred in Gibe bridge to Sekoru town road segment. The method to analyze hotspot areas by IDW includes data importing to ArcGIS 10.4.1 and analyze the data in the geostatistical analyst tool. The analysis is started by clicking on geostatistical analyst tool, after selecting IDW the source data, data field and weighting field data will be selected. Then click on next button and Geostatistical wizard – IDW of step 1 will be displayed. At this stage, the power, neighborhood type (smooth and standard), smoothing factor (increase the value to get increase the weight), angle and axis values will be field. Finally, click on next button and the result will be displayed in tabular and chart format then click on finish to get the map representation.

3.8.3.2. Empirical Bayesian Kriging

Empirical Bayesian kriging (EBK) is a geostatistical interpolation method that automates the most difficult aspects of building a valid kriging model. Other kriging methods in Geostatistical Analyst require you to manually adjust parameters to receive accurate results, but EBK automatically calculates these parameters through a process of sub setting and simulations. Empirical Bayesian kriging also differs from other kriging methods by accounting for the error introduced by estimating the underlying semivariogram. Other kriging methods calculate the semivariogram from known data locations and use this single semivariogram to make predictions at unknown locations; this process implicitly assumes that the estimated semivariogram is the true semivariogram for the interpolation region. By not taking the uncertainty of semivariogram estimation into account, other kriging methods underestimate the standard errors of prediction. The of EBK method is requires minimal interactive modeling; standard errors of prediction are more accurate than other kriging methods; allows accurate predictions of moderately nonstationary data and more accurate than other kriging methods for small datasets. The disadvantage of EBK is processing time rapidly increase as the number of input points, the subset size, or the overlap factor increase. Applying a transformation will also increase processing time; processing is slower than other kriging methods, especially when outputting to raster; cokriging and anisotropic corrections are unavailable. Appendix A2, shows for serious injury using EBK method. The method to analyze hotspot areas by EBK includes data importing to ArcGIS 10.4.1 and analyze the data in the geostatistical analyst tool. The analysis is started by clicking on geostatistical analyst tool, after selecting EBK the source data, data field will be selected. Then click on next button and Geostatistical wizard – EBK of step 1 will be displayed. At this stage, the general properties (like subset size default value of 100, overlap factor of 1, number of simulations of 100, output surface type default prediction, transformation of empirical and semivariogram type of power) and search neighborhood (neighborhood type of standard circular, maximum and minimum neighbors, sector type, angle and radius) options are filled. Clicking on next button will give the result in tabular and chart forms presented. Finally, click finish to get the map representation.

3.8.3.3. Kernel Smoothing Density Estimation

Depend upon the ESRI definition kernel is A weighting function used in several of the interpolation methods offered in Geostatistical Analyst. Typically, higher weights are assigned to sample values that are close to the location where a prediction is being made, and lower weights are assigned to sample values that are further away. Where as in other researchers like Kuo etal.,2011, kernel density method is geostatistical analyst capabilities of ArcGIS program and easy to calculate the risk density for each crash instead of showing the actual location of each crash. It is one of the most common methods of defining hotspots for crash data, because it details smooth and continuous risk targets in the study area (Chainey et al., 2002). Kernel density method uses a density estimation technique instead of parametric methods. The basic principle of this method calculating the density of each point instead of showing the actual location. The density value varies based on the distance from the point, if the distance is closer to the point there is higher density unless the density decreases when the distance faraway. Silverman, 1986 and Wand, Jones, 1995 stated that Kernel density analysis is implemented by applying data usually on a regular grid. The densities of the observations within a set radius are calculated for each event located on the grid, and the contributions of each observation are weighted by its proximity to the center of the moving window. Thus, the result of applying kernel smoothing is a density map (raster format). The values of each pixel represent the relationships between the concentrations of the events per unit area. In addition, kernel smoothing can be used to calculate the density of punctual events (i.e., the density of traffic accidents in a region). It is important to highlight the simplicity, satisfactory properties and good results of the kernel smoothing density method. The last performed method of network points patterns analysis is kernel function analysis.

The kernel density can be defined as:

$$K = \sum_{d<\tau} \frac{3}{\pi\tau^2}\left(1-\frac{d^2}{\tau^2}\right)^2 \quad \textbf{(7)}$$

Where, K is Kernel density value; d is The distance from event; and τ is Bandwidth.

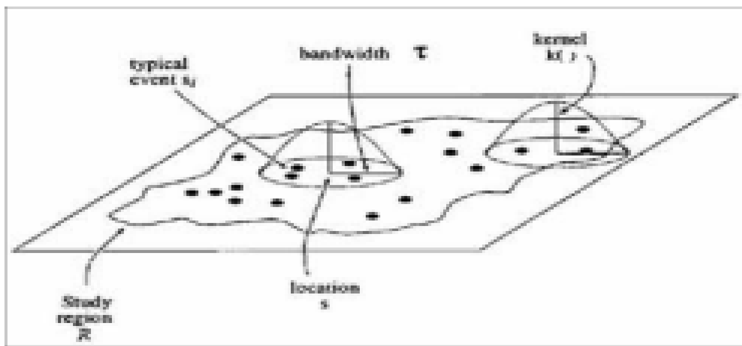


Figure 3.10: Kernel Smoothing Density method (Erdogan, et al., 2008)

Kernel Interpolation uses the following radially symmetric kernels: Exponential, Gaussian, Quartic, Epanechnikov, Polynomial of Order 5, and Constant. The bandwidth of the kernel is determined by a rectangle around the observations. The Epanechnikov kernel usually produces better results when the first-order polynomials are used. However, depending on the data, the cross-validation and validation diagnostics may suggest another kernel, Fan and Gijbels (1996). Their formulas for all kernel functions are presented on appendix B10. As shown in Appendix A2, The method to analyze hotspot areas by KSDE includes data importing to ArcGIS 10.4.1 and analyze the data in the geostatistical analyst tool. The analysis is started by clicking on geostatistical analyst tool, after selecting KSDE the source data, data field will be selected. Then click on next button and Geostatistical wizard – KSDE of step 1 will be displayed. At this stage, the general properties (kernel function of Epanechnikov, output surface type default prediction, ridge and band width) then predicted values and weight are calculated based on the general properties. Clicking on next button gives results in tabular and chart format. Finally, click finish to get the map representation.

3.8.3.4. Modeling Spatial Relationships

Modeling spatial relationships are supportive methods to the spatial statistics and geostatistical methods in order to identify their spatial relationship and accident severity using exploratory regression, generate network spatial weights, generate spatial weight matrix, geographical weighted regression and ordinary least square methods. The traditional ordinary least square multiple regression model is relatively quick, simple and suitable for analyzing punctual events, such as road accidents. As a global regression method, an important assumption of the ordinary least square model is that all variables are stationary across the study area. In addition, the ordinary least square method can involve potential issues related to spatial and temporal autocorrelations, and the relationship between the dependent variable (amount of road accidents) and independent variables (such as, radius of curve, coefficient of friction, grade, vehicle types, drivers age, and sight distance) is possible (Wang et al., 2015).

CHAPTER FOUR

4. RESULTS AND DISCUSIONS

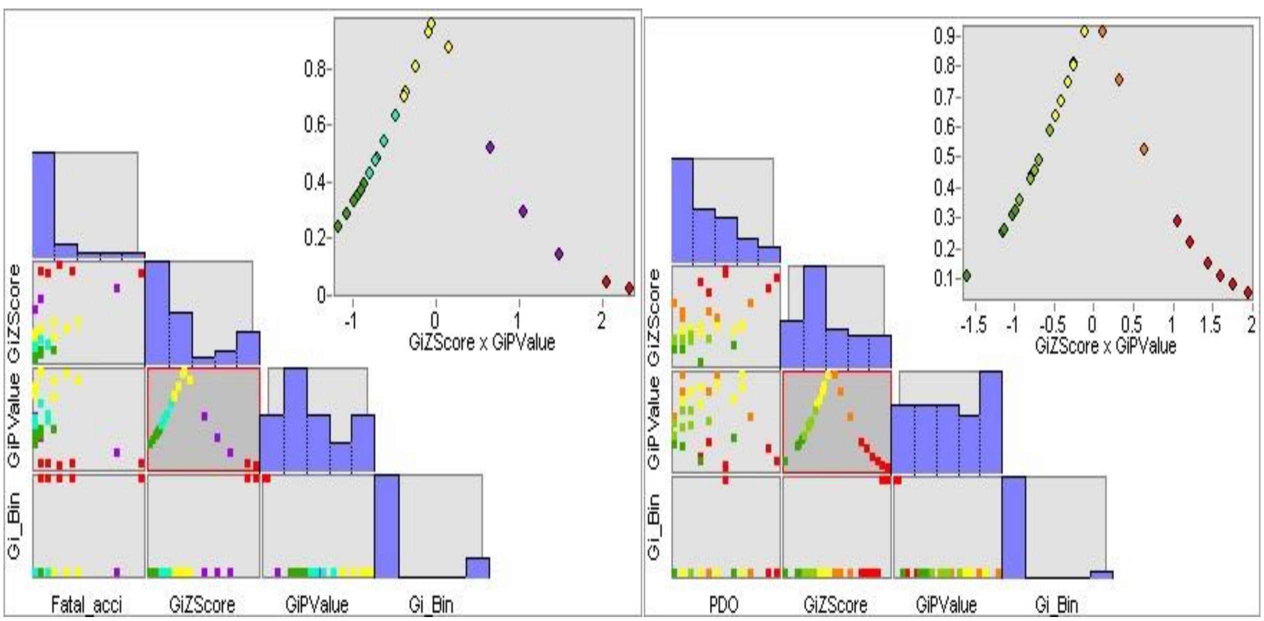
4.1. Traffic Accident Distribution

A total of 222 accidents occurred during the year 2013-2017. Table 3.2 shows a trend in number of accidents occurring in Gibe bridge- Sekoru town road segment, there is an urgent need to carry out proper traffic accident management studies in order to regulate the hot spot analysis.

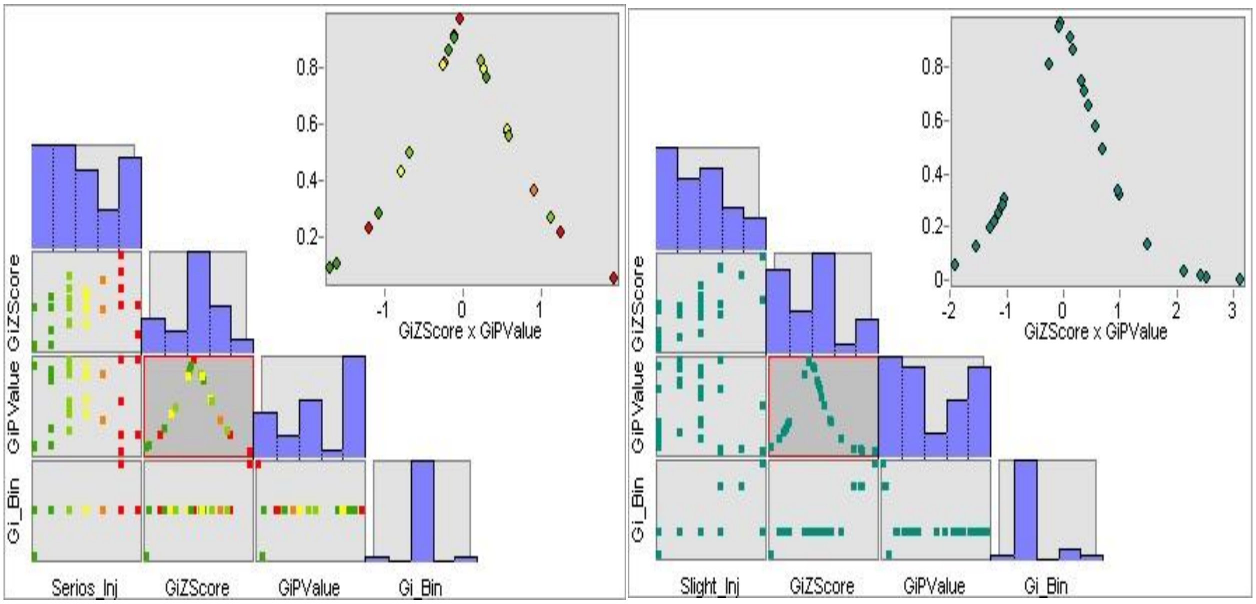
4.2. Scatter Plot Matrix

Figures 4.1a, b, c and d are scatter plot matrix which shows scatter plot matrix results of fatal, serious, slight and PDO accidents in the form of X-Y point coordinates and histograms respectively. To show the results clearly, a normal distribution curve of GiZscore and GiPValue is shown in top right corner of the scatter plot matrix. The values on the Y axis which starts from 0.1 to 0.9 are P values and the values on the X axis which starts from <-1 to >+2 values are Z score values or standard deviation values. From the normal distribution curve of Z score values and P values, the points consisting a Z score values ranging from <-1 to 0 and P values ranging from 0.1 to 0.2 are the cold spot areas. These points are located at (37.442, 8.013 m; 37.557, 8.216 m; 37.459, 8.087 m; 37.436, 8.021 m) The points consisting Z score values ranging from +1 to >2 and P values of < 0.1 are hot spot areas which are located at (37.572, 8.181 to 37.578, 8.173 m; 37.505,8.130 to 37.519, 8.150 m; 37.485, 8.114 to 37.435, 7.933 m) The points in between cold spot and hot spot areas are random which are located at (37.443, 7.962 m; 37.438, 7.956 m; 37.489, 8.125 m) for fatal accidents; From the normal distribution curve of Z score values and P values, the points consisting a Z score values ranging from <-1 to 0 and P values ranging from 0.1 to 0.2 are the cold spot areas. These points are located at (37.433, 7.926 up to 37.435, 7.933 m) The points consisting Z score values ranging from +1 to >2 and P values of < 0.1 are hot spot areas which are located at (37.470, 8.099 m; 37.459, 8.087 m; 37.438, 7.956 m; 37.485, 8.115 m) The points in between cold spot and hot spot areas are random which are located at (37.519, 8.150 m; 37.436, 7.987 m; 37.484, 8.125 m) for serious accidents; cold spot points are located at (37.531, 8.170 up to 37.459, 8.087 m) The points consisting Z score values ranging from +1 to >2 and P values of < 0.1 are hot spot areas which are located at (37.442, 8.013; 37.436, 8.039, 37.431, 8.028 m) The points in between cold spot and hot spot areas are random which are located at (37.043, 8.013 up to 37.538, 7.795 m) for slight accidents; cold spot points are located at (37.436, 8.039; 37.438, 8.007m) The points consisting Z score values ranging from +1 to >2 and P values of < 0.1 are hot spot areas which are located at (37.560, 8.197; 37.570, 8.186; 37.579, 8.173; 37.572, 8.181; 37.578, 8.173 m; 37.505,8.130; 37.519, 8.150 m; 37.485, 8.114; 37.435, 7.933 m) The points in between cold spot and hot spot areas are random which are located at (37.557, 8.216 m; 37.438, 7.956 m; 37.489, 8.125 m) for PDO accidents. Z score values are standard deviation values or Gi* values which are associated with statistical significance. Z score values of <-1.65 and >+1.65 indicate a confidence level or Gi_Bin of 90% and P values of < 0.1; Z score values of < -1.96 and > +1.96 indicate a confidence level of 95% and P values of < 0.05, and Z score values of < -2.58 and > 2.58 indicate a confidence level of 99% and P values of < 0.01.

As Getis and Ord 1992 showed that a point is said to be hot spot if it has a higher positive Z score values and low P values, and it is said to be cold spot if it has a lower negative Z score value and a low P value. Thus, the points which are described above are hot spot or black spot areas for fatal, serious, slight and PDO accidents. At these locations accidents clustered or occurred frequently.



a, fatal accidents b, PDO accidents



C, Serious accidents

d, Slight accidents

Figure 4.1: Scatter plot matrix of all accident types (GiZScore vs GiPValue)

All other Scatter plot matrix are presented on appendix B2

4.3. Vertical Line Graph Representation

Figure 4.6 summarizes figure 4.2, 4.3, 4.4 and 4.5, and is a plot of horizontal distance, which starts from Gibe bridge (0+000 m) to Sekoru town (64+400 m), versus number of fatal, serious, slight and PDO accidents respectively. The figure shows that fatal accidents clustered maximum amount occurred at 7 to 10 km which at the cluster of eastern Abelti town specifically at the area of Shen Debitu, 20 to 25 km which at cluster of Kumbi town, and 50 to 60 km which was clustered at the south west of Saja to north east of Sekoru specifically between Daka and Birilea stream; serious accidents frequently occurred maximum values are happened at segment of 8 to 10 km, 18 to 22 km, at 30 km and at distance between 58 to 63 km..; slight accidents maximum accidents are occurred at kilometers of 8 to 10, 40 to 49 and at distance of 60 to 63 km and PDO accidents occurred at maximum value at km of 8 to 10, 20 to 30 km, 45 to 63 km. Thus, the locations from the maximum accidents happened at distance of 8 to 10 km, 18 to 25 km, 45 to 49 km, 55 to 57 km, and 61 to 64 km along the Gibe bridge to Sekoru town highway section are high accident risk areas.

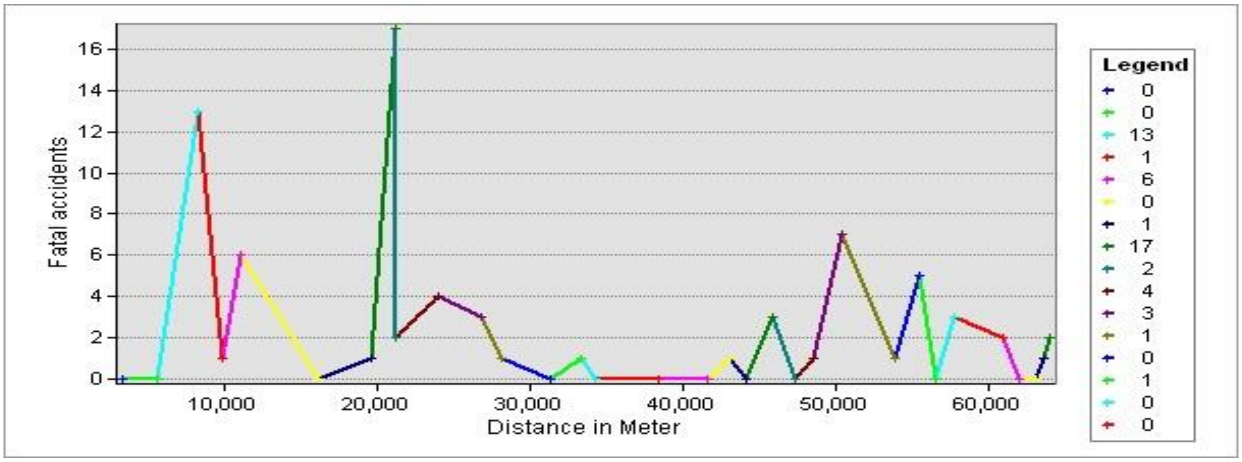


Figure 4.2: Graph of fatal accidents versus distance

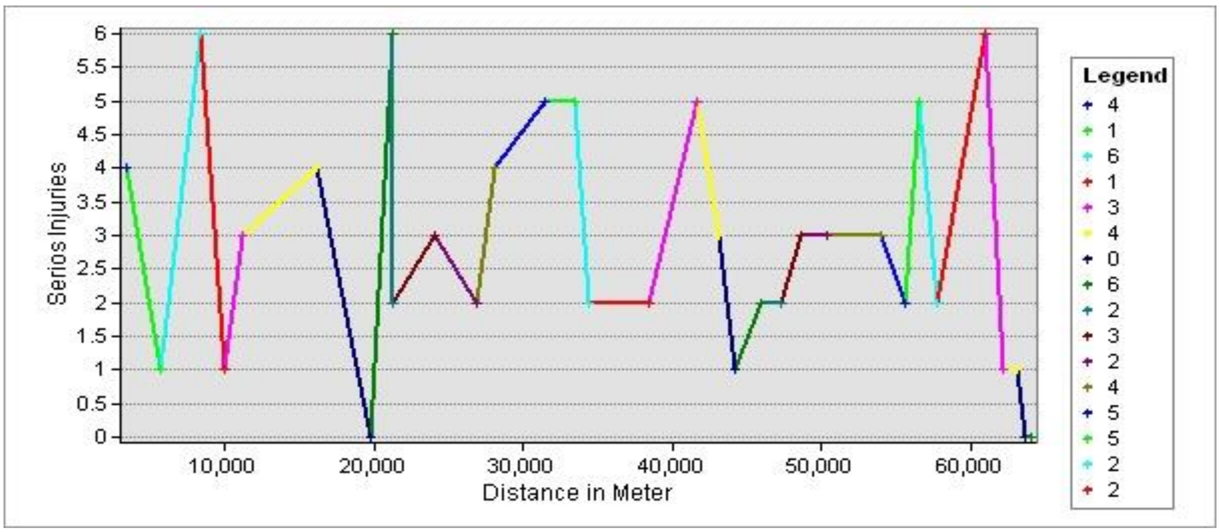


Figure 4.3: Graph of Serious injury versus distance

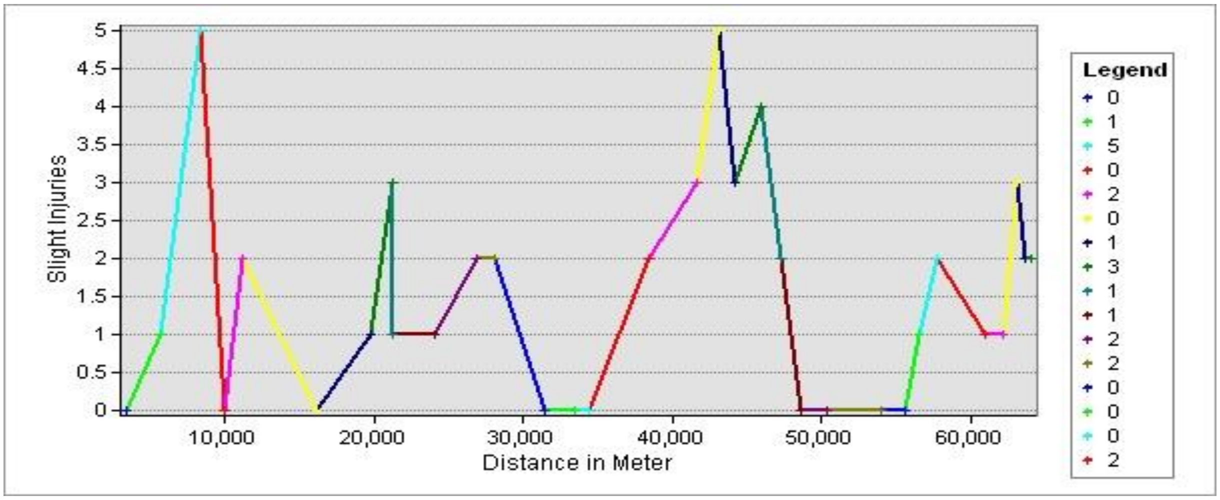


Figure 4.4: Graph of Slight injury versus distance

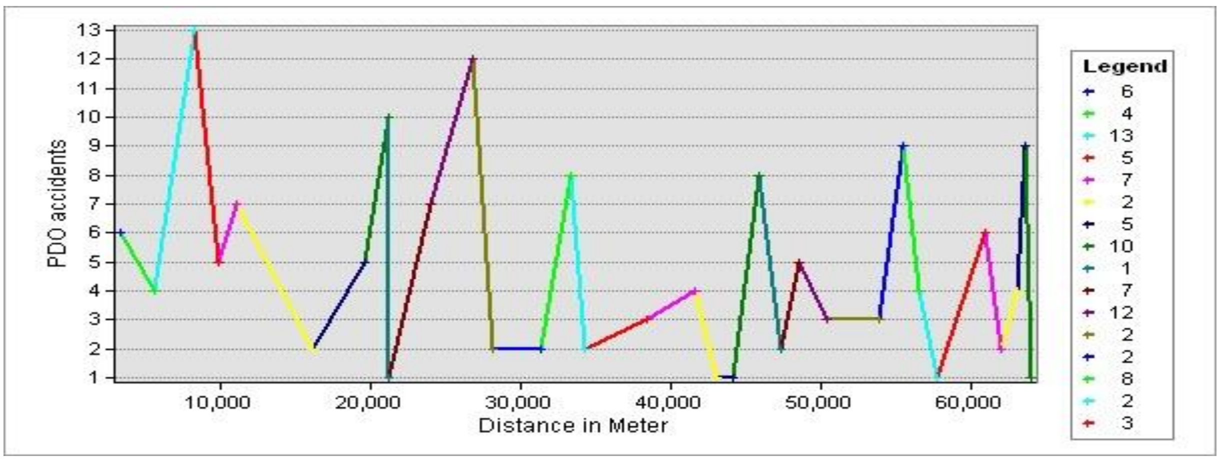


Figure 4.5: Graph of PDO versus distance

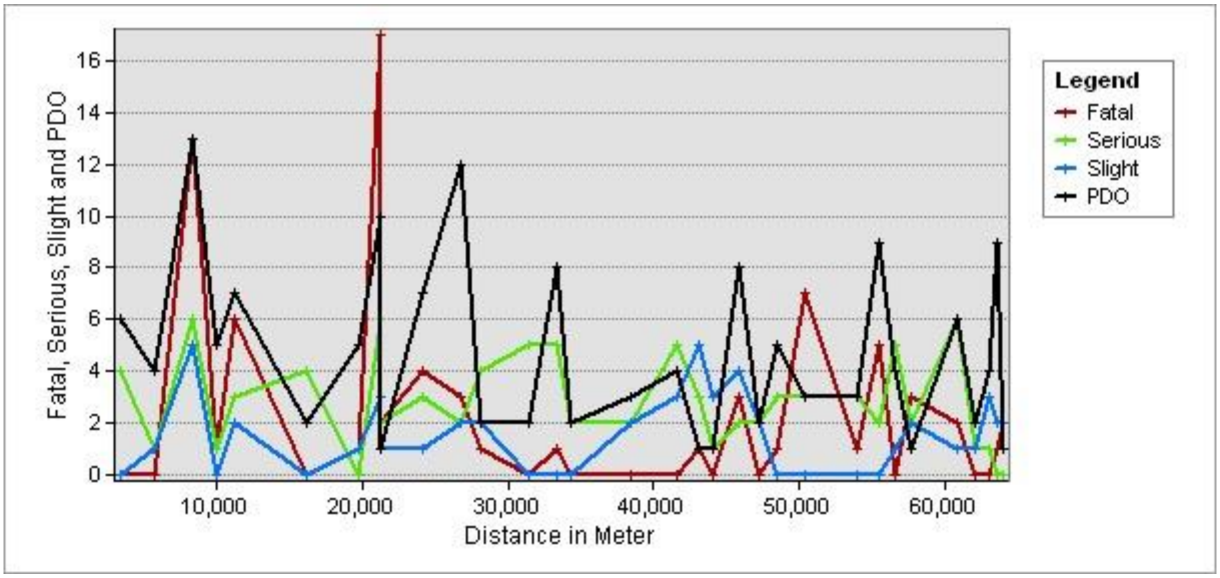


Figure 4.6: Graph of All accidents versus distance

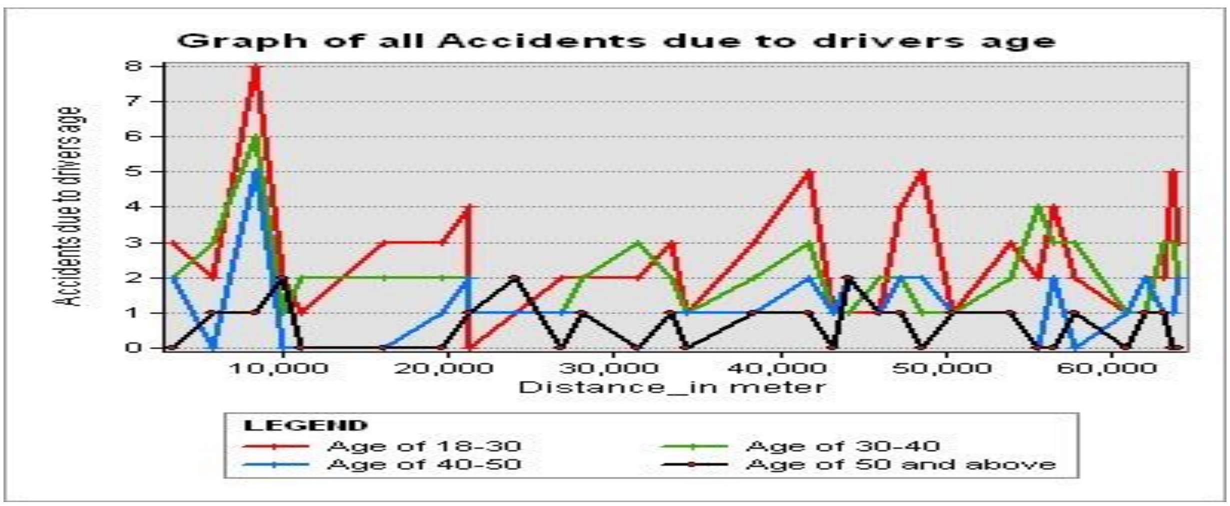


Figure 4.7: Graph Showing the Accident Variation due to Age of Drivers

Figure 4.7 is a plot of horizontal distance versus number of accidents for different drivers are groups who caused the accidents. The result showed that Drivers at the of 50 and above caused less number of accidents compared to that of the others drivers age groups. The number of accidents caused by the drivers meaningfully increases as the age of drivers decreased. Thus, drivers with experience and patience caused lesser number of accidents, and young and young adult (at the age from 18 to 30) caused higher number of accidents especially at distance from 7 to 10 km; 18 to 23 km; 40 to 50 km and 58 to 62 km.

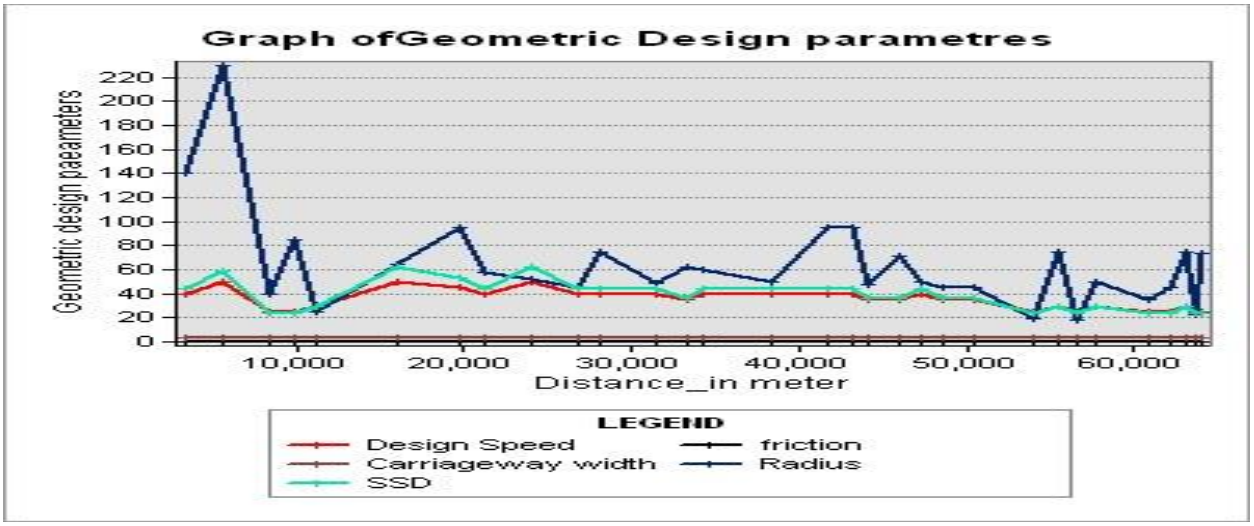


Figure 4.8: Graph Showing the Geometric design parameters

Figure 4.8 is a plot of horizontal distance versus geometric design parameters (design speed, carriage way width, stopping sight distance, lateral friction and horizontal curve radius). The design speed of the segment from Gibe bridge to Sekoru town ranged from 35 up to 40 Km/hr and the horizontal curve radius had a range from 18 up to 100 meters. The stopping sight distance varies from 20 up to 60 meters. the radius of curve is large on the rolling part only at location of (37.577, 8.228 meters to 37.561, 8.197 meters) or from the graph at distance (00.00 meter to 6+500 meters), and from 6+500 up 64+500 m the radius is between 15 m to 60 meter.

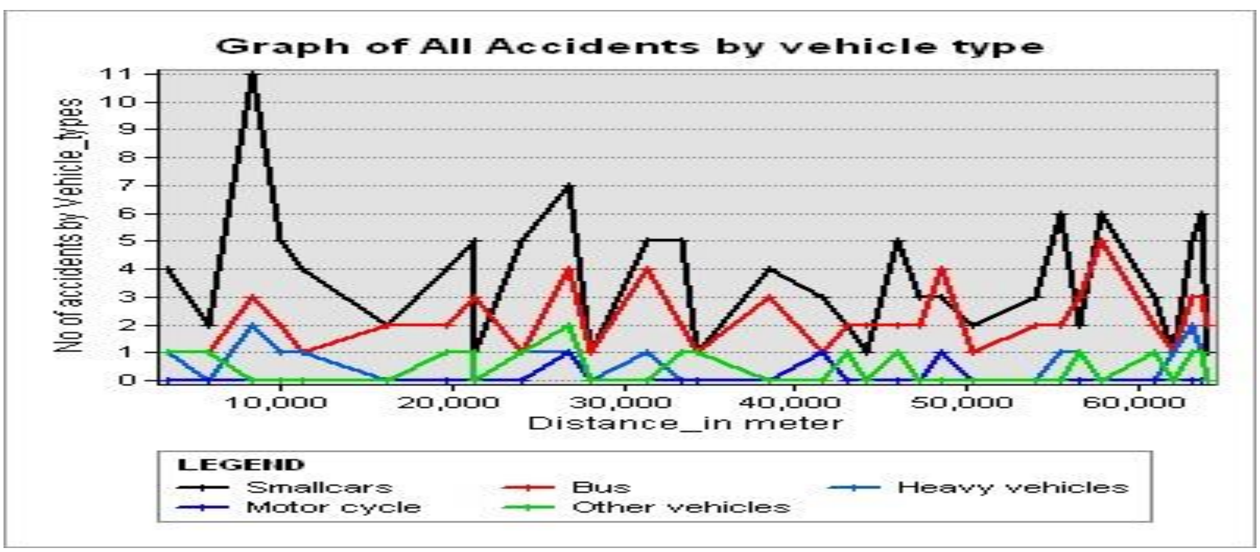


Figure 4.9: Graph Showing the Accident Variation by type of Vehicles

Figure 4.9 is a plot of horizontal distance versus number of accidents for different vehicle types (motor cycle, small vehicles, bus, heavy vehicles and others). From the result it is observed that higher number of accidents are caused by small cars followed by buses, other vehicles and heavy vehicles. Heavy vehicles caused accidents at location which have relatively small turning radius. From Figure 4.9 indicates the variation of accidents versus some vehicle types, on this graph small cars (vehicles which include pickups, minibus and vehicles carry up to 24 passengers) had high contribution as these vehicles are a day to day service contributor and followed by bus (vehicles which carry more than 24 passengers), other vehicles (which are not classified when recording) and heavy vehicles have relatively contributing small accidents. All these accidents happened due to small amount of radius, so many number of curves, restricted visibility. For more explanation, results are presented using statically analysis following presents

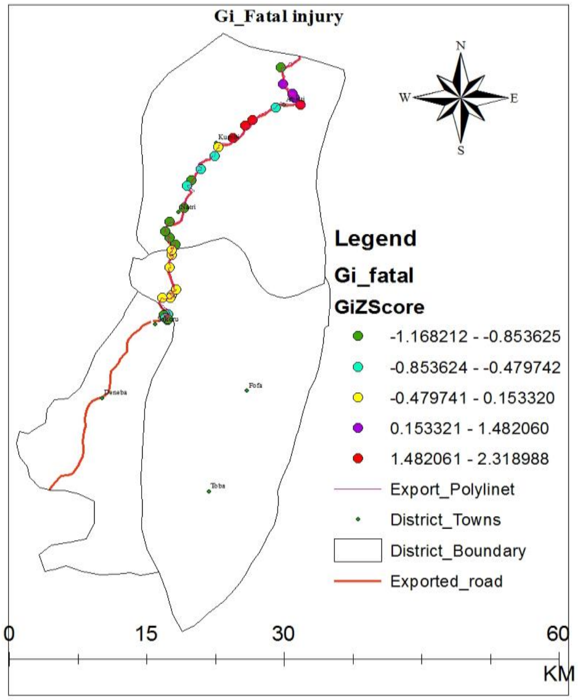
Thus, it is concluded the figures from 4.6 to 4.9 showed that as locations Shendebit, Kumbi, Natri, Saja and around Sekoru town higher number of accidents are observed. When the drivers are young and young adults, the design speed is low, curve radius is relatively small, and the vehicles are small cars and buses higher number of accidents occurred.

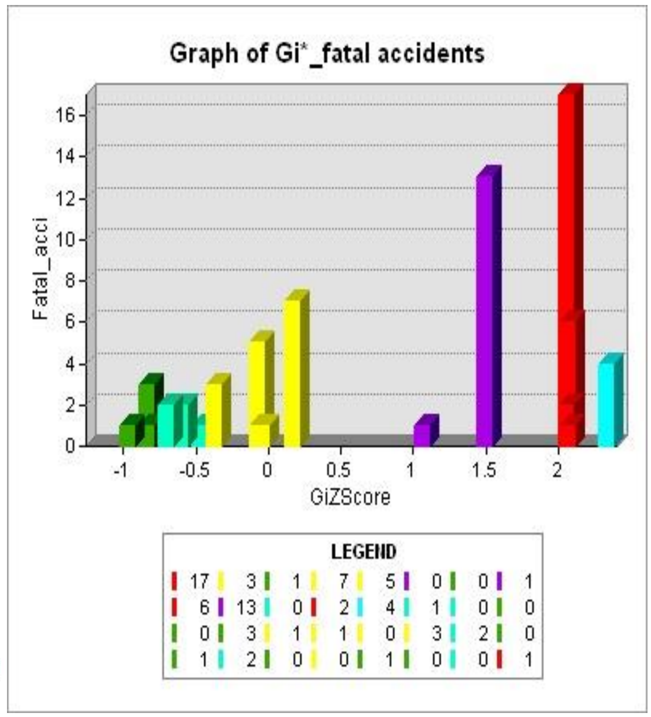
As McLean, 1996 showed that accident rates decrease as the curve radius increases. Also other literatures showed that accident rate increases due to lack of driving skills, poor knowledge of traffic rules and regulations, violation of speed limits, insufficient law enforcement, presence of animal drawn carts and animals, lack safety conscious design of road networks, lack of general safety awareness by pedestrians, restricted sight distance and inconsistency in design speed.

4.4. Spatial Statistical Analysis Tools

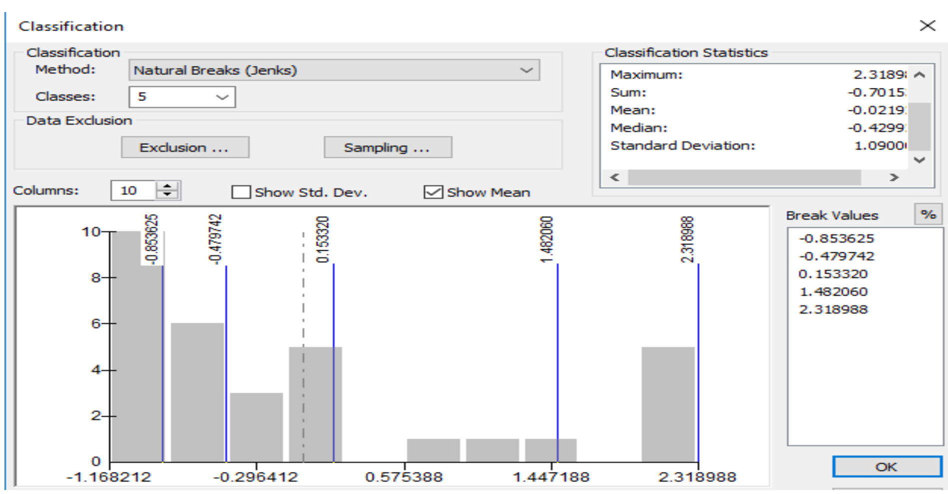
4.4.1. Getis-Ord Gi*

This section presents the description of accidents which occurred during the analysis period. Next, the result of the identification of hotspots by the conceptualization of spatial relationships, mapping and categorization results for Gi* statistic are presented and discussed. Appendix A1 presents the summary statistics of the crash frequency for Sekoru-Gibe valley highway from 2013 to 2017. The major contributing factors are also presented. Among the 222 crashes, 75% occurred on curve and the rest 25% occurred on straight roadway profile. Gi* results are Z scores, Z scores indicate the place of a particular value in a dataset relative to the mean, standardized with respect to the standard deviation Z equals to zero is equivalent to the sample/data mean, Z less than zero is a value less than the mean, Z greater than zero is a value greater the mean. Gi* compares local averages to global averages.



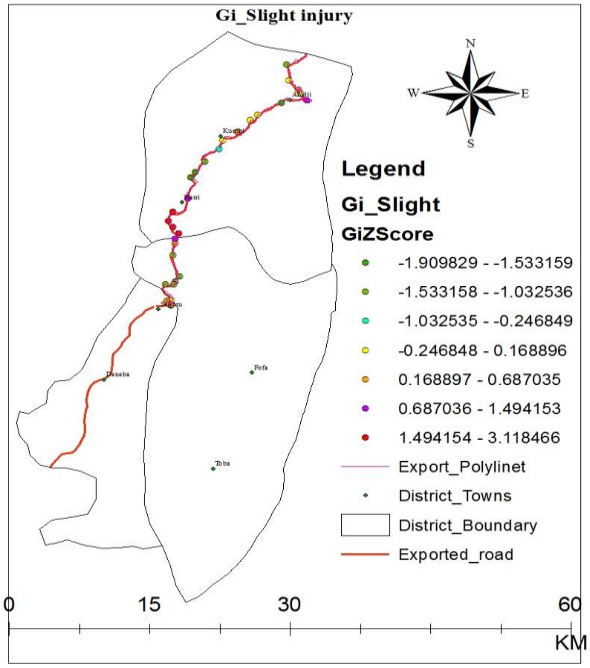


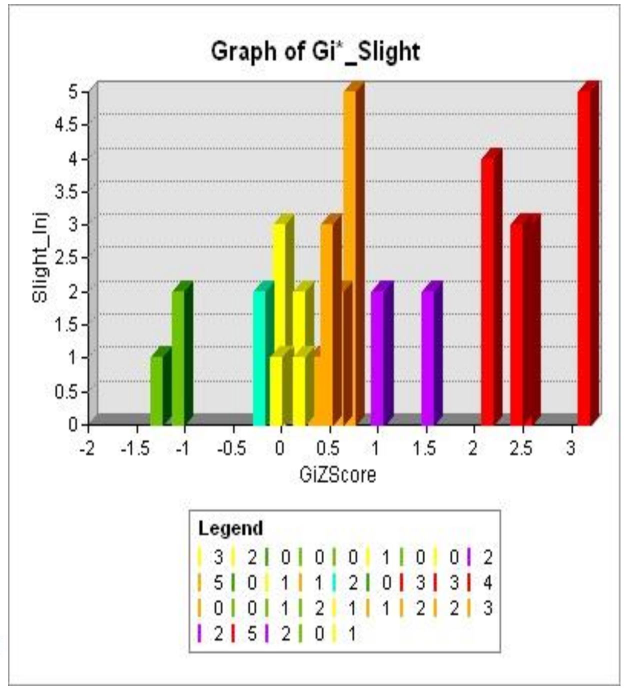
(a) mapping result **(b).** Vertical Graph



(c). Jenks classification with GiZScore

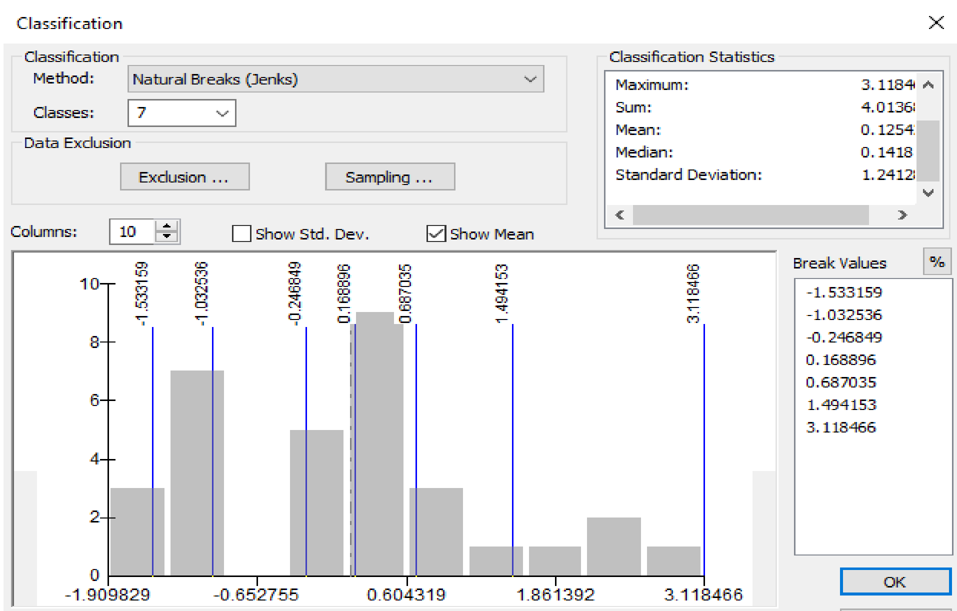
Figure 4.10: . Hotspot identification of Fatal accidents using Getis-Ord Gi*





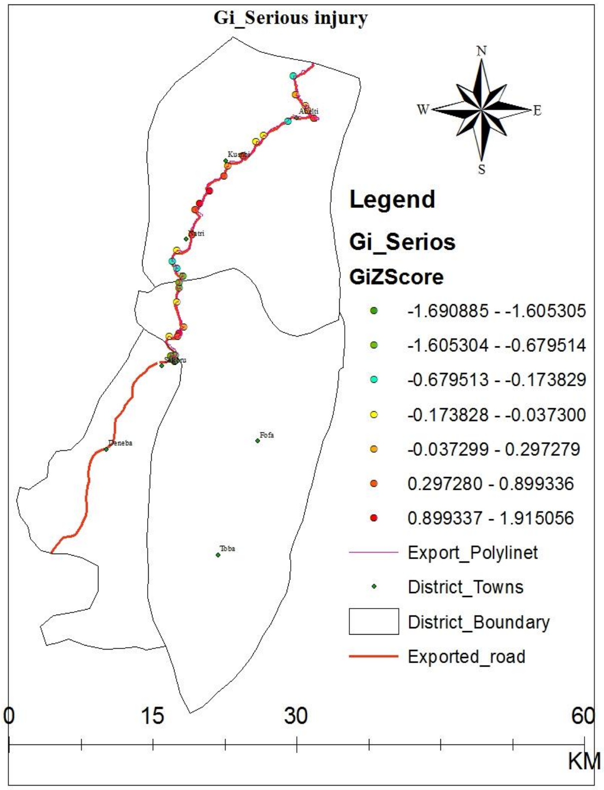
(a) mapping result

(b). Vertical Graph

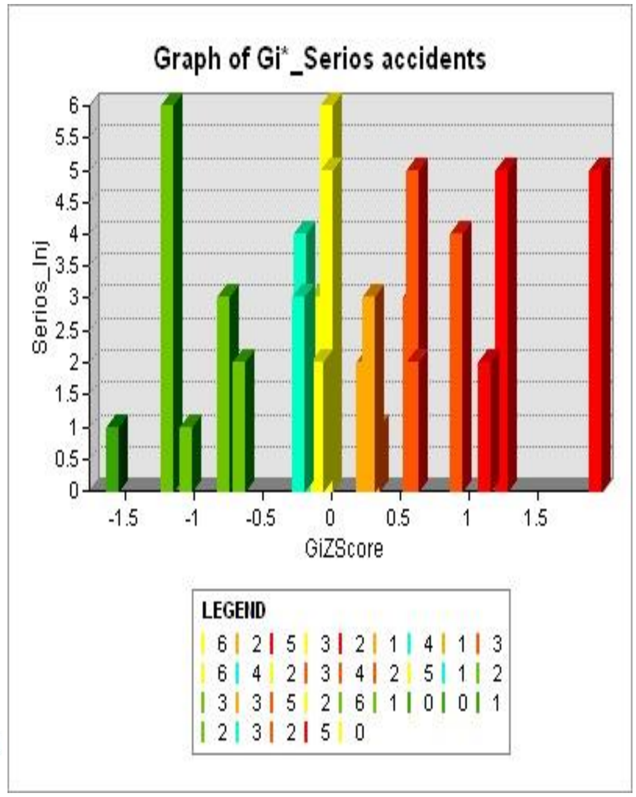


(c). Jenks classification with GiZScore

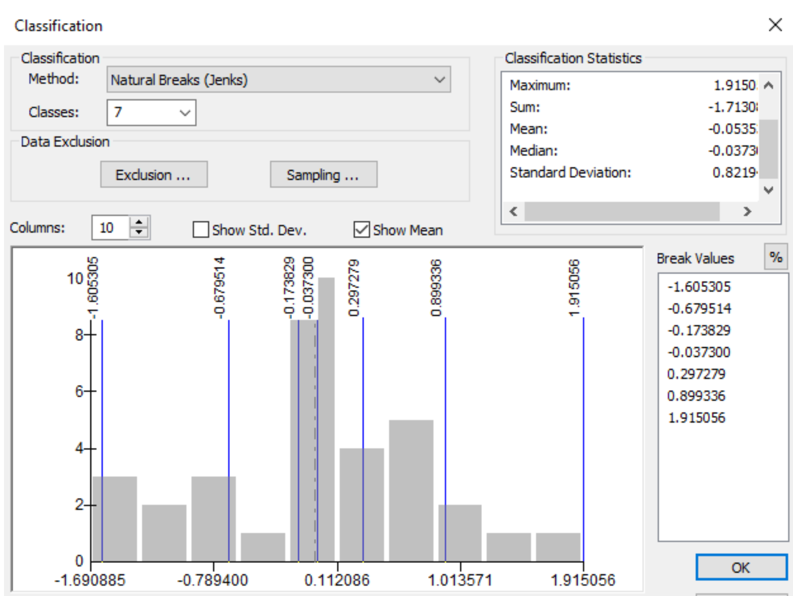
Figure 4.11: Hotspot identification of Slight accidents using Getis-Ord Gi*



(a) mapping result

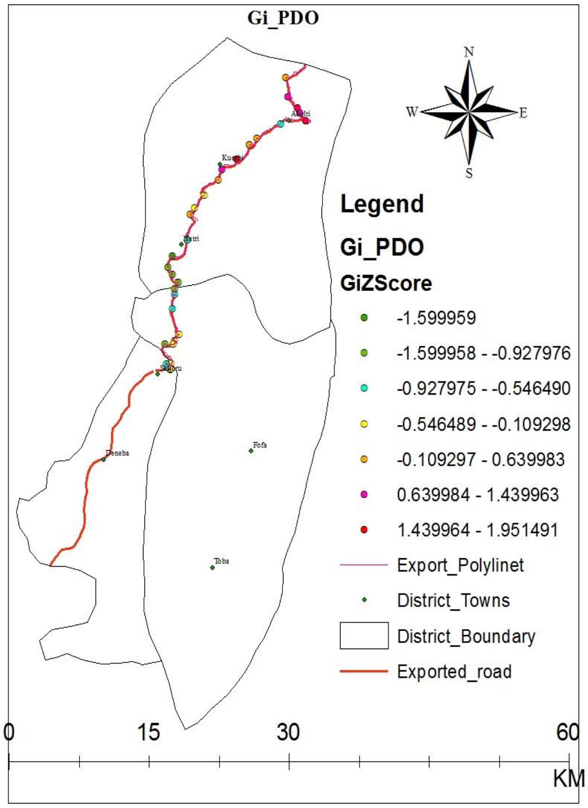


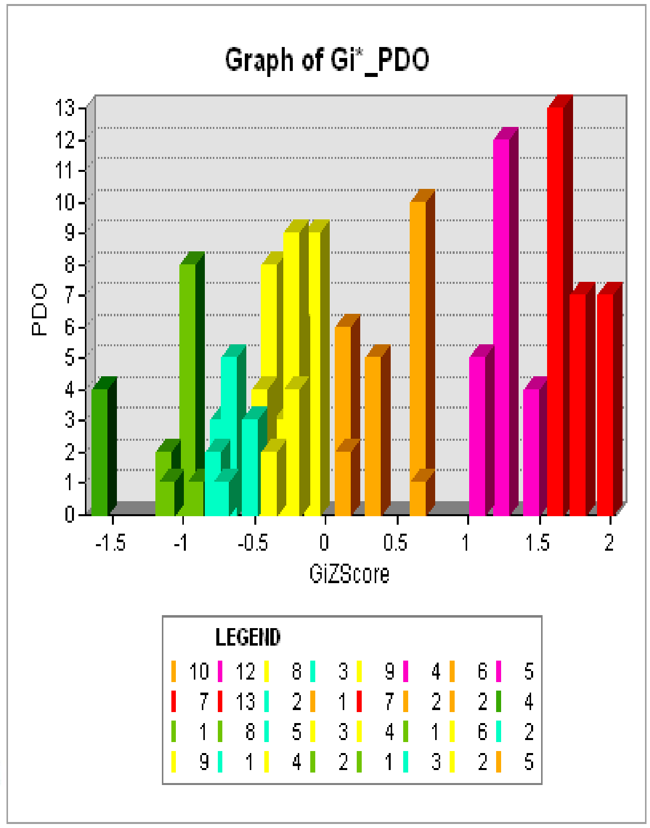
(b). Vertical Graph



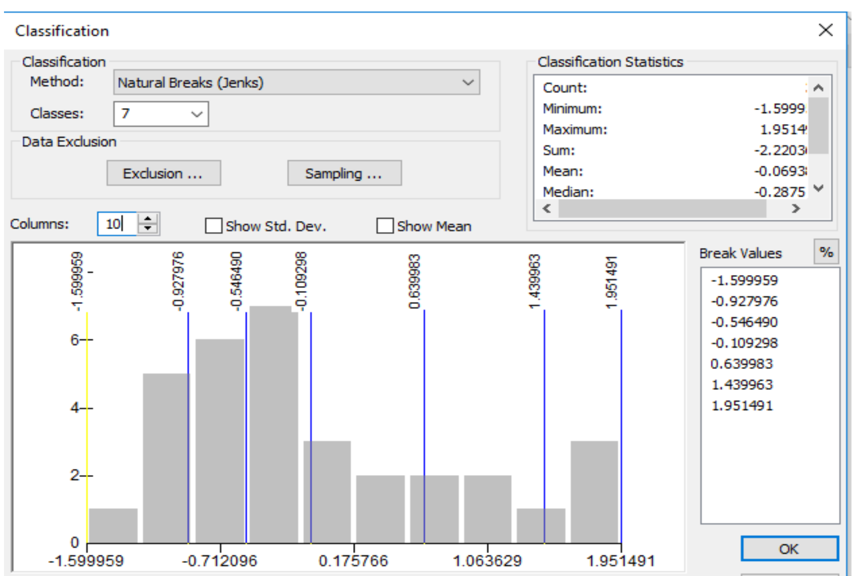
(c). Jenks classification with GiZScore

Figure 4.12: Hotspot identification of Serious accidents using Getis-Ord Gi*





(a) mapping result **(b)** Vertical Graph



(c). Jenks classification (c) with GiZScore

Figure 4.13: Hotspot identification of PDO accidents using Getis-Ord Gi*

After the completion of analysis on ArcGIS following results were identified. To be statistically significant hot spot, the tract needs to be surrounded by tracts with high, positive values and have a much higher, positive values than its neighbors. The inverse is true for the cold spots. Figure 4.10 to 4.13 shows the hotspot identified across the district using Gi* statistics for all year from 2013 to 2017. From figure 4.10 to 4.13 (a) represents the use of fixed distance, (b) represents the use inverse distance, and (c) represents the use categorical representations as the CSR for Getis-Ord Gi* statistics for the identification of hotspots. In the figures 4.10 to 4.13 presented in the map results (a) labels the red color (99% confidence values of standard deviation of >=2.576 means 1 in 100 chance that the observation would have just occurred naturally. i.e. what we are observing is extremely unusual), magenta(95% confidence means values of standard deviation of >=1.96) and dark yellow (90% confidence means values of standard deviation of >=1.645) colors represented for highly accident or hot spot areas based on their degree of freedom by the Getis-Ord Gi* statistics those points or locations have high value of GiZScore and low value of GiPValue points; deep yellow represents areas which are not significant which means based on the spatial autocorrelation those points have not enough neighborhood or they are clustered low areas; whereas the colors of white blue (90% confidence), dark green (95% confidence) and deep green (99% confidence) areas are represented the low accident areas or cold spot areas in other means those points or locations have high GiPValue and low GiZScore points. In the figures 4.10 to 4.13 presented in (b) are shows the fatal accident, slight accident, serious accident and PDO vs GiZScore respectively. Colors are similarly presented as on figures of 4.10 to 4.13 (a) here the x-axis represents the critical values or the GiZScore values, these critical values are the points which are presented the standard deviation of locations or the accident points. As described on the graph GiZScore values are presented as negative, zero and positive values, the negative values are cold spot areas or low accident areas as compared to all of the section whereas the zero value are not statically significant and the positive values are highly accidents areas or hot spot areas. From the figure shown in 4.10 to 4.13 (a) the highly hot spot areas for fatal are at the location of (37.568, 8.189 meters), (37.581, 8.176 meters), (37.525, 8.155 meters), and (37.432,7.921 meters). for slight accidents hot spot areas are at location of (37.455, 8.058 meters), (37.427, 8.029 meters), (37.441,8.019 meters) , and (37.432, 7.933

meters); for serious accidents hot spot areas are presented at location of (37.576, 8.173 meters), (37.507, 8.138 meters), (37.467, 8.101 meters), (37.451, 8.055 meters), (37.438, 7.958 meters), (37.436, 7.937 meters), (37.431, 7.933 meters) and for PDO the highly hotspot areas are presented at location of (37.561, 8.198 meters), (37.569, 8.188 meters), (37.573, 8.182 meters), (37.578, 8.174 meters), (37.564, 8.137 meters), (37.438, 7.978 meters), (37.435, 7.933 meters), (37.452, 8.081). one can notice that the network Gi* identifies more hotspots clustered close to the north east of the district for fatal, serious and property damage only accidents on the other hand Gi* identifies more hotspot clustered in the middle for slight accidents. The northeast gate of the district area where a mix of accident location with in high hotspot frequency.

4.4.2. Anselin Local Moran Index

The left hand side of Figure 4.14, 4.15, 4.16 and 4.17 are map representation of the result of Anselin local Moran Index for fatal, serious, slight and PDO accidents and the right hand side is the graphical representation of the result of Anselin local Moran Index for fatal, serious, slight and PDO accidents. The map representation shows the relative clustering of fatal accidents to other types of accidents. Anselin local Moran Index values which have a confidence level of 95% or less are statistically non-significant. Cluster-high, HH, points are points in which the clustering of fatal accidents is high and the clustering of other types of accidents surrounding the fatal accidents is also high. High- outlier, HL, points are points in which the clustering of fatal accidents is high and the clustering of other types of accidents surrounding the fatal accidents is low. Low-outlier, LH, points are points in which the clustering of fatal accidents is low and the clustering of other types of accidents surrounding the fatal accidents is high. Cluster-low, LL, indicated\s a low clustering of fatal accidents and other types of accidents surrounding the fatal accident at specific point. Thus, HH values occur at locations at Shendebitu (Abelti cluster), and kumbi cluster. at the locations of (37.5, 8.2) around Abelti shen debitu; for serious injury accidents their statistical outlier based on the Anselin Local Moran Index is laid at the locations of (37.434,7.933 meters north eastern of Sekoru town on the gate of Yem district; 37.57, 8.16meters eastern part of Abelti specifically shen debitu; 37.518, 8.15 meters; 37.434, 7.933 meters which north eastern part of kumbi) and for slight injury the highest clustered areas are most of them are located at (37.442,8.013 meters; 37.430,8.029 meters; 37.430, 8.039 meters; 37.583, 8.174 meters zones covered from south west

Natri town to north east kumbi town) all other points clustered low or not significant due to their neighborhood.

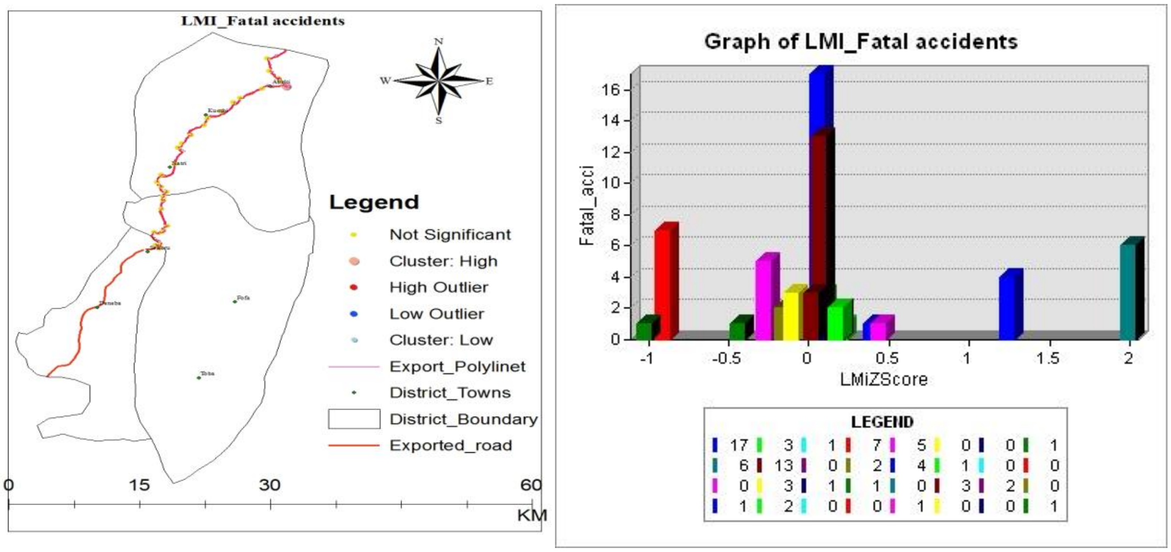


Figure 4.14: Fatal accidents using LMiIndex and Vertical graph vs LMiZScore

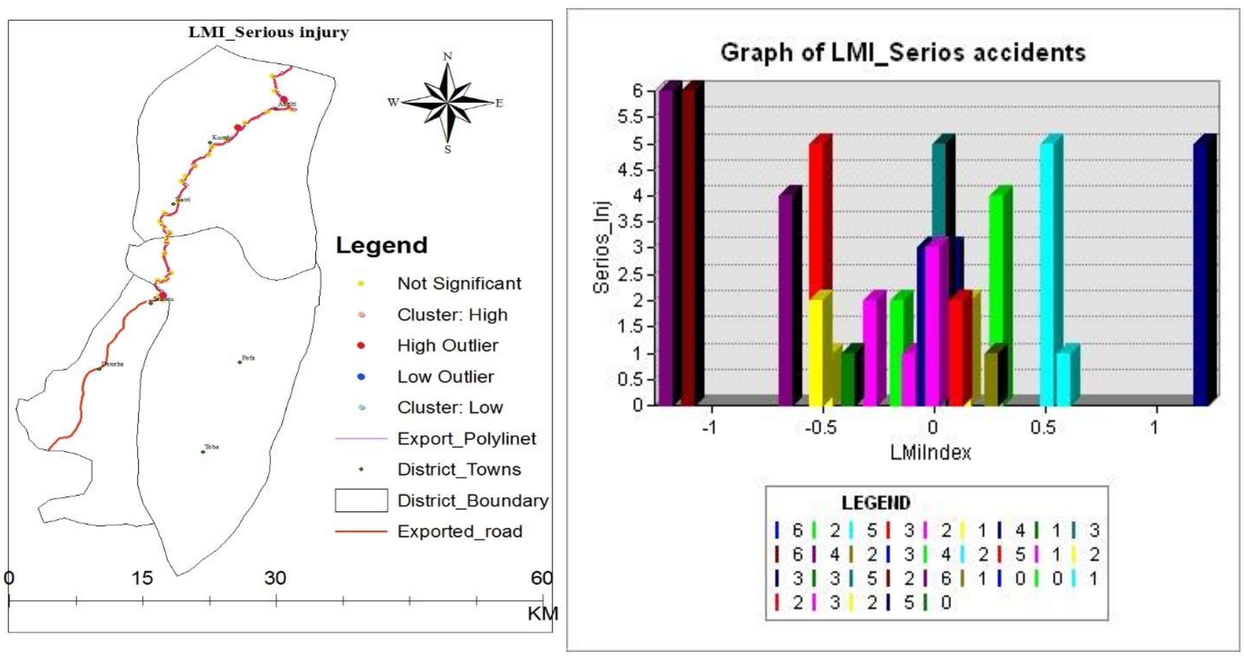


Figure 4.15: Serious accidents using LMiIndex and vertical graph vs LMiIndex

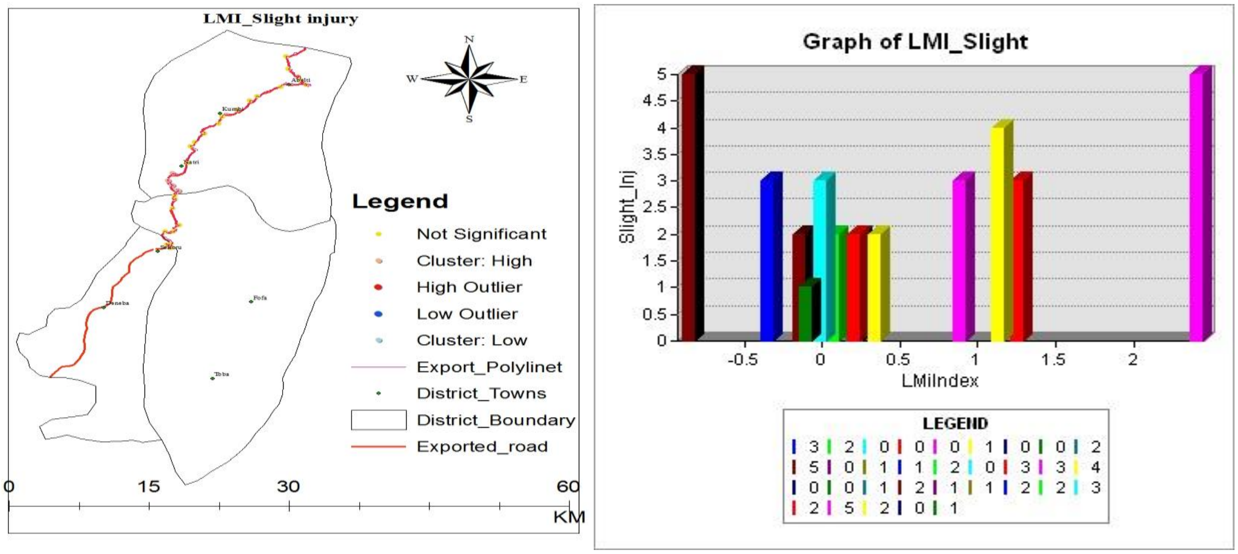


Figure 4.16: Slight accidents using LMiIndex and Vertical graph vs LMiIndex

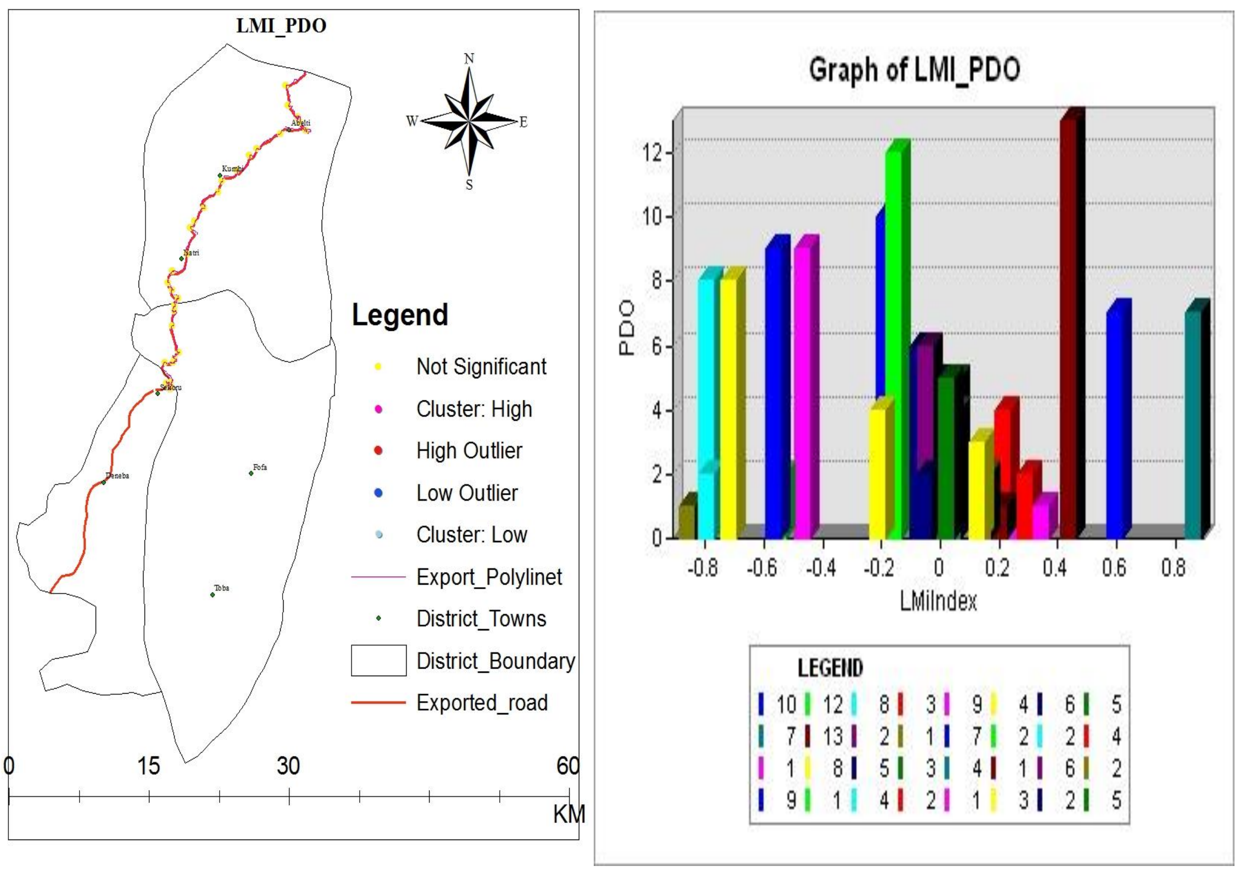
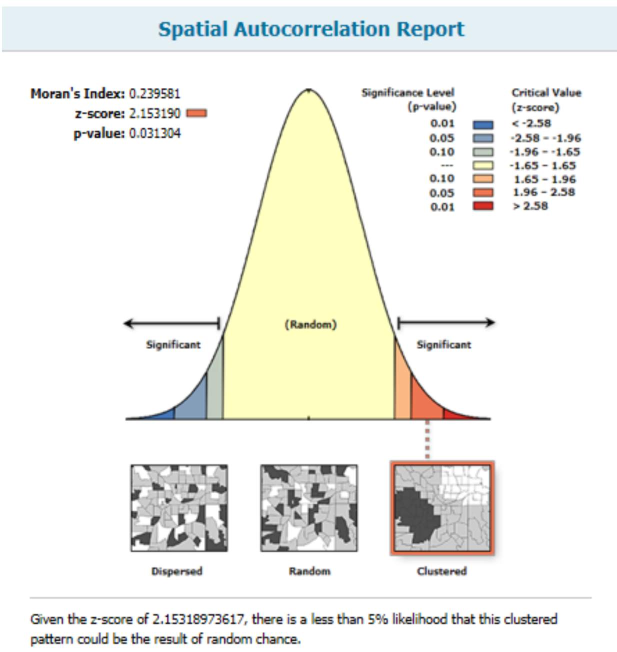


Figure 4.17: PDO accidents using LMiIndex and vertical graph vs LMiIndex

4.4.3. Global Moran Index

Figure 4.18, 4.19, 4.20 and 4.21 shows results of Global Moran Index values for slight, serious, PDO and fatal, accidents. As indicated in the figure, in addition to Z score and P values, Global Moran Index analysis output includes Moran index, expected index and variance. The spatial autocorrelation report is presented in the form of normal distribution curve. The normal distribution curve presents significant and random values based on P-values. Z-score values of less than -1.65 and greater than 1.65 are significant, and Z-score values in between -1.65 and 1.65 are random. Global Moran Index analysis presents Z-score values of less than -1.65 as dispersed or cold spot, Z-score values of greater than 1.65 as clustered or hotspot, and Z-score values in between -1.65 and +1.65 are random values.



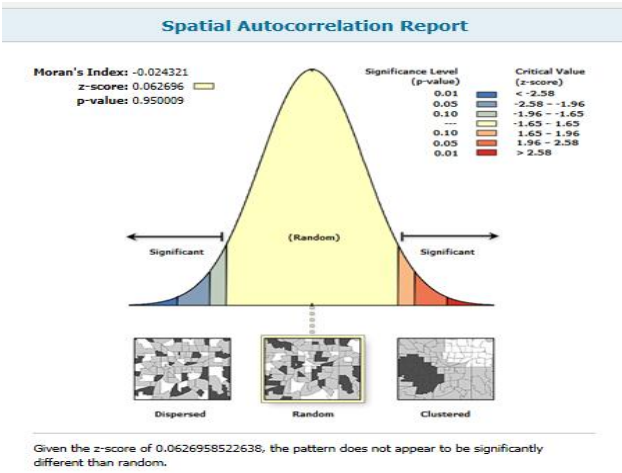
Global Moran's I Summary

Moran's Index:	0.239581
Expected Index:	-0.032258
Variance:	0.015939
z-score:	2.153190
p-value:	0.031304

Dataset Information

Input Feature Class:	Export_Alldata
Input Field:	SLIGHT_INJ
Conceptualization:	FIXED_DISTANCE
Distance Method:	EUCLIDEAN
Row Standardization:	False
Distance Threshold:	0.0236 Meters
Weights Matrix File:	None
Selection Set:	False

Figure 4.18: Slight injuries using Global Moran Index



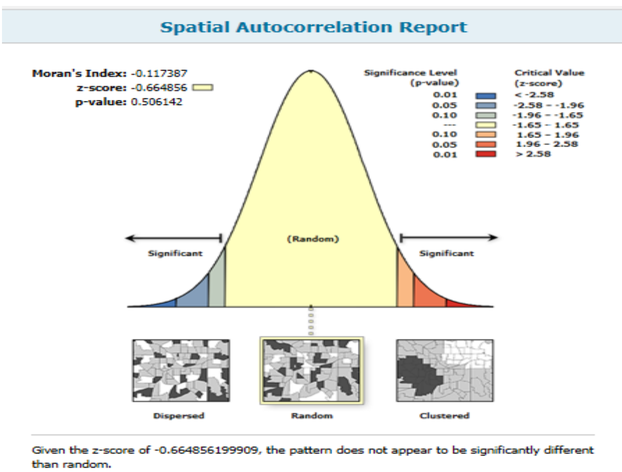
Global Moran's I Summary

Moran's Index:	-0.024321
Expected Index:	-0.032258
Variance:	0.016026
z-score:	0.062696
p-value:	0.950009

Dataset Information

Input Feature Class:	Export_Alldata
Input Field:	PDO
Conceptualization:	FIXED_DISTANCE
Distance Method:	EUCLIDEAN
Row Standardization:	False
Distance Threshold:	0.0236 Meters
Weights Matrix File:	None
Selection Set:	False

Figure 4.19: PDO accidents using Global Moran Index



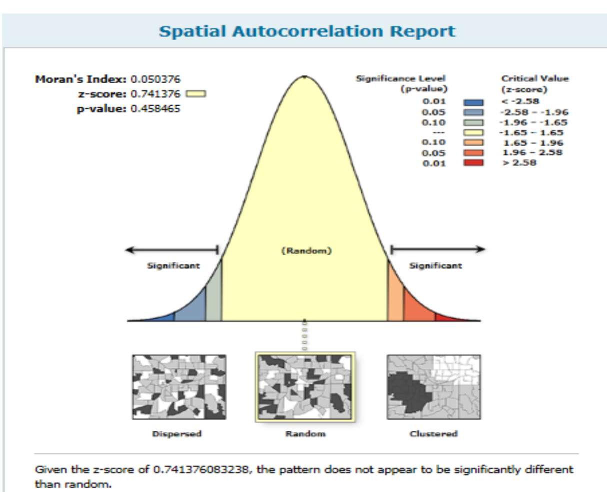
Global Moran's I Summary

Moran's Index:	-0.117387
Expected Index:	-0.032258
Variance:	0.016394
z-score:	-0.664856
p-value:	0.506142

Dataset Information

Input Feature Class:	Export_Alldata
Input Field:	SERIOS_INJ
Conceptualization:	FIXED_DISTANCE
Distance Method:	EUCLIDEAN
Row Standardization:	False
Distance Threshold:	0.0236 Meters
Weights Matrix File:	None
Selection Set:	False

Figure 4.20: Serious injuries using Global Moran Index



Global Moran's I Summary

Moran's Index:	0.050376
Expected Index:	-0.032258
Variance:	0.012423
z-score:	0.741376
p-value:	0.458465

Dataset Information

Input Feature Class:	Export_Alldata
Input Field:	FATAL_ACCI
Conceptualization:	FIXED_DISTANCE
Distance Method:	EUCLIDEAN
Row Standardization:	False
Distance Threshold:	0.0236 Meters
Weights Matrix File:	None
Selection Set:	False

Figure 4.21: Fatal accidents using Global Moran Index

4.5. Geostatistical Analyst

4.5.1. Inverse Distance Weighting

After the completion of analysis on ArcGIS following results were identified. Figure 4.22 to 4.25 shows results using Inverse Distance Weighting for fatal accidents, serious injury, slight injury and property damage only accidents respectively. In this method results are analyzing depend on their distance neighbor to each other. Based on field contour, the symbology is grouped in to five classes each identified by unique colors. The red color showed high number of fatal accident, followed by green, yellow, black and blue (which is the least). Thus, high number of fatal accidents occurred at Abelti, Kumbi, Natri, Saja and around eastern part of Sekoru town. As shown in Figure 4.21 for fatal accident the thematic map layer of red color represented for higher accident areas which is the segment around kumbi and slightly decreases to shen debitu place around Abelti. In Figure 4.22 for slight injury major accidents have shown on the segment of Saja to Natri and kumbi to gibe valley. In Figure 4.23 in the case of serious accidents the segment from Natri to kumbi is dominant. Lastly in Figure 4.24 for property damage only the segments around kumbi and from Abelti to Gibe bridge are dominantly represented.

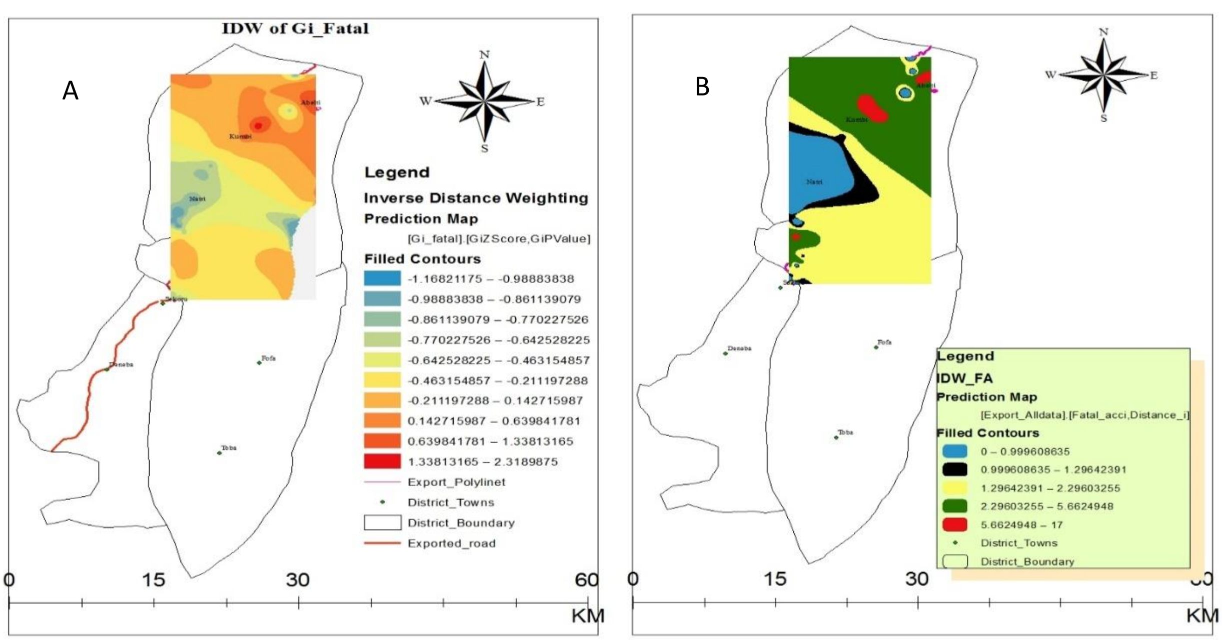


Figure 4.22: Hotspot mapping result of Fatal accidents (A) Gi*, (B) using IDW

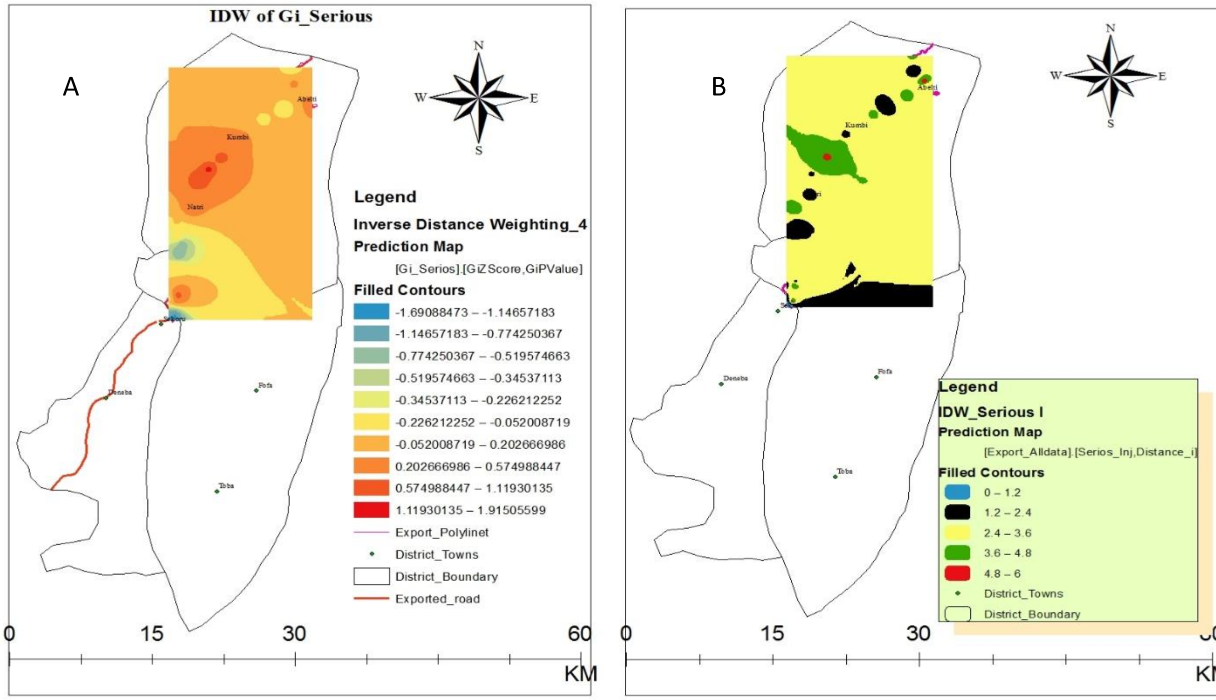


Figure 4.23: Hotspot mapping result of Serious accidents (A) Gi*, (B) using IDW

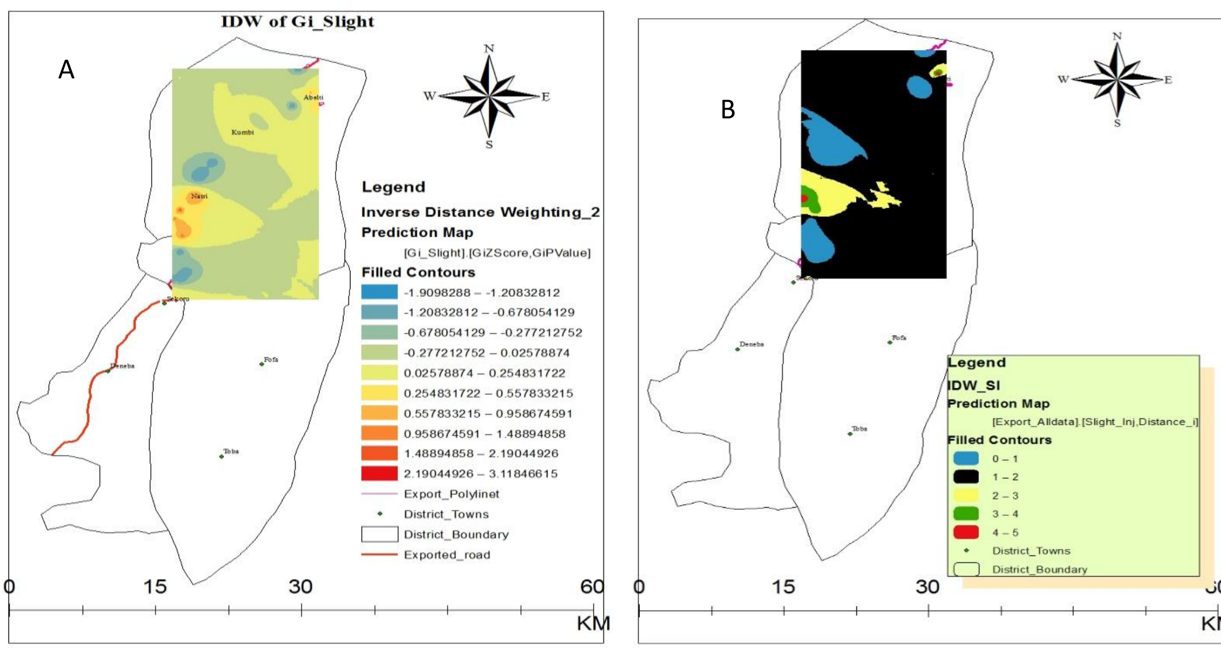


Figure 4.24: Hotspot mapping result of Slight accidents (A) Gi*, (B) using IDW

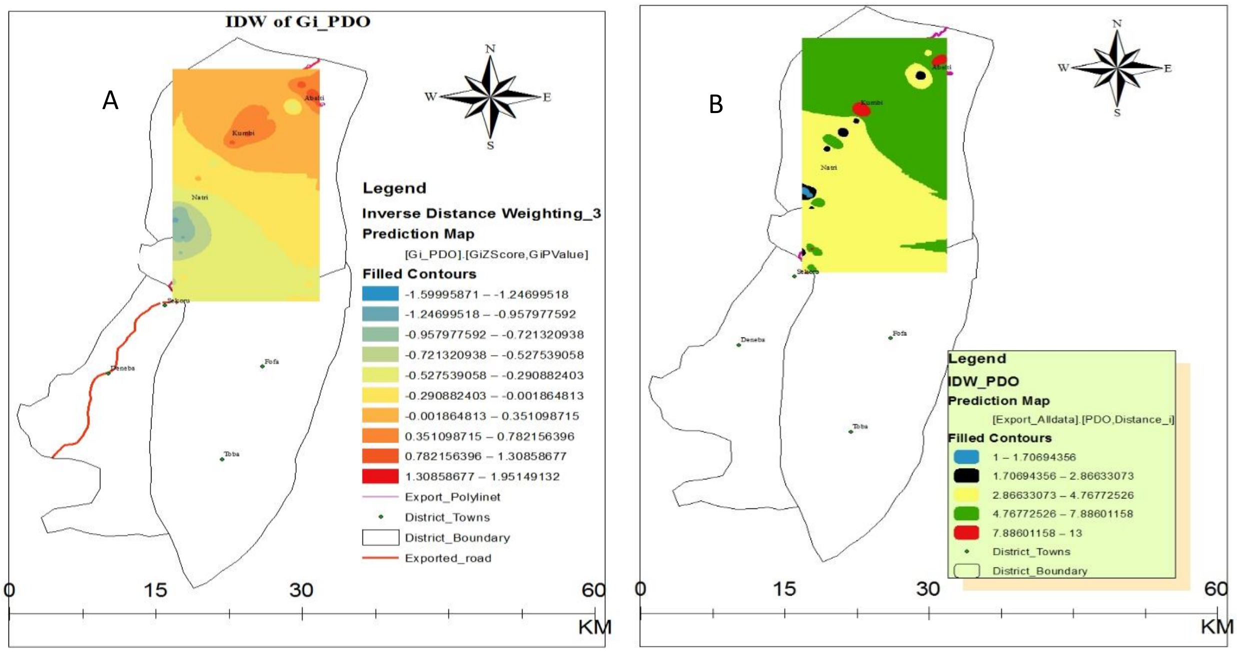


Figure 4.25: Hotspot mapping result of PDO accidents (A) Gi*, (B) using IDW

4.5.2. Empirical Bayesian Method

Figure 4.26, 4.27, 4.28 and 4.29 shows the result of analysis of number of fatal, serious, slight and PDO accidents using Empirical Bayesian Kriging method of Geostatistical analysis. This tool maps number of fatal accidents through the process of sub setting and simulations. The symbology is grouped into five classes each corresponding a unique color.

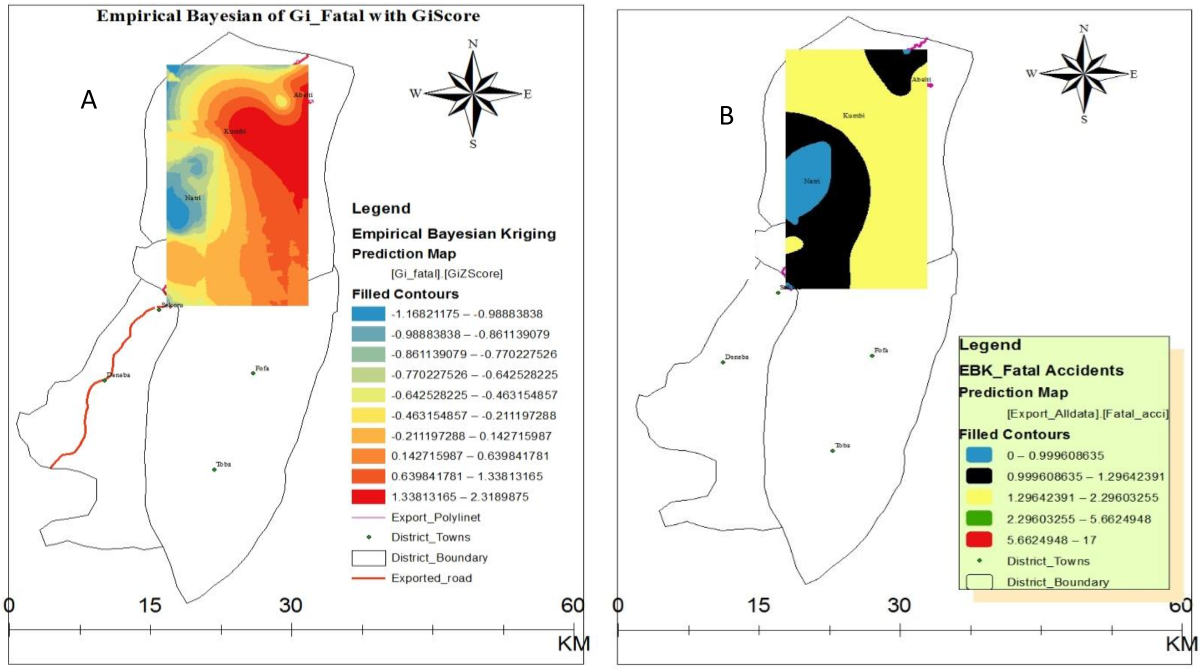


Figure 4.26: Hotspot mapping result of Fatal accidents (A) Gi*, (B) using EBK

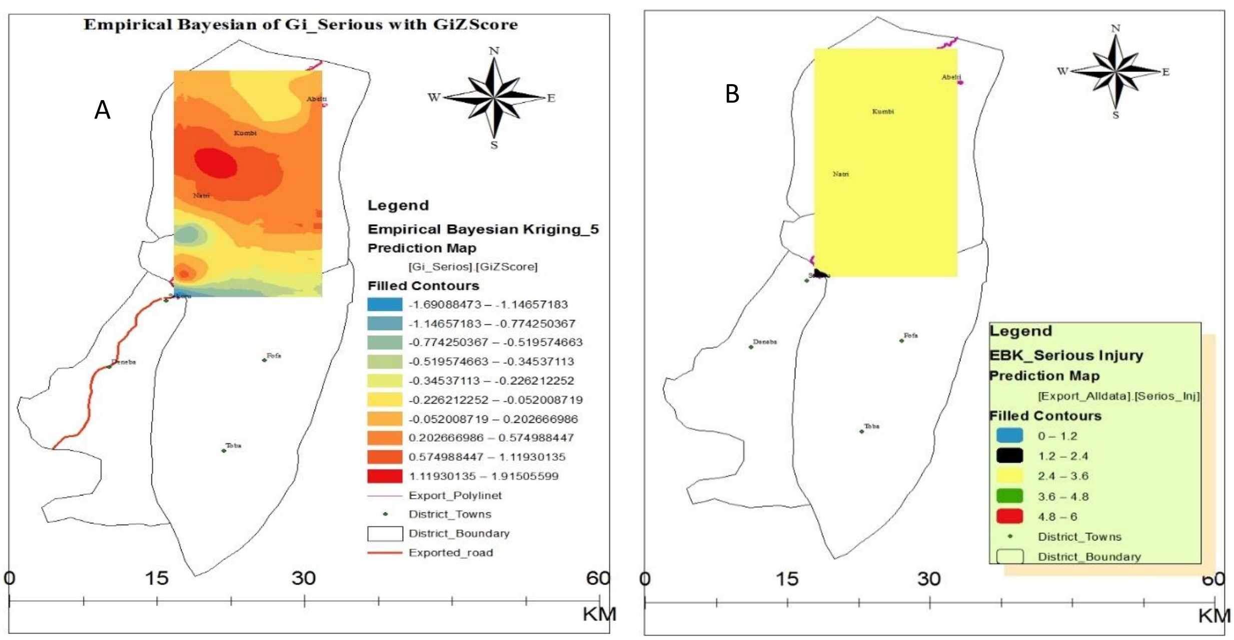


Figure 4.27: Hotspot mapping result of Serious accidents (A) Gi*, (B) using EBK

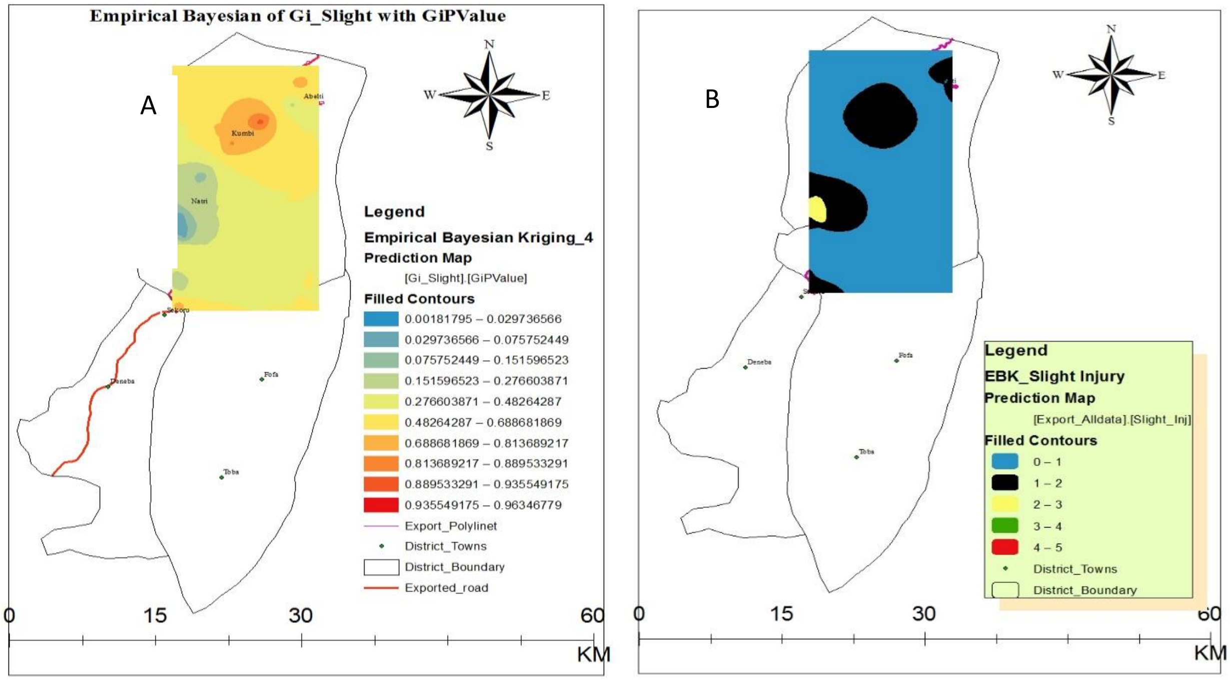


Figure 4.28: Hotspot mapping result of Slight accidents (A) Gi*, (B) using EBK

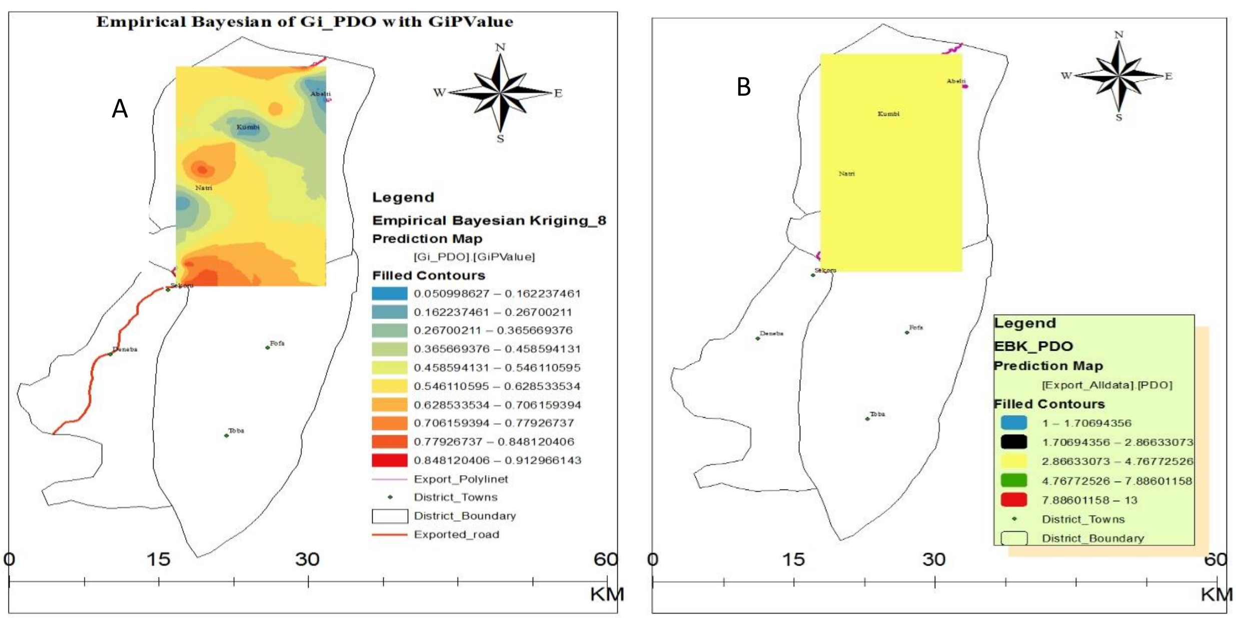
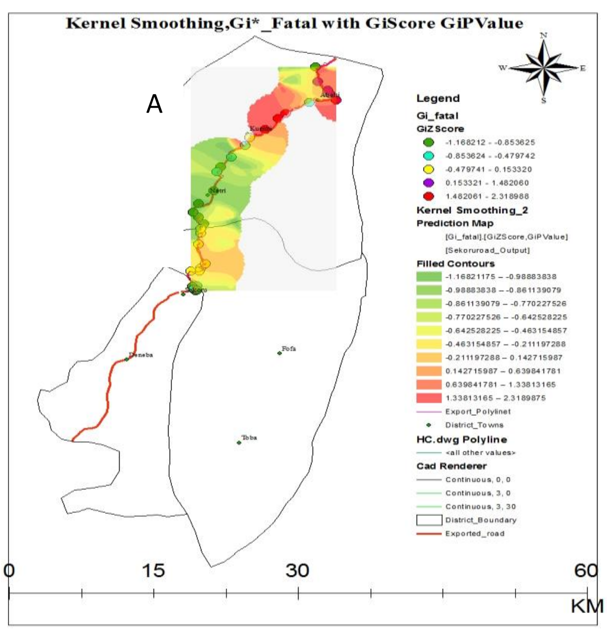


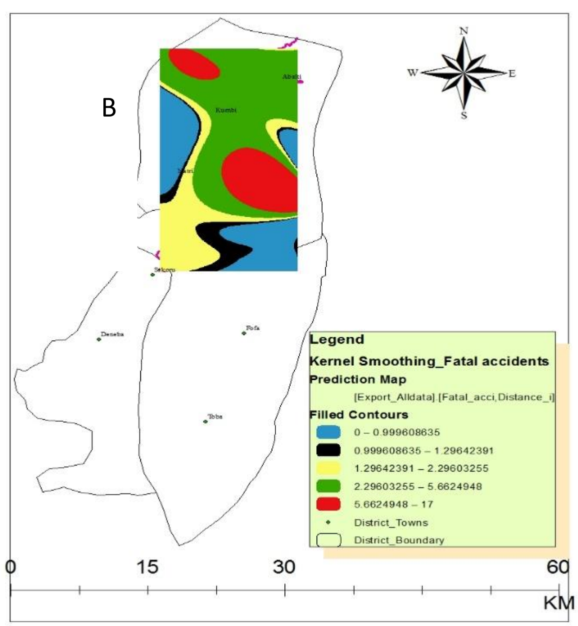
Figure 4.29: Hotspot mapping result of PDO accidents (A) Gi*, (B) using EBK

4.5.3. Kernel Smoothing Density Method

This analysis was used in a GIS environment to identify hotspot areas of traffic accidents such as fatal accidents, slight injury accidents, serious injury accidents and property damage only accidents happened on the district. From Figure 4.30 to 4.33 below shows hot spot identification of all accidents using kernel smoothing density. For kernel smoothing density analysis method the most suitable density level is bandwidth (Bailey and Gatrell, 1995; Fortheringham et al., 2000; Silverman, 1986). Selecting bandwidth affects output of hotspots. For instance, large bandwidth shows hotspot area in large form and small bandwidth in small form. In this analysis, five categories were performed based on their field contours which consider that, one can prioritize investigation to solve safety problem of hotspot and this classification is shown in Figure 4.30 to 4.33. Due to the high number of accidents in Sekoru and Yem district as shown in the Figures, the kernel smoothing density generated by different colors representative density classes for the accidents from low (blue color), medium (yellow color), high (dark red color) and very high density (red color) accident areas.

Figure 4.30 to 4.33 shows results of kernel smoothing density level for accidents such as fatal, injury, slight and property damage in the district from 2013 to 2017. All these computations are computed through ArcMap. The result showed that the high accident zones are concentrated in the area of Abelti (shen debitu), kumbi, Natri, Saja and Sekoru towns. This is intuitive as the higher level of traffic interaction engenders more safety problems. As the highways are extending out ward from the Sekoru district, the risk level is decreasing.





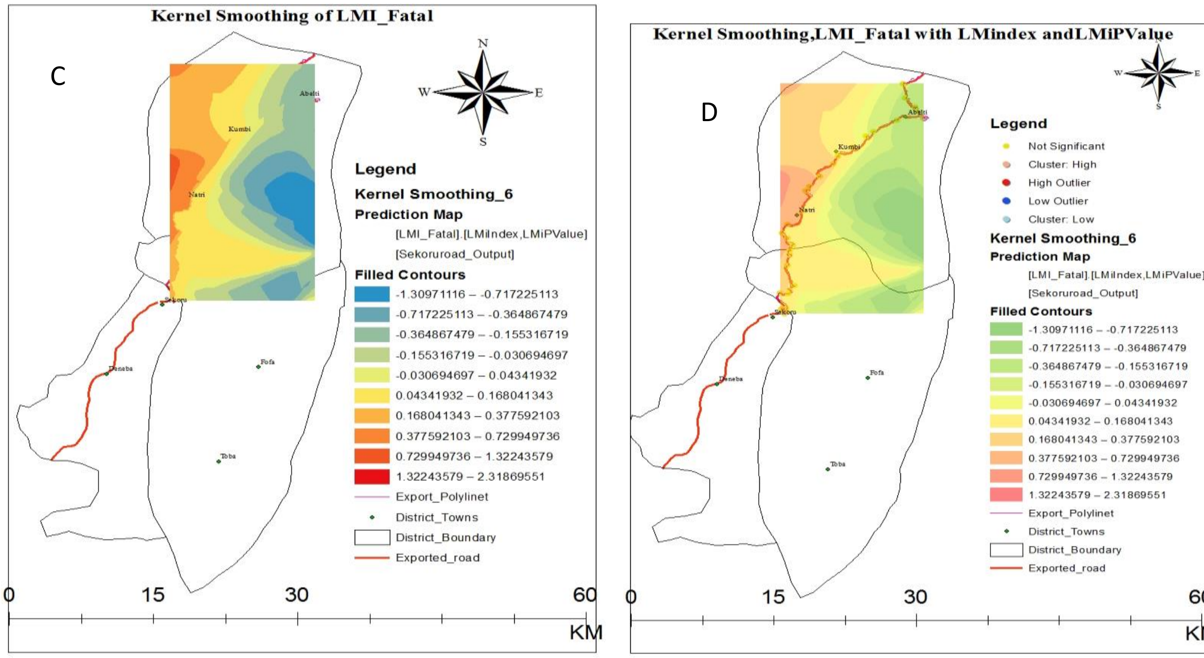


Figure 4.30: Hotspot mapping result of Fatal accidents (A) Gi*, (B) using KSDE

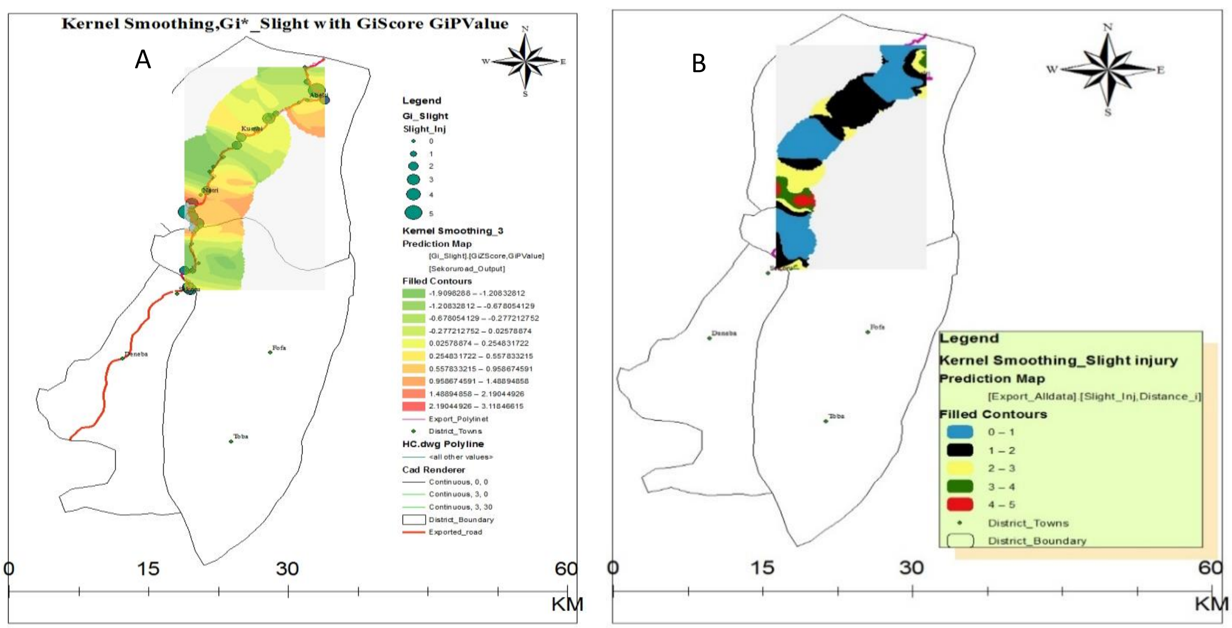


Figure 4.31: Hotspot mapping result of slight accidents (A) Gi*, (B) using KSDE

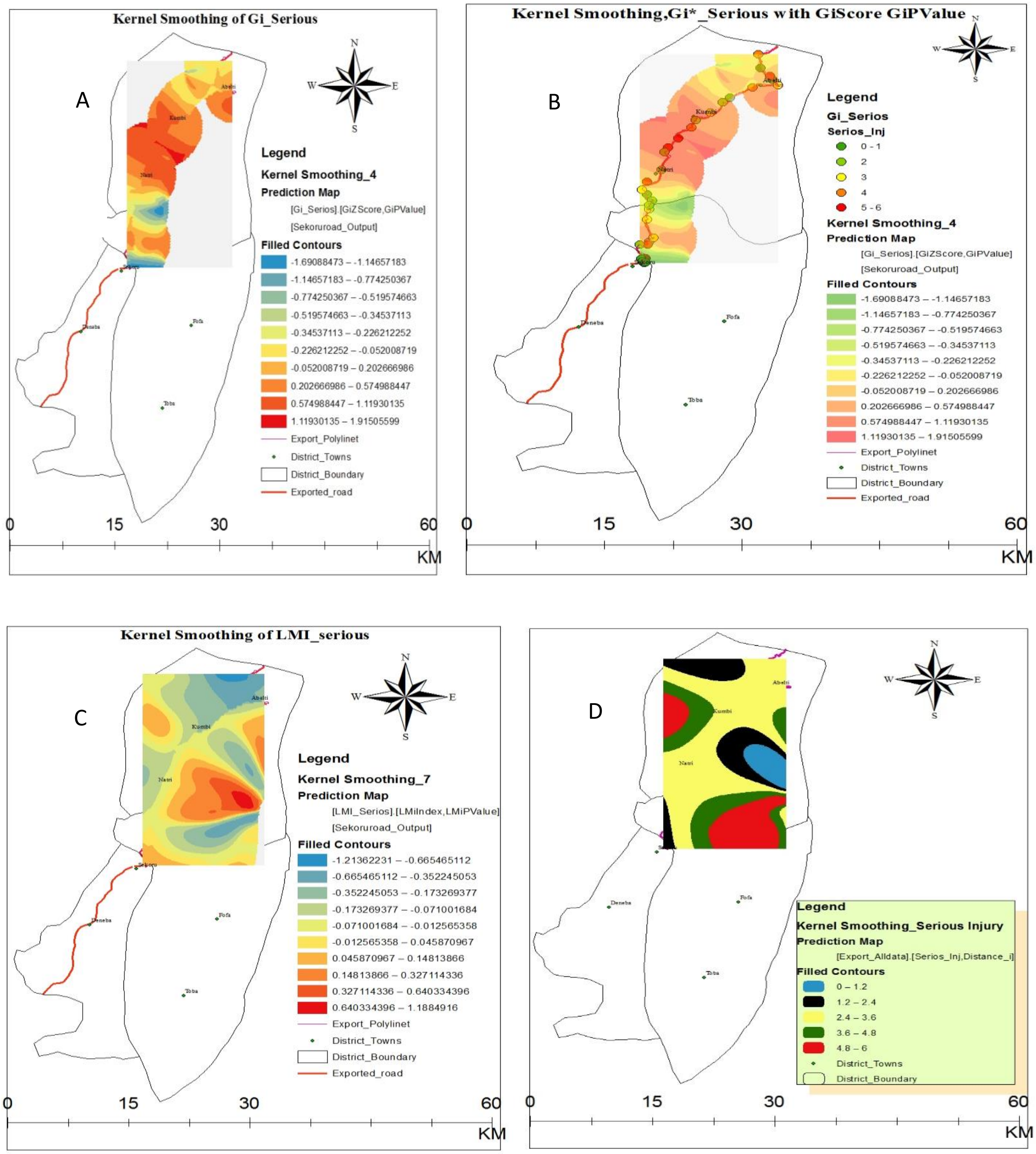


Figure 4.32: Hotspot mapping result of Serious accidents (A) and (B) Gi*, (C) LMI, (D) using KSDE

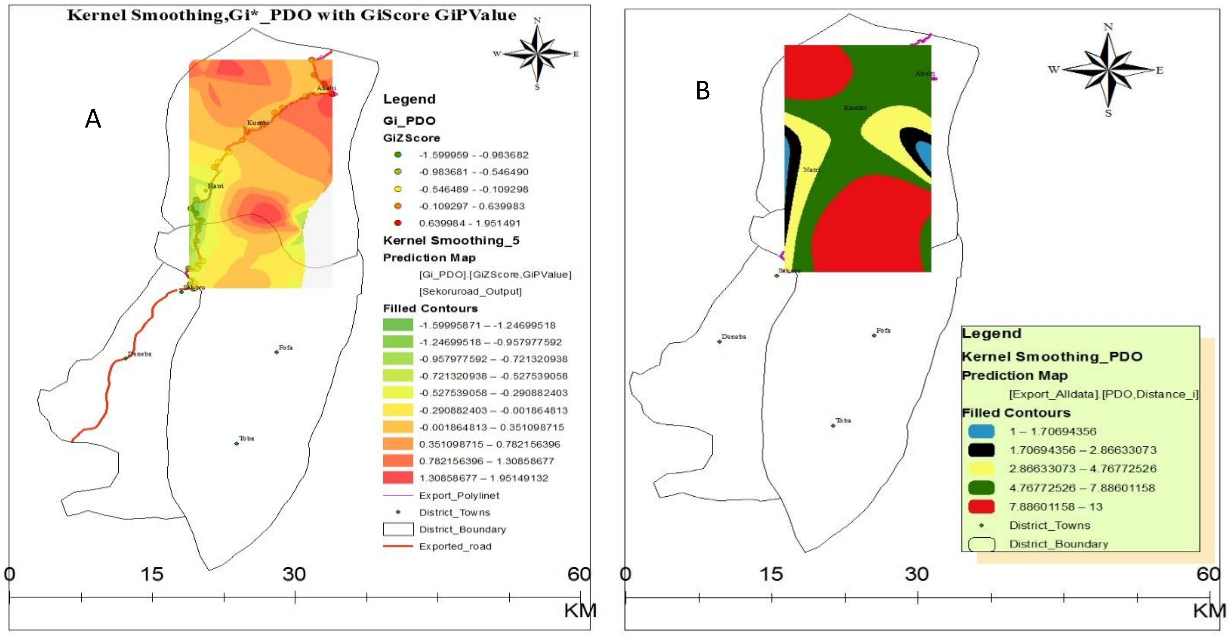


Figure 4.33: Hotspot mapping result of PDO accidents (A) Gi*, (B) using KSDE

4.6. Exploring Results

Figure 4.34, 4.35, 4.36 and 4.37 shows the result of exploring data for fatal, serious, slight and PDO accidents. The histogram is a plot of number of accidents versus frequency of accidents. In addition, the number of counts, minimum and maximum values, a mean value of 2.34, 2.78, 1.53 and 4.75; a standard deviation of 3.83, 1.79, 1.46 and 3.34; skewness of 2.53, 0.3, 0.8 and 0.82; kurtosis value of 9.30, 2.13, 2.95 and 2.79; 1st quartile value of 0, 1.5, 0 and 2; a median of 1, 2.5, 1 and 4; and 3rd quartile value of 3, 4, 2 and 7 are found respectively. The count is the total number of accident points for fatal, serious, slight and PDO accidents with in the research road segment. Minimum and maximum values are the minimum and maximum number of accident occurred at a point. A mean value of 2.34, 2.78, 1.53 and 4.75 are the arithmetic average of values. The measure of the symmetry of distribution of the data for serious injury is found to be 2.53, 0.3, 0.8 and 0.82 respectively. the tendency of the data to produce outliers or kurtosis is found to be 9.30, 2.13, 2.95 and 2.79 which is thick or leptokurtic for fatal accidents and thin or platykurtic for serious, slight and PDO accidents. The 25th percentile (1st quartile) and 75th percentile (3rd quartile) values are found to be 0, 1.5, 0 and 2; 3, 4, 2 and 7 respectively with a median value of 1, 2.5, 1 and 4 respectively.

the QQ plot (Figure 4.34 to 4.37 b) is a plot of standard normal value versus number of accidents to show the closeness of observations from the standard normal line. The data indicated the existence of global trends or local variation. Each vertical sticks in (Figure 4.34 to 4.37 b) showed each location, representing district, and the height of each stick represents the scale measurement of serious accidents. Best fit polynomial lines were also added into (Figure 4.34 to 4.37 b). The green and blue lines were not flat, which represents the existence of directional variation. Initially, the green lines start at a point, then increase over distance when it moves over the x-axis and, similarly, blue lines maximize at the middle. Therefore, it can be inferred that a strong trend in serious accident is seen over all the edges and center part of the district. The possible reason behind this description is high densely curved and small radius of those areas.

The Semivariogram or covariance is the difference between serious injury accident data separated by varying distance. The semivariogram presented in (Figure 4.34 to 4.37 c) shows that adjacent data points are statistically correlated. Since the distance is increasing, the likelihood of the observations becomes smaller. Therefore, directional influence might be considered for the interpolated surface creation. Trend analysis (Figure 4.37 to 4.40 d) is the three dimensional perspective of serious injury accident data. Thus, from the results showed on Figure 4.37 to 4.40 concluded that values of skewness approaches to 0 are the govern analysis (fatal accident value of 2.53, serious accident value of 0.3, slight accidents value of 0.8 and PDO accident value of 0.82); kurtosis value of equal to 3 means proportion of the input data is exact, as results shown in Figure 4.37 to 4.40 (fatal accident value of 9.30, serious accident value of 2.13, slight accidents value of 2.95 and PDO accident value of 2.79).

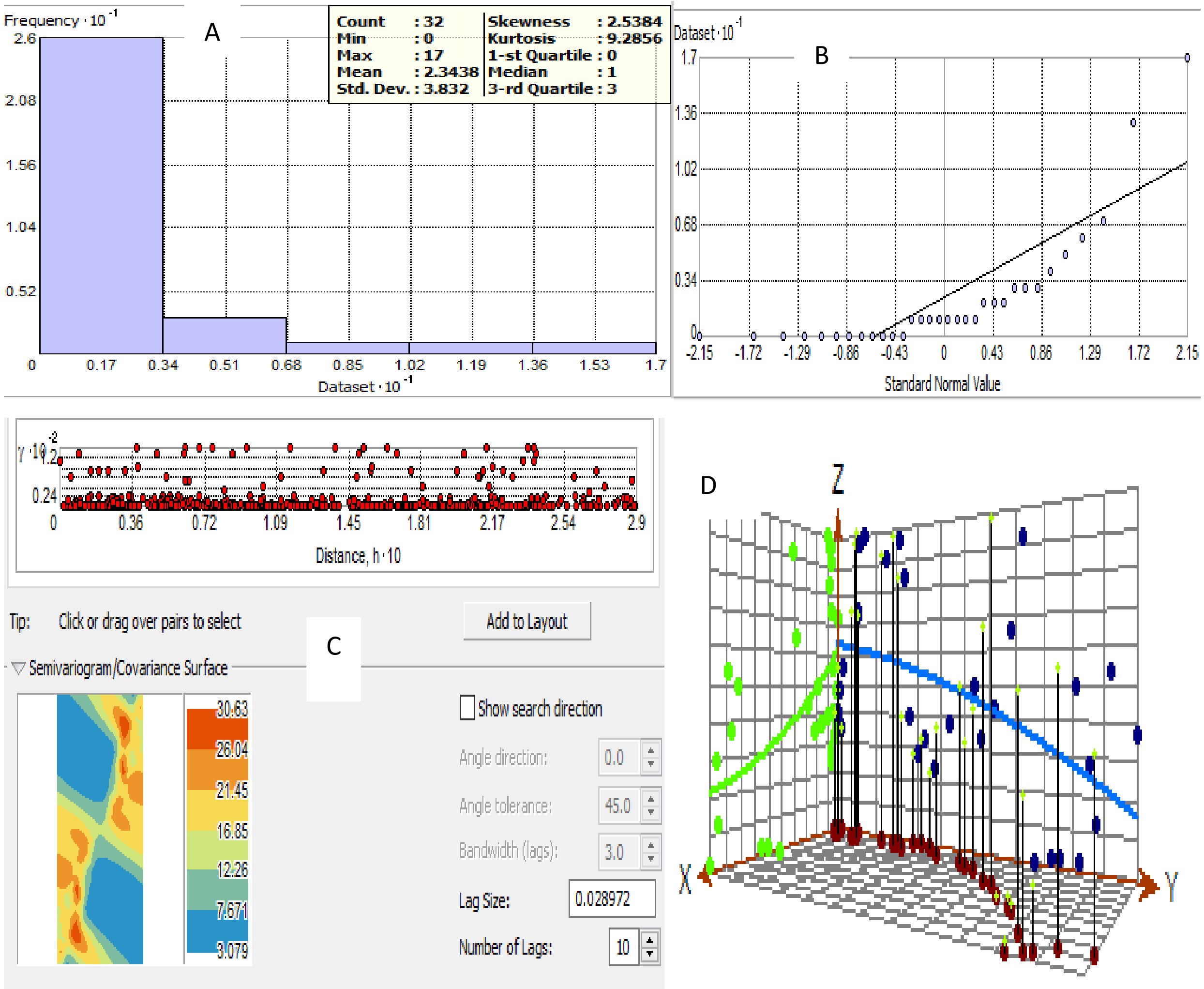
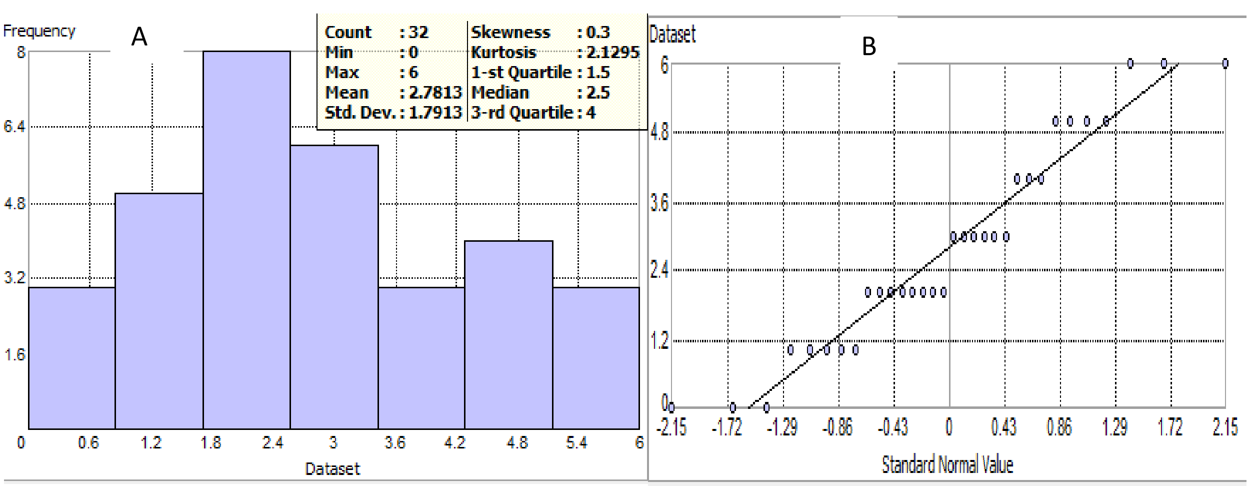


Figure 4.34: Explore results of Fatal accidents



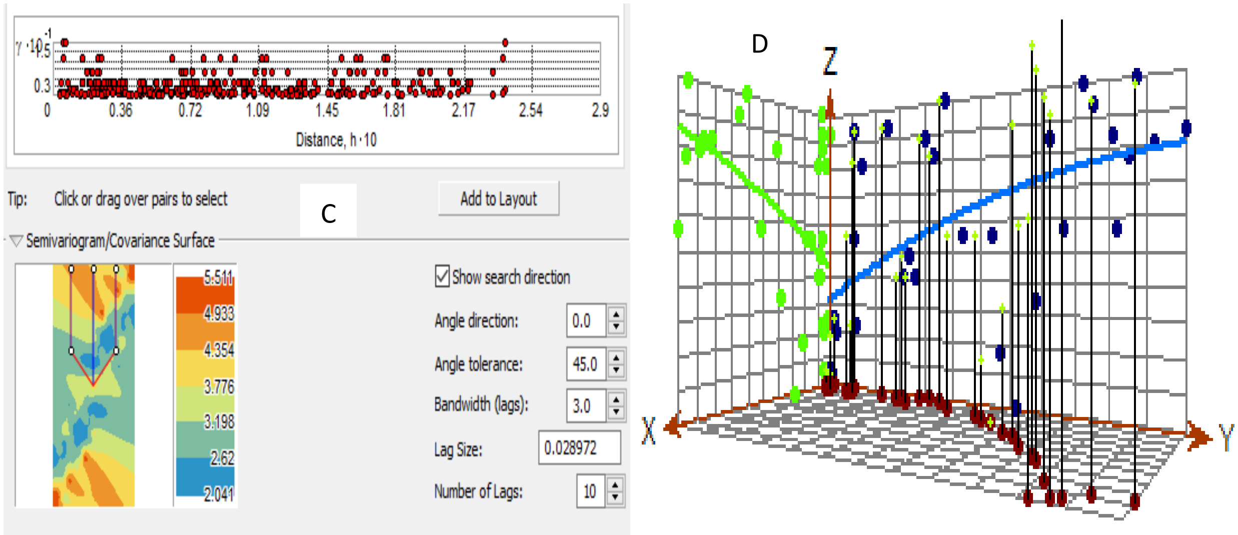
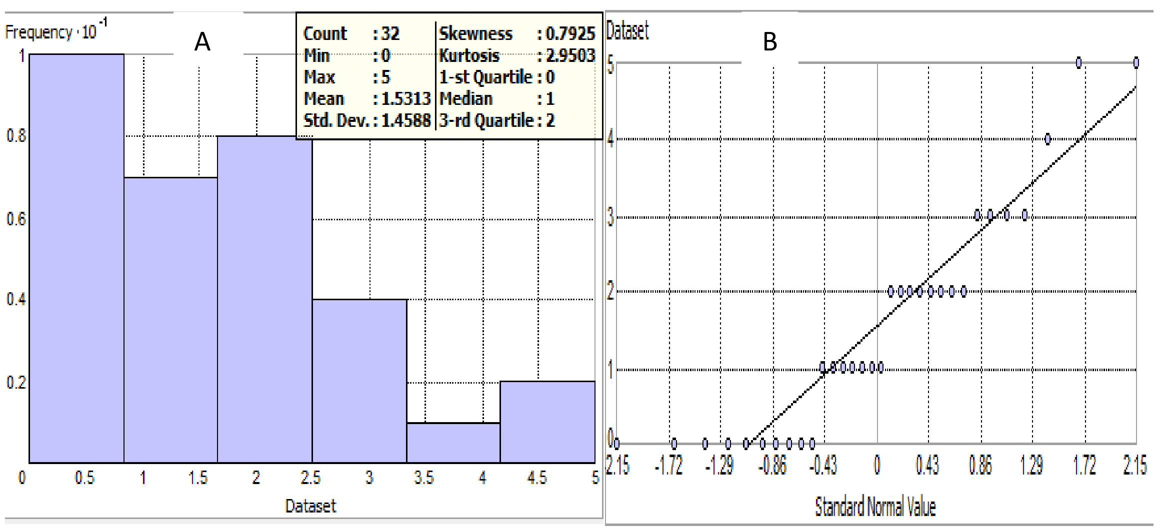


Figure 4.35: Explore results of Serious accidents



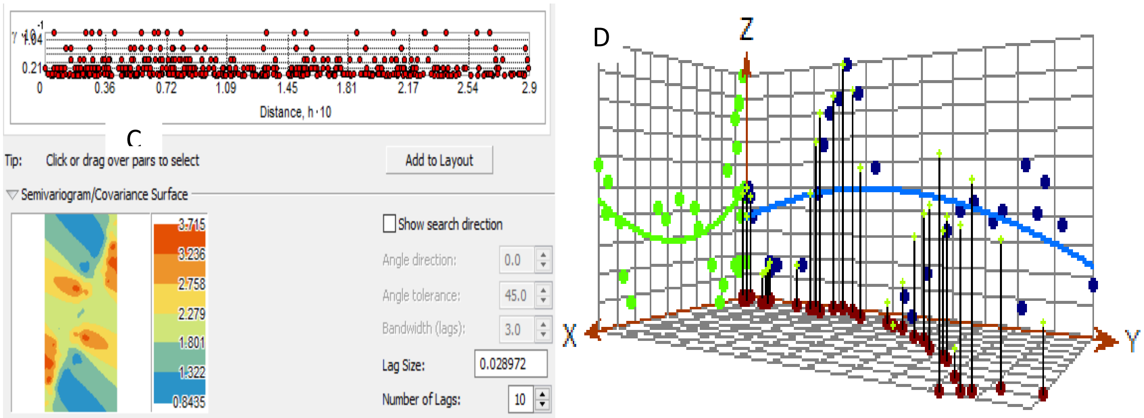


Figure 4.36: Explore results of Slight accidents

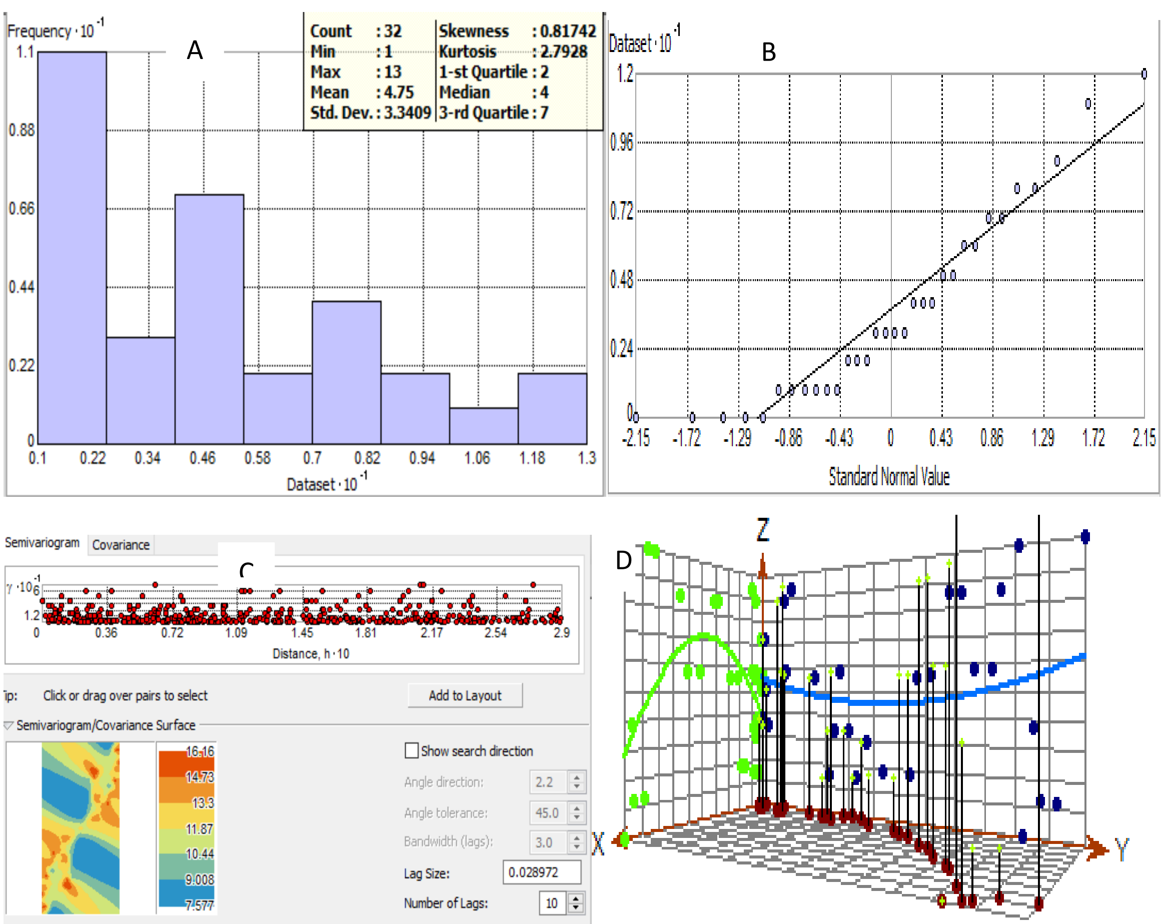


Figure 4.37: Explore results of PDO accidents

4.7. Comparison of Spatial Statistical and Geostatistical Methods

The study began with brief introduction of traffic accident, traffic safety, usage of Geographic Information System and related parameters for the traffic accident. The effect of traditional statistical analysis was also recognized. Generally, it has also described the general similarity and difference between the spatial statistical analysis and the geostatistical analysis with examples and an overview of the statistical information, the articles, maps, graphs and tables that are used in both statistical methods were also identified.

The Spatial and Geostatistical analysis methods involves collection of traffic accident data, accident where happened which is the location, parameters related to traffic accident which also involves dividing data in to fatal accident, serious injury, slight injury and property damage only (PDO). Georeferencing, geocoding and digitizing the input data using Geographic Information System (GIS) tools, data modeling, and exporting data are all components of the statistical approach which are available on the Geographic Information System tools and incorporated in the analysis system.

Finally, six investigations were conducted by using the collected data set to identify and analysis the hot spot area using two different statistical approaches. The first three namely Getis-Ord Gi*, Moran Index and Anselin Local Moran Index are conducted by using Spatial Statistics approach and it has got a better result using the Gi* method rather than Moran (which shows whether is clustering or dispersion, but it does not inform about the location of a cluster) and Anselin local Moran Index (results takes into account by their similarity). The second three experiments namely Inverse Distance Weighting, Empirical Bayesian Kriging, and Kernel Smoothing Density method are carried out by the Geostatistical analyst method and it has a better result using the Kernel Smoothing Density rather than Inverse Distance Weighting (which represents results by assumptions of accidents that are closed to one are more alike than those that are farther apart) and Empirical Bayesian method (which calculates of parameters through a process of sub setting and simulation and it is difficult to understand). after comparing the Gi* and Kernel Smoothing Density method, the Gi* method was the better result.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

The following conclusions are drawn from the results of the study. As observed from the GIS spatial auto correlation analysis results for accident data in the study area, the GiPValue of Shen Debitu curves (around Abelti) and Kumbi are $P < 0.05$ and $P < 0.1$ which is in between 90 and 95% confidence level (Gi_Bin) with GiZScore values >1.96 and >1.65, respectively. The GiPValue of Natri, western Saja, Simini and Birilea stream (around eastern part of Sekoru town also comes P <0.1 which is 90% confidence level (Gi_Bin) with GiZScore values >1.65.

The hot spot areas are located at (37.560, 8.197); (37.570, 8.186); (37.579, 8.173); (37.572, 8.181); (37.578, 8.173); (37.505,8.130); (37.519, 8.150 m); (37.485, 8.114); (37.435, 7.933 m) coordinates in the study segment which have z score values [1 to 2, and >2] and p values < 0.1. Global Moran Index analysis shows that Z-score values of < -1.65 as dispersed or cold spot, Z-score values of greater than 1.65 as clustered or hotspot, and Z-score values in between -1.65 and +1.65 were random values.

The hot spot areas for fatal were at the location of (37.568, 8.189 m), (37.581, 8.176 m), (37.525, 8.155 m), and (37.432,7.921 m) which is around Kumbi and Shen Debitu curves. The hot spot areas for slight accidents were also at location (37.455, 8.058 m), (37.427, 8.029 m), (37.441,8.019 m), and (37.432, 7.933 m) which is from Saja to Natri and Kumbi to Gibe valley; for serious accidents hot spot areas were at location (37.576, 8.173 m), (37.507, 8.138 m), (37.467, 8.101 m), (37.451, 8.055 m), (37.438, 7.958 m), (37.436, 7.937 m), and (37.431, 7.933 m) which is from Natri to Kumbi. PDO observed at location (37.561, 8.198 m), (37.569, 8.188 m), (37.573, 8.182 m), (37.578, 8.174 m), (37.564, 8.137 m), (37.438, 7.978 m), (37.435, 7.933 m), and (37.452, 8.081 m) coordinates which is around Kumbi and from Abelti to Gibe bridge.

Cluster high (HH) values occurred at locations Shendebitu (Abelti cluster), Saja, Natri, eastern part of Sekoru town and Kumbi cluster. The Anselin Local Moran Index was laid at the locations of (37.434,7.933 m north eastern of Sekoru town on the gate of Yem district; 37.57, 8.16 m eastern part of Abelti specifically Shendebitu; 37.518, 8.15 m and 37.434, 7.933 m which north eastern part of kumbi) for serious injury accidents based on their statistical outlier. The location of slight injury was also located at (37.442,8.013 m; 37.430,8.029 m; 37.430, 8.039 m; 37.583, 8.174 m zones covered from south west Natri town to north east kumbi town) all other points clustered low or not significant due to their neighborhood based on Anselin Local Moran Index.

The Kernel Smoothing Density estimation, Inverse Distance Weighting and Empirical Bayesian Kriging method result showed that the high accident zones concentrated in the area of Abelti (shen debitu), kumbi, Natri, Saja and Sekoru towns based on hot and cold spot represented by different colors. The hot and cold spot areas were presented at their geostatistical map by red and blue colors, respectively. As the highways are extending out ward from the Sekoru town to Gibe bridge, the risk level was increased.

The skewness value of fatal accident, serious accident, slight accidents and PDO accident were 2.53, 0.3, 0.8 and 0.82, respectively. The kurtosis value of fatal accident, serious accident, slight accidents and PDO accident were 9.3, 2.31, 2.95, and 2.79, respectively. the values of skewness and kurtosis, show that the value of fatal accident was thick and the others were thin.

Generally, high accident zones were concentrated in the area of Abelti (Shen Debitu), Kumbi, Natri, Saja and Sekoru cluster due to high speed, many number of curves, teenage drivers, night driving, design defects, and improper sight distance.

5.2. Recommendations

Depending on the findings of this study, the following recommendations are given;

- The curves along the Gibe bridge to Sekoru town road segment can be redesigned, reconstructed, limit the speed and develop road infrastructure to minimize road traffic accident on the identified hot spot areas.
- Road design elements like lane width, shoulder, bridges on curves, small radius of horizontal curve especially at escarpment area, right of way, clear sight distance, the shoulder type around rural area, road marking, traffic sign, and sign inventory are improper and insufficient in the road segment. Thus, redesign and reconstruction are the primary recommended points.
- It would be better to use GIS to determine the hot spots, accident rates, curve widening, passing sight distance and stopping sight distance of the road segment than the other software to investigate accidents rate on the study area.

- Safety barriers should be provided at adverse road sections to increase the psychological confidence of both drivers and passengers.
- Night driving has to be restricted.
- Traffic accident data need to be stored using GIS technology.
- Further investigation is required using GIS tools focus on geospatial and clustering analysis of road network components and root causes of traffic accidents, including vehicle types and weather conditions present when accidents occurred. Factors affecting the spatial distribution on those locations might also be investigated.

REFERENCES

1. American Association of State Highway and Transportation Officials (AASHTO) (2004). A Policy on Geometric Design of Highways and Streets. Washington, DC, USA.
2. A. Persson. (2008). Road Traffic Accidents I n Ethiopia: magnitude, causes and possible interventions. Advances in Transportation Studies an international Journal Section A 15.
3. AfDB, OECD, UNDP, & UNECA. (2012). Economic Outlook: Ethiopia. Retrieved fromhttp://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Ethiopia%20Ful l%20PDF%20Country%20Note.pdf
4. Akloweg, Y., Hayshi, Y., & Kato, H. (2011). The Effect of Used Cars on African Road Traffic Accidents: A Case Study of Addis Ababa, Ethiopia. International Journal of Urban Sciences. 15(1): 61-69.
5. American Association of State Highway and Transportation Officials (2011). A Policy on Geometric Design of Highways and Streets. 4th edition, ISBN: 1-56051-156-7.
6. Anderson TK (2009) Kernel density estimation and K-means clustering to profile road accident hotspots. Accid Anal Prev. 41(3):359–364
7. Anitha SD Selvasofia, and Prince G Arul raj (2016). Accident and Traffic Analysis Using GIS, SNS College of Technology, Coimbatore, Tamil Nadu, India, Biomedical Research; Special Issue: S103-S106.
8. Anselin, L., & Getis, A. (1992). Spatial statistical analysis and geographic information systems. Registers of Regional Science. 26: 19-33.
9. Anselin, L. (1988). Spatial econometrics, methods, and models. Dordrecht: Kluwer.
10. Arndt O.K, Cox R.L, Lennie SC, Whitehead. T, (2011). Provision of sight distance around concrete barriers and structures on freeways and interchanges. Transport Research Rec 2262:22–30.
11. Bello, T., 2005, A Stratified Traffic Accident Analysis Case Study: City of Richardson, Masters in Geographic Information Sciences, University of Texas at Dallas, (http://charlotte.utdallas.edu/mgis/prj masters/index.htm).
12. Berhanu, G, Effects of Road and Traffic Factors on Road Safety in Ethiopia, Trondheim, Norway, 2000.
13. Brunsdon, C. Estimating probability surfaces for geographical point data: An adaptive kernel algorithm. Compute. Geosci. 1995, 21, 877–894.
14. Cafiso S., Lamberti R., Marchionna A., Perco P. (2010). "Country Report of Italy", Proceedings of 4th International Symposium on Highway Geometric Design, Valencia, Spain.
15. Castro M, Anta JA, Iglesias L, Sánchez JA, (2014). GIS-based system for sight distance analysis of highways. Journal of Computing in Civil Engineering. 28(3).
16. Castro M, Iglesias L, Sánchez JA, Ambrosial L, (2011). Sight distance analysis of highways using GIS tools, Transportation Research Part C: Emerging Technologies. 19(6): 997-1005.
17. Chainey, S.; Thompson, L.; Uhlig, S. The utility of hotspot mappering for predicting spatial patterns of crime. Secure. J. 2008, 21, 4–28
18. Cheng W, Washington S (2008) New criteria for evaluating methods of identifying hot spots. Transport Res Rec 2083(1):76–85
19. Christo J. baste & JosterA.Makrajep, (1999). The effect of rural road geometry on safety in Southern Africa, university of Stellenbosch, Malawi.

20. Council, F.M., (1998), Safety Benefits of Spiral Transitions on Horizontal Curves on two-lane.
21. Deepjyoti Kalita, Saurabah Sarma, Tina Gogoi, and Kildeep Kalita, (November 2015). Design of Overtaking Sight distance in highways using probabilistic approach, Don Bosco University, Guwahti781017, Assam, India. Vol 1(2).
22. Demirel, A., Akgungor, A.P. (2002). The Importance of Reports in The Accident Analyses, Problems in Application and Recommendations for Solution, Gazzi University, International Traffic and Road Safety Congress, Ankara, Turkey (in Turkish). (www.trafik.gov.tr).
23. Erdogan, E., I. Yilmaz, T. Baybura and M. Gullu. Geographical Information Systems Aided Traffic Accident Analysis System Case Study: City of Afyonkarahisar. Accident Analysis Prevention, Vol. 40, No. 1, 2008, pp. 174-181.
24. Ericson, M., & Kim, P. (2011). How Road Traffic Injuries Affect Household Welfare in Cambodia Using the Millennium Development Goals Benchmarks. Asian Studies Review. 35(2): 209-234.
25. Ethiopian Road Authority manual (ERA). (2001). A Policy on Geometric Design of Highways and Streets. A.A, Ethiopia.
26. F. Maltinti, (2012). Works on horizontal alignment of a rural road to increase sight distance. A case study: thesis. No. 292, SIIV- 5th International Congress- Sustainability of Road Infrastructure Cagliari, Italy. 53: 580-589
27. FHWA (Federal Highways Administration), 2006. Interactive Highway Safety Design Model (IHSDM), website: http://www.tfhrc.gov/safety/ihsdm/ihsdm.htm
28. Flahaut, B.; Mouchart, M.; San Martin, E.; Thomas, I. The local spatial autocorrelation and the kernel method for identifying black zones—A comparative approach. Accid. Anal. Prev. 2003, 35, 991–1004.
29. Garber N. J., Hoel L. A., 2001. Traffic & Highway Engineering, Third Edition.
30. Gerber, M.S. Predicting crime using twitter and kernel density estimation. Deci’s. Support Syst. 2014, 61, 115–125.
31. Getis, A.; Ord, J.K. The analysis of spatial association by use of distance statistics. Georg. Anal.1992, 24, 189–206.
32. Getis, A.; Ord, J.K. Local spatial statistics: An overview. In Spatial Analysis: Modelling in GIS Environment; Longley, P., Batty, M., Eds.; John Wiley & Sons Inc.: New York, NY, USA, 1996; pp. 261–278.
33. Getu S. Tulu, Simon Washington, Mark J. King (2011). Characteristics of police-reported road Traffic Crashes in Ethiopia over a six-year period.
34. Geurts, K. and G. Wets. Black Spot Analysis Methods: Literature Review. Report, Flemish Research Center for Traffic Safety, Belgium, 2003.
35. Glennon, J. (1985). “Effect of Alignment on Highway Safety”, “Relationship between Safety and Key Highway Features. State of the Art Report 6, Transportation Research Board, Washington, D.C., pp. 48-63.
36. Gundogdu, I. B. Applying Linear Analysis Methods to GIS-supported Procedures for Preventing Traffic Accidents: Case Study of Konya. Safety Science, Vol. 48, No. 6, 2010, pp. 763-769.
37. Hauer E (1992) Empirical Bayes approach to the estimation of ‘‘unsafety’’: the multivariate regression method. Accid Anal Prev. 24(5):457–477

38. Hauer, E, Draft Safety of Horizontal Curves, March 24, 2000, 2001
39. Hauer, E., J. Kononov, B. Allery and M. S. Griffith. Screening the Road Network for Sites with Promise. In Transportation Research Record: Journal of the Transportation Research Board, No. 1784, Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 27-32
40. Haworth and Vulcan [29] Haworth N., Vulcan P., 1991. Testing of Commercially Available Fatigue Monitors. Report No. 15. Melbourne, Australia: Monash University Accident Research Centre.
41. Highway Safety Manual, First Edition, American Association of State Highway and Transportation Officials 2010.
42. Hummer J., Zegeer C.V, Reinfurt D., Herf L., Hunter W., 1987. Safety Effects of Cross-Section Design for Two-Lane Roads. Report No. FHWA-RD-87/008. (US) Federal Highway Administration (FHWA), TRB National Research Council, Washington D.C., Oct. 1987.
43. Iyinam, A.F., Iyinam, S. & Ergun, M. (2000), Analysis of Relationship between Highway Safety and Road Geometric Design Elements, Technical University of Istanbul, Turkey.
44. Jha MK, Karri GAK, Kuhn W. (2011). New three-dimensional highway design methodology for sight distance measurement. Transportation Research Record. 2262(1):74–82.
45. Jones, A.P., Langford, I.H., Bentham, G., 1996, The Application of K-Function Analysis to the Geographical Distribution of Road Traffic Accident Outcomes in Norfolk, England, Social Science and Medicine, Vol. 42, Issue: 6, pp. 879–885.
46. Karl Kim, and Ned Levine (1996). Using GIS to Improve Safety; Compute., Environ. and Urban Systems. Vol. 20, No. 4/5: pp. 289-302, PII: S0198-9715(96)00022-1.
47. Keskin F, Yenilmez F, C ¸olak M, Yavuzer I, Du ¨zgu ¨n HS (2011) Analysis of traffic incidents in METU campus. Procedia Social Behave Science 19:61–70
48. Khan, G., Qin, X., Noyce, D.A. (2006). Spatial Analysis of Weather Crash Patterns in Wisconsin, 85th Annual Meeting of the Transportation Research Board, Washington, USA, (www.topslab.wisc.edu/publications).
49. L. Altamira, J.E., Marcet, A. B., Graffigna, A. M. Gómez (2010). "Assessing available sight distance: an indirect tool to evaluate geometric design consistency" Proceedings of 4th International Symposium on Highway Geometric Design, Valencia, Spain.
50. Lalita Thakai, Tae J. Kwon, LIPING Fu. (2015). Identification of crash hotspots using kernel density estimation and kriging methods: a comparison. Vol. 23(2):93-106, DOI 10.1007/s40534-015-0068-0
51. Levine, N., Kim, K.E., Lawrence, H.N., 1995. Spatial analysis of Honolulu motor vehicle crashes: 1. spatial patterns. Accident Anal. and Prev. 27 (5), 663–674.
52. Lieberman E. B. (1982). Model for calculating safe passing sight distances on two-lane rural roads. Transport Res. Rec. 869: 70-76.
53. Lim Yu LIANG, Law Teik HUA (2005). Traffic Accident Application using Geographic Information System. Journal of the Eastern Asia Society for Transportation Studies. Vol. 6: pp. 3574 – 3589.
54. Machado, C.A.S.; Giannotti, M.A.; Shinohara, E.J.; Nishisaki, H.; Quintanilha, J.A. Characterization of the sites of traffic accidents involving vulnerable road users (VRU) in São Paulo city. In Proceedings of the Transport Research Arena (TRA 2014)—5th

Conference Transport Solutions: From Research to Deployment-Innovative Mobility, Mobilise Innovation, Paris, France, 14–17 April 2014; p.10.
55. Masayuki HIRASAWA, Motoki Asano. Development of traffic accident Analysis System Using GIS. Civil Engineering Research Institute of Hokkaido 1-3 hiragishi, Toyohira-ku, Sapporo, Japan.
56. McLean, J. (1996). Review of accidents and urban arterial cross-section treatments, research report ARR 309, ARRB Transport Research, Vermont South, Victoria.
57. Mekonnen, H. (2007). Design Consistency of Horizontal Alignments on Two - Lane Trunk. MSc Degree Thesis, Addis Ababa, Ethiopia.
58. Meyer, M., & Sarasua, W. (1992). Geographic information system-based transportation program management system for county transportation agency. Transportation Research Record, 1364.
59. Milton J., Mannering F., 1998. The Relationship among Highway Geometrics, Traffic-Related Elements and Motor-Vehicle Accident Frequencies. Transportation 25, pp. 395-413
60. Miranda-Moreno, L. F. Statistical Models and Methods for Identifying Hazardous Locations for Safety Improvements. PhD thesis, University of Waterloo, 2006
61. Mitchell, A. Spatial Measurements & Statistics. In The ESRI Guide to GIS Analysis; ESRI Press: Redlands, CA, USA, 2005; Volume 2.
62. Nie, K.; Wang, Z.S.; Du, Q.Y.; Ren, F.; Tian, Q. A network-constrained integrated method for detecting spatial cluster and risk location of traffic crash: A case study from Wuhan, China. Sustainability 2015, 7, 2662–2677.
63. O'Day, J. (1993). Accident data quality. NCHRP Synthesis 192. Washington, DC: National Academy Press.
64. Oxley J., Corben B., Koppel S., Fildes B., Jacques N., Symmons M., Johnston I., 2004. Cost-Effective Infrastructure Measures on Rural Roads. Monash University Accident Research Centre, Report No. 217.
65. Peled, A., & Hakkert, A. (1993). A PC oriented GIS application for road safety improvements. Traffic Engineering and Control. pp. 355-361.
66. Peeters, A.; Zude, M.; Käthner, J.; Ünlü, M.; Kanber, R.; Hetzroni, A. Getis-Ord's hot- and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. Compute. Electron. Agric. 2015, 111, 140–150
67. Pei-Fen Kuo, Dominique Lord, Troy Duane Walden, using geographical information systems to effectively organize police patrol routes by grouping hot spots of crash and crime data Submitted to the Third International Conference on Road Safety and Simulation, September 14-16, 2011, Indianapolis, USA
68. R.A. G anorkar, P.I. Rode, and Ashtashil V. Bhambulkar (March -April 2013). Application of GIS in Transportation Engineering. International Journal of Engineering Research and Applications (IJERA). ISSN: 2248-9622 www.ijera.com. Vol. 3, Issue 2: pp.540-542.
69. Sabel, C.E., Kingham, S., Nicholson, A., Barite, P., 2005. Road traffic accident simulation modelling-a Kernel estimation approach. In: Presented at SIRC 2005 (November), The 17th Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand.
70. Saccomanno FF, Grossi R, Greco D, Mahmoud A (2001) Identifying black spots along highway SS107 in southern Italy using two models. J Transp Eng. 127(6):515–522

71. Santos, G.; Behrendt, H.; Maconi, L.; Shirvani, T.; Teytelboym, A. Part I: Externalities and economic policies in road transport. Res. Transp. Econ. 2010, 28, 2–45.
72. Sarbaz Othman & Robert Thomson (2000), influence of road characteristics on traffic safety, Sweden.
73. Saravade S. B., 2005. A Study of Fatal Rollover Crashes in the State of Florida - a Study of Fatal Traffic Crashes in Florida from 1998-2000. Master Thesis, Department of Civil Engineering, Florida State University.
74. Silverman, B.W. Density Estimation for Statistics and Data Analysis—Monographs on Statistics and Applied Probability, 1st ed.; Chapman and Hall: London, UK, 1986; p. 200.
75. Silyanov VV. (1973). Comparison of the pattern of accident rates on roads of different countries. Traffic Engineering and Control. 14(9): 432–435.
76. Sparks WJ. (1968). The influence of highway characteristics on accident rates. Public Works. 99(3): 101–103.
77. Steenburgen T, Dufay's 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):169–181.
78. Terje, A, Road Safety in Africa Appraisal of Road Safety, (1998).
79. Thomas, I., 1996. Spatial data aggregation: exploratory analysis of road accidents. Accident Anal. and Prev. 28 (2), 251–264.
80. Trinca, G. W., Johnson, I. R., Campbell, B. J., I-Iaight, F. A., Knight, P. R., Mackay, G. M., McLean, A. J., & Petrucelli, E. (1988). Reducing traffic injury-A global challenge. Melbourne: Royal Australian College of Surgeons.
81. Utoyo B., Agus P. Saido, and Syafi. Traffic Accident Blackspot Identification and Ambulance Fastest Route Mobilization Process for the city of Surakarta. Journal Transportation Vol. 12 No. 3 December 2012: 237-248
82. Wand, M.P.; Jones, M.C. Kernel Smoothing—Monographs on Statistics and Applied Probability, 1st ed.; Chapman and Hall: London, UK, 1995; p. 212.
83. Wang, Wuhong, Wang and Sun, Fengchun. (2001). Road Traffic System Theory and Method of Driving Behaviour: Beijing, China, Science Press.
84. Wang, Y.; Potoglou, D.; Orford, S.; Gong, Y. Bus stop, property price and land value tax: A multilevel hedonic analysis with quantile calibration. Land Use Policy 2015, 42, 381–391.
85. WHO (2004). Global Status on Road Safety. Geneva.
86. Wilson, C., & Burtch, T. (1992). Traffic accidents and highway safety. In J. Pline (Ed.), Traffic engineering handbook, 4th Edition. ITE, Prentice Hall, Englewood Cliffs, pp. 94-116.
87. Zegeer, C.V., Twomey. J. M., Heckman, M.L., and Hyward. J.C. (1992). Safety effectiveness of Highway Design Features, Vol 2, Alignment, FHWA, RD-91-045.Federal Highway Administration, Washington, DC.
88. Zimmermann M, Roos R. (2005). Increased safety resulting from quantitative evaluation of sight distances and visibility conditions of two-lane rural roads. 3rd International Symposium on Highway Geometric Design, In: Proceedings of the 3rd International Symposium on Highway Geometric Design, Transportation Research Board; Chicago, USA. pp. 1–14.

APPENDIX

APENDIX A: Details of procedure of importing data using ArcGIS10.4.1

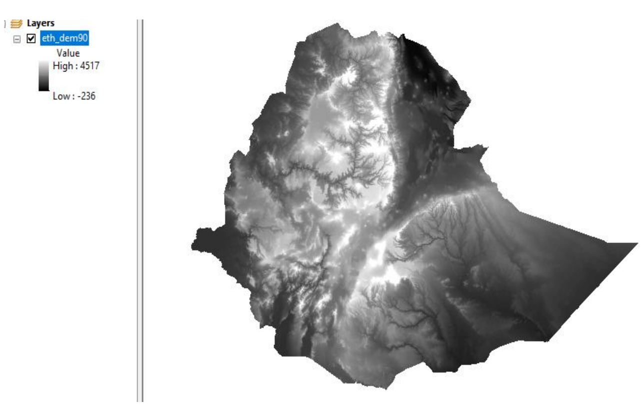
Select on the sign on the ArcGIS10.4.1 menu bar then browse folder location add eth_dem90 then click on add. In order to adjust the layer properties there are different ways to adjust some of these are method click on the menu bar then click on view then data frame properties then click on coordinate system under this there are two opportunities the first one is Geographic Coordinate System(GCS) and the other is Projected Coordinate System(PCS). Here in this research the PCS is applied under the PCS click on Universal Transvers Mercator(UTM) then click on World Geodetic System 1984(WGS1984) then click on northern hemisphere then click on zone 37 finally click ok.

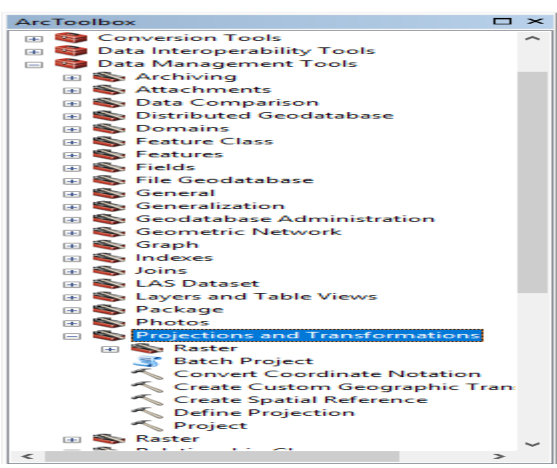
Clipping Area of Interest

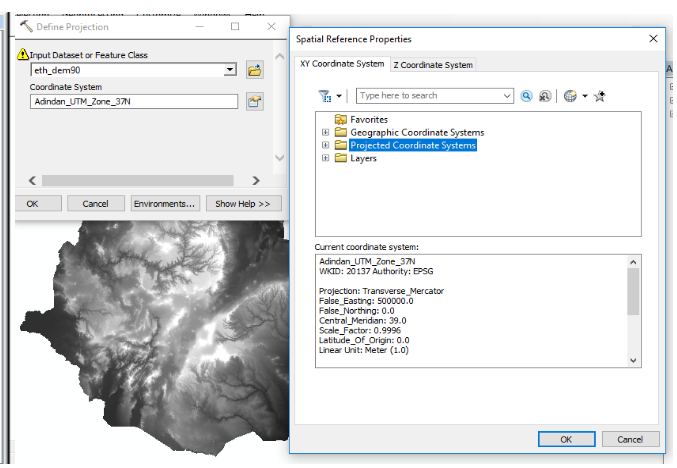
The first thing going to do is clip the part of area of interest. Using the input files such as DEM 90, Ethio shape, Ethio boundary, Ethio wereda, Ethio towns, Ethio roads shape files from Ethio GIS data files and AutoCAD horizontal alignments from Ethiopian Road Authority data files. Then using these files create a folder that is easily available in the desk top or any other local disks and connect them with the ArcGIS using the folder connection in Arc Catalog tool . After adding the required inputs in to ArcGIS using the add data sign then from the menu bar click on selection which presents on the menu bar from the drop down menu click on Selection by attributes then on the displayed window fill the respective information's for example in my case on the layer drop down choose Ethio wereda →method:(create a new selection) → double click on "W_NAME" →double click on = sign→click on Get Unique Values from the displayed window click on Sekoru→ click Or→ double click on "W_NAME" → double click on = sign → click on Get Unique Values from the displayed window click on Yem→Verify→ok. After selection the displayed area of interest is as shown below in color of tourmaline green, then to clip the area of interest follow these steps

Click search→ type clip→ clip (Data Management) (tool) then on the displayed window fill the respective information's. before that right click on Ethio wereda in the table of contents→ selection→create layer from selected feature→now there is one input on the layer named by Ethio wereda selection.

Now clip window displays fill the respective data is next step: in the input raster fill the raster that we use (DEM_90); in the output extent (optionally) use the selected layer; on the rectangle fill maximum and minimum latitudes and longitudes then check the box use input feature for clipping geometry; choose output raster dataset the location where we save the result then finally click ok after running the process displays the clipped area of interest. To clip road of area of interest from menu bar click on Geoprocessing →clip→fill the respective information as stated below. On the input feature use Ethio road shape file; on the clip feature we use the Sekoru selection; output feature class is the location of output and on the XY tolerance we choose meter then finally run. After adding and clipping all data the next step we can export all the data in order to save as shape files to do this we use the following steps on the layer under the table of contents right click on each layer then select data then click on export data in this case we can modify only the output feature class that is the location of the result or shape file placed.







Universal transverse Mercator's coordinate system (UTM)

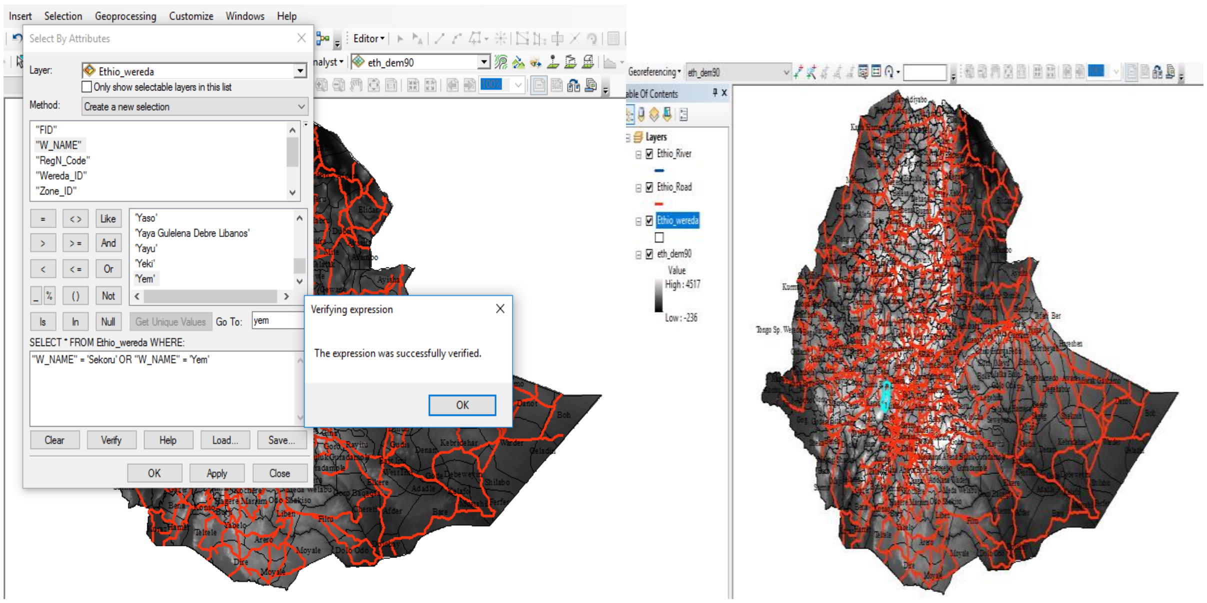
The UTM system divides the earth into 60 zones each 60 of longitude wide. These zones define the reference point for UTM grid coordinates within the zone. UTM zones extend from latitude of 80° S to 84° N, are numbered 1 through 60, starting at the International Date Line, longitude 180°, and proceeding east. Zone 1 extends from 180° W to 174° W and is centered on 177° W ()

Geographic Coordinate System

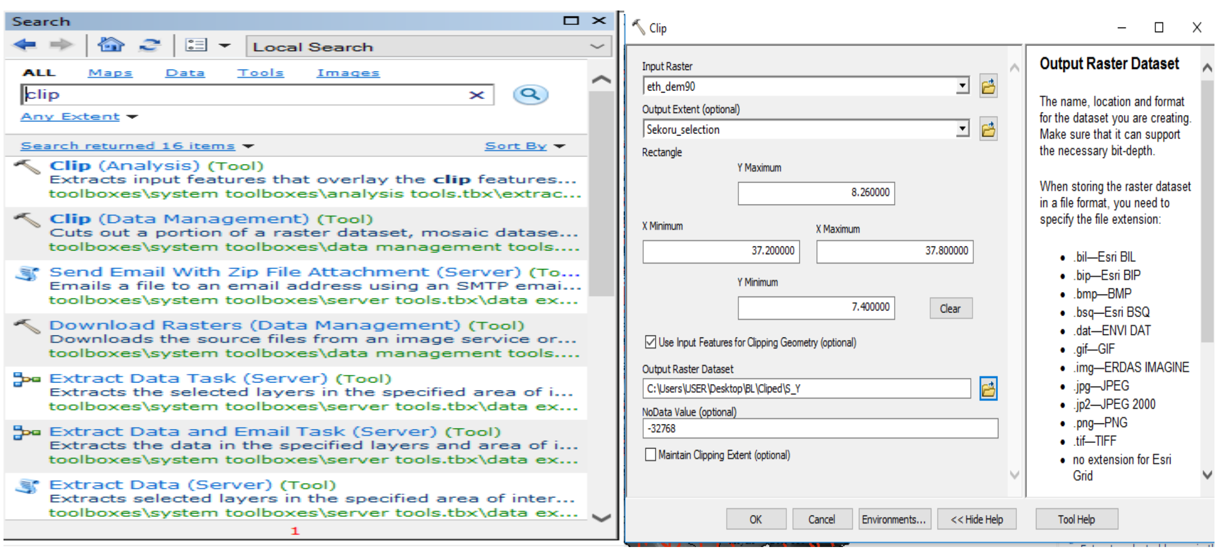
Geographic coordinate systems consist of latitude, which varies from north to south, and longitude, which varies from east to west by convention, the equator is taken as zero degrees' latitude, and latitudes increase in absolute value to the north and south. Latitudes are thus designated by their magnitude and direction, for example 35°N or 72°S. An international meeting in 1884 established a longitudinal origin intersecting the Royal Greenwich Observatory in England. Known as the prime or Greenwich meridian, East or west longitudes are specified as angles of rotation away from the Prime Meridian. ()

$$f = \frac{(a - b)}{a}$$

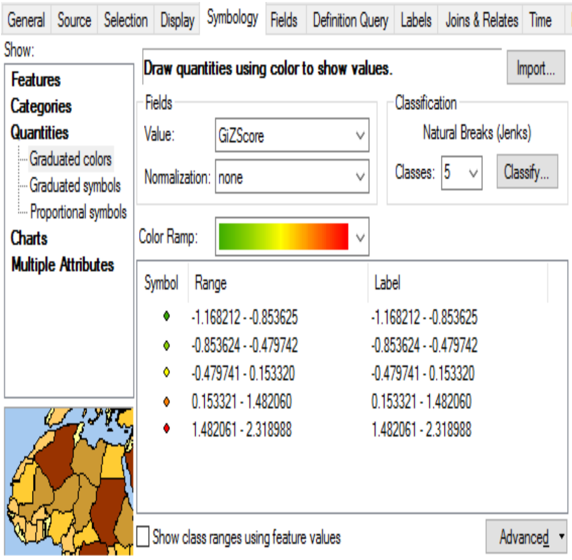
S/No	Ellipsoid	Semi-major axis	1/flattening
1	Airy 1830,	6377563.396	299.3249646
2	Modified Airy	6377340.189	299.3249646
3	Australian National	6378160	298.25
4	Bessel 1841 (Namibia)	6377483.865	299.1528128
5	Bessel 1841	6377397.155	299.1528128
6	Clarke 1866	6378206.4	294.9786982
7	Clarke 1880,	6378249.145	293.465
8	Everest (India 1830)"	6377276.345	300.8017

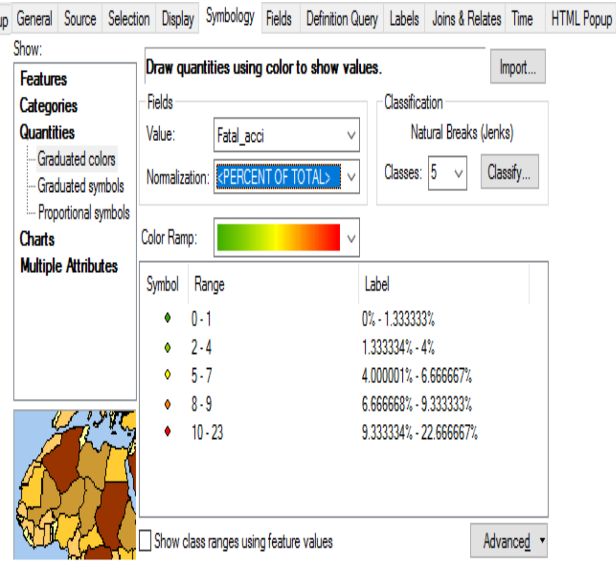


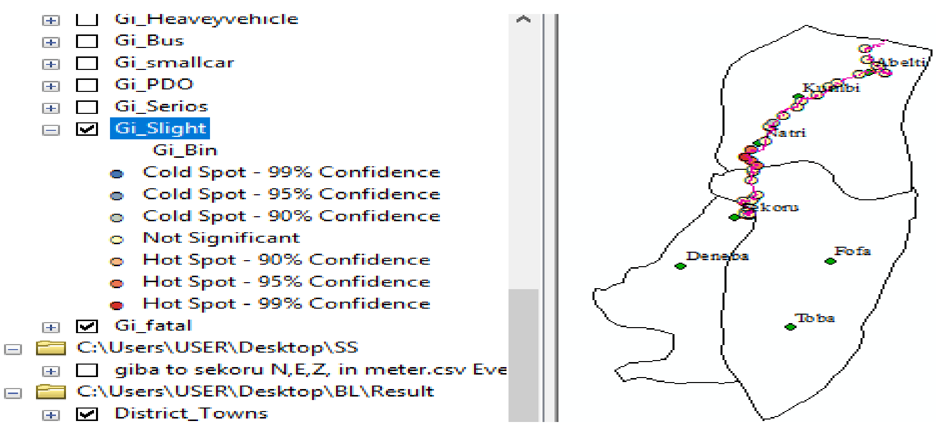
Inorder to select the clip process there are two methods the first one is clunk the search tool and transcribe clip on the given window then choosing the clip(Data Management)(Tool) or using the second method by picking ArcToolBox

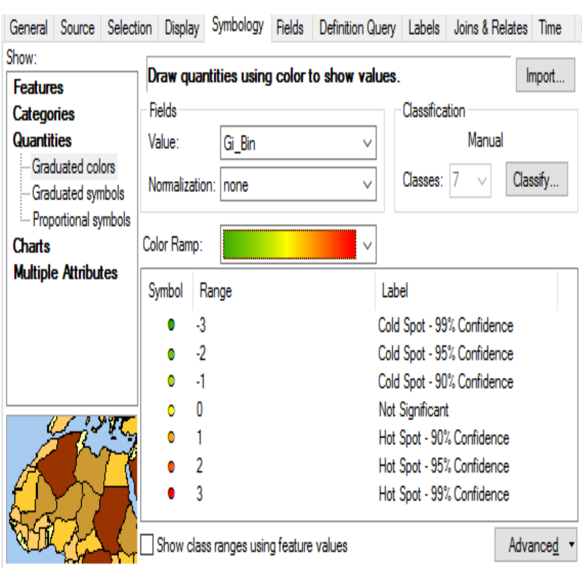


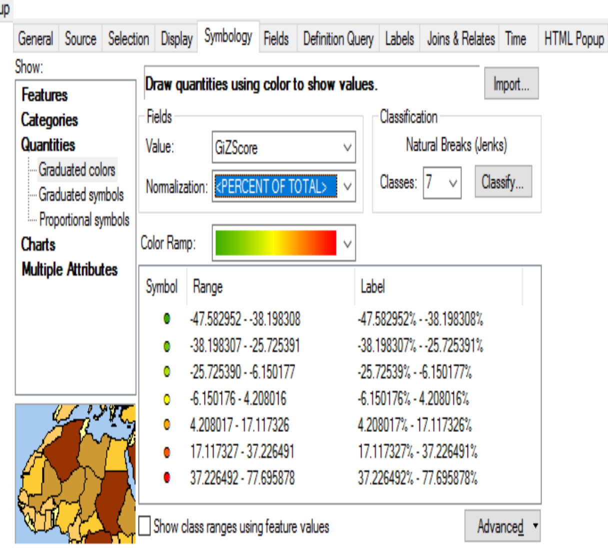
APPENDIX A1: Symbological Representation of Thematic map layers Gi* samples



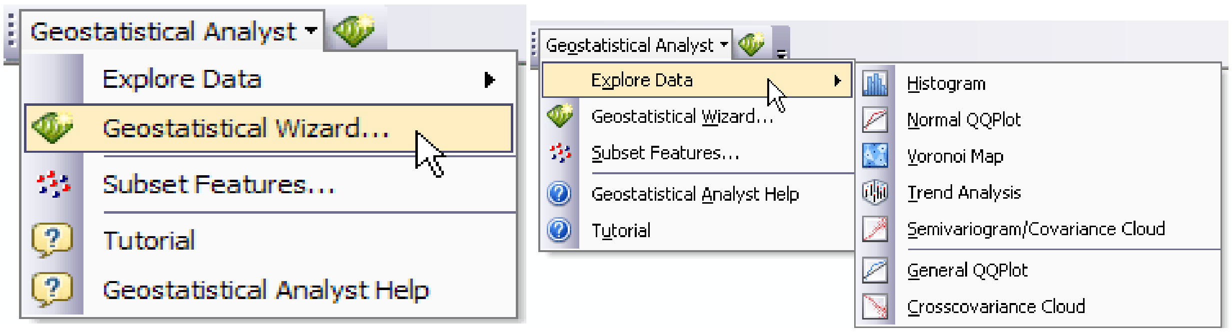




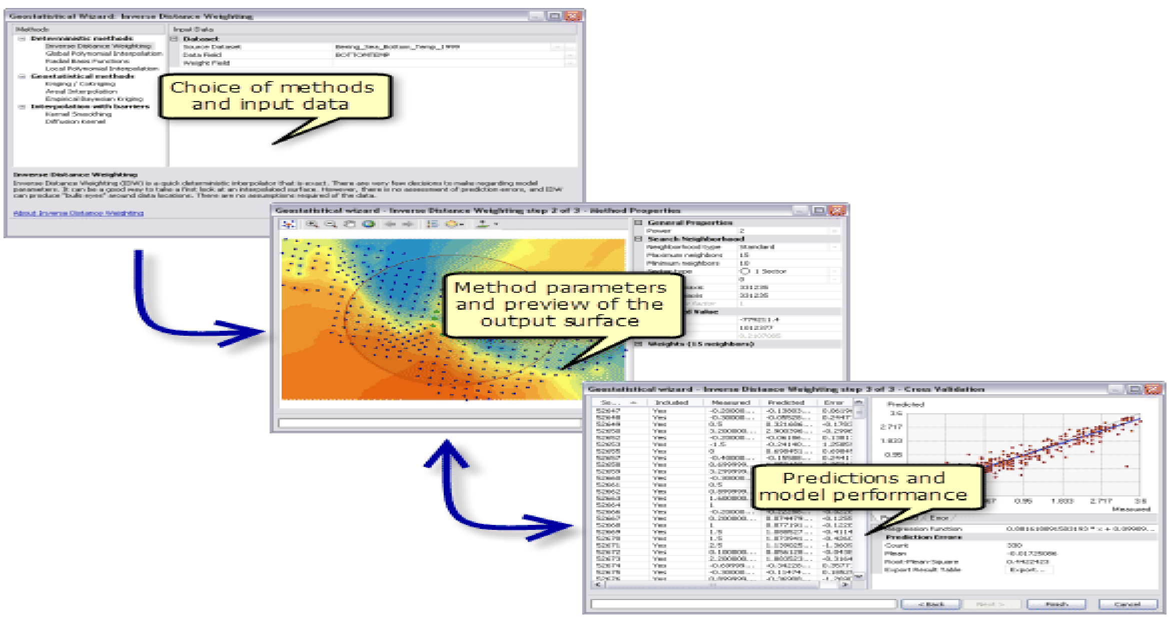




APPENDIX A2: Location of Explore data and Geostatistical wizard workflow

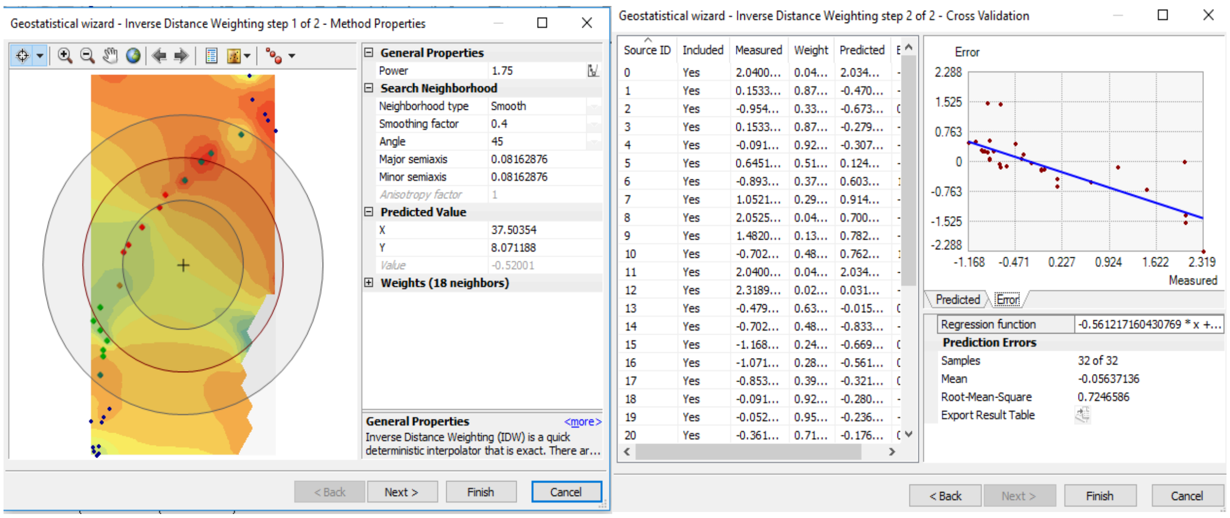


Geostatistical Analyst (ArcGIS10.4.1)

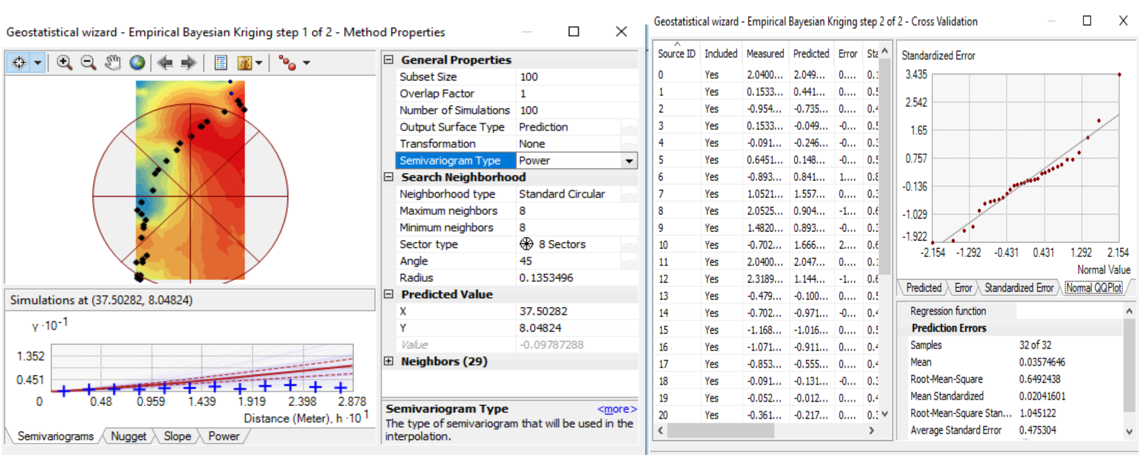


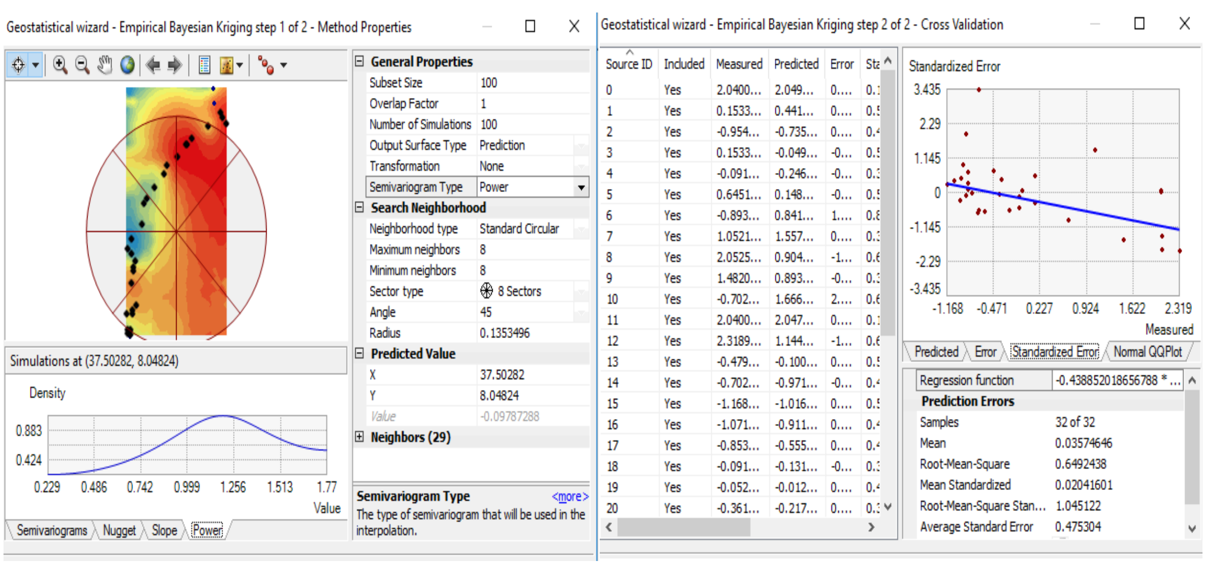
Geostatistical wizard work flow (ArcGIS 10.4.1)

APPENDIX A3: Methodology for analysis of traffic accident using Geostatistical method

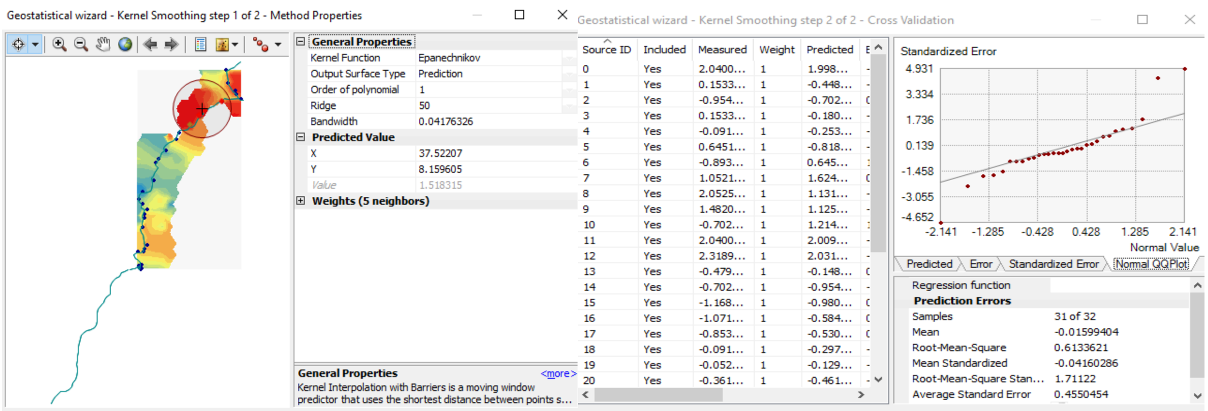


Fatal accidents represented by IDW





Serious accidents represented by EBK



Fatal accidents represented by kernel smoothing density

APPENDIX B: Exploratory spatial data analysis

Histogram: examine the distribution and summary statistics of dataset and provides a univarite(one variable)ndescription of a data.this tool dialog box displays the frequency distribution for the dataset of interest and calculates summary statistics.

The mean is the arithmetic average of the data. The mean provides a measure of the center of the distribution.

The median value corresponds to a cumulative proportion of 0.5. If the data was arranged in increasing order, 50 percent of the values would lie below the median, and 50 percent of the values would lie above the median. The median provides another measure of the center of the distribution.

The first and third quartiles correspond to the cumulative proportion of 0.25 and 0.75, respectively. If the data was arranged in increasing order, 25 percent of the values would lie below the first quartile, and 25 percent of the values would lie above the third quartile. The first and third quartiles are special cases of quantiles. The quantiles are calculated as follows:

quantile = (i - 0.5) / N

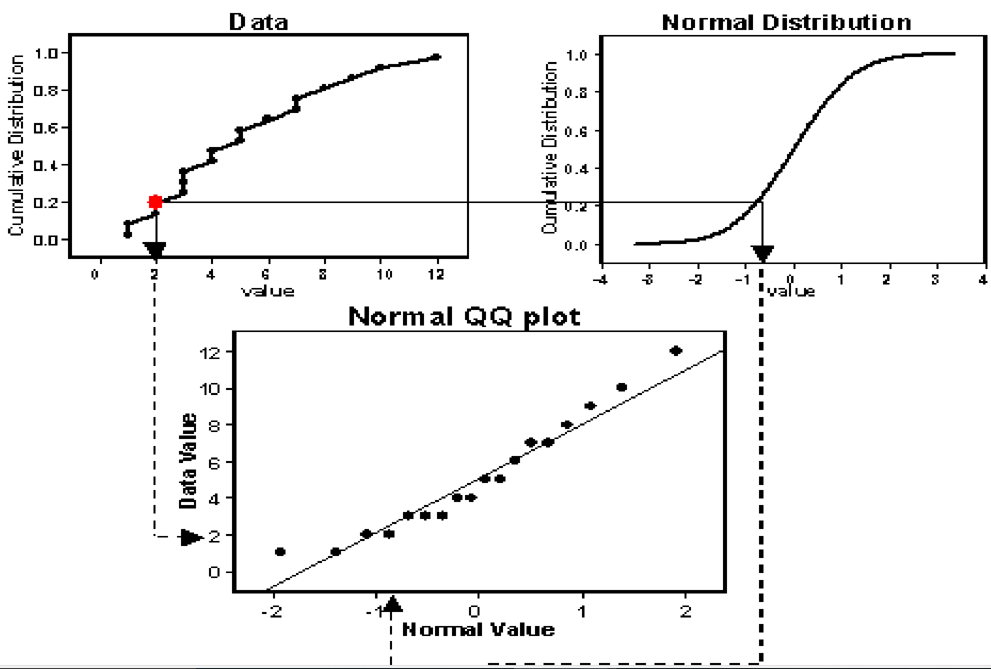
where i is the ith ordered data value.

The coefficient of skewness is a measure of the symmetry of a distribution. For symmetric distributions, the coefficient of skewness is zero. If a distribution has a long right tail of large values, it is positively skewed, and if it has a long left tail of small values, it is negatively skewed. The mean is larger than the median for positively skewed distributions and vice versa for negatively skewed distributions. The image below shows a positively skewed distribution.

Kurtosis is based on the size of the tails of a distribution and provides a measure of how likely it is that the distribution will produce outliers. The kurtosis of a normal distribution is equal to three.

Distributions with relatively thick tails are termed leptokurtic and have kurtosis greater than three. Distributions with relatively thin tails are termed platykurtic and have a kurtosis less than three. In the following diagram, a normal distribution is given in red, and a leptokurtic (thick-tailed) distribution is given in black.

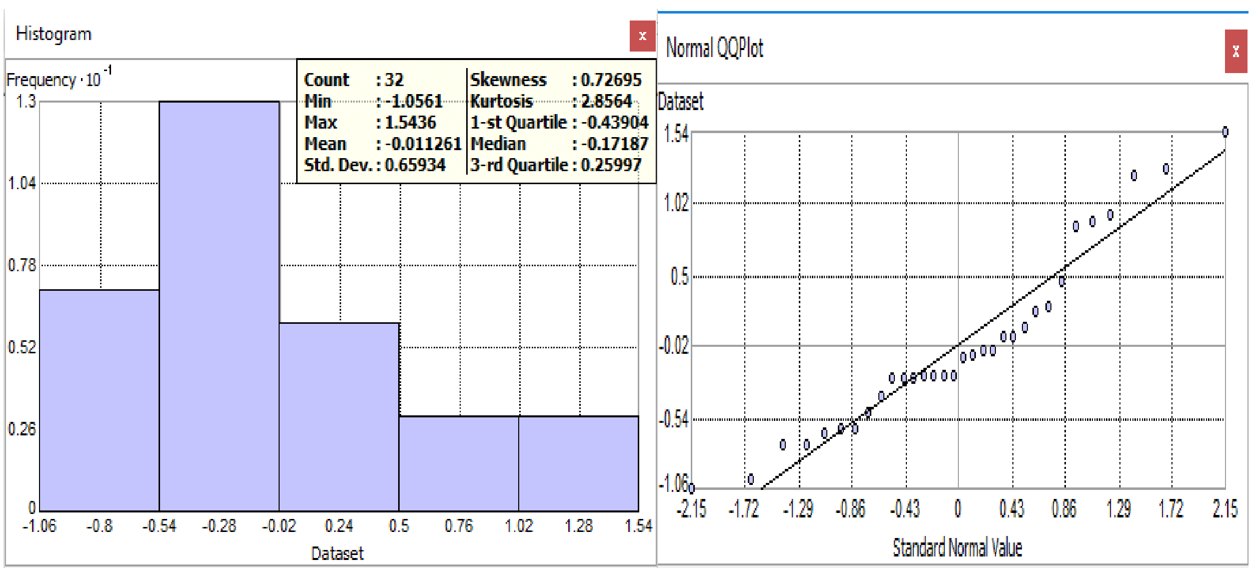
QQ Plot: First, the data values are ordered and cumulative distribution values are calculated as $(i-0.5)/n$ for the ith ordered value out of n total values (this gives the proportion of the data that falls below a certain value). A cumulative distribution graph is produced by plotting the ordered data versus the cumulative distribution values (graph on the top left in the figure below). The same process is done for a standard normal distribution (a Gaussian distribution with a mean of 0 and a standard deviation of 1, shown in the graph on the top right of the figure below). Once these two cumulative distribution graphs have been generated, data values corresponding to specific quantiles are paired and plotted in a QQ plot (bottom graph in the figure below).

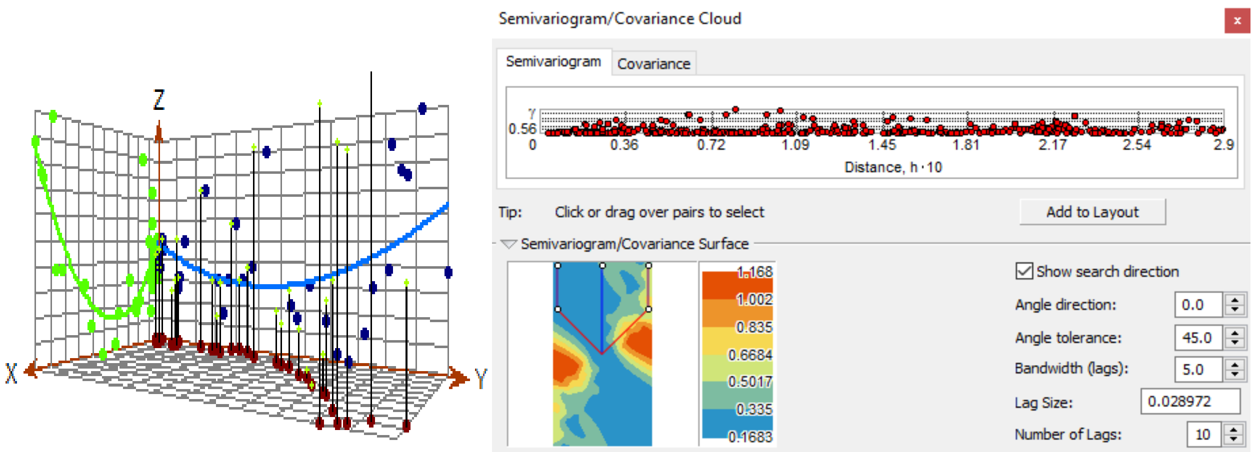


Trend Analysis: Visualize and examine spatial trends in a data set. The Trend Analysis tool provides a three-dimensional perspective of the data. The locations of sample points are plotted on the x, y plane. Above each sample point, the value is given by the height of a stick in the z-dimension. A unique feature of the Trend Analysis tool is that the values are then projected onto the x, z plane and the y, z plane as scatterplots. This can be thought of as sideways views through the three-dimensional data. Polynomials are then fit through the scatterplots on the projected planes. An additional feature is that you can rotate the data to isolate directional trends. The tool also includes other features that allow you to rotate and vary the perspective of the whole image, change size and color of points and lines, remove planes and points, and select the order of the polynomial that is to fit the scatterplots. By default, the tool will select second-order polynomials to show trends in the data, but you may want to investigate polynomials of order one and three to assess how well they fit the data.

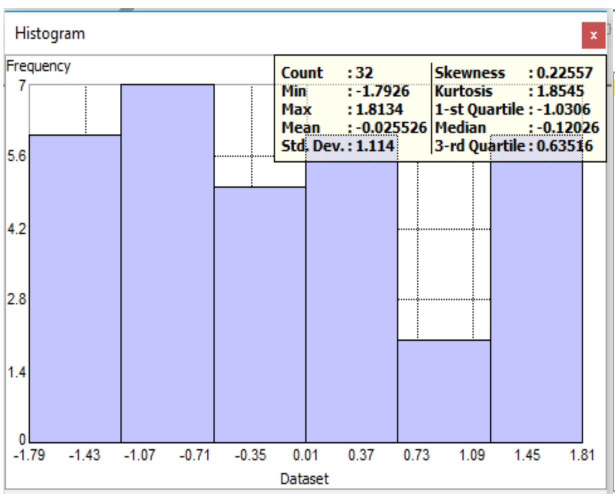
Semivariogram/Covariance Cloud: Evaluate the spatial dependence (Semivariogram and covariance) in a dataset. A function that describes the differences (variance) between samples separated by varying distances. Typically, the semivariogram will show low variance for small differences and larger variances at greater separation distances, indicating that the data is spatially autocorrelated. Semivariogram estimated from sample data are empirical semivariogram. They are represented as a set of points on a graph. A function is fitted to these points and is known as a semivariogram model. The semivariogram model is a key component in kriging (a powerful interpolation method that can provide predicted values, errors associated with the predictions, and information about the distribution of possible values for each location in the study area).

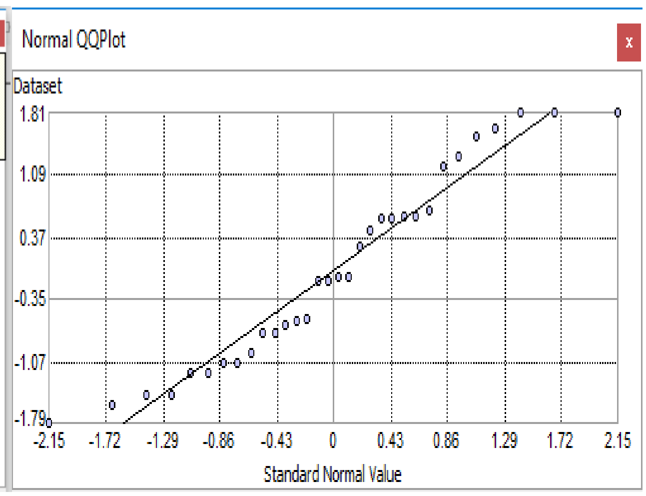
APPENDIX B1: Geostatistical Analyst Explore Data

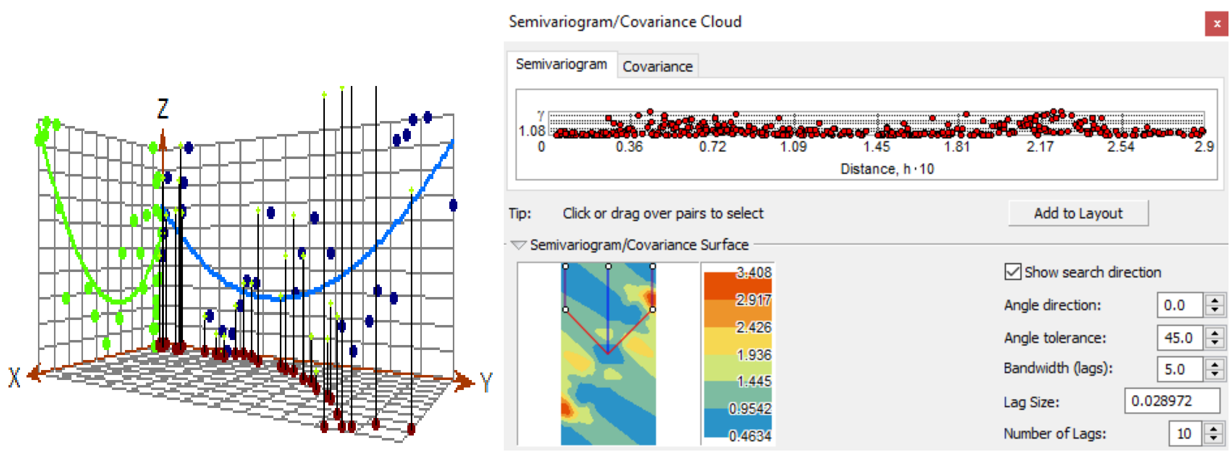




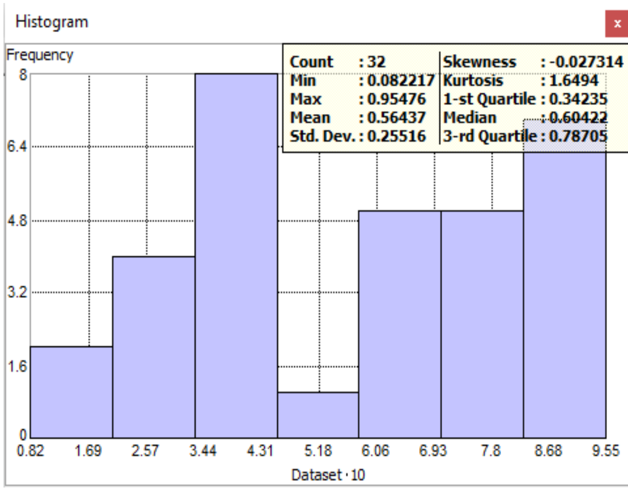
Drivers Age of18-30 vs GiZScore

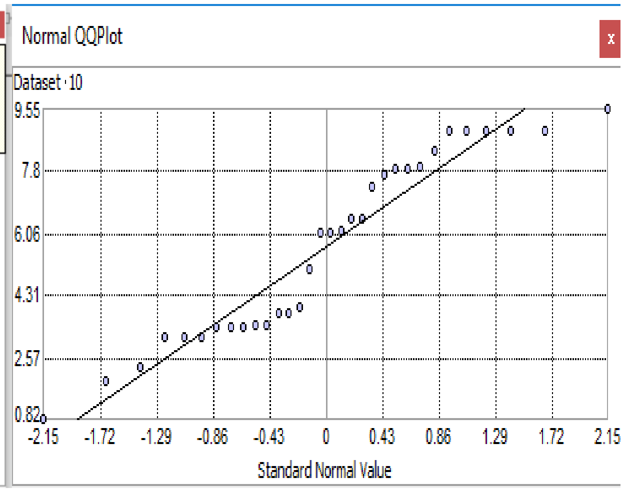


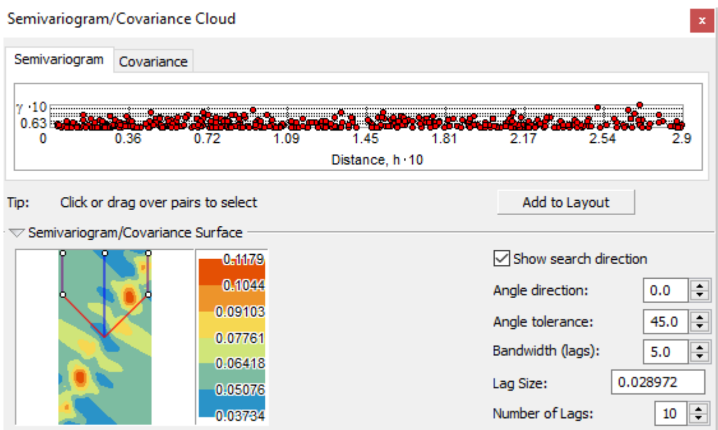


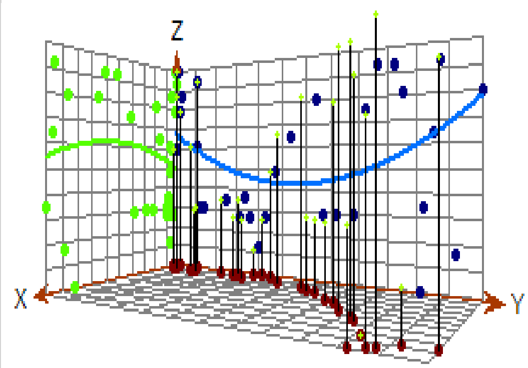


Drivers Age of 30-40 vs GiZScore

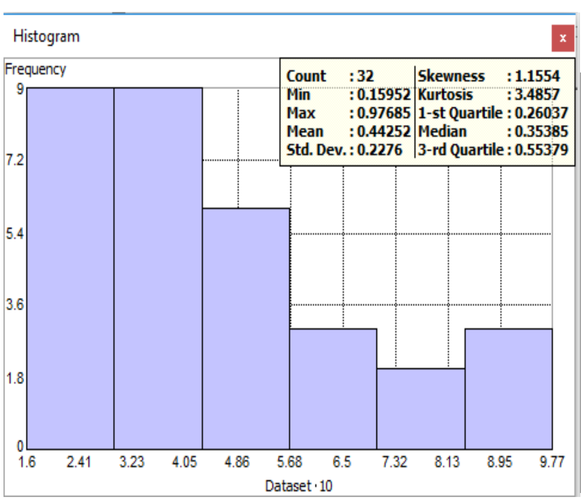


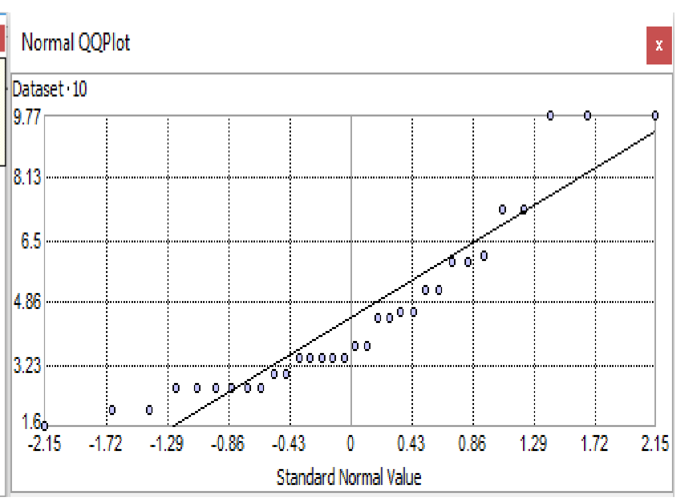


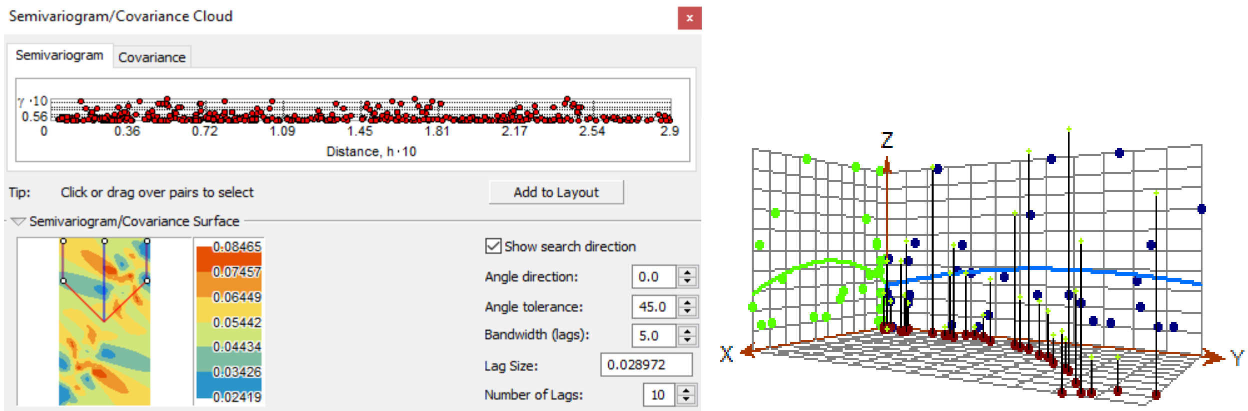




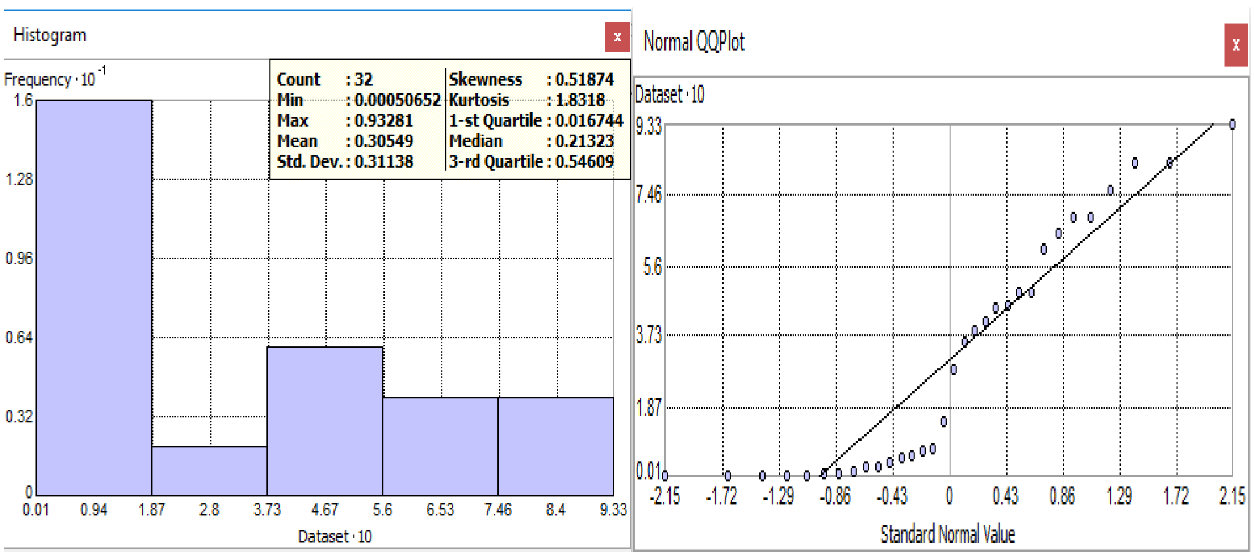
Drivers Age of 40-50 vs GiPValue

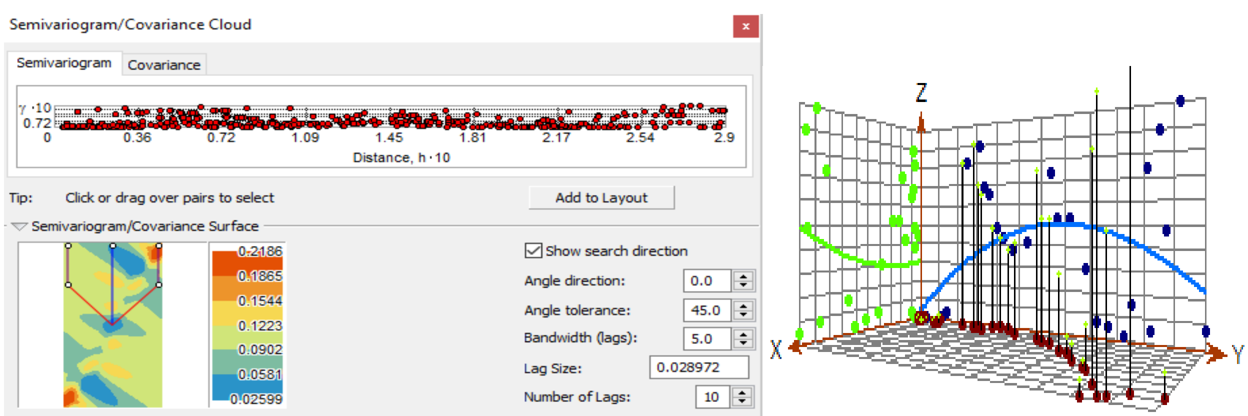




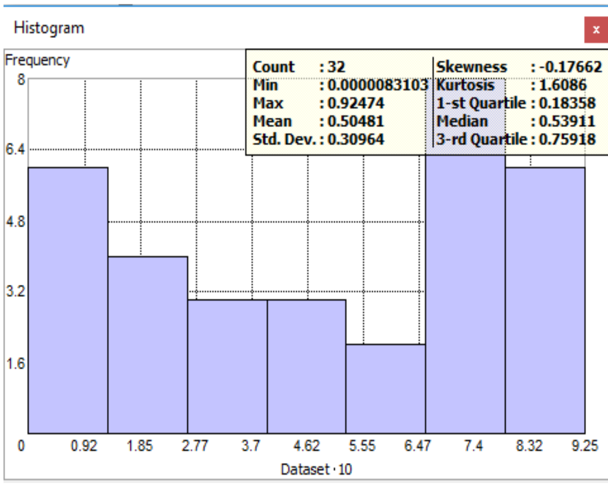


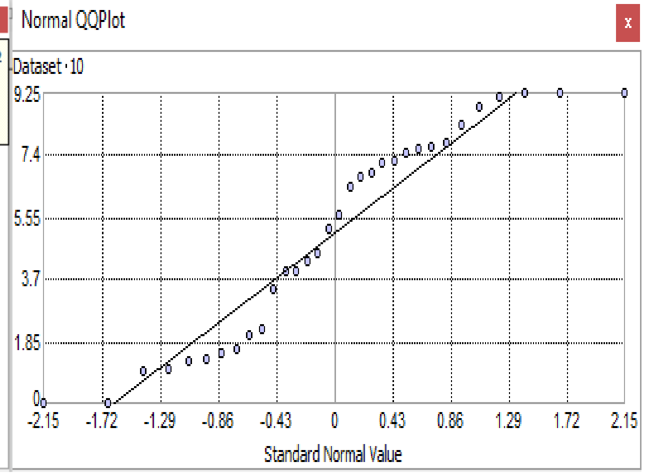
Drivers Age of 50 and above vs GiPValue

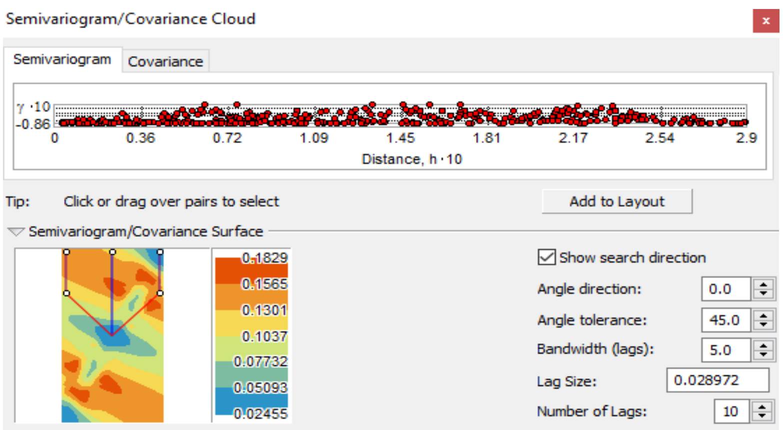


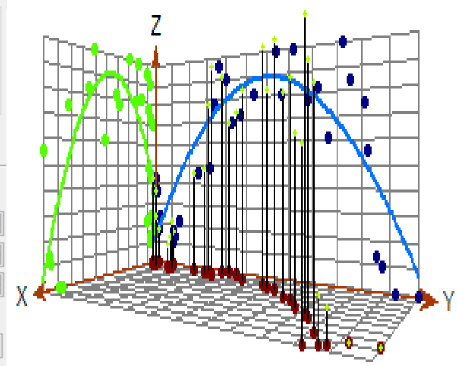


Gi_Velocity vs GiPValue

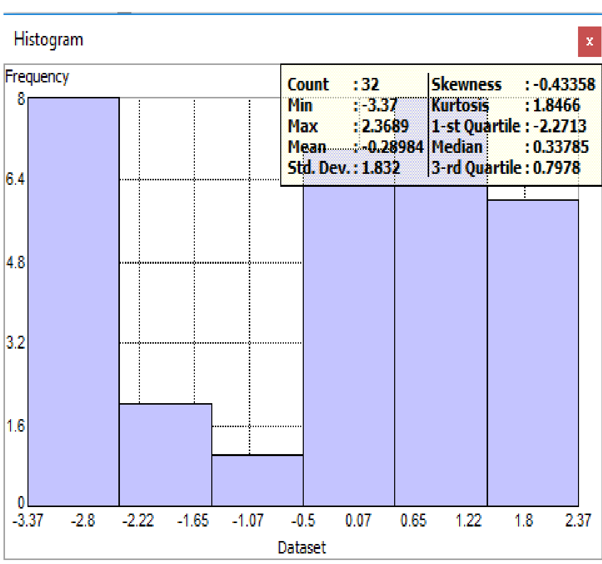


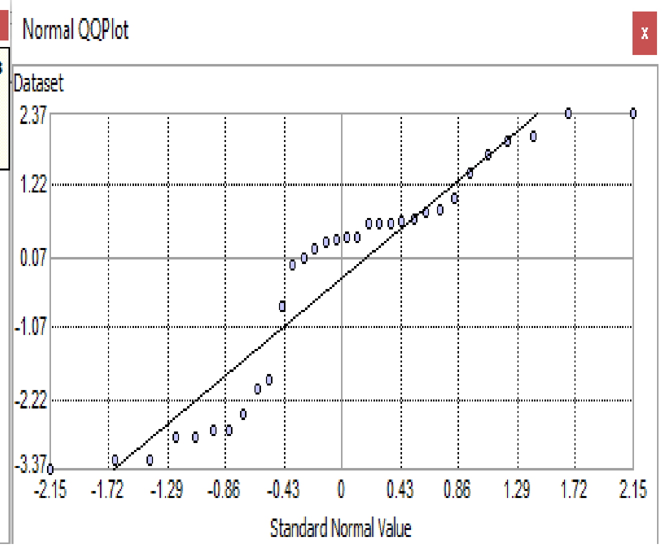


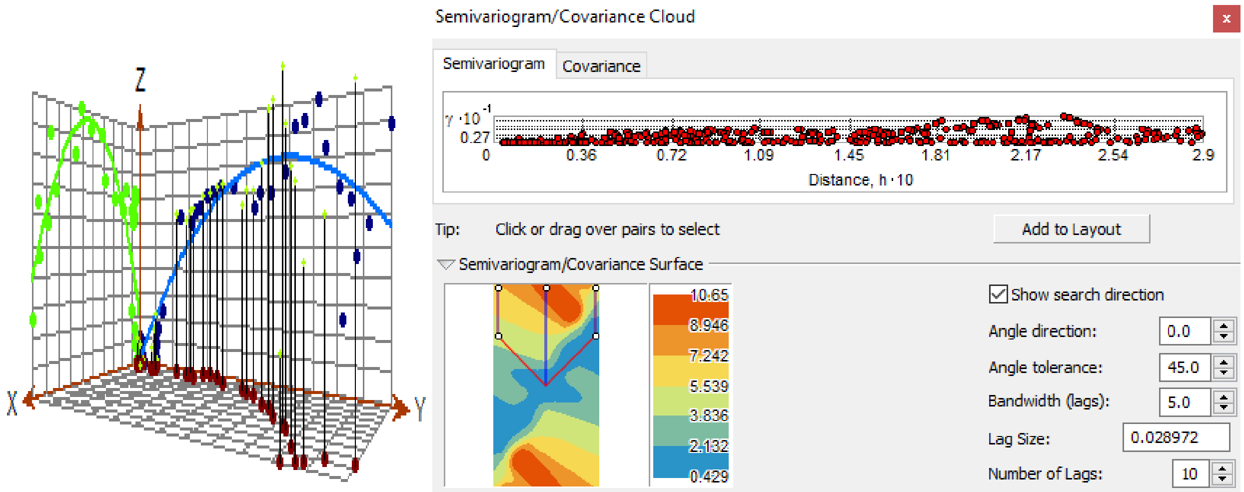




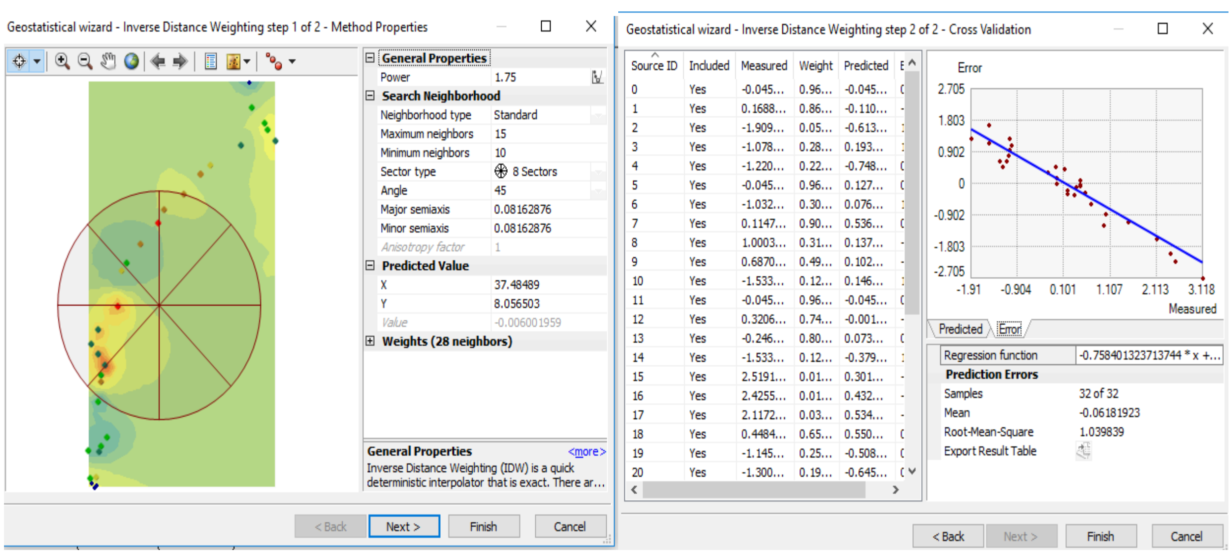
Gi_Radius vs GiPValue



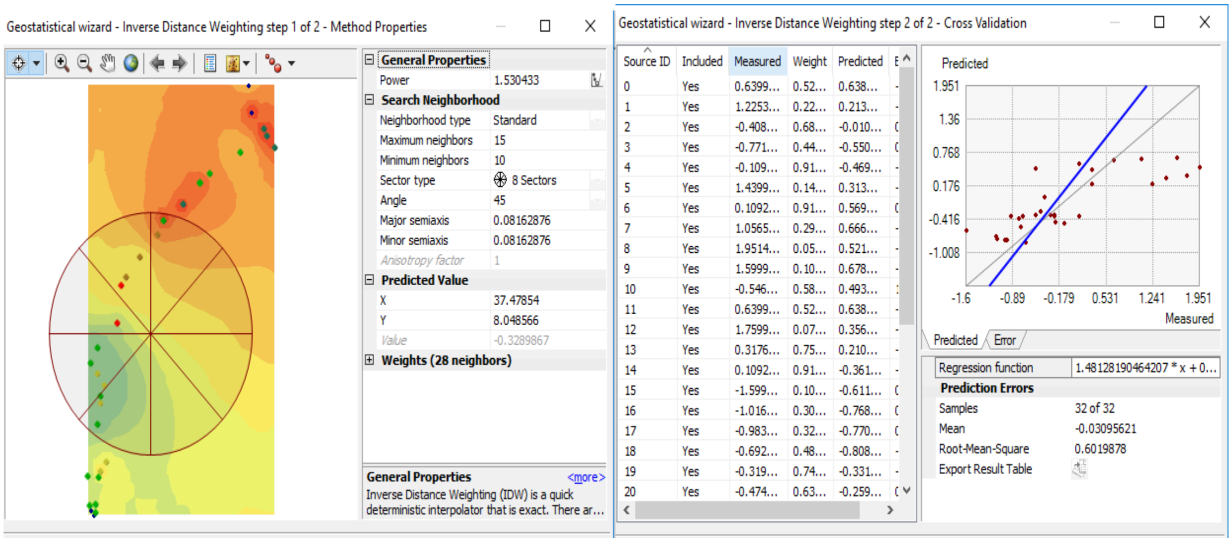




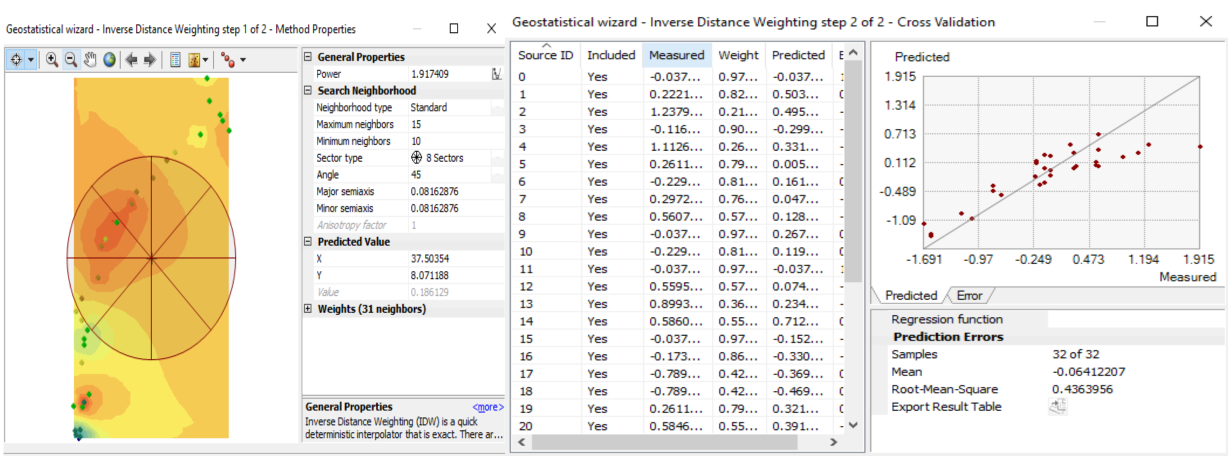
Gi_SSD vs GiZScore



IDW slight

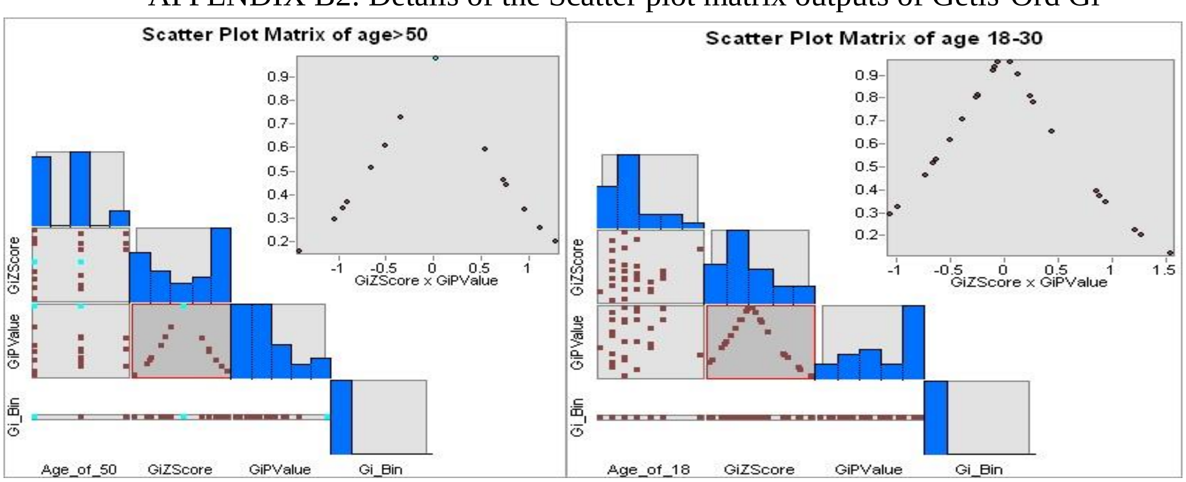


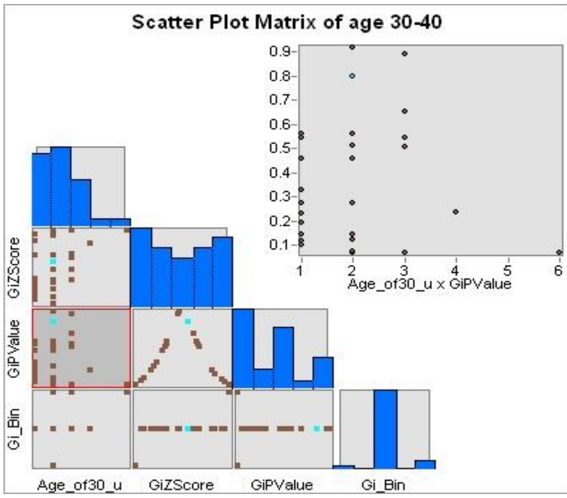
IDW PDO

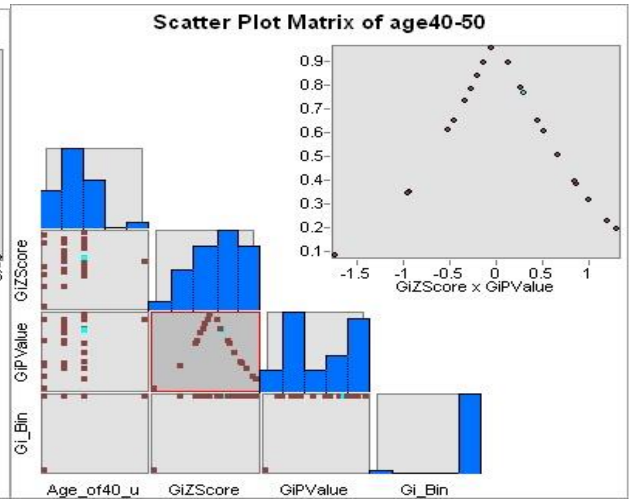


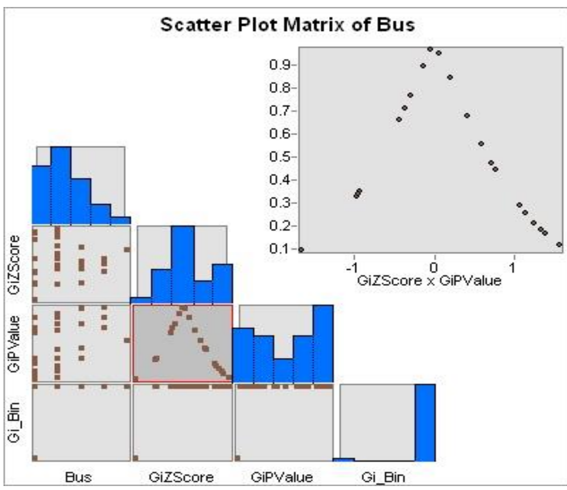
IDW Serious

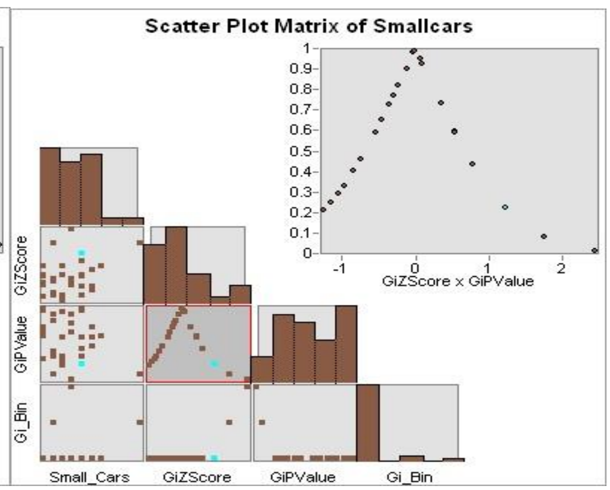
APPENDIX B2: Details of the Scatter plot matrix outputs of Getis-Ord Gi*



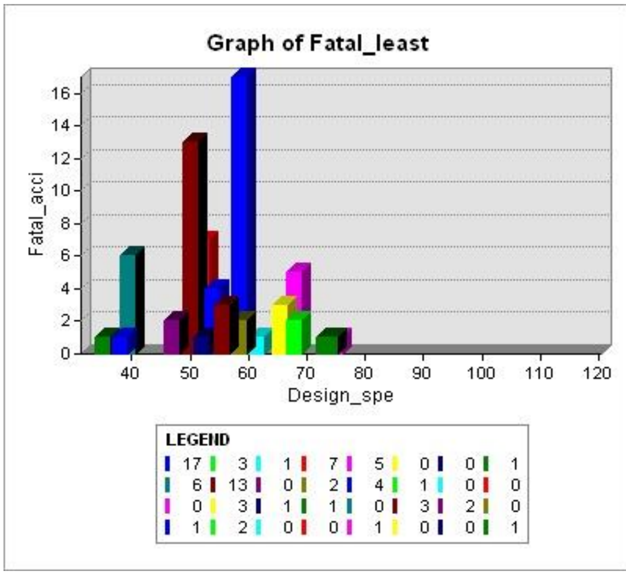

Scatter Plot Matrix of age 30-40
Age_of30_u x GiPValue
GiZScore
GiPValue
Gi_Bin
Age_of30_u

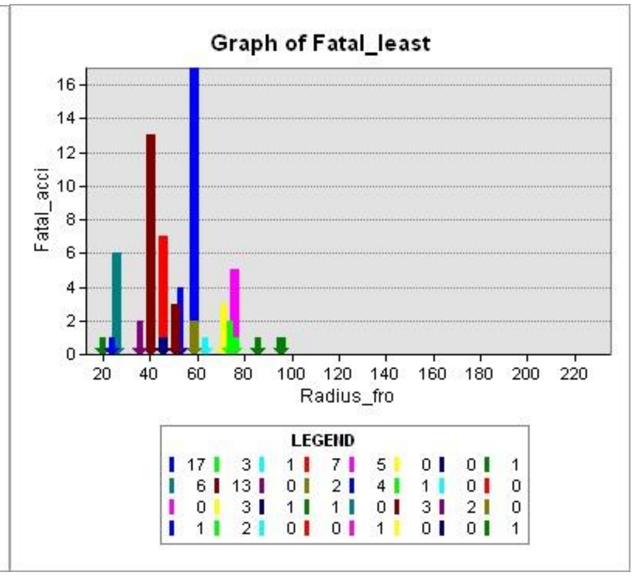

Scatter Plot Matrix of age40-50
GiZScore x GiPValue
GiZScore
GiPValue
Gi_Bin
Age_of40_u

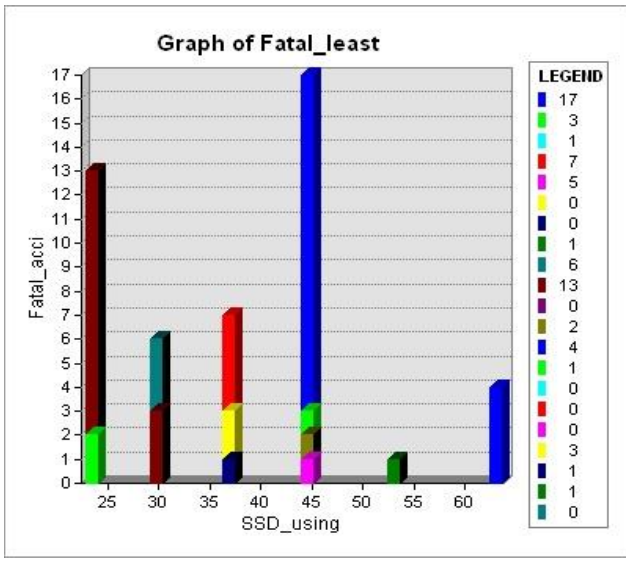

Scatter Plot Matrix of Bus
GiZScore x GiPValue
GiZScore
GiPValue
Gi_Bin
Bus

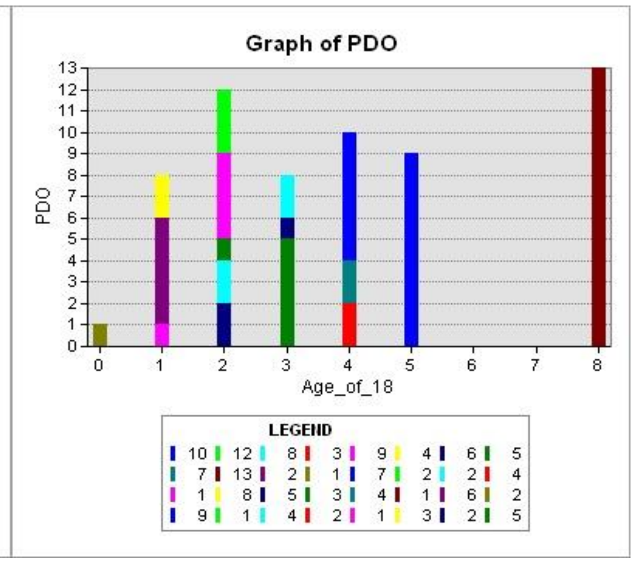

Scatter Plot Matrix of Smallcars
GiZScore x GiPValue
GiZScore
GiPValue
Gi_Bin
Small_Cars

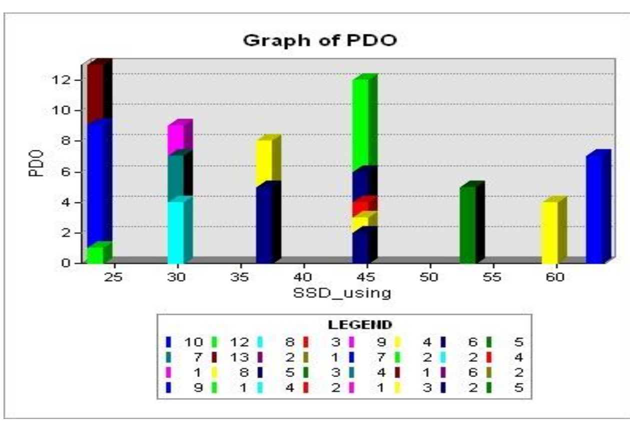
APPENDIX B3: Vertical Graph Results of Least Square

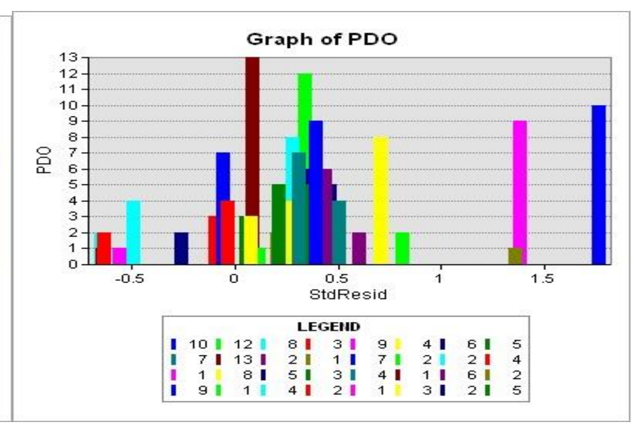


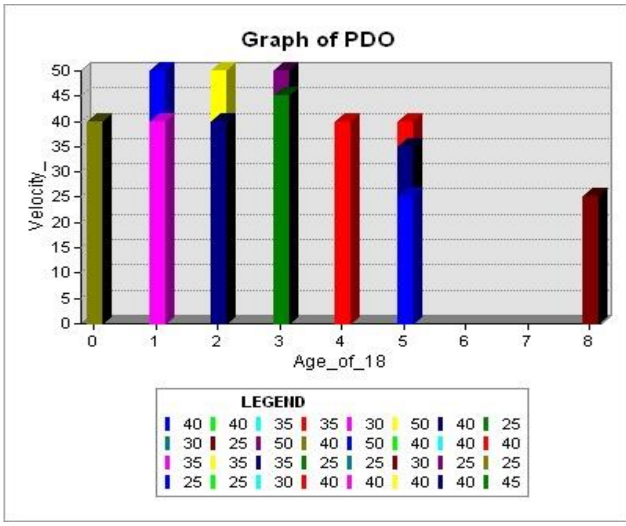


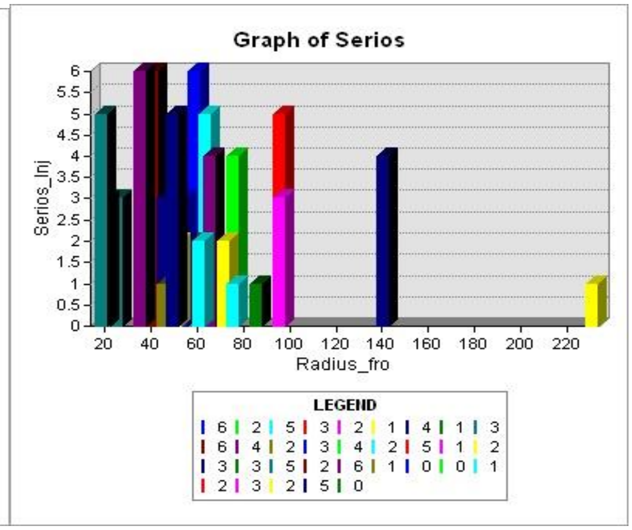


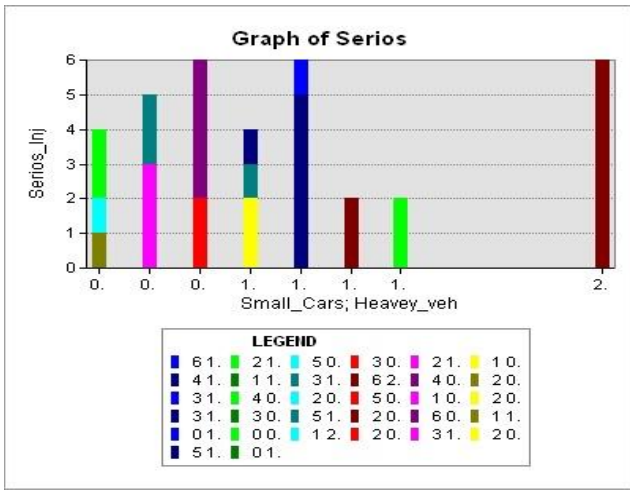


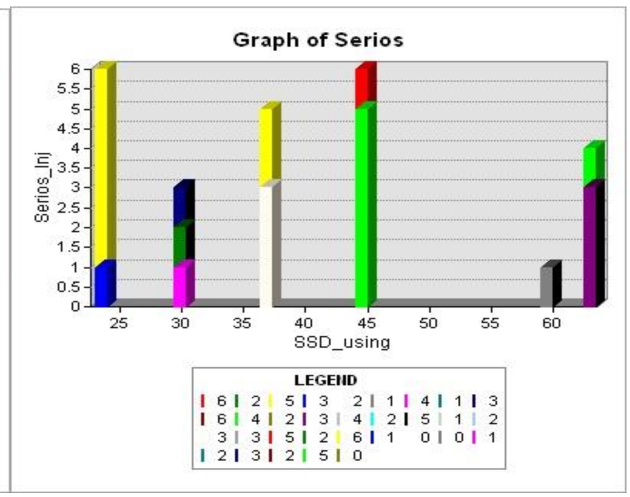


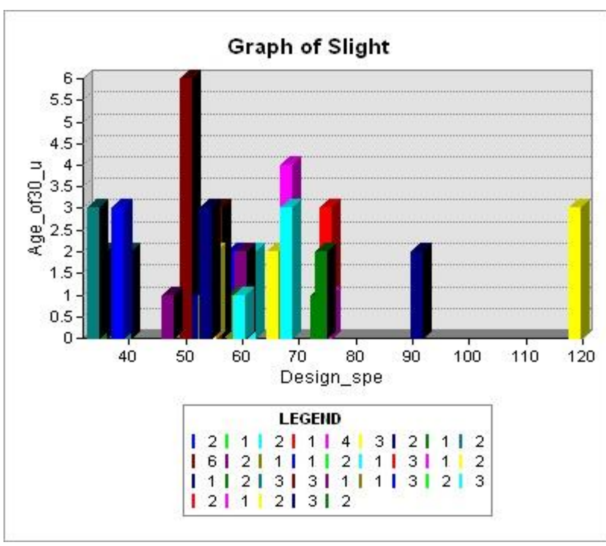


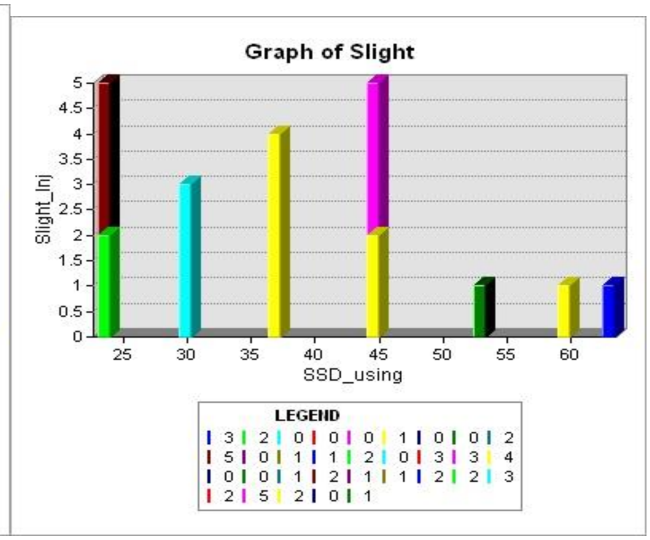

Graph of PDO
Velocity_
Age_of_18
LEGEND


Graph of Serios
Serios_Inj
Radius_fro
LEGEND

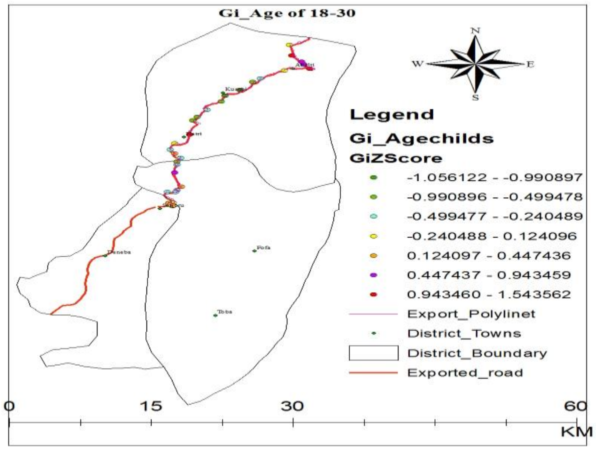

Graph of Serios
Serios_Inj
Small_Cars; Heavey_veh
LEGEND

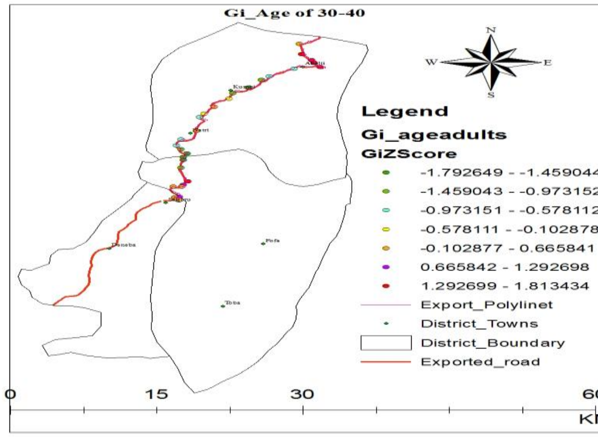

Graph of Serios
Serios_Inj
SSD_using
LEGEND

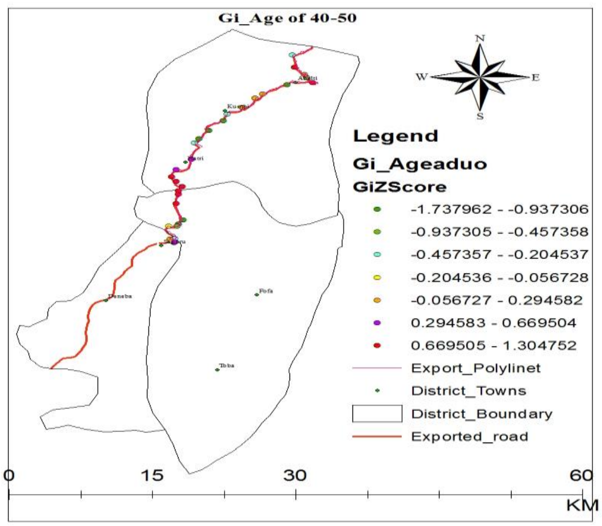

Graph of Slight
Age_of30_u
Design_spe
LEGEND

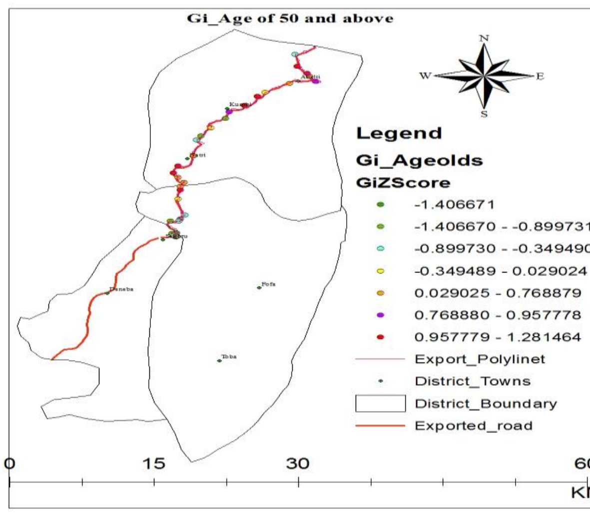

Graph of Slight
Slight_Inj
SSD_using
LEGEND

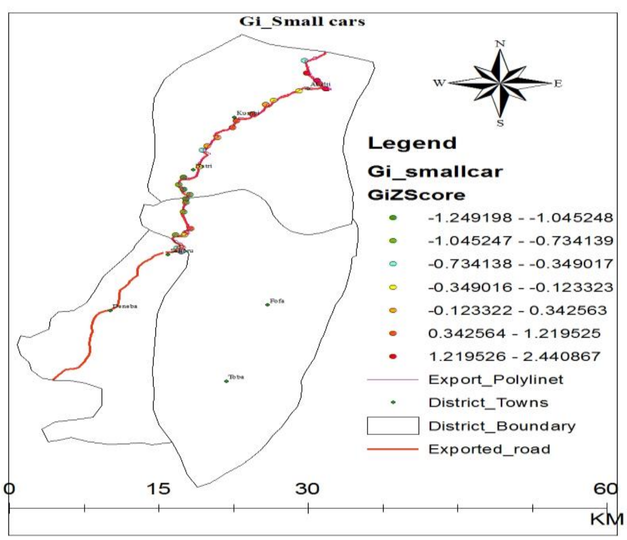
APPENDIX B4: Mapping results of the Getis-Ord (Gi* statistics)

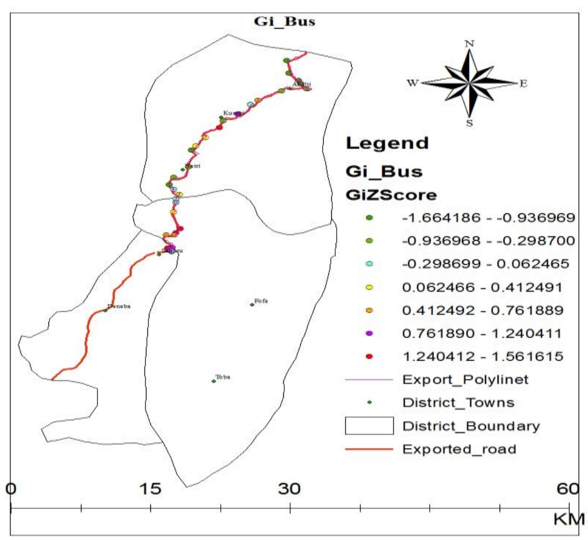


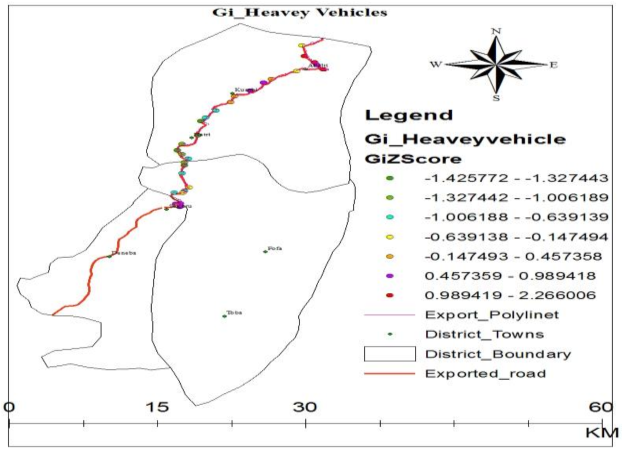


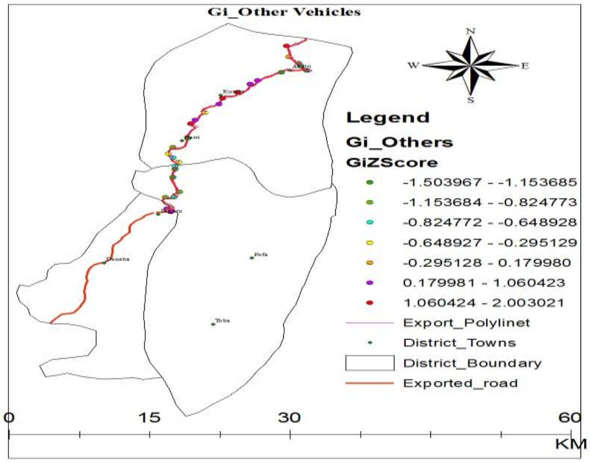












APPENDIX B5: Anselin Local Moran Index Results

Table

LMI_SSDAuto

FID	Shape *	SOURCE_I	SSD_using	LMiIndex	LMiZScore	LMiPValue	COType
0	Point ZM	0	44.376875	0.700199	1.340902	0.179952	
1	Point ZM	1	44.376875	0.712373	1.094586	0.273698	
2	Point ZM	2	36.691243	-0.058698	-0.038866	0.968997	
3	Point ZM	3	36.691243	-0.023435	0.01297	0.989652	
4	Point ZM	4	29.708268	0.766763	1.718517	0.085702	
5	Point ZM	5	59.356299	-1.157071	-2.059184	0.039477	HL
6	Point ZM	6	44.376875	0.973396	1.028529	0.303701	
7	Point ZM	7	23.376536	-0.60796	-1.238207	0.215639	
8	Point ZM	8	29.708268	0.859978	1.31156	0.189668	
9	Point ZM	9	23.376536	0.051971	0.154197	0.877454	
10	Point ZM	10	62.871485	-2.587752	-2.613624	0.008959	HL
11	Point ZM	11	44.376875	0.700199	1.340902	0.179952	
12	Point ZM	12	62.871485	1.133466	2.134079	0.032836	HH
13	Point ZM	13	44.376875	0.291281	0.475593	0.634365	
14	Point ZM	14	44.376875	-0.058698	-0.027041	0.978427	
15	Point ZM	15	44.376875	0.174621	0.378732	0.704887	
16	Point ZM	16	36.691243	-0.030487	0.004335	0.996541	
17	Point ZM	17	36.691243	-0.023435	0.018977	0.984859	
18	Point ZM	18	36.691243	-0.005803	0.056899	0.954626	
19	Point ZM	19	23.376536	1.079398	2.035098	0.041841	LL
20	Point ZM	20	23.376536	1.254935	3.151195	0.001626	LL
21	Point ZM	21	29.708268	0.797834	2.267872	0.023337	LL
22	Point ZM	22	23.376536	1.236128	3.815913	0.000136	LL
23	Point ZM	23	23.376536	1.298819	3.636599	0.000276	LL
24	Point ZM	24	23.376536	1.353674	2.980834	0.002875	LL
25	Point ZM	25	23.376536	1.254935	3.151195	0.001626	LL
26	Point ZM	26	29.708268	0.859978	1.919002	0.054984	
27	Point ZM	27	44.376875	0.011298	0.106629	0.915083	
28	Point ZM	28	44.376875	0.116291	0.319497	0.74935	
29	Point ZM	29	44.376875	0.291281	0.330898	0.740721	
30	Point ZM	30	44.376875	0.116291	0.218363	0.827146	
31	Point ZM	31	52.822138	0.675851	1.0409	0.297922	

1 (0 out of 32 Selected)

LMI_SSDAuto

Table

LMI_BUS

FID	Shape *	SOURCE_I	Bus	LMiIndex	LMiZScore	LMiPValue	COType
0	Point ZM	0	3	-0.4481	-0.215883	0.829079	
1	Point ZM	1	4	-3.794769	-2.761584	0.005752	HL
2	Point ZM	2	2	-0.107544	-0.031854	0.974588	
3	Point ZM	3	1	-1.664372	-1.184407	0.236252	
4	Point ZM	4	2	-0.597467	-0.253551	0.799843	
5	Point ZM	5	1	0.699036	0.489011	0.624834	
6	Point ZM	6	1	1.29821	1.370983	0.17038	
7	Point ZM	7	2	0.35848	0.263877	0.791875	
8	Point ZM	8	1	-0.599174	-0.395818	0.692239	
9	Point ZM	9	3	-1.813739	-1.055043	0.291406	
10	Point ZM	10	2	0.041823	0.076337	0.939151	
11	Point ZM	11	3	-0.4481	-0.215883	0.829079	
12	Point ZM	12	1	-3.561757	-2.129167	0.03324	LH
13	Point ZM	13	1	-3.794769	-2.761584	0.005752	LH
14	Point ZM	14	1	0.233012	0.273348	0.784586	
15	Point ZM	15	1	-0.366162	-0.165534	0.868524	
16	Point ZM	16	2	0.017924	0.088302	0.929637	
17	Point ZM	17	2	-0.215088	-0.04658	0.962848	
18	Point ZM	18	4	-2.919053	-1.510162	0.131002	
19	Point ZM	19	2	-0.63929	-0.333366	0.738858	
20	Point ZM	20	3	0.618805	0.384366	0.700707	
21	Point ZM	21	5	-3.190474	-1.372768	0.169824	
22	Point ZM	22	2	-0.663188	-0.189026	0.850072	
23	Point ZM	23	1	-4.993117	-2.198485	0.027915	LH
24	Point ZM	24	3	-0.597467	-0.253551	0.799843	
25	Point ZM	25	2	-0.555644	-0.194305	0.845937	
26	Point ZM	26	3	-0.597467	-0.253551	0.799843	
27	Point ZM	27	2	0.017924	0.088302	0.929637	
28	Point ZM	28	2	0.35848	0.263877	0.791875	
29	Point ZM	29	3	-0.832186	-0.824287	0.409776	
30	Point ZM	30	4	-2.237941	-1.609032	0.107609	
31	Point ZM	31	2	-0.298733	-0.173396	0.86234	

1 (0 out of 32 Selected)

LMI_BUS

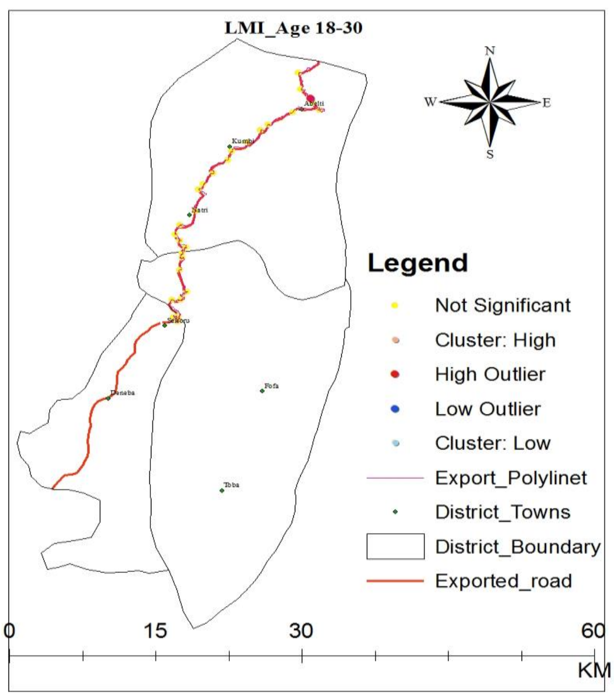
Table

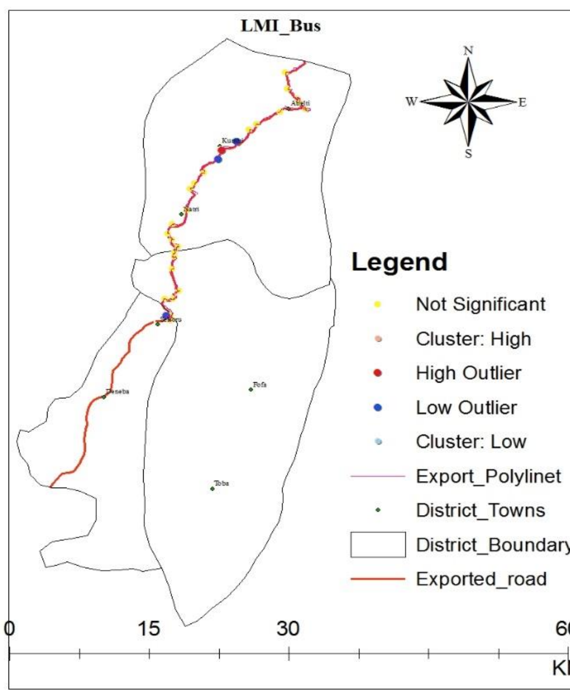
LMI_Serios

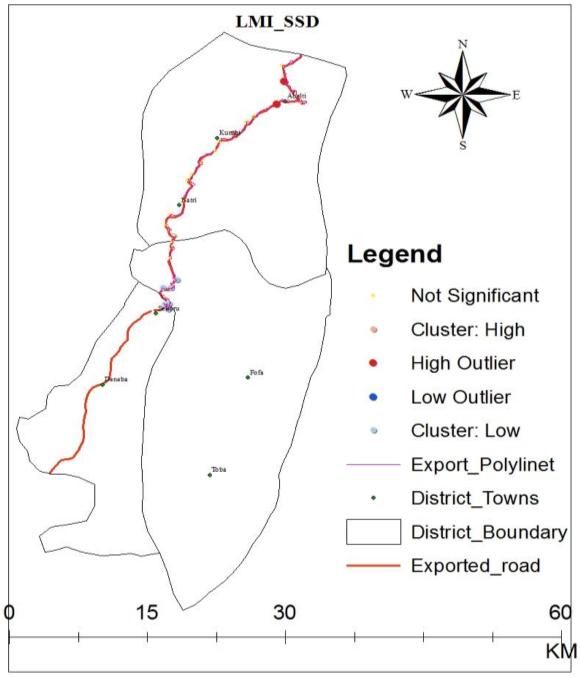
FID	Shape *	SOURCE_I	Serios_Inj	LMiIndex	LMiZScore	LMiPValue	COType
0	Point ZM	0	6	-1.118085	-1.979532	0.047756	HL
1	Point ZM	1	2	-0.175002	-0.208919	0.834511	
2	Point ZM	2	5	0.497006	0.774627	0.43856	
3	Point ZM	3	3	-0.019174	0.01915	0.984722	
4	Point ZM	4	2	-0.296742	-0.566582	0.570998	
5	Point ZM	5	1	-0.491527	-0.837277	0.402437	
6	Point ZM	6	4	-0.676573	-0.656007	0.51182	
7	Point ZM	7	1	-0.399004	-0.78565	0.432073	
8	Point ZM	8	3	0.049001	0.11893	0.905331	
9	Point ZM	9	6	-1.118085	-1.979532	0.047756	HL
10	Point ZM	10	4	-0.676573	-0.656007	0.51182	
11	Point ZM	11	2	-0.053261	-0.038291	0.969456	
12	Point ZM	12	3	0.037638	0.127425	0.898604	
13	Point ZM	13	4	0.273003	0.446779	0.655035	
14	Point ZM	14	2	-0.540223	-0.517183	0.605028	
15	Point ZM	15	5	-0.540223	-0.926053	0.354418	
16	Point ZM	16	1	-0.121436	-0.217492	0.827825	
17	Point ZM	17	2	0.129349	0.346198	0.729194	
18	Point ZM	18	3	-0.053261	-0.044994	0.964112	
19	Point ZM	19	3	0.014913	0.085996	0.931469	
20	Point ZM	20	5	0.012965	0.110293	0.912177	
21	Point ZM	21	2	-0.012681	0.053294	0.957498	
22	Point ZM	22	6	-1.213622	-3.542232	0.000397	HL
23	Point ZM	23	1	0.248655	0.764734	0.44443	
24	Point ZM	24	0	0.677182	1.519772	0.128568	
25	Point ZM	25	0	0.677182	1.730216	0.083592	
26	Point ZM	26	1	0.572485	1.295489	0.195152	
27	Point ZM	27	2	0.092827	0.305064	0.760317	
28	Point ZM	28	3	-0.019174	0.028029	0.977639	
29	Point ZM	29	2	-0.540223	-0.517183	0.605028	
30	Point ZM	30	5	1.188492	1.786683	0.073989	
31	Point ZM	31	0	-1.056403	-1.498933	0.133891	

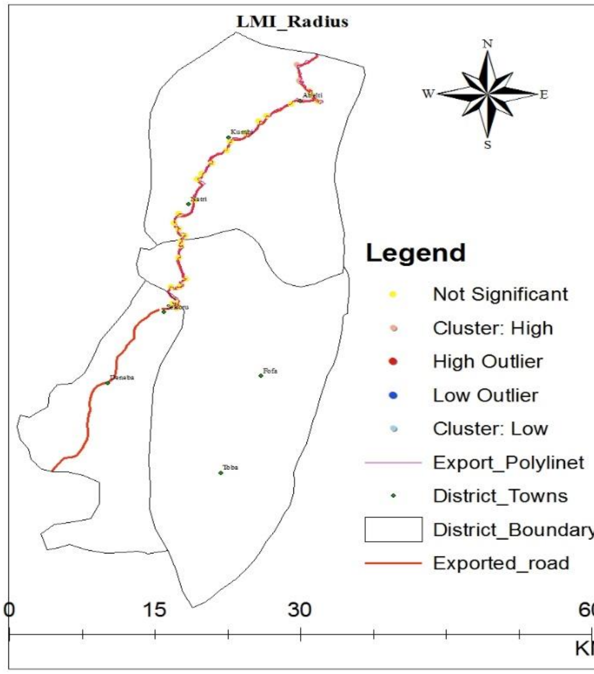
1 (0 out of 32 Selected)

LMI_Serios

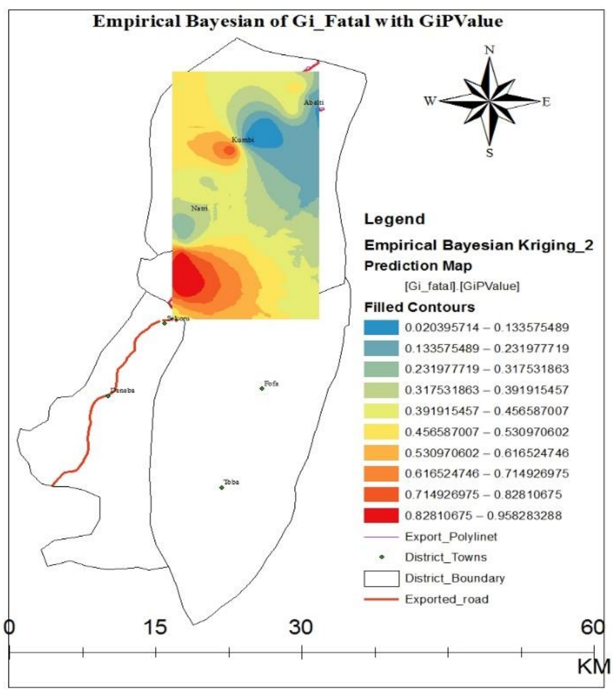


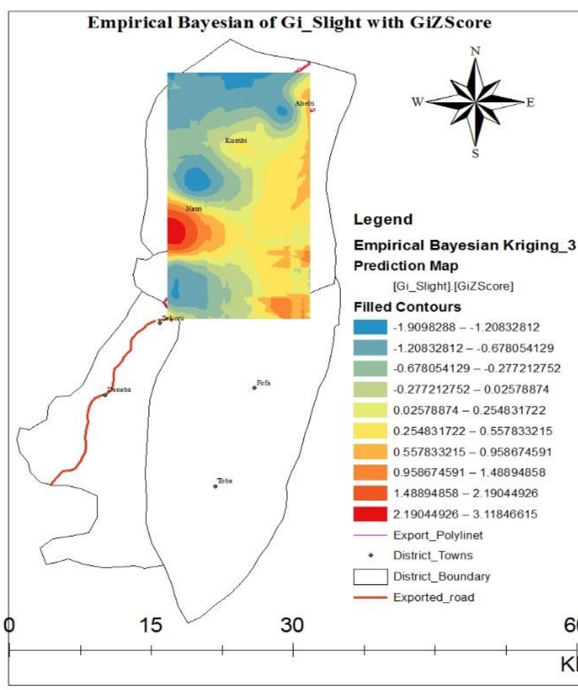


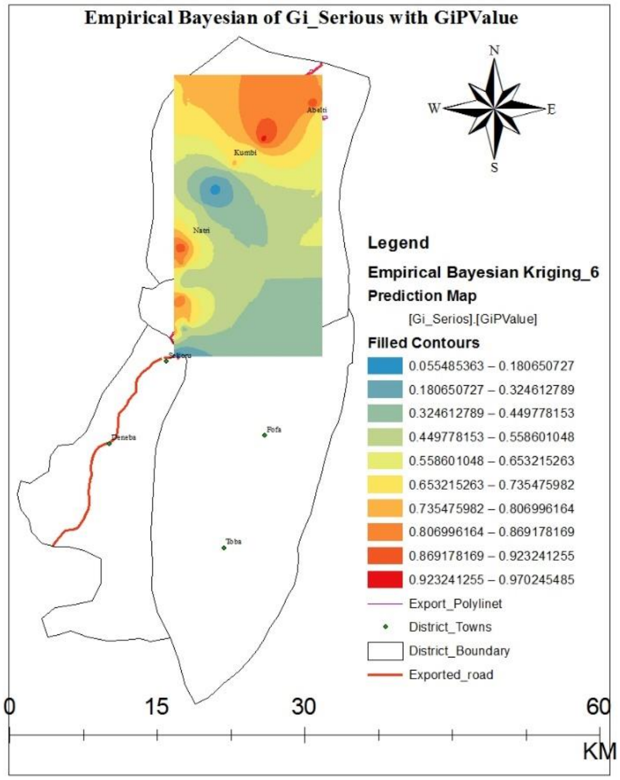


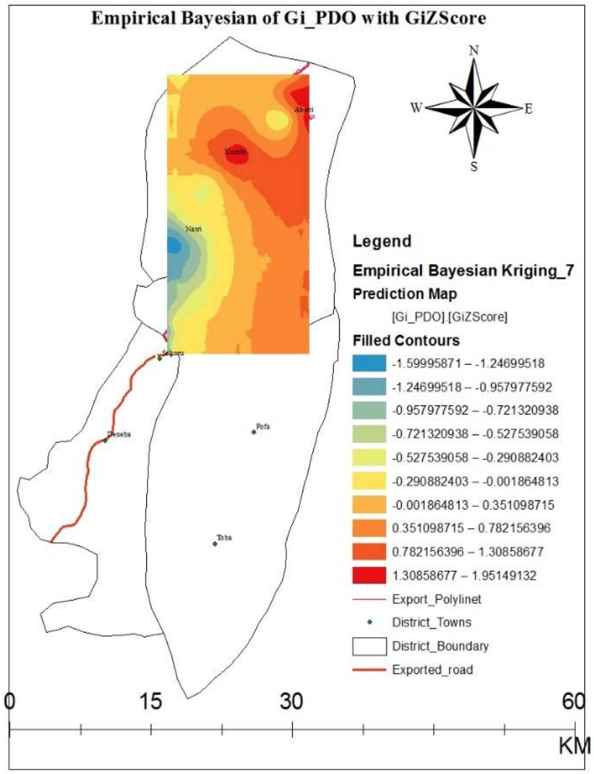


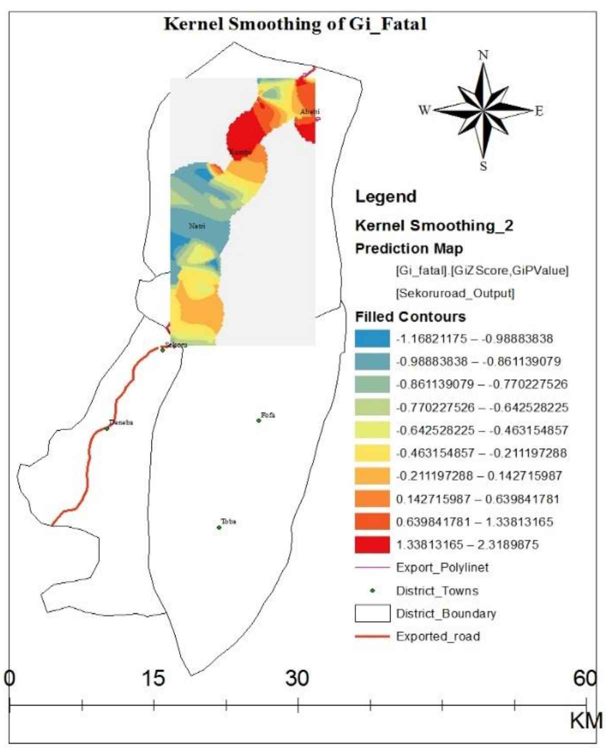
APPENDIX B6: CoKriging Geostatistical Analyst Results

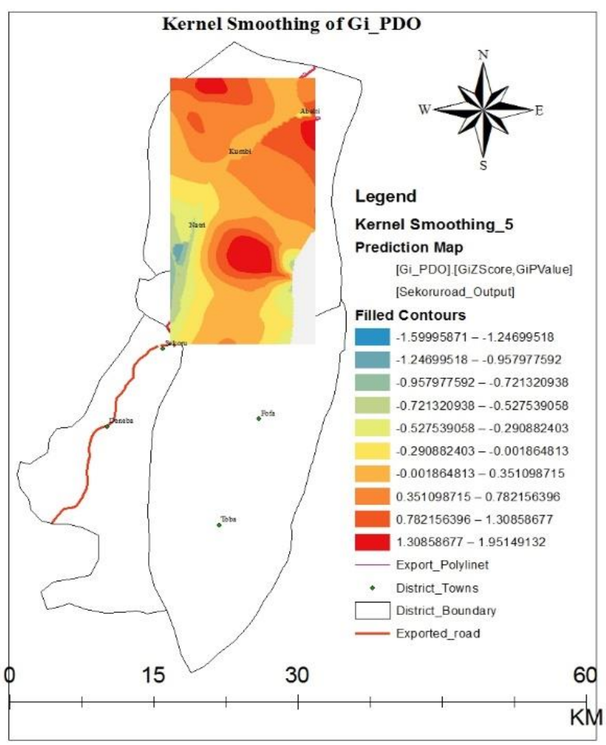


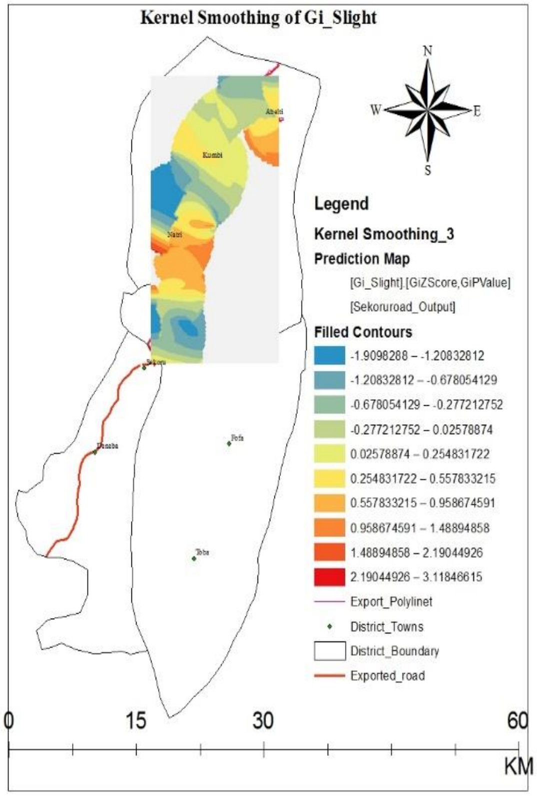


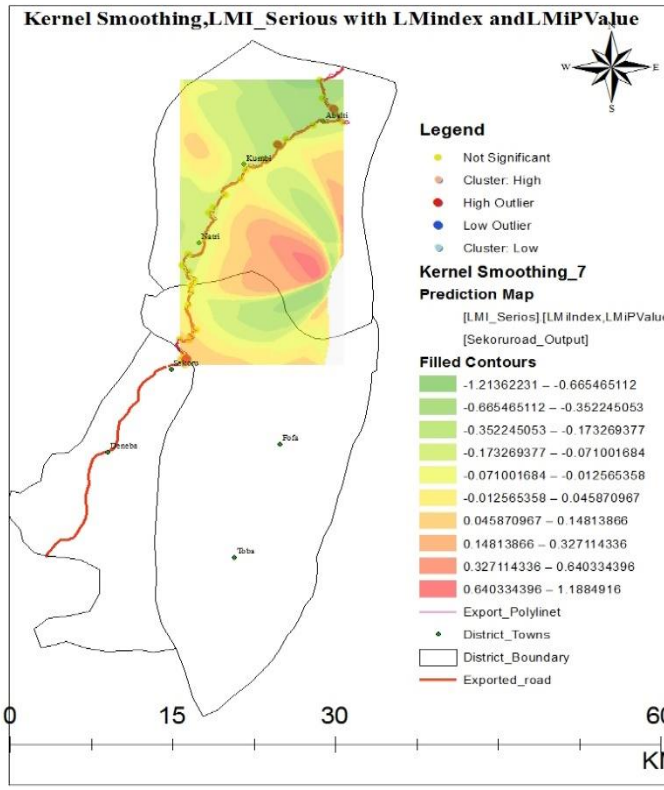

Empirical Bayesian of Gi_Serious with GiPValue
N
W
E
S
Legend
Empirical Bayesian Kriging_6
Prediction Map
[Gi_Serios].[GiPValue]
Filled Contours
0.055485363 – 0.180650727
0.180650727 – 0.324612789
0.324612789 – 0.449778153
0.449778153 – 0.558601048
0.558601048 – 0.653215263
0.653215263 – 0.735475982
0.735475982 – 0.806996164
0.806996164 – 0.869178169
0.869178169 – 0.923241255
0.923241255 – 0.970245485
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM

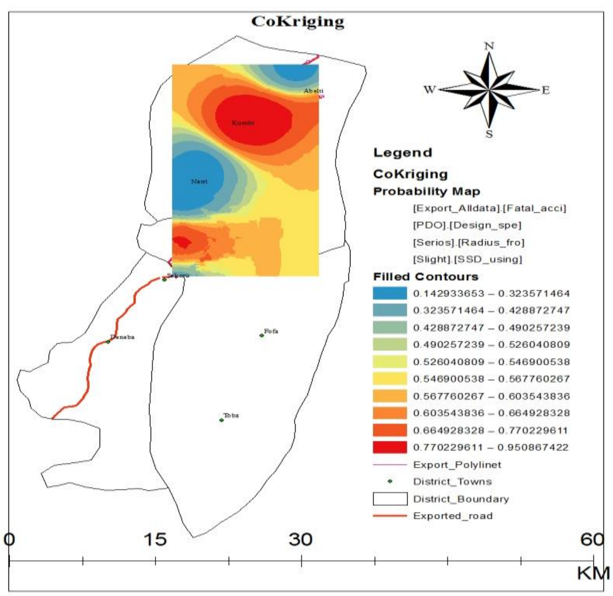

Empirical Bayesian of Gi_PDO with GiZScore
N
W
E
S
Legend
Empirical Bayesian Kriging_7
Prediction Map
[Gi_PDO].[GiZScore]
Filled Contours
-1.59995871 – -1.24699518
-1.24699518 – -0.957977592
-0.957977592 – -0.721320938
-0.721320938 – -0.527539058
-0.527539058 – -0.290882403
-0.290882403 – -0.001864813
-0.001864813 – 0.351098715
0.351098715 – 0.782156396
0.782156396 – 1.30858677
1.30858677 – 1.95149132
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM

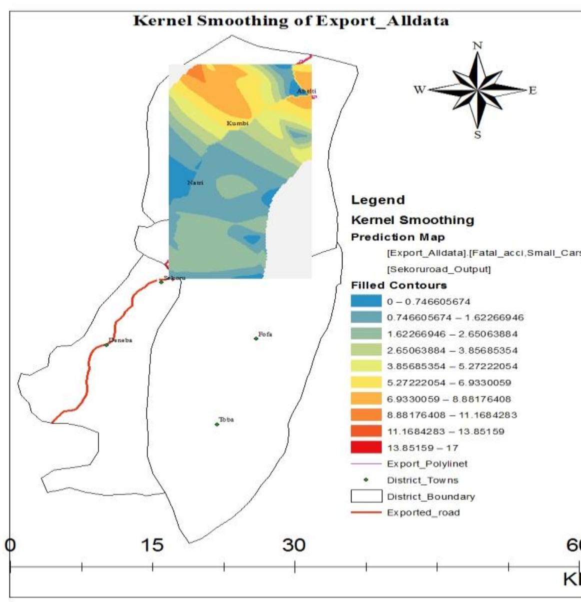

Kernel Smoothing of Gi_Fatal
N
W
E
S
Legend
Kernel Smoothing_2
Prediction Map
[Gi_fatal].[GiZScore,GiPValue]
[Sekoruroad_Output]
Filled Contours
-1.16821175 – -0.98883838
-0.98883838 – -0.861139079
-0.861139079 – -0.770227526
-0.770227526 – -0.642528225
-0.642528225 – -0.463154857
-0.463154857 – -0.211197288
-0.211197288 – 0.142715987
0.142715987 – 0.639841781
0.639841781 – 1.33813165
1.33813165 – 2.3189875
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM


Kernel Smoothing of Gi_PDO
N
W
E
S
Legend
Kernel Smoothing_5
Prediction Map
[Gi_PDO].[GiZScore,GiPValue]
[Sekoruroad_Output]
Filled Contours
-1.59995871 – -1.24699518
-1.24699518 – -0.957977592
-0.957977592 – -0.721320938
-0.721320938 – -0.527539058
-0.527539058 – -0.290882403
-0.290882403 – -0.001864813
-0.001864813 – 0.351098715
0.351098715 – 0.782156396
0.782156396 – 1.30858677
1.30858677 – 1.95149132
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM


Kernel Smoothing of Gi_Slight
N
W
E
S
Legend
Kernel Smoothing_3
Prediction Map
[Gi_Slight].[GiZScore,GiPValue]
[Sekoruroad_Output]
Filled Contours
-1.9098288 – -1.20832812
-1.20832812 – -0.678054129
-0.678054129 – -0.277212752
-0.277212752 – 0.02578874
0.02578874 – 0.254831722
0.254831722 – 0.557833215
0.557833215 – 0.958674591
0.958674591 – 1.48894858
1.48894858 – 2.19044926
2.19044926 – 3.11846615
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM


Kernel Smoothing,LMI_Serious with LMindex andLMiPValue
N
W
E
S
Legend
Not Significant
Cluster: High
High Outlier
Low Outlier
Cluster: Low
Kernel Smoothing_7
Prediction Map
[LMI_Serios].[LMiIndex,LMiPValue]
[Sekoruroad_Output]
Filled Contours
-1.21362231 – -0.665465112
-0.665465112 – -0.352245053
-0.352245053 – -0.173269377
-0.173269377 – -0.071001684
-0.071001684 – -0.012565358
-0.012565358 – 0.045870967
0.045870967 – 0.14813866
0.14813866 – 0.327114336
0.327114336 – 0.640334396
0.640334396 – 1.1884916
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM


CoKriging
N
W
E
S
Legend
CoKriging
Probability Map
[Export_Alldata].[Fatal_acci]
[PDO].[Design_spe]
[Serios].[Radius_fro]
[Slight].[SSD_using]
Filled Contours
0.142933653 – 0.323571464
0.323571464 – 0.428872747
0.428872747 – 0.490257239
0.490257239 – 0.526040809
0.526040809 – 0.546900538
0.546900538 – 0.567760267
0.567760267 – 0.603543836
0.603543836 – 0.664928328
0.664928328 – 0.770229611
0.770229611 – 0.950867422
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM


Kernel Smoothing of Export_Alldata
N
W
E
S
Legend
Kernel Smoothing
Prediction Map
[Export_Alldata].[Fatal_acci,Small_Cars]
[Sekoruroad_Output]
Filled Contours
0 – 0.746605674
0.746605674 – 1.62266946
1.62266946 – 2.65063884
2.65063884 – 3.85685354
3.85685354 – 5.27222054
5.27222054 – 6.9330059
6.9330059 – 8.88176408
8.88176408 – 11.1684283
11.1684283 – 13.85159
13.85159 – 17
Export_Polylinet
District_Towns
District_Boundary
Exported_road
0
15
30
60
KM

APPENDIX B7: Kernel function formulas

For all formulas below, r is a radius centered at point s and h in the b

andwidth.

1. Exponential:

$$e^{-3\left(\frac{r}{h}\right)}$$

2. Gaussian:

$$e^{-3\left(\frac{r}{h}\right)^2}$$

3. Quartic:

$$\left(1-\left(\frac{r}{h}\right)^2\right)^2 \text{ for } \frac{r}{h} < 1$$

4. Epanechinkov:

$$1-\left(\frac{r}{h}\right)^2, \text{ for } \frac{r}{h} < 1$$

5. PolynomialOrder5:

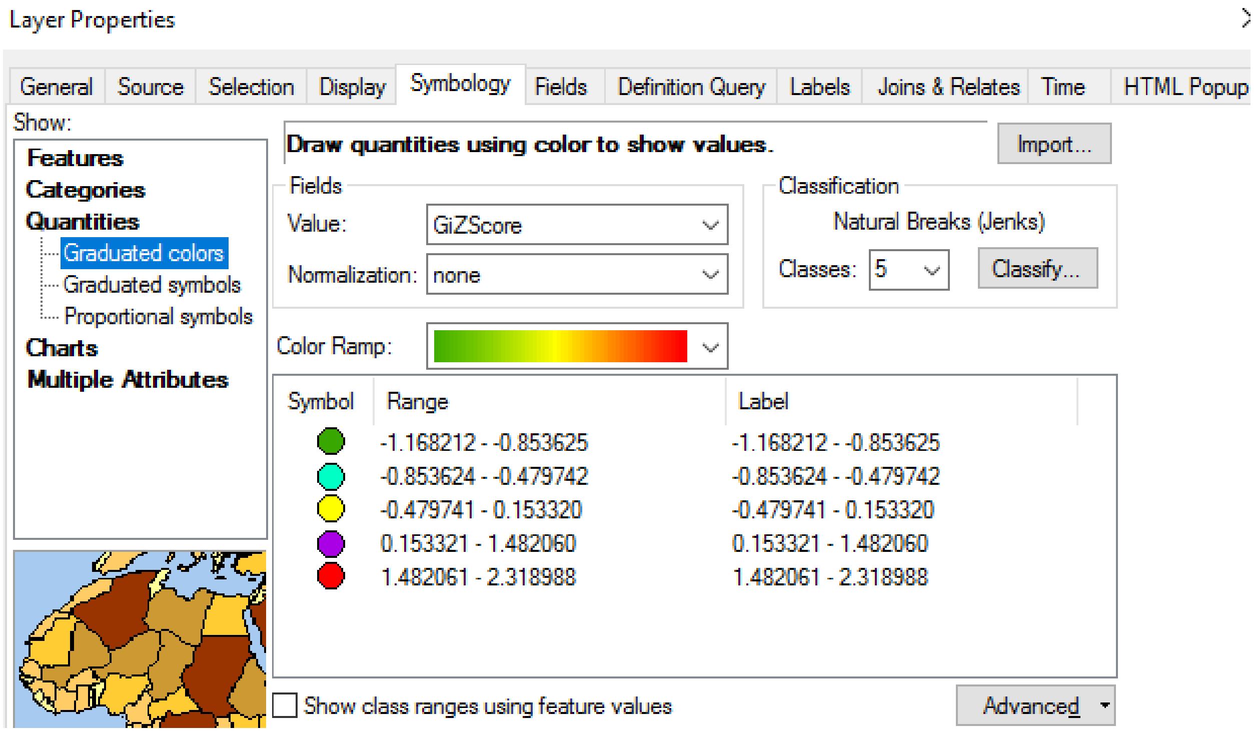
$$1-\left(\frac{r}{h}\right)^3\left(10-\left(\frac{r}{h}\right)\left(15-6\left(\frac{r}{h}\right)\right)\right), \text{ for } \frac{r}{h} < 1$$

6. Constant:

$$I(s-h < s_i < s+h)$$

Where I(expression) is an indicator function that takes a value of 1 if expression is true and a value of 0 if expression is false.

APPENDIX C: Thematic map layer of Gi*ZScore for Fatal accidents



APPENDIX C1: Hotspot areas selected from the whole segment

No	Northing	Easting	Elevation	Distance in m	Fatal	Slight Injur	Serios Injur	PDO	Small_Ca	Bus	HV	Motor	Other	Age18-30	Age30-40	Age 40-	Age of>50	V	V_cal	g%	Carriage	e	f	R	SSD
1	8.1498	37.5187	1792	21246	17	3	6	10	5	3	1	0	1	4	2	2	1	40	58.2	0	3.25	0.08	0.38	58	45
2	8.12471	37.4889	1931	26811	3	2	2	12	7	4	1	1	2	2	1	1	0	40	51.3	0.1	3.25	0.08	0.38	45	45
3	8.08662	37.4593	1949	33400	1	0	5	8	5	2	0	0	1	3	2	1	1	35	61.1	0.1	3.25	0.08	0.39	62.5	37.5
4	7.98704	37.4362	1895	50398	7	0	3	3	2	1	0	0	0	1	1	1	1	35	51.8	0.1	3.25	0.08	0.39	45	37.5
5	7.95552	37.4381	1813	55528	5	0	2	9	6	2	1	0	0	2	4	0	0	30	67.6	0.1	3.25	0.08	0.4	75	30
6	8.1974	37.5599	1350	5702.3	0	1	1	4	2	1	0	0	1	2	3	0	1	50	118	0.1	3.25	0.08	0.4	230	30
7	8.21605	37.5574	1241	3395.2	0	0	4	6	4	1	1	0	1	3	2	2	0	40	90.4	0	3.25	0.08	0.38	140	45
8	8.18117	37.5723	1575	9940.4	1	0	1	5	5	2	1	0	0	2	1	0	2	25	72.7	0	3	0.08	0.41	85	25
9	8.17338	37.5788	1644	11184	6	2	3	7	4	1	1	0	0	1	2	0	0	30	39	0	3	0.08	0.4	25	30
10	8.18617	37.5702	1506	8338.7	13	5	6	13	11	3	2	0	0	8	6	5	1	25	49.9	0.1	3	0.08	0.41	40	25
11	8.1702	37.5514	1680	16129	0	0	4	2	2	2	0	0	0	3	2	0	0	50	59.6	0.1	3.25	0.08	0.35	65	55
12	8.1498	37.5186	1792	21260	2	1	2	1	1	3	0	0	0	0	1	1	1	40	58.2	0.1	3.25	0.08	0.38	58	45
13	8.13535	37.5052	1841	24104	4	1	3	7	5	1	1	0	1	1	1	1	2	50	53.5	0.1	3.25	0.08	0.35	52.5	55
14	8.11497	37.4847	1883	28114	1	2	4	2	1	1	0	0	0	2	2	1	1	40	66.2	0.1	3.25	0.08	0.38	75	45
15	8.08073	37.4548	1936	34343	0	0	2	2	1	1	0	0	1	1	1	1	0	40	59.2	0.1	3.25	0.08	0.38	60	45
16	8.03896	37.436	1928	41649	0	3	5	4	3	1	0	1	0	5	3	2	1	40	74.5	0.1	3.25	0.08	0.38	95	45
17	8.02144	37.4363	1853	44149	0	3	1	1	1	2	0	0	0	1	1	2	2	35	53.2	0	3.25	0.08	0.39	47.5	37.5
18	8.01293	37.4419	1788	45925	3	4	2	8	5	2	0	0	1	1	2	1	1	35	65.1	0	3.25	0.08	0.39	71	37.5
19	8.00106	37.4385	1909	48552	1	0	3	5	3	4	1	1	0	5	1	2	0	35	51.8	0.1	3.25	0.08	0.39	45	37.5
20	7.96171	37.443	1904	53931	1	0	3	3	3	2	0	0	0	3	2	1	1	25	34.8	0.1	3.25	0.08	0.41	19.5	25
21	7.95165	37.4371	1780	56543	0	1	5	4	2	3	1	0	1	4	3	2	0	25	33.5	0.1	3.25	0.08	0.41	18	25
22	7.95199	37.4282	1806	57733	3	2	2	1	6	5	0	0	0	2	3	0	1	30	55.2	0.1	3.25	0.08	0.4	50	30
23	7.93287	37.4346	1751	60896	2	1	6	6	3	2	0	0	1	1	1	1	0	25	46.7	0	3.25	0.08	0.41	35	25
24	7.93232	37.4297	1802	62083	0	1	1	2	1	1	1	0	0	2	1	2	1	25	52.9	0.1	3.25	0.08	0.41	45	25
25	7.92633	37.4332	1894	63606	1	2	0	9	6	3	1	0	1	5	3	1	0	25	37.8	0.1	3.25	0.08	0.41	23	25
26	7.92874	37.4309	1926	64023	2	2	0	1	1	2	0	0	0	3	2	2	0	25	67.6	0.1	3.25	0.08	0.41	73.5	25
27	7.92743	37.4343	1859	63103	0	3	1	4	5	3	2	0	1	2	3	1	1	30	67.6	0	3.25	0.08	0.4	75	30
28	8.00646	37.4383	1859	47304	0	2	2	2	3	2	0	0	0	4	2	2	1	40	54	0.1	3.25	0.08	0.38	50	45
29	8.02813	37.4307	1873	43112	1	5	3	1	2	2	1	0	1	1	1	1	0	40	74.5	0	3.25	0.08	0.38	95	45
30	8.0556	37.4516	1913	38399	0	2	2	3	4	3	0	0	0	3	2	1	1	40	54	0.1	3.25	0.08	0.38	50	45
31	8.09939	37.47	1918	31401	0	0	5	2	5	4	1	0	0	2	3	0	0	40	53.2	0.1	3.25	0.08	0.38	48.5	45
32	8.15585	37.5263	1738	19698	1	1	0	5	4	2	1	0	1	3	2	1	0	45	73.7	0.1	3.25	0.08	0.37	95	50

APPENDIX C2: Sight distance values (ERA, 2001)

Design Speed (km/hr)	Coefficient of Friction (f)	Stopping sight distance(m)	Passing Sight distance (m) from formula	Reduced Passing Sight Distance for design (m)
20	0.42	20	160	50
30	0.4	30	217	75
40	0.38	45	285	125
50	0.35	55	345	175
60	0.33	85	407	225
70	0.31	110	482	275
85	0.3	155	573	340
100	0.29	205	670	375
120	0.28	285	792	425

Clearance Distance (ERA, 2001)

Speed Group (km/hr.)	50-65	66-80	81-100	101-120
d3 (m)	30	55	80	100

Minimum provision of PSD (AASHTO, 2004)

Design Standard	PERCENT PASSING OPPORTUNITY vs TERRAIN				
	Flat	Rolling	Mountainous	Escarpment	Urban/peri-urban
DS2	50%	50%	25%	0%	20%
DS3	50%	33%	25%	0%	20%
DS4	25%	25%	15%	0%	20%
DS5	25%	25%	15%	0%	20%
DS6	20%	20%	15%	0%	20%
DS7	20%	20%	15%	0%	20%

APPENDIX D: Curves along the Gibe bridge to Sekoru town road segment

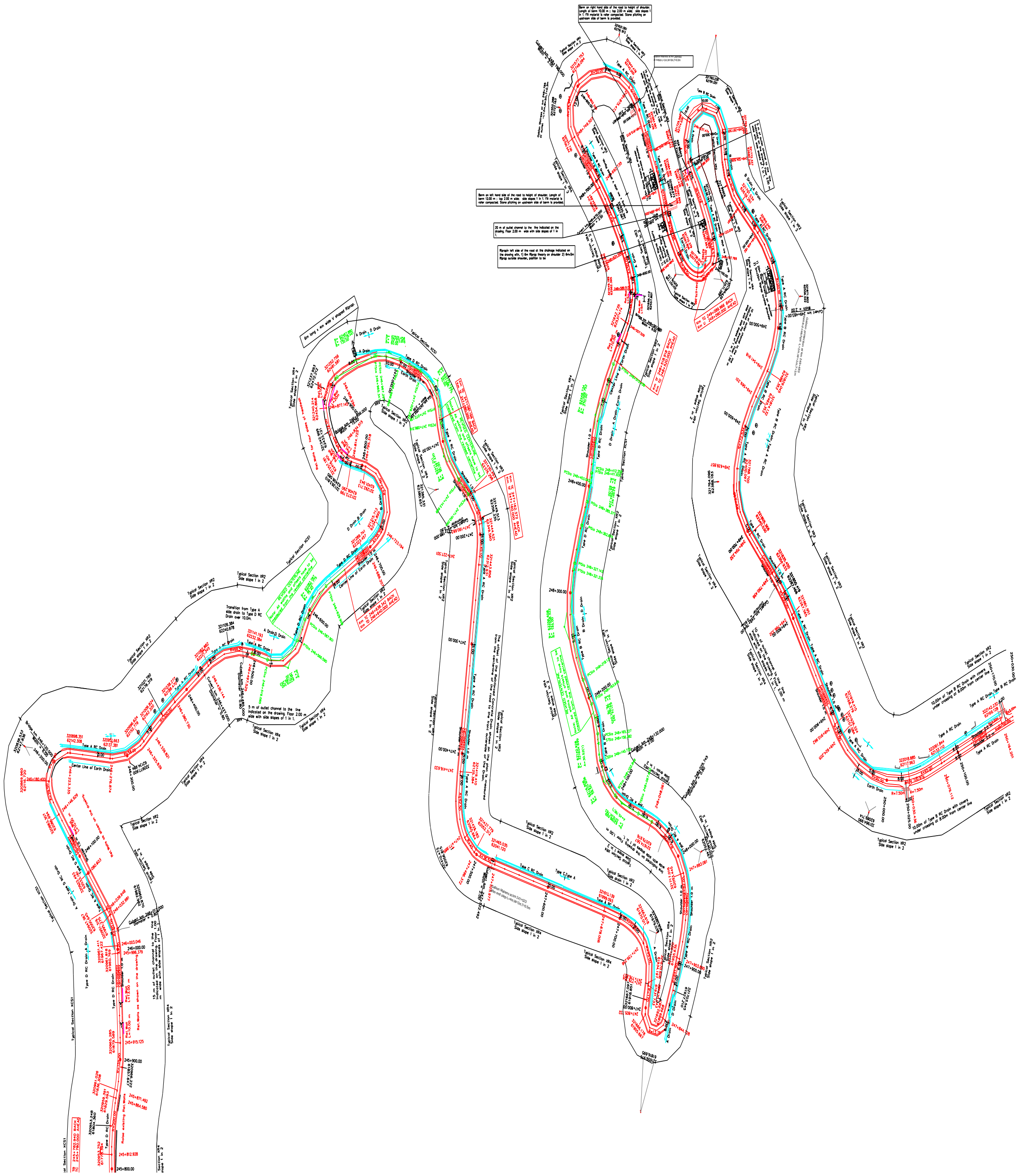


Figure: curves around Eastern part of Sekoru town (Simini river)

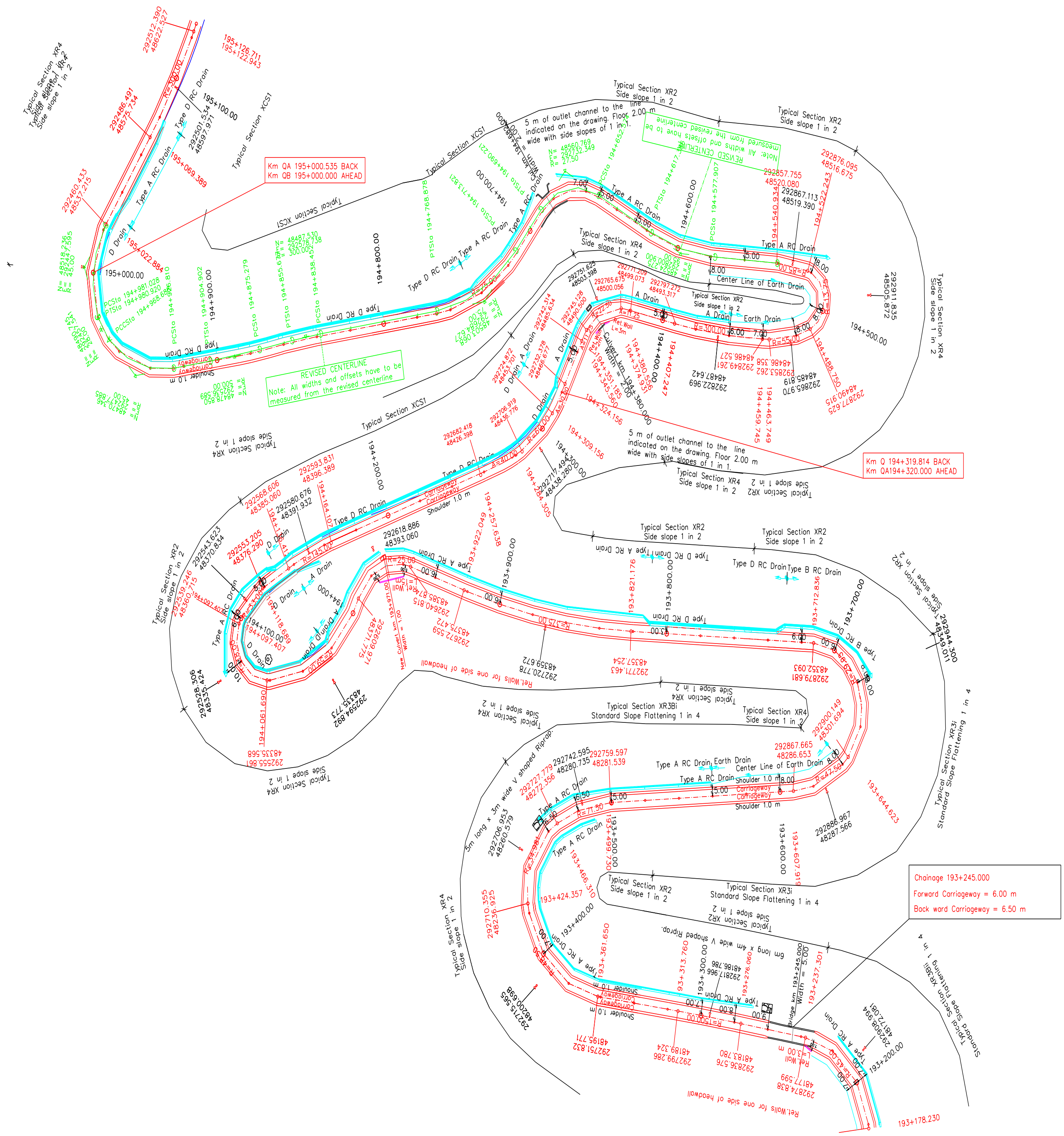


Figure: Shen Debitu curves (around Abelti)

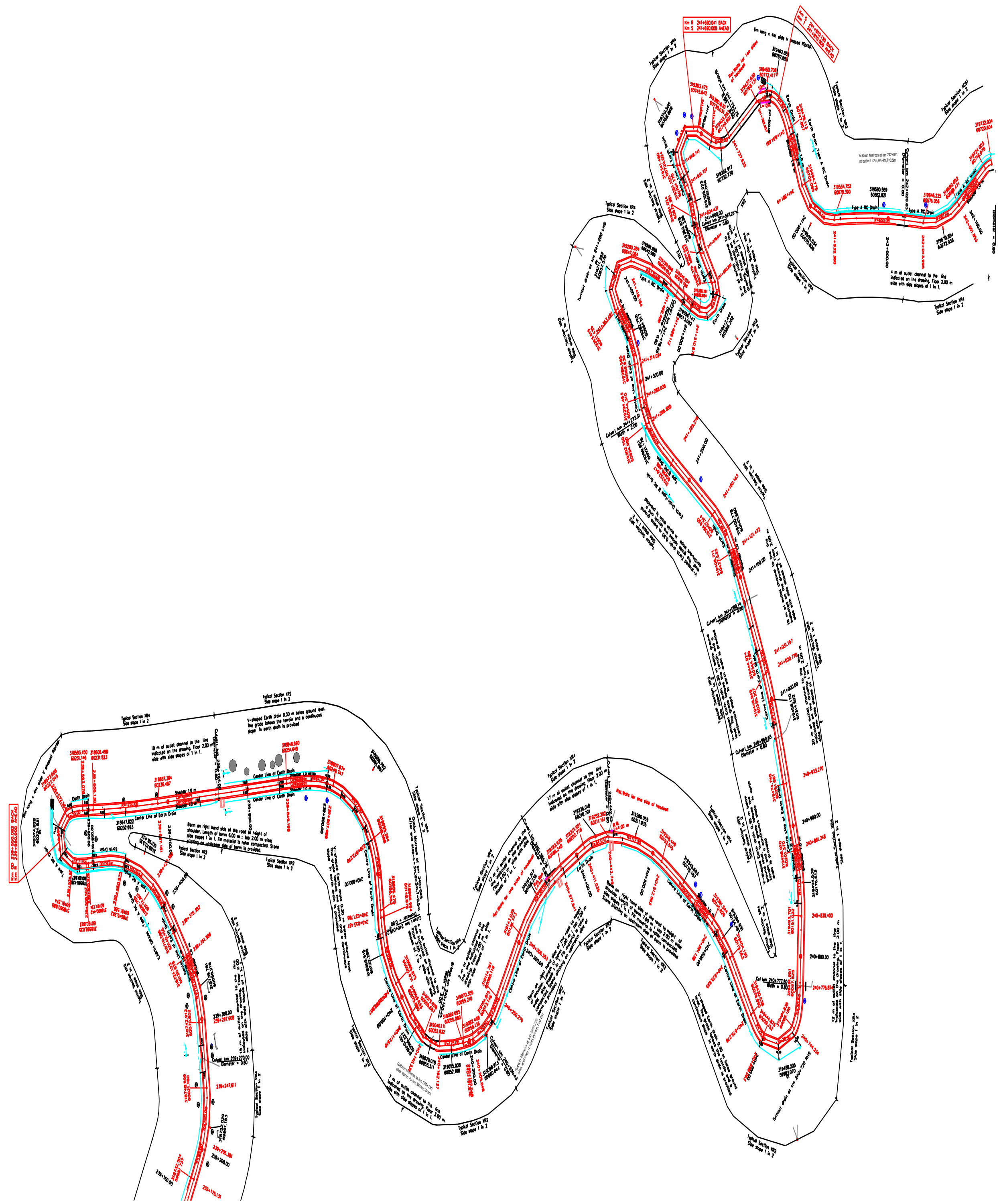


Figure: Curves around Saja to Natri

APPENDIX E: Photos

Figure: curves with insufficient sight distance, steep slope and sharp curves

Figure: field photos for location recording