AN INNOVATIVE MODEL INTEGRATING SPATIAL AND STATISTICAL ANALYSES FOR A COMPREHENSIVE TRAFFIC ACCIDENT STUDY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

İPEK NEŞE ŞENER

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
CIVIL ENGINEERING

JUNE 2004

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Erdal Çokça
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Ayhan İnal
Supervisor

Examining Committee Members

Prof. Dr. Özdemir Akyılmaz	(METU, CE)	
Prof. Dr. Ayhan İnal	(METU, CE)	
Asst. Prof. Dr. Murat Güler	(METU, CE)	
Asst. Prof. Dr. Hikmet Bayırtepe	(Gazi Unv., CE)	
Inst. Dr. Osman Acar	(METU, CE)	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : İpek Neşe Şener

Signature :

ABSTRACT

AN INNOVATIVE MODEL INTEGRATING SPATIAL AND STATISTICAL ANALYSES FOR A COMPREHENSIVE TRAFFIC ACCIDENT STUDY

ŞENER, İpek Neşe

M.S., Department of Civil Engineering

Supervisor: Prof. Dr. Ayhan İnal

June 2004, 165 pages

The negative social and economic results of traffic accidents are the most serious problems within the concept of traffic safety. Every year, unfortunately, a huge number of traffic accidents result in destructive losses. Especially, when the holiness of human life is concerned, traffic safety has an invaluable role for the traffic improvement strategies. In this manner, Turkey places one of the highest ranks regarding the growing rate and severity of traffic accidents that should be immediately taken under control.

In this study, an innovative model that constructs a hybrid between the spatial and statistical analyses is developed in order to examine the importance of enhancing statistical analysis with georeferenced data and so location-based

studies in traffic accident analysis. Meanwhile, the effects of road characteristic and environment are considered for exploring the integral role of roadway factor to the occurrence of accidents, and consequently for emphasizing easily applicable and controllable engineering safety measures.

Because of the rare and random distribution of traffic accident data, logistic regression is used for the statistical part of the study in order to find the pairwise risk factors among the roadway and environmental parameters. After unifying these relative risk factors with the logic of Analytic Hierarchy Process, the finalized accident risk factors are attached to the digitized road characteristics map through Geographic Information Systems (GIS).

The abilities of GIS in mapping, displaying and overlaying different data sets ensure to visualize high risked accident areas with their corresponding potential causal factors. The integration of statistical and spatial analyses is essential for developing appropriate and effective precautions in addition to its easily understandable, applicable and modifiable structure. Finally, the model is proven to be appropriate for both interpreting the existing traffic accident problem or potential future accidents and also developing comprehensive and reliable location-based safety studies.

Keywords: Traffic Safety, Traffic Accident Analysis, Roadway Effect, Logistic Regression, Analytic Hierarchy Process, Accident Risk Factor, Geographic Information Systems

ÖZ

KAPSAMLI BİR TRAFİK KAZA ÇALIŞMASI İÇİN COĞRAFİ VE İSTATİSTİKSEL ANALİZLERİN ENTEGRASYONU İLE OLUŞTURULAN YENİ BİR MODEL

ŞENER, İpek Neşe

Yüksek Lisans., İnşaat Mühendisliği Bölümü

Tez Yöneticisi: Prof. Dr. Ayhan İnal

Haziran 2004, 165 sayfa

Trafik güvenliği kavramı içerisindeki en ciddi problemi trafik kazalarının olumsuz sosyal ve ekonomik sonuçları oluşturmaktadır. Ne yazık ki, her yıl, oldukça fazla sayıdaki trafik kazası ölümcül kayıplara yol açmaktadır. Özellikle, insan yaşamının kutsallığı göz önüne alındığında, yol güvenliği trafiği iyileştirme stratejileri için çok değerli bir role sahiptir. Bu bakımdan, sayısı ve şiddeti giderek artan ve bir an önce kontrol altına alınması gereken trafik kazaları düşünüldüğünde, Türkiye en riskli ülkelerden biri konumundadır.

Bu çalışmada, trafik kaza analizinde istatistiksel analizin coğrafi bilgi ile pekiştirilmesi ve dolayısıyla mekansal çalışmaların öneminin incelenebilmesi için

istatistiksel ve coğrafi analizler arasında entegrasyon oluşturan yeni bir model geliştirilmiştir. Bununla beraber, yol faktörünün trafik kazalarının meydana gelmesindeki etkin rolünün anlaşılabilmesi ve mühendisliğe dayalı, kolayca uygulanabilir ve kontrol edilebilir güvenlik önlemlerinin öneminin vurgulanabilmesi amacıyla, yolun karakteristik ve çevresel etkileri dikkate alınmıştır.

Trafik kaza verisinin seyrek ve rastgele dağılımından dolayı, yola ait ve çevresel parametrelerin göreli risk değerlerinin bulunabilmesi için, çalışmanın istatistiksel kısmında lojistik regresyon analiz modeli kullanılmıştır. Göreli risk değerlerinin Analitik Hiyerarşi Yöntemi mantığı ile önem ağırlıklarının bulunmasından sonra, gerçek risk değerleri Coğrafi Bilgi Sistemleri (CBS) aracılığında, sayısallaştırılmış yol karakteristiği haritasıyla eşleştirilmiştir.

CBS'nin farklı veri setlerini haritalama, görüntüleme ve üst üste koyarak analiz edebilme yeteneği, yüksek riskli kaza alanlarının birebir olası kaza nedenleriyle ilişkilendirilmesini ve görselleştirilmesini sağlar. İstatistiksel ve coğrafi analizlerin entegrasyonuyla ortaya çıkan model, kolay anlaşılan, uygulanabilen ve değiştirilebilen yapısının yanında, probleme uygun ve etkili güvenlik önlemlerinin geliştirilebilmesi için de büyük önem taşır. Sonuç olarak, oluşturulan modelin varolan ya da gelecekte varolabilecek trafik kazalarının anlaşılması ve aynı zamanda kapsamlı ve etkili mekansal güvenlik çalışmalarının geliştirilmesi için uygulanabilir olduğu saptanmıştır.

Anahtar Kelimeler: Trafik Güvenliği, Trafik Kaza Analizi, Yol Etkisi, Lojistik Regresyon, Analitik Hiyerarşi Yöntemi, Kaza Risk Katsayısı, Coğrafi Bilgi Sistemleri

ACKNOWLEDGMENTS

“Fear can hold you prisoner
Hope can set you free**”**

This thesis, first and foremost, is an invaluable proof of a special friendship, which opens my mind to believe in the existence of pure and timeless affection. Although it was quite tiring and wearying especially while closing to its end, I always felt a faithful support that pushed me to go on and to hope. This study has also a different meaning for me because of its role in ending up the first curve of my life. In fact, I am walking towards a new world with all my love and grateful to some precious people. Now, I would like to send my special thanks to the ones who had a definite contribution throughout this study.

I would like to express my special thanksgiving to my family; my dear lovely mother *Aynur*, my beloved father *Erdinç* and my honey little brother *Hakan* for their never-ending love and encouragement that teach me to live and to go after my ambitions; and for being always with me even in my intolerable times. Even though I cannot show my love or appreciation for all time, please be sure that I love you very much and I know that you do everything for my goodness. You are my heart, my soul and my life; you are in my happiness, in my peace.

I would like to special thanks to my supervisor, *Prof. Dr. Ayhan* İnal, for his continuous understanding, patience, support and trust throughout this study in addition to his valuable advices, which guide me to clarify and realize my goals.

I, also, express my gratefulness to *Prof. Dr. Özdemir Akyılmaz, Asst. Prof. Dr. Murat Güler, Asst. Prof. Dr. Hikmet Bayırtepe* and *Dr. Osman Acar*, for their valuable ideas and contributions on this study.

I would also like to thank to General Directorate of Security and General Directorate of Highways of Turkey for their significant technical support on this study.

I would like to particularly show my deep thanks to my dear teacher, *Ali İhsan Ünay*, for his unconditional and sincere friendship, endless guidance and support, unlimited patience and especially for his precious kindness. You know that you are one of my best friends and I am in great debt to you for your beauties to make my life more livable in good spirits.

Finally, sincere thanks to all of *my dear friends* with whom I always found the power to keep alive even in my worst hopelessness times. I am grateful to you for being always on my side, for your forever love, patience and thoughtfulness that encourage me not to give up and to see the light of the life, but "especially for his full-existence during this hard-working study" many special thanks to;

Musa Yılmaz, who continuously believes me more than I do, for being always just beside me at every single stage throughout this study and my life, for his tireless encouragement and endless trust when I lost all my hopes, for his unbelievable patience even in my unbearable or irritating moods that considerably helps me to regain my consciousness and to fight with my fears and surely for his sincere-kindhearted friendship in all happy and worry times. You should definitely know that I can never show enough my appreciation for your effort in reviewing this study, for your brilliant advices and ideas that prevents me to go in a deep darkness, for being my best fried ad infinitum and especially for your beauty that enliven my life, enlighten my world.

To my ***angel*** blood brother

with every single piece of my heart

to the end of time...

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF ABBREVIATIONS

AHP	Analytic Hierarchy Process
CAD	Computer Aided Design
CI	Confidence Interval -in Logistic Regression-
DBMS	Database Management System
EU	European Union
GIS	Geographic Information Systems
GIS-T	Integration of Geographic Information Systems with Transportation
GNP	Gross National Product
MCDA	Multicriteria Decision Analysis
OR	Odds Ratio -in Logistic Regression-
PDO	Property Damage Only Accident

CHAPTER 1

INTRODUCTION

1.1. Seriousness of Traffic Accidents

Every year, several destructive traffic accidents, which threaten human life and property, occur in different regions of the world. Unfortunately, these accidents are the by-products of the modern transportation and they consequently give rise to one of the leading problems being faced by modern societies. The injury rates, human deaths and serious economic losses caused by road accidents demand a continuous and special attention to this problem. Especially, in developing countries, preventing the accelerating rate in the occurrence of traffic accidents require an additional care.

From an analysis carried out for the entire globe, it has been estimated that over 1 million people die and 15-20 million people are injured every single year in road accidents (1). Every year, the total number of people died in Europe's roads is particularly equal to a medium-haul plane crash. In 2000, about 40,000 people died and 1.7 million people injured in road accidents in European Union. The approximate cost of road accidents to society is nearly 160 billions euros (2), but it is surely not possible to state a value for human life cost. Sadly, traffic accidents have higher rates and cause more serious losses in developing countries in spite of their lower level of motorization (3).

When the traffic density at highways is taken into consideration, it will be not surprising to see that the situation in Turkey is not so different. Since the major transportation mode is highway and there is an increasing trend in the number of motor vehicles with respect to inadequate improvement in the standards and low increase in the length of road network, Turkey is among the countries having the highest rate of traffic accidents.

In order to take the necessary precautions against to this critical traffic accident rate, it is very important to define the main factors that lead to traffic accidents and cause congestion in roads. In addition, special attention should be given to the technologies and methods, which help to decrease the number and severity of traffic accidents.

1.2. Role of Geographic Information Systems in Traffic Accident Analysis

Many applications have been suggested and implemented within the transportation field, which mainly cover the management of urban transport infrastructure and the assessment of road and traffic impact on the environment. The integration of Geographic Information Systems (GIS) with transportation (GIS-T) has an essential role for the development of an efficient information system and for managing database information (4). GIS-T, in fact, covers a broad area of disciplines, which refers to the application and adaptation of GIS to research, planning and management in transportation (5).

Especially for developing countries, collecting and organizing transport-related data is a chronic problem. Because of non-georeferenced data collection, maintenance activities and narrow concept of data use and application, it is generally not possible to perform highly accurate analysis or to use the data for more than one purpose. Complexities in developing, updating and processing of transport-related data also necessitate new approaches, by which data are required to be identified, collected, stored, retrieved, managed, analyzed, communicated and presented. In this respect, GIS represents a new paradigm, which uses the concept of location as the basis for structuring of its information system (6).

The main aim of the traffic accident analysis is to investigate the influences of potential causal factors on the occurrence of traffic accidents, to determine hazardous locations and to develop suitable countermeasures in order to decrease the frequency of traffic accidents. From this respect, a good data analysis system with well-defined data retrieval and analytic capabilities should be constructed in order to improve traffic safety.

GIS is often adopted as a powerful mapping tool to facilitate the process of preparing geographic distribution of the accident data and to perform spatial analysis. GIS-T can be directly implemented in order to determine spatial relationships associated with crashes (5). Since GIS enables a true, comprehensive and clear picture of accidents by collecting georeferenced data and by giving an enhanced visualization of traffic accident patterns easily and accurately, it has a considerable importance in roadway safety and traffic accident analysis systems. There have been many GIS-based developments in accident analysis. The power of this technology primarily comes from its ability to link different data sets together spatially, i.e. geographically.

Briefly, GIS-based accident analysis is an interdisciplinary study, which includes identifying historical trends and relationships in accident causation; such as volumes, road conditions, weather, lightening, data comparison and analysis and development of effective precautions by integrating different types of data (7, 8).

1.3. A Significant Causal Factor to the Occurrence of Traffic Accidents; Roadway

Road safety involves three major components: human factor, road system and vehicle element. Although human factors -in Turkey- constitute approximately 97 percent of the traffic accidents, it does not mean that these accidents can be prevented by solely concentrating on the precautions related with human factor. On the other hand, the most effective remedy can lie in a different area. For instance, accidents resulting from high speed may directly be related with roadway characteristics instead of driver. Although speed should be decreased in the case of low visibility, high volume or poor surface conditions, it will mostly not create any problem under normal traffic and environmental conditions on the improved roads.

Even though the traffic accident statistics published in Turkey show that traffic accidents resulting from inadequacies in road characteristics have a very low ratio compared with the human-caused accidents, in reality, roadway factor has much more serious impact and a high ratio on the occurrence of traffic accidents. As stated in a forum “Traffic Safety in Highways” (9), when the realistic accident analyses are performed, it can be observed that up to 50 percent of all accidents are caused by poor roadway infrastructure, construction and operation deficiencies in roads and insufficient traffic control.

Since human factor requires long term social and educational remedies and needs strictly applied enforcement laws, it generally cannot be controlled in an easy way. Alternatively, it is possible to easily remove the problems related with roadway characteristics in short run. Road improvements can play a significant role in reduction of traffic accidents especially where the road user fails to cope with the road environment. Even, the increasing effect of human factor on traffic accidents can be decreased particularly regarding unconsciousness and inexperienced drivers and variable uncontrolled parameters.

1.4. Purpose and Objectives

The increasing rate of traffic accidents is one of the most severe problems of the technology world of today. As it was emphasized before, it is very important to decrease the number and severity of these accidents in order to increase the efficiency and safety of the road traffic. In this manner, there exist many researches undertaken by different researchers, associations and countries.

In our country, as in the whole world, traffic accidents have been mostly studied by statistical methods. However, it is not possible to find the most effective solution by using only the general statistical distributions; i.e. without specializing the problematic location and the dominating cause of that accident location. Identification of risky locations has a considerably essential role for undertaking of remedial measures because of its high economic and social returns.

Although human factor has a quite high percentage in the occurrence of accidents -as it seems to be the most critical factor-, it should be a serious mistake to neglect the effects of other factors, i.e. road and vehicle factors. From the transportation engineering point of view, traffic accidents occurring due to poor road characteristics or road environment constitute a high risk, which should urgently be taken under control. The important point, requiring an extra attention, is the ability of engineers or planners to change and control these kinds of defects immediately and successfully. Therefore, studies investigating the methods used for the minimization of road effects deserve much more attention since they directly contribute to an increase-improvement in traffic safety.

From these respects, integration of statistical and spatial analyses has an invaluable role within the consideration of traffic accident studies. While statistical analysis shows the numeric distribution of accidents and explores related risk factors according to their causes, spatial analysis supports this numeric distribution with a location-based implementation, improves the visualization and interpretation and permits any modification in the system for future researches, which increase the efficiency of the proposed system.

In this study, a model covering a well-developed traffic accident analysis is presented. The primary purpose of this thesis is to develop an innovative methodology in order to better understand the traffic safety problem and potential causal factors of traffic accidents. This will be also helpful for finding suitable countermeasures to the risky areas in an accurate and quick manner. For this reason, the importance of identifying risk factors related with the road characteristics and environment are taken as the logic behind this study. After determining the risk factors based on the traffic accident data of the year 2003 in Turkey, an algorithm is constructed to identify the associating risky zones or areas with respect to the roadway characteristics and environment. This model is applied to the state highways and motorways of the Ankara County as a case study.

Transportation systems continuously interact with the world they serve. When the information gathered from multiple sources are integrated through Geographic Information Systems (GIS), a better understanding and deeper insight can be provided for the transportation related issues. Consequently, a model including both statistical and spatial analyses is developed in order to effectively investigate the potential causes of traffic accidents.

In this study, a brief introduction is given in the first chapter, the importance of road safety, traffic accident problem and accident analysis are discussed in the second chapter and in the third chapter, traffic safety problem is clarified for Turkey. The basic concepts of GIS and its importance in transportation and traffic engineering are covered within the body of the fourth chapter.

For the main part of the study, traffic accident data of the state highways and motorways for the year 2003, which are taken from the Ministry of Interior-General Directorate of Security of Turkey, are used. In the fifth chapter, these traffic data are analyzed from a statistical point of view using logistic regression. Logistic regression, actually, is one of the most suitable regression analyses methods since it assumes random and rare data distribution rather than homogeneity. With the help of Multicriteria Decision Analysis (MCDA) methods (through Analytic Hierarchy Process - AHP), the obtained relative risk values from the logistic regression are converted to a form, from which direct effects of the selected parameters to the accidents are calculated. After finding the final form of risk values (weights) with respect to the causes, each parameter that have a contributing effect on the occurrence of accidents, is attached to different weights accordingly, so different severity.

Finally, spatial analyses of the traffic accidents are performed by using GIS in the sixth chapter. Initially, the data taken from the General Directorate of Highways are digitized and a GIS layer -covering roadways and their attributes- is created. Then, several queries are carried out in order to identify the spatial pattern of traffic accidents according to the characteristics of roadway and road environment. Subsequently, an algorithm that determines the risk zones with respect to the attribute data is created by integrating the risk factors/weights obtained in the fifth chapter into GIS. Afterwards, several maps are produced showing the effects of constant road parameters with the combination of environmental parameters and also the most risky zones with respect to those parameters. As a last thing, the attribute data of the actual traffic accidents are compared with the obtained accident risk maps for the studied year.

Conclusively, the model incorporating too different disciplines; i.e. Statistics, Multicriteria Decision Analyses and Geographic Information Systems, is discussed for its efficiency, suitability and ability in order to find effective solutions in the prevention and reduction of traffic accidents.

CHAPTER 2

GENERAL CONCEPTS IN TRAFFIC SAFETY ANALYSIS

2.1. Road Safety

Transportation, which is the movement of people and goods, is one of the basic human needs and a key factor in modern economies. In order to satisfy the increasing demand for mobility, various transportation systems have been developed. Different aspects of these systems have been studied within the concept of transportation engineering in order to plan, construct and manage them safely and effectively (10).

Safety and efficiency are the two primary goals of traffic and transportation engineering. When the high traffic density concentrated on roads is considered, road safety especially requires a considerable research that incorporates traffic safety knowledge into the road planning and the design process with an aim of preventing traffic accidents. Unfortunately, the increasing rate and severity of accidents on roads, resulting in severe social and economic losses, significantly indicate the necessity for the development of efficient road safety studies. From the traffic engineering point of view, for instance, it is essential to identify the factors that influence accident frequency and severity in order to improve highway design and also to provide a safer driving environment (11, 12, 13).

Between different modes of transportation, there occurs a significant imbalance, which results in congestion on dominant routes; i.e. on highways. Figure 2.1 and Figure 2.2 show this growing imbalance in passenger and goods transport in the European Union. In reality, most frequently used transportation mode through the world is highways. As an example, in Europe, road haulage increased by 19.4 % between 1990 and 1998; on the other hand, rail haulage decreased by 43.5 % during the same period. Most of the passenger and goods traffic flow through roads. In 1998, 44 % of the goods traffic and 79 % of the passenger traffic were accounted through road transport (2).

Since it is generally seemed as a symbol of personal freedom in modern society, in European countries, nearly two households in three own a car. According to the studies performed by European Union, if effective countermeasures cannot be implemented by 2010, even heavy goods vehicle traffic alone will increase by nearly 50 % over its 1998 level. This, unfortunately, will result in extremely congested traffic on European roads (2).

The greatest strength of road transportation comes from its capacity to carry goods and passenger with unequalled flexibility and at lower price than other transportation roads. Although it is irreplaceable, road transport sector has problems even regarding its economic position. Due to its considerable fragmentation and the pressure exerted on prices by consignors and industry, margins of road transport are narrow. This fact, for instance, mostly forces some road haulage companies to reduce the prices and to neglect the social and safety legislation in order to make up this concluding handicap (2).

Among all modes of transportation, road transport is the most dangerous and most costly one in terms of human lives. Traffic accidents have become a serious problem that threatens human life and property at an accelerating rate. Traffic safety problems, particularly, have been rapidly growing and they constitute a major health problem especially in developing countries.

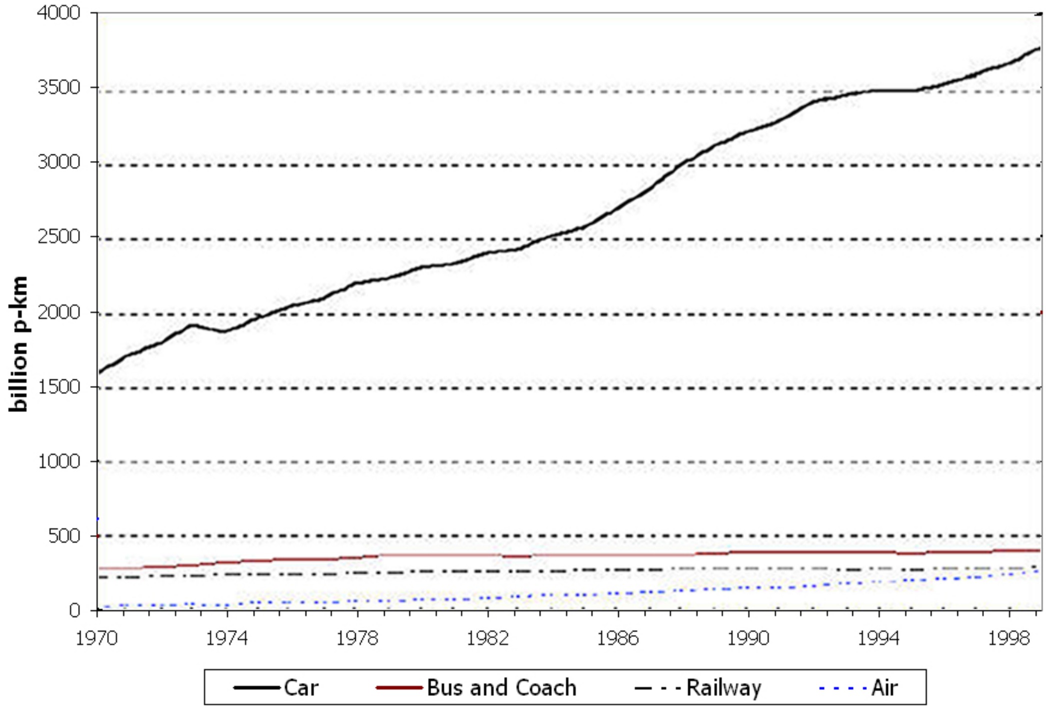


Figure 2.1. Passenger transport in EU between the years 1970 and 1999 (2)

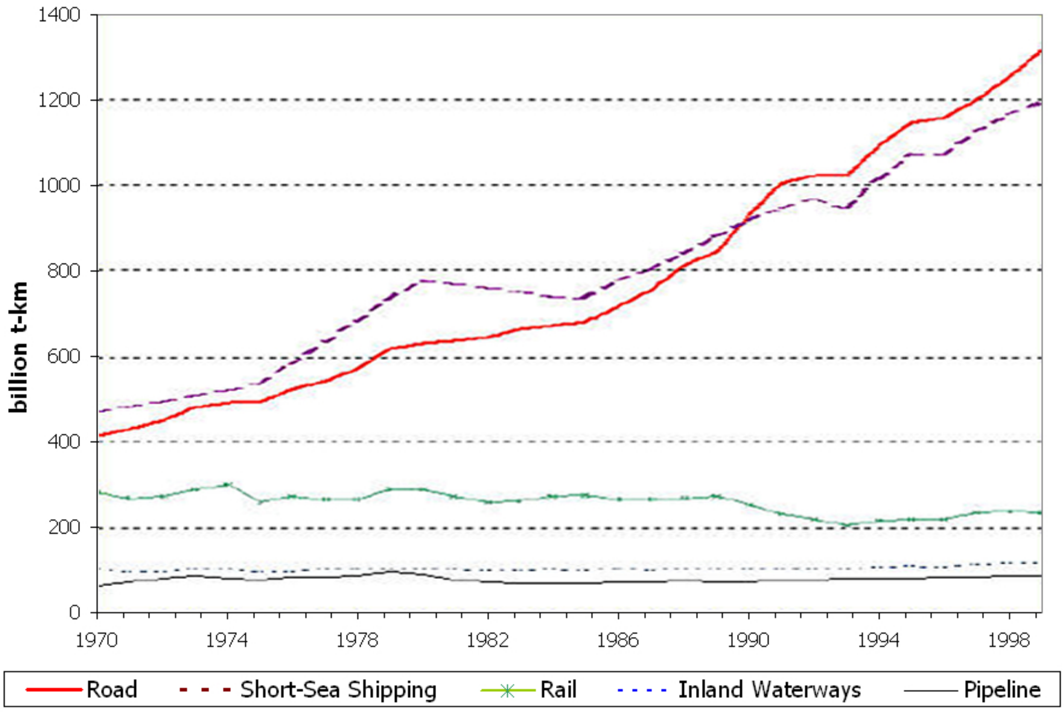


Figure 2.2. Goods transport in EU between the years 1970 and 1999 (2)

According to the World Bank (14), every year more than 1.17 million people die in road crashes around the world while the majority of these deaths, about 70 percent, occur in developing countries. The same statistics also show that over 10 million are crippled or injured each year. Moreover, it has been estimated that at least 6 million more will die and 60 million will be injured during the next 10 years in developing countries unless urgent actions and precautions are taken. The economic cost of road accidents is approximately 1 to 3 percent of a country's annual Gross National Product (GNP). Undoubtedly, these losses seriously inhibit the economic and social development of developing countries.

Economic component of road accidents in the form of economic losses due to injured and death, and property damage, social losses and high numbers of injuries and death are the destructive impacts of traffic accidents, which can be overcome by determining the problem locations and by developing corresponding safety improvement programs (13). Although it is not possible to prevent traffic accidents completely, their number and severity can be reduced since they truly do not happen accidentally; but then they are the results of caused occurrences, which can be discovered and controlled by effective safety measures. Drunken driving, speeding, narrow streets, dangerous curves on highways, lack of needed traffic signs and signals are some of the examples resulting in serious accidents. These examples, actually, reveal the fact that a good hybrid of engineering, enforcement and education will particularly prevent the occurrence of traffic accidents (15).

2.2. Basic Principles of Traffic Accident Analysis

Traffic accidents occur in different places at varying times with different participants and under numerous circumstances. However, some specific patterns in the overall accident picture can be determined from the assessment of extensive accident statistics in a general sense (16).

In addition to defining specific trends, the problem areas can be identified and a first global impression about the road safety can be obtained by the analyses including the following circumstances (16):

- problem groups**;** specific homogeneous population groups of road users and means of transport, such as the elderly, motor cycle riders, etc.
- problem situations**;** specific homogeneous traffic situations, for a certain problem group, such as the ease of crossing roads for the elderly, wet road surfaces for motor cycle riders, etc.
- problem sites**;** individual or specific homogeneous groups of sites, indirectly linked to specific problem groups, such as black spots, trunk roads through small centers of population, etc.

2.2.1. Causes of Traffic Accidents

Understanding causes and characteristics of accidents is the most important stage in order to design proper countermeasures. Traffic accidents are mostly caused from the combination of human errors and failings, poor road design or adverse weather conditions and vehicle defects (17).

Traffic accident statistics generally indicate that the most dominating factor in accidents is human factor. According to a research made in USA, it is put forward that human factors solely take place approximately 57 percent of all accidents while the road and vehicle factors, alone, constitute 4 percent and 2 percent, respectively. However, when the role of the interaction between these contributing factors are investigated, it is seen that accidents caused by road faults reach a value of 35 %, and 12 % by vehicle faults (Figure 2.3) (18).

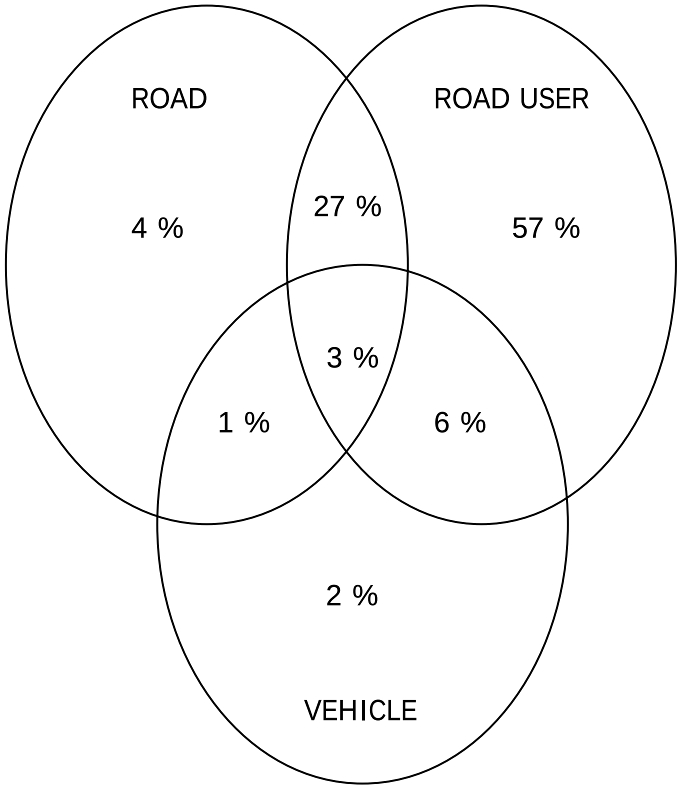


Figure 2.3. The interactive distribution of human, road and vehicle factors causing traffic accidents (18)

In fact, traffic accidents are complex phenomena with varying multiple causes. The construction of traffic accident taxonomy is significant in order to describe the accidents or define the categories of their causes. Completing any taxonomy of traffic accidents is actually not possible since new categories might be continuously added. The decision of including or excluding a variable mainly depends on four basic criteria:

- the importance of the variable in the occurrence of accidents
- the availability of the required data for coding the related variable

- the usefulness of the category from accident analysis and prevention point of view
- the borders of the study (for instance, if too many categories are used, too small samples will be obtained) (19)

2.2.2. Accident Prevention and Reduction

Accident prevention concerns the application of safety principles to road improvements or traffic management schemes that are initiated to satisfy traffic or environmental demands, and so it is not only justified on the basis of accident savings. This complex work usually involves carrying out of safety checks on designs in order to prevent any problematic feature and also to identify any safety measures, which would be added to lower their accident potential. For instance, if an intersection is required to be redesigned to improve the traffic flow, in addition to applying the best safety principles of layout, accident prevention checks should be carried out for preventing the existence of any hazards due to a change in carriage surface, new road markings or additional new road furniture. In summary, all implementations should be developed for enhancing safety (20).

The process for accident prevention, investigation and reduction as shown in Figure 2.4 (21) includes four major planning steps (12, 20):

1. *Data collection, storage and retrieval:* A reliable database should be constructed for identifying the problems faced and objectively assessing the measures that have been implemented to treat.
2. *Identification of hazardous locations for further study:* Identification phase has a great importance to draw the analyst's attention on network

locations that exhibit an abnormal concentration of accidents. This phase also ensures to enhance the efficiency of safety treatments.

3. *Diagnosis of accident problem(s):* High hazard locations of accidents are investigated by conducting engineering studies in order to propose main causes of them. Accident data and location characteristics are the two main types of information that should be carefully concerned for hazardous location analysis.

4. *Final selection of sites to be included in remedial implementation program and assessment:* After identifying a set of locations for analysis and diagnosis, a priority setting process should be launched by considering a large set of factors: economic, environmental, political and social. Multicriteria decision analysis methods are particularly used to evaluate the quantitative and qualitative criteria and to check the potential impacts of some variables. After all, the value of the activity, which is measured by its success or failure for successfully achieving the predetermined set of goals or objectives, is assessed.

Improving traffic safety and so preventing traffic accidents depend on three basic elements that also influence traffic operations: road user, vehicle and roadway. However, traffic engineer has an effective and direct control over only one of these elements: the roadway. This does not mean that the other two elements are strictly uncontrollable; in fact, traffic engineer can also play an important role to ensure an improvement in driver education and licensing procedures in enforcement and for the required incorporation of safety features in vehicle design. The main point is that these are usually subject to the political and legislation processes and so they are not under the direct control of the engineer (22).

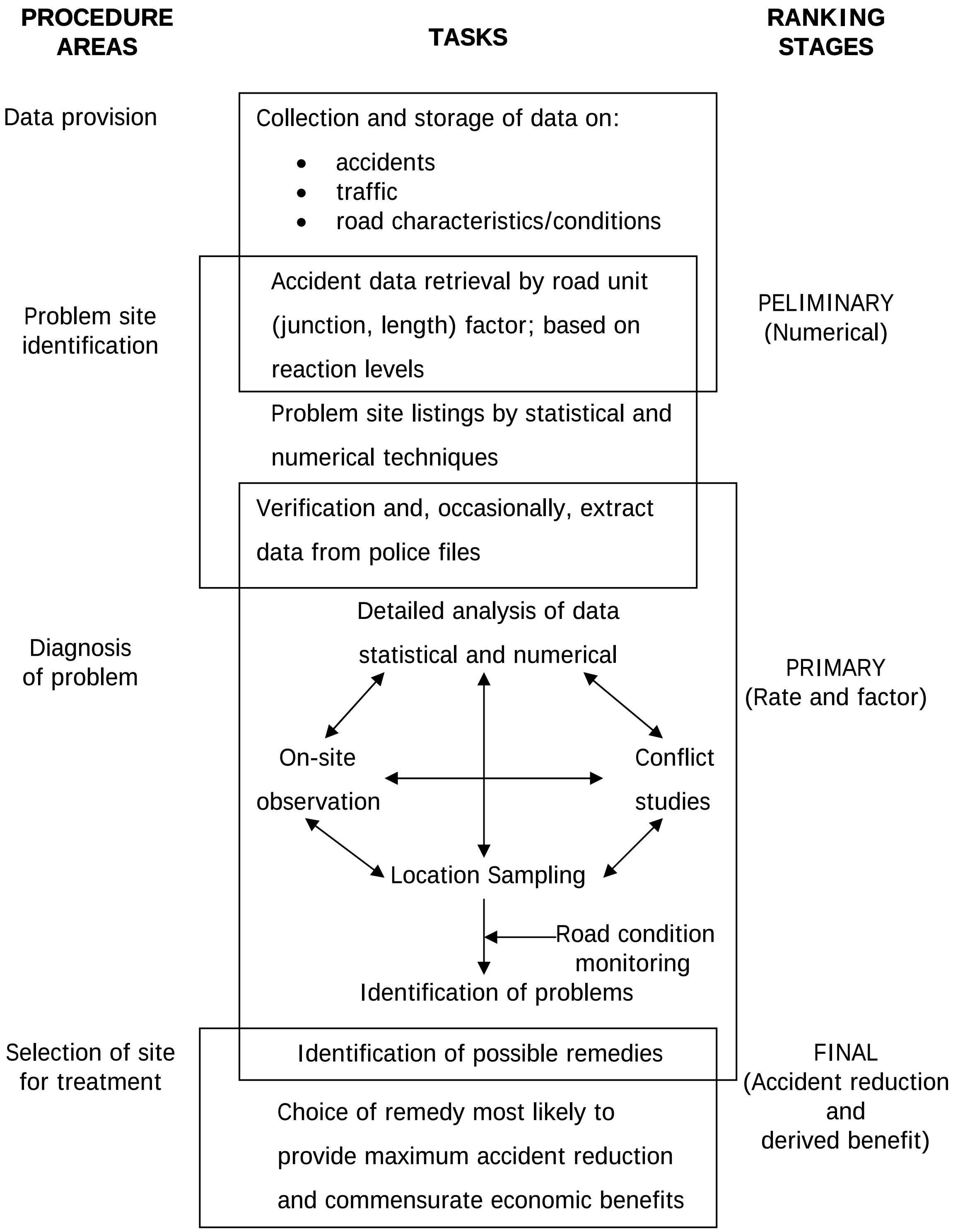


Figure 2.4. Outline of the process in accident prevention, investigation and reduction (21)

The road factor especially becomes to gain significance when it is incorporated with the road user factor. Indeed, identification of the contributory factors in accidents is not same as identification of the efficient measures. While one of the contributory factors is defined to be related with road user, the measures taken against it might be an improvement to the infrastructure. It is also possible to reduce the complexity of the road and traffic system and possible misinterpretation and misuse of it by road users with the improvement on infrastructure (16).

Improvements in road design should be placed on taking driving behavior better into account, incorporating the road into its general environment and enabling the route to be adopted to its function and to driver perception. This, principally, leads to the safest possible solution with the predetermined socio-economic standards. The aim should be to minimize the risk of accident and their consequences by providing the driver with a clear, unambiguous perception, which conforms to road configuration and operation (16).

In this manner, -as an example case- United Kingdom has recommended relaxing some standards (such as lowering design speeds) and has developed some new design concepts in order to reduce costs and environmental impacts without adversely affecting safety. This proposal aims to provide (16):

- a greater flexibility in the application of standards
- the design of roads with more straight sections but with lower radii (horizontal & vertical curves) between the straights resulting in better visibility when overtaking on straights and, with appropriate road markings, no overtaking on curves

2.3. Availability of Traffic Accident Data

From planning and design perspectives, accident data has a great importance for identifying problem areas or locations of interest (23). Accurate analysis of traffic accidents primarily depends on the knowledge of characteristics of drivers, vehicles and roadways; their interrelationships, and thus, a uniform and accurate accident reporting (24).

In general, the validity of the database refers to how well it reflects reality as a whole; i.e. the actual number of accidents and their features. The reliability of the study is strictly related with the correctness of the available data and the accuracy of the process of transferring these data into the database (19).

2.3.1. Traffic Accident Reporting

Precisely filled police accident reports have an invaluable importance for collecting, storing and retrieving traffic accident information in a usable and efficient format. Complete and correct reporting of accidents are vital for a properly integrated accident prevention program; i.e. for effectively developing related countermeasures by utilizing the principles of engineering, education and enforcement (22, 24, 25).

The traffic accident report form should be filled out carefully through investigating the accident, questioning participants and witnesses, making physical measurements, examining for evidence of intoxication and reconstructing the accident successfully. Detailed information on *time of accident, location of accident, drivers, vehicles, persons injured, extent of damage of vehicles, location and description of traffic control devices, regulations in force, roadway and weather conditions, diagram of accident, possible violations and probable causes* should be included in a systematic manner. Uniformity of accident recording system should also be achieved in order to attain uniform accident statistics (24).

The information collected from these forms is needed for a variety of purposes, which cover (22):

- identification of locations at which unusually high number of accidents occur
- detailed functional evaluation of high-accident locations to determine contributing causes of accidents
- development of general statistical measures of various accident-related factors to give insight into general trends, common causal factors, driver profiles and other factors
- development of procedures that would identify hazardous locations before large number of accidents occur

Unfortunately, traffic accident reports vary in respect of completeness and clarity of their descriptions of accidents. Different handwriting, spelling and grammar due to manually entered data by police officers at the accident scene, misunderstandings in location, time, cause and involved people cause these problems. Incompleteness of data is also a common problem in most of the accident analysis studies. From majority of researches, it is understood that severe accidents, involving casualties and injuries, cover a much greater extent than expected (19). For instance, in the case of guardrail accidents, drivers involved in minor crashes may leave the accident scene instead of reporting it, mainly for not facing with charges or costs of damage to public property (26).

Ideally, all accidents regardless of the seriousness of the injury or the cost of the property damage should be reported in order to make the information more useful. Detailed accident reporting would allow researchers to develop effective safety programs, which are based on more comprehensive picture of traffic accidents (25). However, it is estimated that only 50 % of all accidents are recorded by police officers (22).

In addition to inaccuracy and incompleteness, police accident records might show strong variations between districts. Because of the effects of social desirability and forgetting, self reporting of accidents is also not sufficiently valid. As it is pictured clearly, reporting system is not perfectly reliable since all accidents are not always reported or those reported are sometimes incomplete or erroneous, which decrease the quality and validity of the traffic accident data (19, 26).

As a definite necessity, accident data should be recorded continuously, completely and accurately. Then, these data should be analyzed effectively in order to measure and predict unsafety and so to reach more reliable, realistic and acceptable solutions (27).

2.3.2. Time Period of Accident Analysis

Selecting the time period of the study is one of the most important stages of traffic accident analysis. From a purely statistical point of view, using as many accidents as possible is the most favorable option; however, one should carefully consider the validity of the data for an accurate analysis, which also directly affects the decision on the range of the traffic accident data used for the study. Apart from the difficulties of storing many years of accident data, there should not be any changes at the spot, not in traffic flows or behavior and not in geometry or surface, etc (9).

It is usually regarded that three years is a reasonable period for a reliable analysis. Generally, results of three years of records present a suitable balance between having a long period, which means of analyzing many accidents and having a short period so that the conditions are unlikely to have changed a great deal. Moreover, a time period of three years eliminates many problems related with discarded data. The limitation size of the time period within the area of study depends on the frequency and size of the change. It is also common to use time frames up to 4 years, but for periods longer than 4 years the analysis should be handled with a special care in order to ensure the changes in the study area to be within the tolerated limits with respect to the scope of the study (9, 15).

2.4. Road Safety Statistics

Road safety is generally measured in terms of number of accidents, injuries and fatalities, either in absolute numbers or in relative figures; accidents per 10,000 vehicles, per kilometer of road or per kilometer traveled, etc. In order to compare road safety of different countries, it is needed to relate the number of accidents to an equally comparable and available parameter for each country; the number of accidents per kilometer traveled (the risk of being exposed to a traffic accident) is an example for this parameter (16).

Accident statistics are estimates or measures based on the number and severity of accidents. It is essential to present these statistics in a systematic and clear way for providing an insight into the general state of road safety and contributing potential causal factors of accident occurrence and for developing effective safety programs, policies and specific site improvements (22).

Accident statistics generally address and describe one of the three principal informational elements (22):

- *Accident involvement*; relates to the numbers and types of vehicles and drivers involved in accidents.
- *Accident occurrence*; concerns the numbers and types of accidents that occur, often described in terms of rates.
- *Accident Severity*; numbers of fatalities and fatality rates are often used to measure the seriousness of accidents.

Within the margins of these elements, the hazardous locations (black spots) are identified from historical accident data by using six most common approaches listed below (13, 28):

- *Accident Frequency* method is used to identify and rank locations on the basis of the number of accidents. A critical threshold can be determined for the selection of hazardous locations. This method is adaptable for highway systems of 2500 miles or less.
- *Accident Rate* method combines the accident frequency with the volume of traffic and defines the dangerous location as the one having a number of accidents more than a predetermined number. This method is good for small scale low volume networks.
- *Frequency Rate* method combines the accident rate and accident frequency methods and a location is defined as critical if both of them are satisfied.
- *Accident Severity* method defines a problem location as a section of the network having a number of accidents with a certain amount of fatalities, injuries, property damages or combination of them.

- *Frequency-Quality Control* method requires a statistical test for significance that the accident frequency determined for the location is higher than the average value of similar locations.

- *Rate-Quality Control* method uses a statistical test to determine whether the accident rate at a particular location is significantly higher than a predetermined average rate for locations of similar characteristics or not; the method is based on Poisson distribution.

The general procedures for these methods, actually, do not significantly differ from each other. The main aim, as stated before, is to identify the dangerous section and/or point using historical accident data.

CHAPTER 3

THE CHAOTIC TRAFFIC ACCIDENT PROBLEM

3.1. Traffic Safety in Turkey

Although transportation is inevitably essential for the well being of society and for each individual, it is increasingly becoming dangerous. Traffic safety problem is so serious that road traffic accidents are one of the leading cause of deaths in Turkey. According to the annual abstract statistics made by Ministry of Interior-General Directorate of Security of Turkey in 2004, about 3,082 people were killed in road crashes; 109,681 were injured, and the cost of road crashes was about US $ 653 million to Turkish economy, which is approximately 2.5 % of GNP. These numbers, indeed, cover only the above water part of the ice berg including the ones who die at the time of the accident, i.e. only those who die at the accident scene. Also, erroneous structure of recording system causes insufficient representation of the definite reality. In fact, those who die on the way to hospitals or at hospitals are not included within the accident fatality.

As mentioned above, traffic accidents are a foremost cause of death in Turkey. From the global perspective, Turkey places one of the highest ranks in mortality and accident rate from traffic accidents; about 5 death and 750 accidents per 100,000 populations in 2004. In the last years, as in most of the developed and developing countries (Tables 3.1 and 3.2 (29)), traffic on Turkish road network has rapidly increased, which also increases the risk of vehicle collision.

Table 3.1. A general statistical view about remarkable traffic numbers for the year 2002 for different countries (29)

COUNTRIES	Population	Vehicle	Accident (dead and injured)	Death	Injury	Per 100,000 Vehicles		Per 100,000 People	
						Death	Injury	Death	Injury
GERMANY (2)	82,260,000	53,655,835	362,017	6,828	476,420	13	888	8	579
AUSTRIA (2)	8,139,310	5,419,073	43,175	956	56,684	18	1,046	12	696
BULGARIA (4)	7,845,499	3,070,614	6,769	959	8,100	31	264	12	103
CZECH REPUBLIC(4)	10,270,000	4,397,000	26,586	1,431	34,389	33	782	14	335
FINLAND (2)	5,206,295	2,539,953	6,196	415	8,156	16	321	8	157
HOLLAND (1)	16,105,000	8,168,398	33,538	987	40,682	12	498	6	253
SPAIN (2)	40,683,037	25,065,732	98,433	5,347	146,917	21	586	13	361
SWEDEN (2)	8,909,128	4,611,000	16,925	535	24,750	12	537	6	278
ICELAND (2)	288,471	183,698	1,496	29	2,286	16	1,244	10	792
JAPAN (2)	127,435,000	90,106,830	936,721	9,575	1,166,606	11	1,295	8	915
KOREA (3)	47,639,618	15,656,897	230,953	7,090	348,184	45	2,224	15	731
LATONIA (1)	2,328,800	814,470	5,083	518	6,300	64	774	22	271
HUNGARY (3)	10,175,000	3,141,073	19,686	1,429	25,978	45	827	14	255
NORWAY (2)	4,552,200	2,590,078	8,790	310	11,522	12	445	7	253
TURKEY (4)	69,757,000 (5)	9,500,235	65,850	4,146	116,045	44	1,221	6	167
NEW ZELAND (2)	3,939,100	2,709,526	10,103	404	13,826	15	510	10	351

For the Table 3.1, the people who died on road and in medical establishments are added to the total number of deaths; but, the period of prosecution may vary within countries:

1) 7-day-period

2) 30-day-period

3) indefinite period

4) not performed on account of law

5) 2002 population data refer to the numbers in the "Annual Turkey Statistics Data" prepared by Turkish State Statistic Institute

Table 3.2. The rates of injury and mortality per 1 million/1 billion vehicle-km in some countries for the year 2002 (29)

** "a" represents the values of 2001

Country	Injury per 1 million vehicle-km	Mortality per 1 billion vehicle-km
Austria	0.55	12.3
Canada	0.51	9.3
Czech Republic	0.62	33.1
Denmark	0.15 [a]	9.2 [a]
Finland	0.13	8.5
France	0.19	13.8
Germany	0.59	11.1
Ireland	0.18 [a]	10.9 [a]
Japan	1.18	12.1
Netherlands	0.26	7.6
Norway	0.25 [a]	8.3 [a]
Slovenia	0.83	21.7
Switzerland	0.39	8.4
Turkey	1.05 [a]	73.0 [a]
USA	0.46 [a]	9.4 [a]

Turkey has experienced high traffic rates because of the dominant usage of road transportation (Table 3.3 and Figure 3.1) and sharply increasing number of motorized vehicles. In spite of accelerating increase in population and the number of vehicles, lesser increase in the length and improved standards of highways can be stated as the primary causes for the congestion in roads and increased traffic accidents (Table 3.4, Figures 3.2 to 3.4). Furthermore, the rapid and uncontrolled urban growth have promoted to the occurrence of traffic accidents.

Table 3.3 and Figure 3.1 are arranged according to the data taken from State Institute of Statistics of Turkey while Table 3.4 and Figures 3.2, 3.3 and 3.4 are all arranged according to the numbers taken from the annual abstract statistics published by the Ministry of Interior-General Directorate of Security of Turkey.

Table 3.3. The proportion of the usage of transportation modes for interstate passenger and freight transportation in Turkey for the year 2003

Actual Values	***Passenger (passenger-km)***	***Freight (ton-km)***
Highway	164,311	125,163
Railway	5,878	8,669
Waterway	22	5,400
Airway	2,752	276
Percantages	***Passenger (%)***	***Freight (%)***
Highway	95.00	91.38
Railway	3.40	5.21
Waterway	0.01	3.24
Airway	1.59	0.17

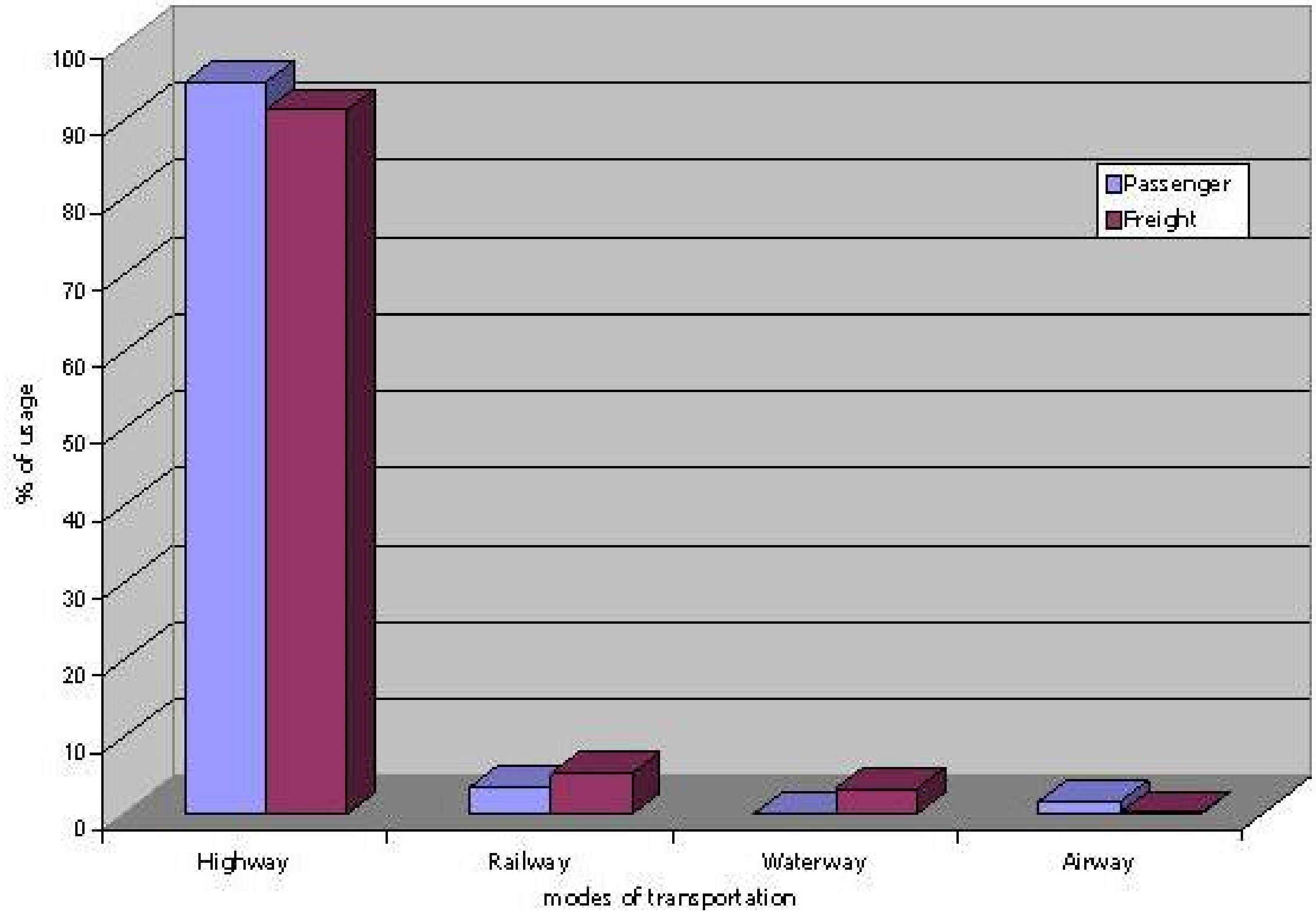


Figure 3.1. Distribution of transportation modes in Turkey in 2003

Table 3.4. A general statistical view about remarkable traffic numbers between the years 1980 and 2004 in Turkey

	Road Length (km)		(Number of)					(Per 100,000,000 veh-km)		
Year	State Highway	Motorway	Vehicle	Driver	Accident	Death	Injury	Accident	Death	Injury
1980	31,976	-	1,684,019	2,619,554	36,914	4,199	24,608	68.552	7.798	45.699
1981	31888	-	1,785,758	2,764,635	40,953	4,441	29,744	71.918	7.799	52.234
1982	31,953	-	1,887,878	2,897,771	46,249	4,884	35,976	76.668	8.096	59.639
1983	31,210	-	2,018,608	3,031,200	55,208	5,201	44,769	87.631	8.255	71.061
1984	30,982	-	2,186,515	3,198,492	60,840	5,731	50,521	89.811	8.460	74.578
1985	30,997	-	2,375,141	3,474,643	65,831	5,680	51,586	89.417	7.715	70.068
1986	30,986	-	2,653,715	3,835,279	92,625	7,315	71,264	112.644	8.896	86.666
1987	31,062	-	2,987,215	4,502,055	110,207	7,530	80,321	118.772	8.115	86.563
1988	30,999	-	3,313,005	4,919,156	107,651	6,846	79,174	104.821	6.666	77.093
1989	31,048	-	3,655,090	5,519,101	103,758	6,332	80,013	91.430	5.580	70.506
1990	31,149	-	4,091,888	6,235,196	115,295	6,286	87,693	90.457	4.932	68.801
1991	31,261	-	4,487,259	6,778,291	142,145	6,231	90,520	101.332	4.442	64.530
1992	31,343	862	5,055,968	7,465,559	171,741	6,214	94,824	105.474	3.816	58.236
1993	31,424	1,070	5,799,718	8,162,959	208,823	6,457	104,330	110.807	3.426	55.360
1994	31,389	1,167	6,228,016	8,794,743	233,803	5,942	104,717	115.311	2.931	51.646
1995	31,422	1,246	6,635,938	9,388,630	279,663	6,004	114,319	129.006	2.770	52.734
1996	31,412	1,405	7,109,926	10,242,628	344,641	5,428	104,599	147.708	2.326	44.829
1997	31,320	1,528	7,776,394	11,297,235	387,533	5,181	106,146	151.713	2.028	41.554
1998	31,345	1,726	8,359,636	12,277,101	440,149	4,935	114,552	159.208	1.785	41.435
1999	31,388	1,749	8,837,403	13,151,950	438,338	4,596	109,899	149.683	1.569	37.528
2000	31,397	1,773	9,554,868	14,109,116	466,385	3,941	155,877	147.155	1.243	49.183
2001	31,376	1,851	9,821,084	14,767,694	409,407	2,954	94,497	125.460	0.905	28.958
2002	31,319	1,851	9,430,147	15,183,033	407,103	2,900	94,225	130.149	0.927	30.123
2003	31,358	1,882	8,903,843	15,488,494	422,302	2,818	95,324	142.687	0.952	32.208
2004	-	-	10,236,358	16,151,623	494,851	3,082	109,681	-	-	-

Figure 3.2. Distribution of accidents, deaths and injuries per 10,000 vehicles between the years 1980 and 2004

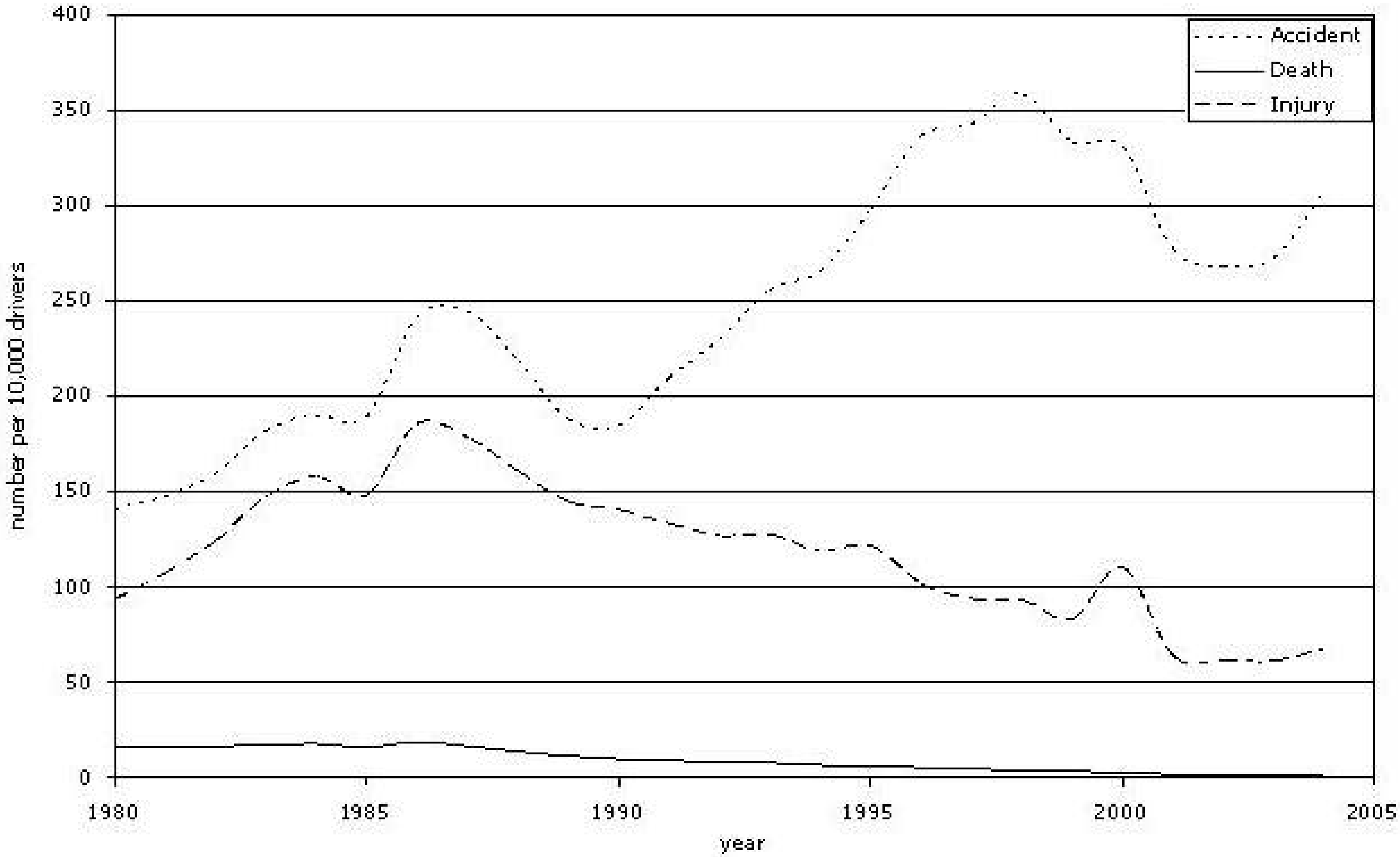


Figure 3.3. Distribution of accidents, deaths and injuries per 10,000 drivers between the years 1980 and 2004

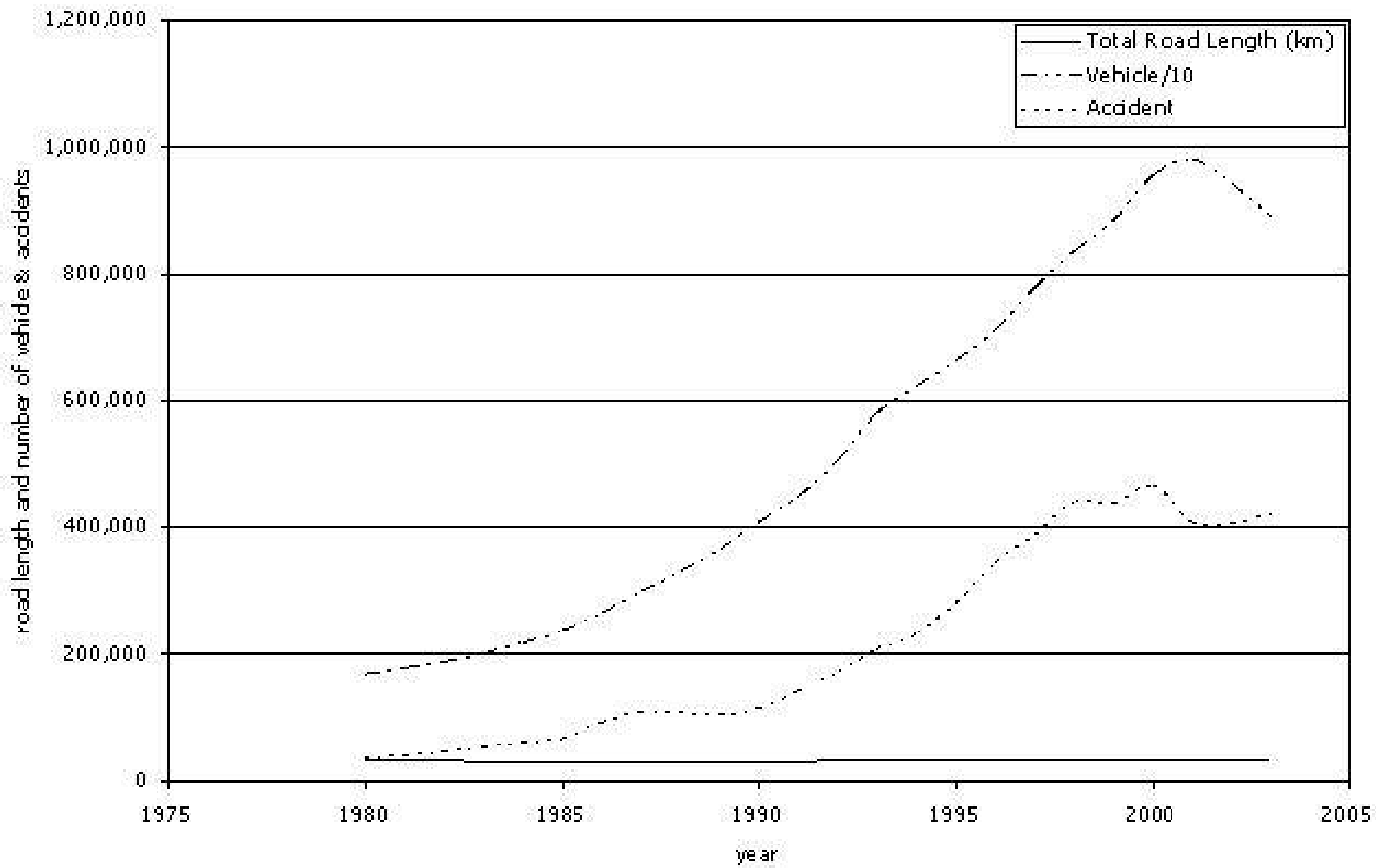


Figure 3.4. Distribution of road length and number of vehicles & accidents between the years 1980 and 2003 in Turkey

As it is seen from the previous tables and figures, there appears a relatively minor increase in the total road length (including state highways and motorways), approximately 1.04 times from 1980 to 2003, but the number of vehicles increased 5.29 times and road accidents 11.44 times. Although there exists a slight decrease in the number of accidents and mortality during the last years, traffic accident problem cannot be judged as to be solved since human life is invaluable. Moreover, statistics show that traffic accident rate starts increasing again with the year 2003. Indeed, there are still many necessitates to overcome this destructive problem immediately and effectively.

Road safety, hence, should be significantly taken as one of the most serious problems for our country. Sustainable solutions should be developed considering economical, environmental and social dimensions of the arousing safety problem by looking from an integrated perspective of engineering and planning approaches.

Within the consideration of priority needs for future roads and highways, three basic aspects should be regarded (6):

- Part or all of the new highway system should be projected carefully by professional planners.
- Proposed plans should be an overall transportation facility.
- Proposed plans should be immune from political influence and other pressures.

A useful framework for analyzing traffic accidents, which indicates the countermeasures to cope with the existing situation, has been provided by Haddon in 1980 with an indication of three contributing factors of accidents: human factor, vehicle factor and road factor (Table 3.5) (28).

Table 3.5. Haddon's countermeasure matrix (28)

	Before Crash	In Crash	After Crash
Driver	Training, education, behavior (for instance, avoidance of drinking and driving), attitudes, conspicuous clothing (pedestrians, cyclists)	In-vehicle restraints fitted and worn	Emergency medical services
Vehicle	Primary safety (for instance, braking, roadworthiness, visibility), speed, exposure	Secondary safety (for instance, impact protection)	Salvage
Road	Delineation, road geometry, surface condition, visibility	Roadside safety (for instance, no hazardous poles)	Restoration of road and traffic devices
Pedestrian	Education, clothing, behavior, attitude	-	Emergency medical services

3.2. Importance of Roadway Factor in the Occurrence of Accidents

According to the statistical analysis of accidents made by Ministry of Interior-General Directorate of Security of Turkey for the year 2004, human factors (including driver, passenger, pedestrian) covered about 99.74 percent of the accidents whereas the effects of the other factors indicated very low values; the road environment has a role by 0.10 percent and the vehicle by 0.16 percent. However, in reality, accidents causing from the deficiency of road characteristics and environment occur more frequently than it has been indicated in the statistics.

The 99 percent of human factor contribution may not be surprising since all accidents involve road users. Hence, it is also possible to think that some remedies taken by road user are sufficient to avoid the accidents. However, the most effective remedy is not necessarily related directly with the main factor, and may lie in a different area. Even though studies on accident causation indicate that influencing the human behavior is the most essential remedy taken for accident reduction, changing the pattern of behavior by persuasion or training is more difficult than applying engineering measures, which are founded on removing hazards or improving vehicles. Furthermore, human behavior may often be influenced more readily by engineering means rather than by education or by enforcement of legislation. For instance, if the road user fails to cope with the road environment, road engineering improvements can play the most essential role in reducing accidents (17).

The roadway may directly contribute to the traffic accidents or an accident may result from a combination of causes, which include roadway imperfections. While accidents that are directly caused by roadway defections are estimated to a smaller percent of all accidents, this number increases considerably if roadway problems are counted as one of the contributing factors (25).

Although various estimates have been made according to the percentage of contributing factors to accidents, unfortunately, present state of traffic accident reporting does not ensure definitive statements for the role of the roadway parameters to the accidents (25). In Turkey, while the existence of the inadequacies in the road network infrastructure is obvious and also accepted by the authorities, a statement proposing to cover up these inadequacies, which was published in a circular by one of the organizations responsible from the road network infrastructure, is quite dramatic from this manner (9):

"Recently, it is seen that highways are given many faults for the occurrence of traffic accidents. Since this, in fact, gives to the citizens the right to litigate for the action of damages, one should be more careful from this respect. "

3.3. Contribution of Roadway Characteristics and Environmental Factors to the Accidents

In order to reduce accidents resulting from roadway inadequacies, it is needed to define the specific weakness of roadways by concerning the occurrence of accidents. As emphasized before, unsafe road conditions and environmental factors may considerably contribute to the traffic accidents.

Numerous studies have been conducted in order to investigate the relationship between the traffic accidents and road characteristics and environment. These studies have indicated that improvements to topography, geometric features, road environment, road surface and meteorologically affected conditions could significantly reduce the number of vehicular accidents. Small radius of curvature, very abrupt variations of superelevation of curvature, adhesion and evenness fault of long steep hills, roadside obstacles, unsealed roads and wet or slippery roads are some of the common potential factors, which have a direct impact on the occurrence of traffic accidents (24, 30).

According to the results of the studies reported by Ferguson et al. (1995), Brughton and Sedman (1989), Meyerhoff (1978) and Joksch and Wuerdemann (1974), the number of fatal crashes is inversely proportional with the intensity of light. It is also commonly known that rain, sleet or snow not only obscures visibility but also contributes to unsafe road conditions, which directly result in an increase on the likehood of traffic accidents (31).

Okamoto and Koshi (1989) studied the relationship between accident rates and geometric design of roads using a linear regression model. In another study, an empirical relationship was established among truck accidents and key highway geometric design components by using a non-linear assumption, Poisson regression approach. In that study, Miaou, Hu, Wright, Rathi and Davis (1992) found that annual average daily traffic per lane, horizontal curve and vertical grade were particularly correlated with truck accident rates. Hadi et al. (1995) performed a negative binomial regression analysis to investigate the effects of cross-sectional design elements and brought up that increasing lane width, center shoulder width and median width were vital in reducing accidents. Shankar, Mannering and Barfield (1995) studied the frequency of rural freeway accidents caused by deficiency on roadway geometry and weather-related environmental factors (30). Ivan and O'Mara (1997) analyzed the Connecticut Department of Transportation's accident data using Poisson regression and found that geometric factors critically affect the occurrence of accidents (11). Al-Ghamdi (2002) studied the impact of the existence of intersections on the occurrence of traffic accidents (32).

According to a research made in Sweden (33), some reduction factors were estimated indicating the effects of countermeasures applied to the roadway geometry and environment. The approximate effects (or effect intervals) of these improvements are indicated in Tables 3.6, 3.7 and 3.8.

Table 3.6. Estimated reduction factors for countermeasures for road sections (33)

Countermeasures on sections	Estimated reduction factors			Comments
	Accidents	**Fatalities**	**Injuries**	
Road widening	-20 %	-10 %	-15 %	no reduction in urban areas
Climbing lanes	-25 %	-15 %	-20 %	
Decrease the number of approaches	-5 % -10 %	-5 % -10 %	-5 % -10 %	
Road side delineators	close to ±0 %	close to ±0 %	close to ±0 %	in darkness on roads with bad alignment increase
Road markings	±0 % -10 %	±0 % -10 %	±0 % -10 %	
General speed limits	-10 % -15 %	-20 % -30 %	-15 % -20 %	depends on decrease in average speed
Lower speed limits during winter	-20 %	-40 %	-30 %	depends on decrease in average speed
Local speed limits	decrease	decrease more than accidents	decrease more than accidents	
Bridge widening	-40 %	-20 %	-30 %	not based on empirical data
Side area improvement	±0 %	-20 % -40 %	-20 % -40 %	
Guardrails	±0 %	-20 % -40 %	-20 % -40 %	

Table 3.6. Estimated reduction factors for countermeasures for road sections (continued) (33)

Median barriers	+20 % +25 %	-15 % -20 %	-10 % -15 %	
Vertical alignment	±0 % -20 %	±0 % -20 %	±0 % -20 %	
Increased horizontal curve radius	-5 % -60 %	-5 % -60 %	-5 % -60 %	
Improved signing in horizontal curves	-10 % -40 %	-10 % -40 %	-10 % -40 %	
Super elevation	-10 % -20 %	-10 % -20 %	-10 % -20 %	
Sight distance	-5 % -15 %	-5 % -10 %	-5 % -10 %	
New surface	around 0 %	around 0 %	around 0 %	no reduction
Increased friction	-5 % -10 %	-5 % -10 %	-5 % -10 %	
Decreased rutting	around ±0 %	around ±0 %	around ±0 %	
Decreased unevenness	±0 % -5 %	±0 % -5 %	±0 % -5 %	
Prohibit overtaking	-5 % -10 %	-5 % -10 %	-5 % -10 %	
Variable message signs	-15 % -20 %	-15 % -20 %	-15 % -20 %	
Improved route guidance	around –2 %	around –2 %	around –2 %	

+ means increase, **-** means decrease, when the countermeasure is applied

Table 3.7. Estimated reduction factors for countermeasures for junctions (33)

Countermeasures in junctions	**Estimated reduction factors**			**Comments**
	Accidents	**Fatalities**	**Injuries**	
Island on the secondary road in 3-leg junction	±0 %	±0 %	±0 %	
Island on the secondary road in 4-leg junction	-5 % -10 %	-5 % -10 %	-5 % -10 %	
Left-turning lane with curbs in 3-leg junction	±0 % -10 %	±0 % -10 %	±0 % -10 %	closer to –10 % in urban areas and closer to 0 % in rural areas
Left-turning lane, painted in 3-leg junction	±0 % -10 %	±0 % -10 %	±0 % -10 %	
Left-turning lane, with curbs in 4-leg junction	-10 %	-10 %	-10 %	
Left-turning lane, painted in 4-leg junction	-10 %	-10 %	-10 %	
Right-turning lane	±0 %	±0 %	±0 %	can increase accidents
Roundabout	+20 % -70 %	-50 % -80 %	±0 % - 50 %	can increase accidents in central areas

Table 3.7. Estimated reduction factors for countermeasures for junctions (continued) (33)

Change one 4-leg junction into two 3-leg junctions	±0 %	±0 % -40 %	±0 % -40 %	higher reduction factor when the percentage of vehicles from secondary road is higher
Modern traffic regulated signals	-15 % -30 %	-15 % -30 %	-15 % -30 %	time regulated signals increase accidents
Interchange 3-leg	-20 % -40 %	-40 % -60 %	-40 % -60 %	
Interchange 4-leg	-60 % -70 %	-60 % -90 %	-60 % -90 %	
Lighting in junctions	-5 % -10 %	-5 % -10 %	-5 % -10 %	
Change yield to stop in rural areas	-10 % -15 %	-10 % -15 %	-10 % -15 %	
Change yield to stop in urban areas	±0 % -5 %	±0 % -5 %	±0 % -5 %	
Flashing yellow in signal during low traffic hours	+50 %	+50 %	+50 %	
Counter measures in railway junctions	-25 % -70 %	-25 % -70 %	-25 % -70 %	

+ means increase, **-** means decrease, when the countermeasure is applied

Table 3.8. Estimated reduction factors for improvements for pedestrians and bicyclists (33)

Improvements for pedestrians and bicyclists	**Estimated reduction factors**			**Comments**
	Accidents	**Fatalities**	**Injuries**	
Sidewalks	-5 % -10 %	-5 % -10 %	-5 % -10 %	
Separate bicycle and pedestrian lanes in rural areas	±0% -5%	±0% -5%	±0% -5%	
Separate bicycle and pedestrian lanes in urban areas	around -4 %	around -4 %	around -4 %	
Grade separated pedestrian and bicycle junctions	around -80 % in pedestrian accidents	around -80 % in pedestrian fatalities	around -80 % in pedestrian injuries	reduction factors depend on the use of separation
Marked pedestrian crossing	+25 % -20 %	+25 % -20 %	+25 % -20 %	can increase accidents
Bus stop	small or no reduction factor	small or no reduction factor	small or no reduction factor	small or no reduction factor

+ means increase, **-** means decrease, when the countermeasure is applied

CHAPTER 4

GENERAL OVERVIEW OF GEOGRAPHIC INFORMATION SYSTEMS AND ITS ROLE IN TRANSPORTATION

4.1. A Brief History of GIS

The earliest antecedent of Geographic Information Systems (GIS) goes back to the University of Washington where quantitative methods in transportation studies were developed by geographers and transportation engineers in the early 1950s (34).

However, the actual development of GIS began in the mid-1960s when it was used as a computerized map measuring system. Roger Tomlinson developed the “Canadian Geographic Information Systems (CGIS)”, which was needed to analyze Canada's national land resources, their existing and potential uses. This system mainly aimed to be used as a measuring tool, for producing tabular information rather than as a mapping tool and it pioneered many aspects of GIS (35, 36).

The second milestone in GIS application was the establishment of the Harvard Laboratory for Computer Graphics and Spatial Analysis by Howard Fisher. It was an important research center led to the creation of software for spatial data handling. Howard Fisher developed a pioneering automated computer mapping program with an exclusively on line printer output, named as Synagraphic Mapping System “SYMAP”, at the Northwestern Technology Institute and completed it at the Harvard Laboratory for Computer Graphics and Spatial Analysis. In the 1970s, the laboratory also produced the first modern vector GIS product called “ODYSSEY GIS” (36).

In the late 1960s US Bureau of Census Dual Independent Map Encoding (DIME) data format was developed by George Farmsworth, which was used to create digital records of all US streets to support automatic referencing and aggregation of census records (35).

Studies related with GIS speeded up in universities and the term of topology was introduced into the system in 1970s. In 1973, Maryland Automatic Geographic Information (MAGI), one of the first state-wide GIS projects began. Minnesota Land Management Information System (MLMIS), another significant state-wide GIS, was implemented in 1976 as a research project at the Center for Urban and Regional Analysis, University of Minnesota. These systems primarily increased the experiences on the creation and management of large geographic information systems (36, 37).

In 1980s, significant computerized programs were developed and this also gave rise to the reliable systems. The concept of the spatial data modeling and data structures was progressed. First commercial systems were created. In these years, a breakthrough occurred in government planning organizations and local levels were also affected from these developments by 1990s. Spatial information science was introduced, object oriented studies started and large databases were better handled with advanced GIS programs in 1990s (37).

Recently, the accelerating development in the computer technology has also accelerated the developments in the GIS and studies in the area primarily become to be concentrated on 3-D and temporal GIS.

4.2. Defining GIS

Geographic Information Systems can be defined in many ways because of its wide variety of applications. In general, GIS is an organized collection of computer hardware, software, geographic data and personnel, designed to efficiently capture, store, update, manipulate, analyze and display all forms of geographically referenced data (38).

GIS has an important role in performing efficient analyses of huge amounts of multiattribute geographic information since it integrates digitized drawings/maps with relational database (39). In addition to its abilities to manage spatially distributed data for decision making process, the analyses results obtained from GIS can be represented in quick, correct and clear way, and several scenarios can be produced and evaluated efficiently and effectively. It is also possible to create a sharable database by which communication increases among various users (individual and/or departments), a common link between two or more previously unrelated databases is provided, redundancy is reduced, productivity is enhanced and overall efficiency is improved (23).

GIS technology integrates common database operations such as query and statistical analysis with the unique visualization and geographic analysis benefits offered by maps. These abilities distinguish GIS from other information systems and make it valuable to a wide range of public and private enterprises for explaining events, predicting outcomes and planning strategies (40).

In summary, a comprehensive GIS requires a means of (41):

- data input from maps, aerial photos, satellites, surveys and other sources
- data storage, retrieval and query

- data transformation, analysis and modeling, including spatial statistics

- data reporting, such as maps, reports and plans

4.3. Components of GIS

A complete GIS integrates five key components (Figure 4.1):

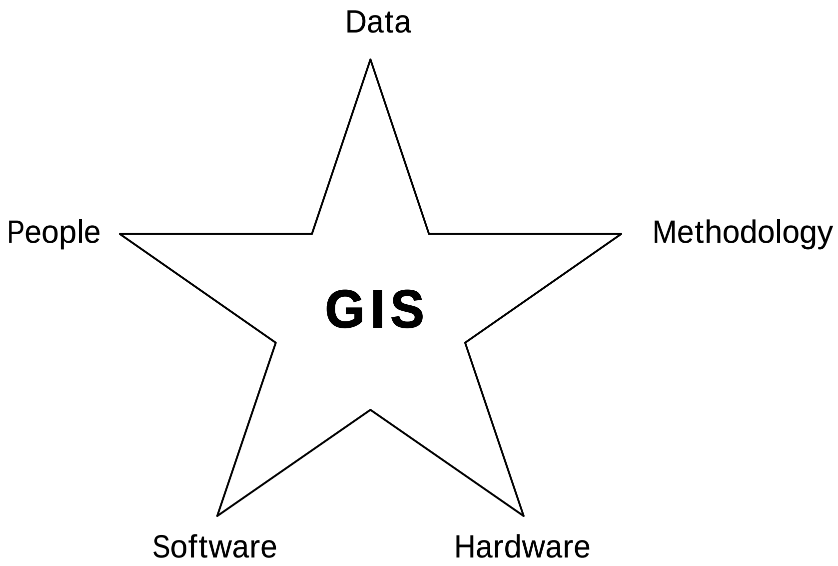


Figure 4.1. The key components of GIS

Data: Data is the most essential component of GIS, which requires two sets of information in order to perform GIS functions: spatial (graphic) data and attribute (non-graphic) data. While spatial data basically indicate positional information, attribute data point property information. In simple, spatial data is “x, y, z” and attribute data is “A, B, C”. By combining spatial and attribute data, data layers are constructed.

People: GIS user has a very crucial role since he/she makes the decisions and manages the systems. User should understand the data, the software and the hardware and should also develop the required methodology successfully. GIS users may range from designers of the system to those that only use that system for performing the analyses.

Methodology: Construction of a reliable and suitable methodology constitutes an important part of a GIS study in order to achieve the desired goals. Methodology may also show variations according to the scope, borders and constraints of the study.

Software: GIS software provides the functions and tools needed to store, query, create, modify, analyze and display data. There are three key parts of a GIS software (Figure 4.2) (35):

1. A graphical *user interface* (GUI) in order to provide easy access to the GIS tools.

2. The *tools*, which define the capabilities and functions that the GIS software has for processing data (for input and manipulation of geographic information; for geographic query, analysis and visualization).

3. A *database management system (DBMS)* that supports the data to be stored, organized and maintained.

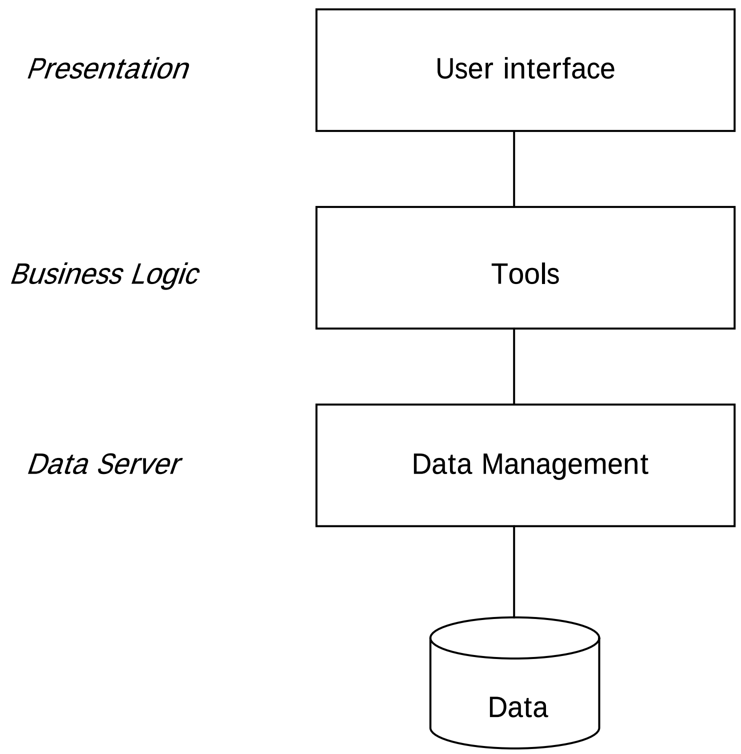


Figure 4.2. Classical three-tier architecture of a GIS software (35)

Hardware: Hardware is the computer on which a GIS operates including the resources available to the computer; by printers, plotters, digitizers, scanners, monitors, network and wide area communications. Today, GIS softwares run on a wide range of hardware types, from centralized computer servers to desktop computers, which are used in stand-alone or network configurations (40).

4.4. GIS Data Types

GIS data, as stated before, can be mainly divided into two groups:

- spatial data
- attribute data

4.4.1. Spatial Data

Spatial data represent features that have a known location on earth. They are defined by information about position, connections with other features and details of non-spatial characteristics. Large volumes of spatial data can be handled with GIS. It is also possible to capture, store and manage spatially referenced data in the form of data models (42, 43).

All GIS are computer representations of the real world. Since real world geographic systems are too complex for even the most advanced information systems, they should be simplified (Figure 4.3 (37)). This simplified view of the real world adopted by GIS is generally termed as a model; i.e. synthesis of data (43, 44).

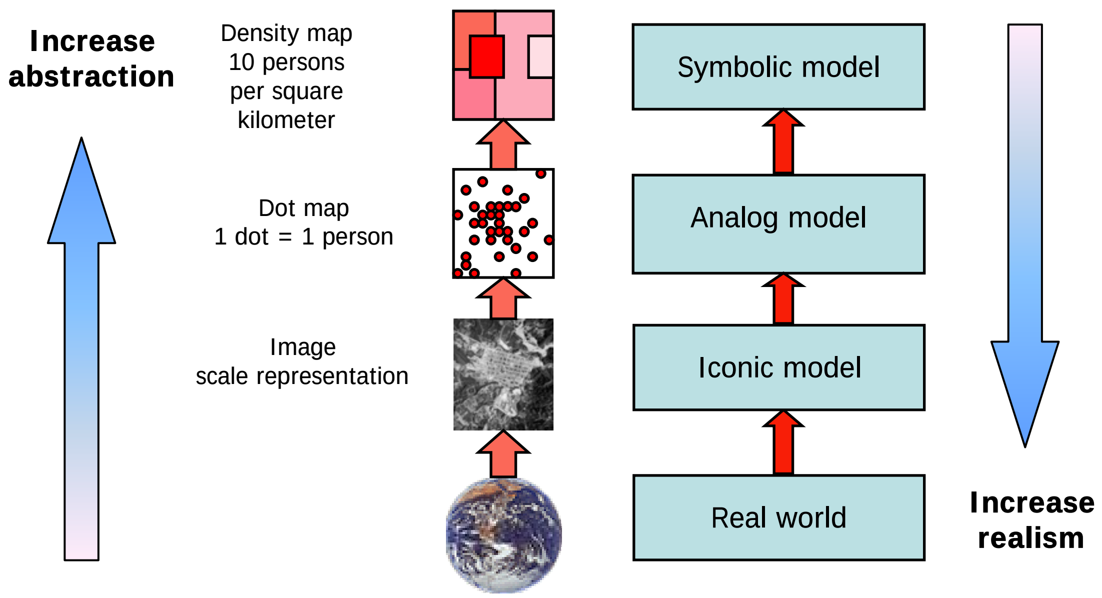


Figure 4.3. Levels of data model abstraction and realism (37)

Spatial data can be represented by one of the two fundamental approaches of GIS: vector data model and raster data model.

Vector data model: It is an object-based model. Space is populated by well distinguishable, discrete and bounded objects. In the vector data model, each object in the real world is classified with string of coordinates; i.e. in the form of points, lines and polygons (37, 44).

Points are recorded as single coordinate pairs and represent objects without shape and size (depending on scale) and having single locality. Lines are defined as a series of ordered coordinate pairs or line segments between nodes and vertices. Lines have 'begin' and 'end' nodes. They may be composed of zero or more 'internal' nodes, which are also named as vertices defining the shape of the line (Figure 4.4). Polygons are characterized with one or more lines that close to form an area. They can also be divided in representative regions using an arbitrary grid or block groups, depending on the data available and the desired analysis (35, 37). Figure 4.5 (45) shows examples of different forms of vector data.

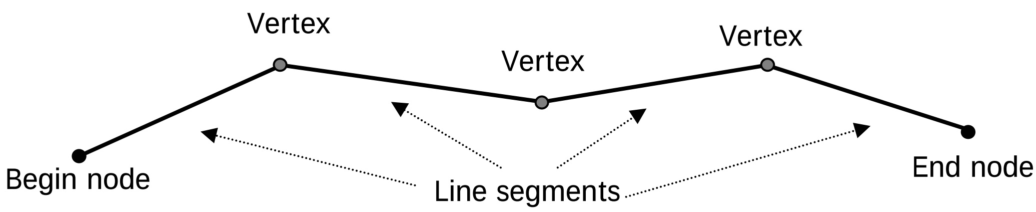


Figure 4.4. Basic structure of a line in vector data model (37)

In a particular transportation application, points may be used to represent the locations of traffic accidents or zonal centroids, traffic lights and so on; lines may be used to describe elements, such as roadway segments, utility lines, etc. and lastly polygons may be used to represent real features, such as traffic analysis zones, political jurisdictions or land subdivisions (46).

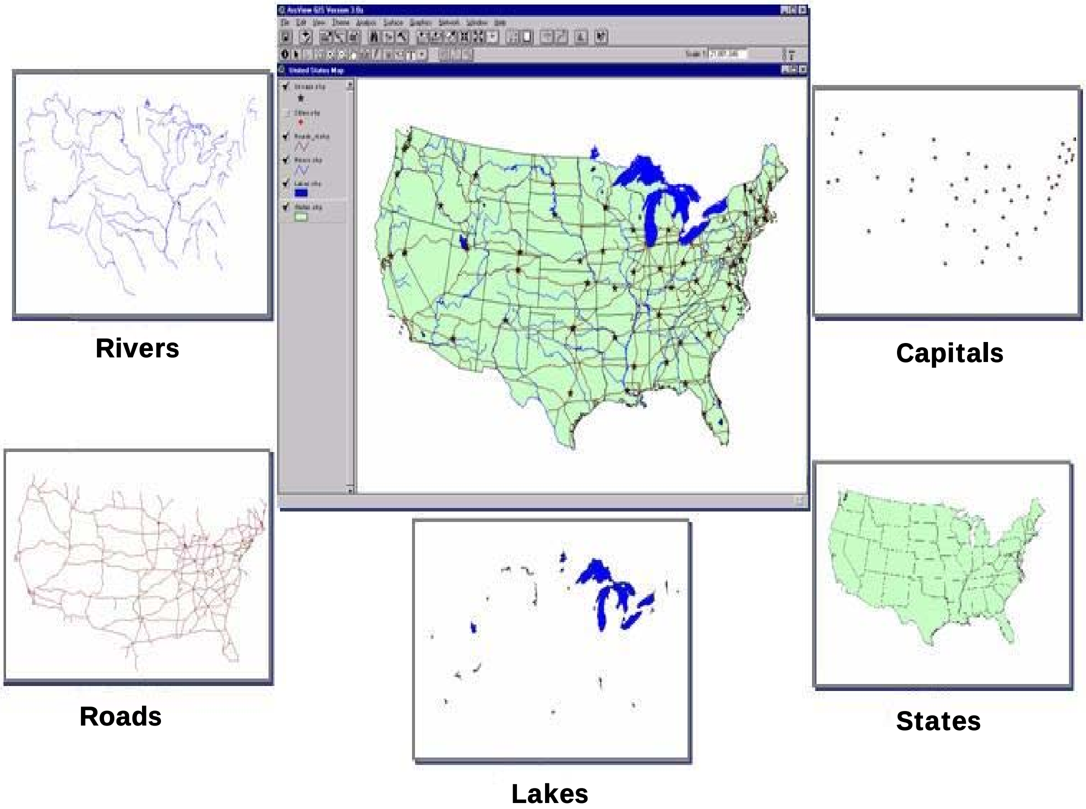


Figure 4.5. Examples of vector data (45)

In vector representation, a spatial relationship exists between geographic objects, named as **topology**. The principles of topology form the basis of GIS and distinguish it from computer aided design (CAD) and mapping tools. Topology is a branch of mathematics that deals with properties of space that remain invariant under certain transformations, such as translation, scaling, rotation and shear. Although some geometric properties (area, perimeter, orientation, etc.) can be affected from these transformations, the topological characteristics, i.e. *contiguity* (adjacency analysis), *enclosure* (containment) and *connectivity* are unchanged (Figure 4.6 (44)). In fact, topological data model stores the part of polygon's boundary as non-looping arcs and indicates which polygon is on the left and which is on the right of each arc (37).

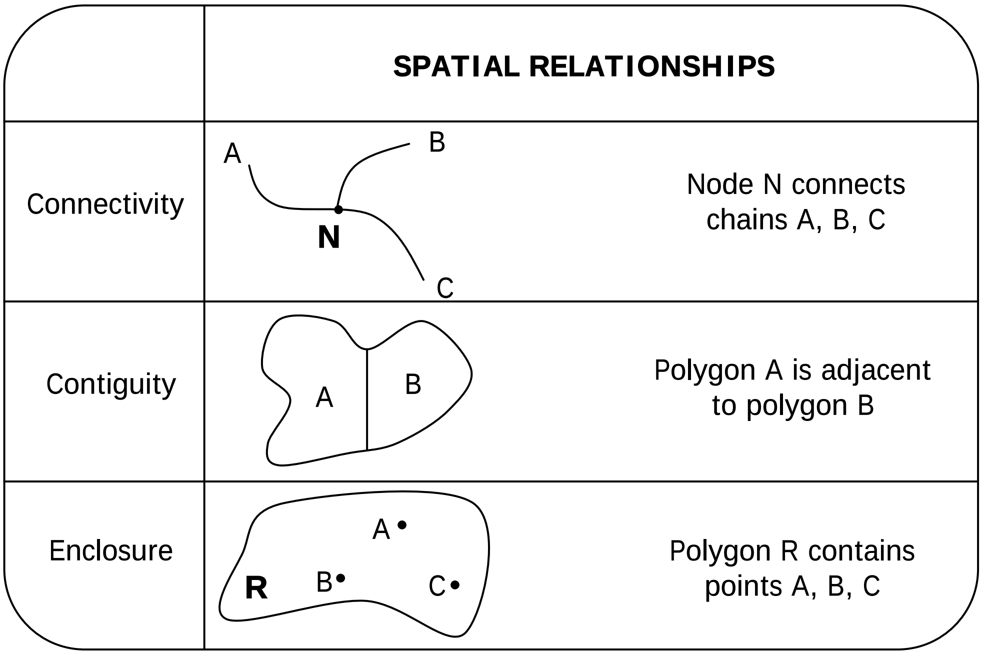


Figure 4.6. Spatial relationships in GIS (44)

Raster data model: It is a field-based spatial data model based upon a regular tessellation of a surface into pixels or grid cells. Tessellation is the partitioning of space into pairwise disjoint cells, which contain information that characterizes that part of space (Figure 4.7). Data in raster format are stored in a two dimensional matrix of uniform grid cells (pixels) generally in the form of square or rectangular regular grid that are filled with measured attribute values (37, 47).

In raster model, points are single cells, lines are formed by connecting cells into a one pixel-thick line and areas are composed of contiguous pixels with the same value (Figure 4.8). Each spatial object has location information inherent to where they lie in the grid on the map, which shows exactly one value (land use, elevation, political division) for each cell. The grid size that can have various values determines spatial resolution of the map. Higher resolution means more detailed images (44).

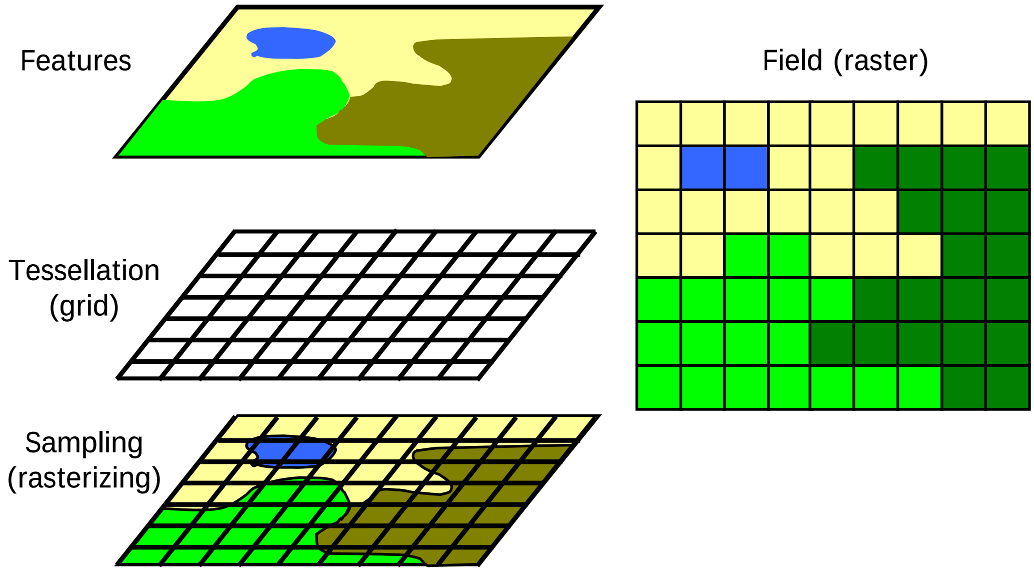


Figure 4.7. Raster data modeling (37)

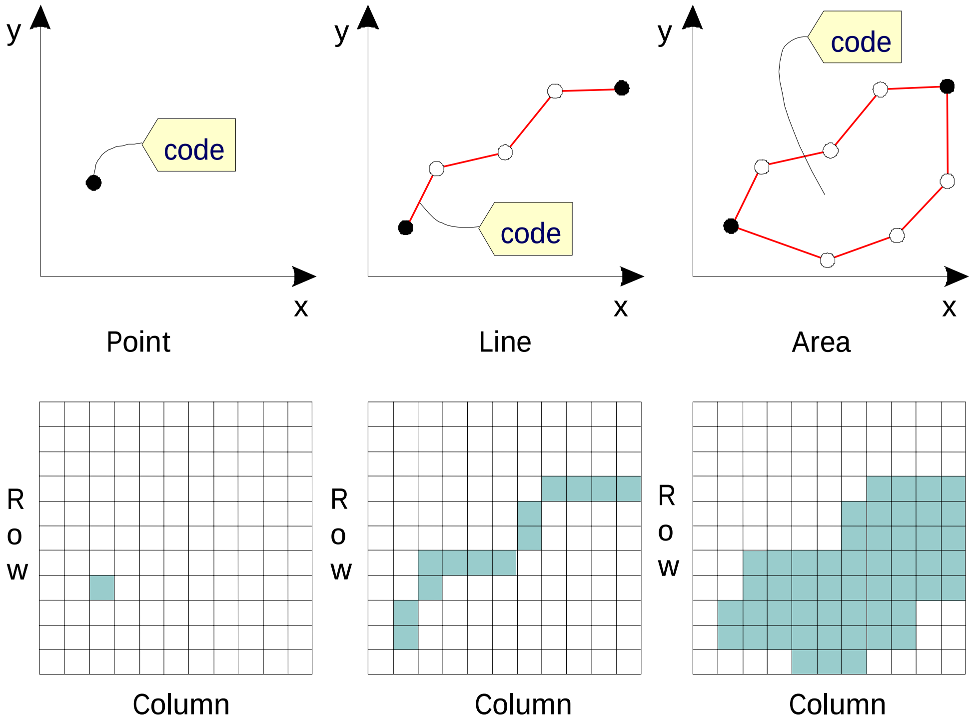


Figure 4.8. Representation of vector and raster models (37)

4.4.2. Attribute Data

Attribute data represents the information linked to the geographic features (spatial data) that describe those features. This means that attribute data are the non-graphic information associated with a point, line or area element in GIS (45) (Figure 4.9).

Attribute data generally take the form of tables and are stored in databases. Soil and vegetation characteristics, health records, city populations, employment numbers, highway route numbers, bus stop names and pavement condition at a specific location can be counted as the examples for attribute data (10).

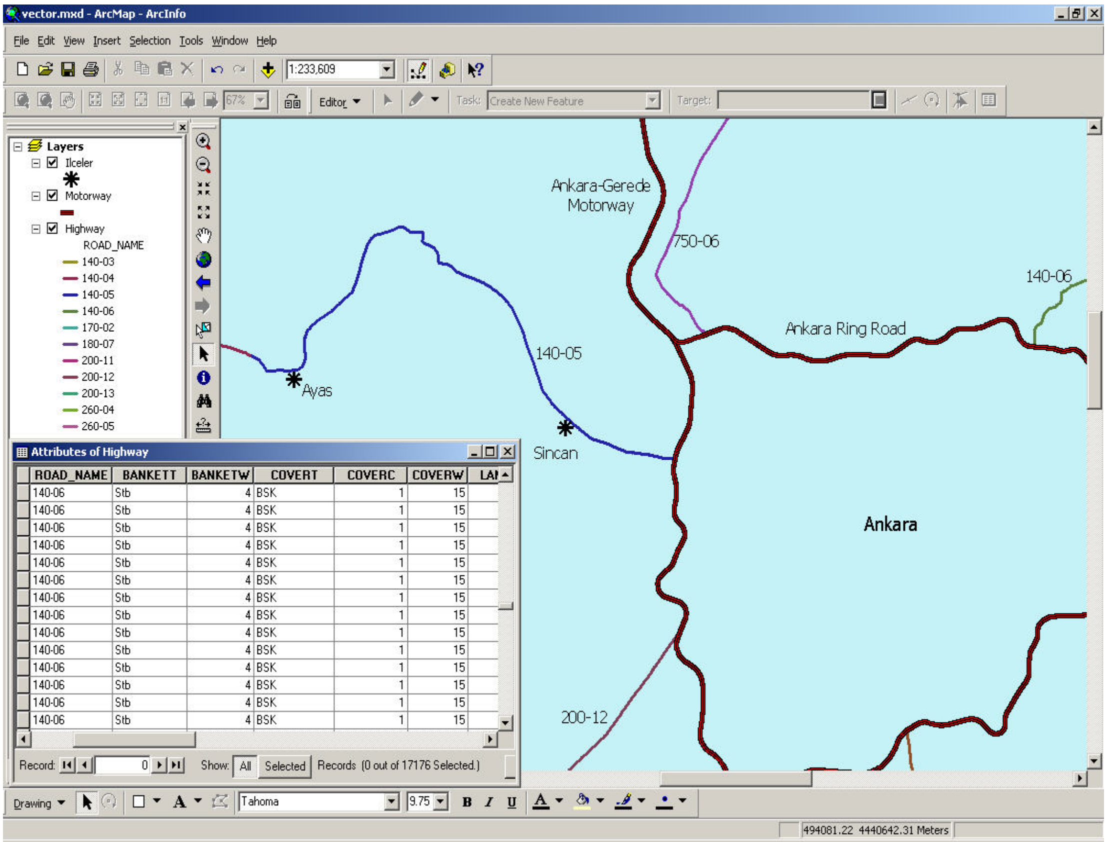
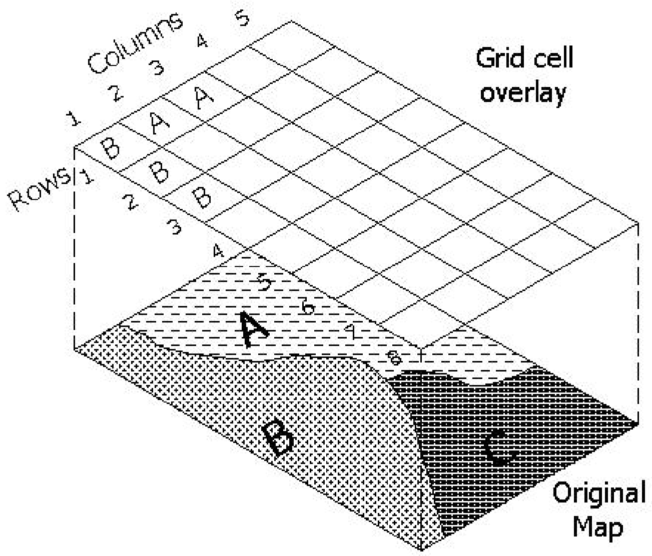


Figure 4.9. Attribute data of a map feature

4.4.3. Linking Spatial and Attribute Data

Construction of the relationship between the spatial and attribute data is one of the most crucial parts of GIS analysis that varies regarding the data model. GIS typically utilizes DBMS strategies in order to link the spatial and attribute data. The way of connecting these two types of data is named as layer type. There are primarily two layer types: *raster* and *vector*.

For a simple raster GIS, a database may not be necessary and attribute values can be held in the same file as the data layer itself, where each cell in a layer of data is associated with a single attribute value. However, this is a very rare case; nowadays, handling the attribute values in a file separate from the image (hybrid approach) is more preferable and efficient. In this case -valid for both raster and vector data model- whereas spatial data are stored as part of the GIS data structure, attribute data are stored in a relational DBMS (Figures 4.10 to 4.12) (43).



Raster data file

Row	Column	Attribute
1	1	B
1	2	A
...	...	...

Figure 4.10. Raster data model in hybrid approach (43)

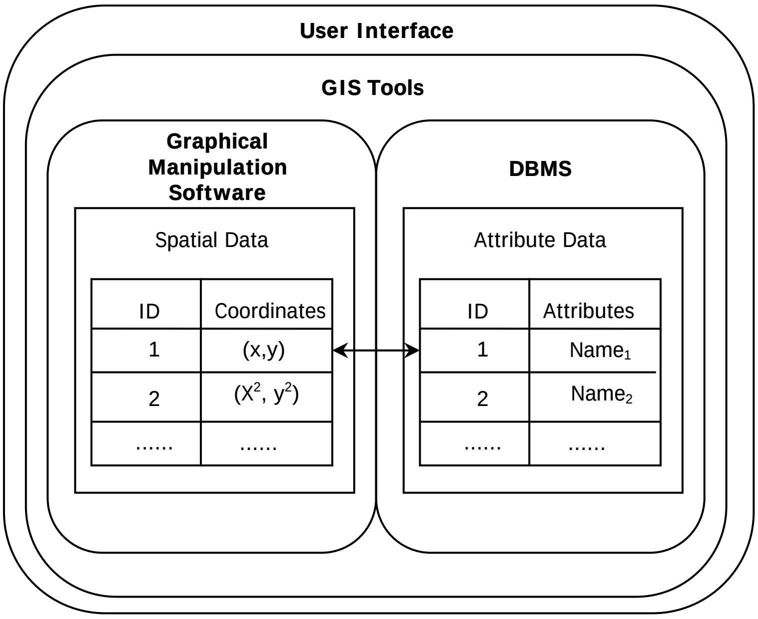
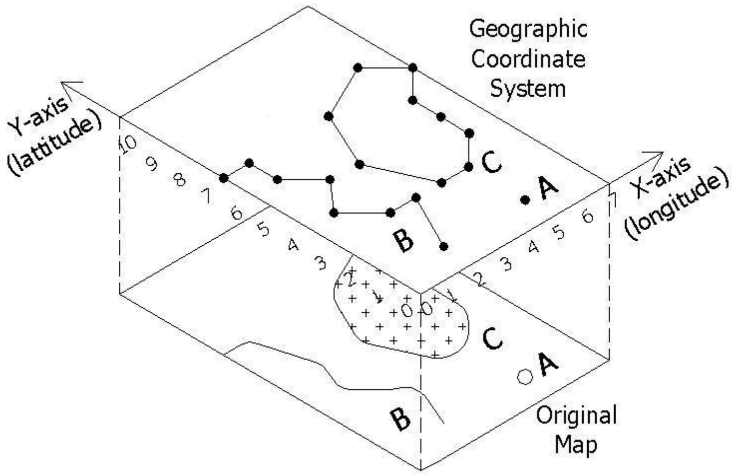


Figure 4.11. Linking spatial and attribute data in vector-based GIS (43)



Vector data file

Entity	X,Y Coordinates	Name
A	5,1	Well
B	0,7; 1,7; 1,6; 2,5; 1,4; 2,3; 3.3; 2,1	Stream
C	4,3; 5,3; 6,4; 6,5; 6,6; 7,7; 6,8; 4,7; 3,5; 4,3	Lake
...	...	...

Figure 4.12. Vector data model in hybrid approach (43)

4.5. Functional Structure of GIS

GIS functionality based on its definition can be divided into four main components or subsystems as indicated in Figure 4.13 (44):

- data input
- data storage and management
- data manipulation and analysis
- data output

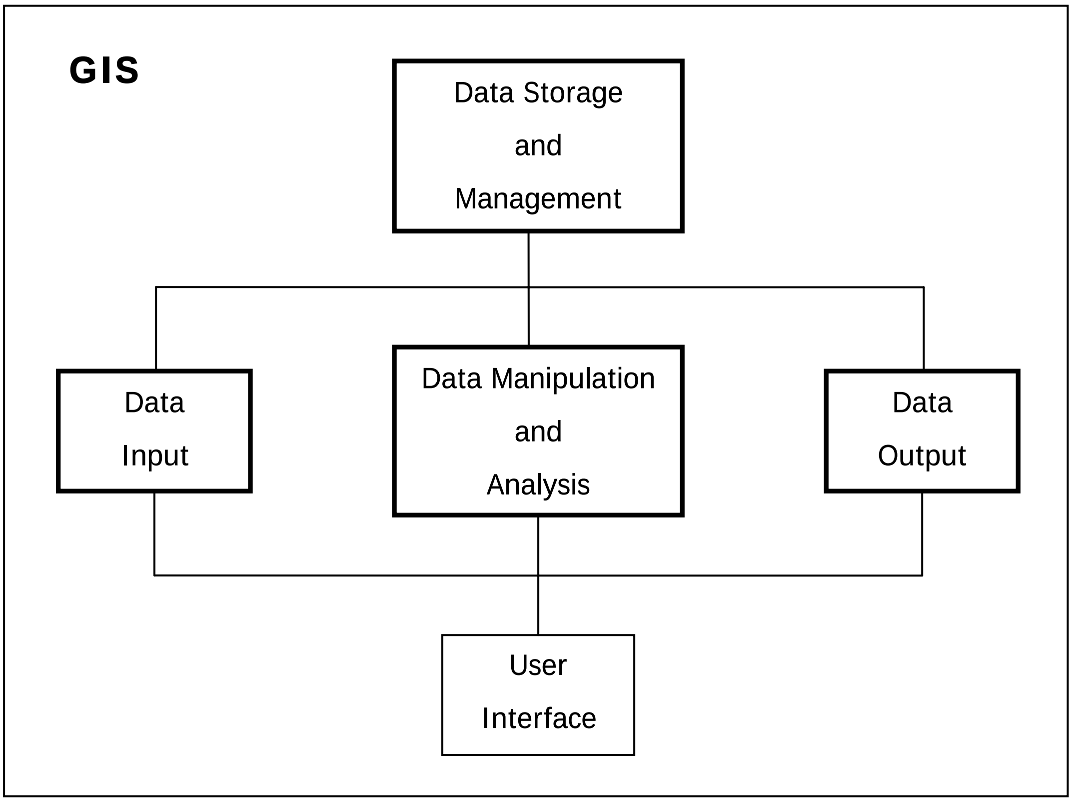


Figure 4.13. Basic Structure of GIS (44)

4.5.1. Data Input

Data input refers to the process of identifying, gathering and converting the required data from their raw or existing form into one that can be used by GIS for a specific application. This process involves acquisition, reformatting, georeferencing, compiling and documenting the data. The data required for the analysis can be provided from different sources in different forms including digitizing paper maps, tables, charts and direct digitizing from aerial photos, satellite images and surveys (44).

Because of its ability to store and maintain large data sets and its potential to integrate wide range of data and information sources into a compatible format, GIS is unique in its capacity for spatial analysis of multi-source data sets, such as data on population, topography, hydrology, climate, vegetation, transportation network and public infrastructure, effectively. It also provides facilities to understand one to many, many to many and many to one relationships existing in spatial data (8, 44).

In creating a GIS database, quality of data should be as good as possible in order to ensure accurate and precise data sets. Accuracy is the degree to which information on a map or digital database matches true or accepted values. It particularly indicates the quality of data and number of errors contained in a data set or map. Precision refers to the level of measurement and exactness of description in a GIS database. The level of precision shows a great variety according to the aim of the study. In fact, highly precise data can be difficult, time-consuming and costly to collect (5).

The accuracy of a GIS mainly depends on the scale and resolution of the maps, satellites and any other data sources and techniques used to create the GIS data set. Scale is the ratio between dimensions of the map and those of reality. Maps can be in small scale (less detailed), large scale (very detailed) or medium scale. The decision of this limitation controls how and what entities are shown on the maps (44).

Map resolution refers to how accurately the location and shape of map features can be depicted for a given map scale. While high resolution implies large scale, a low resolution map shows course detail (data). Resolution is also defined as the minimum mapping unit or pixel size; i.e. the level of detail obtained in each piece of data (37, 44).

Data quality standards should be well evaluated before creating the GIS data in order to attain a reliable study; but also an optimum solution should be developed within the consideration of the scope of the study.

4.5.2. Data Storage and Management

The data storage and management functions include the storage, organization and retrieval of data using a database management system (DBMS). The methods used to implement these functions affect the efficiency of the system in performing operations with the data. Most GIS systems are database oriented. The database can be defined as a large collection of non-redundant data in a computer, organized so that it can be expanded, updated, retrieved and shared by various individuals. The GIS databases can be thought of as a representation or model of real world geographic system rather than a simple system for storing data (44).

A DBMS is a computerized record keeping system (a software package) used for creating, organizing and maintaining a database on demand. An ideal GIS DBMS should provide support for multiple users and multiple databases (data sharing), allow efficient updating , avoid unnecessary duplication (redundancy) of data, have the ability to store and manipulate large data sets without any loss, allow data independence, security and integrity (37, 43).

4.5.3. Data Manipulation and Analysis

The great distinguishing feature of GIS is its ability to perform an integrated analysis of spatial and attribute data. The data are manipulated and analyzed in order to obtain desired and so useful information for a particular study (44). Transformation of the spatial data, the process of changing the representation of a single entity or a whole set of data, may be needed before performing the spatial analysis. In GIS, transformation may involve changing the projection of a map layer or correction of systematic errors caused by digitizing. Moreover, data converting (raster to vector or vice versa) may be needed. These potentials make GIS advantageous and appealing from other types of information systems (43).

According to Aronoff (1989) analysis procedures of GIS can be classified into three types (48):

1. Those used for storage and retrieval. For instance, presentation capabilities may allow the display of a soil map of the area of interest.

2. Constrained queries that allow the user to look at patterns in their data. Using queries, only sandy soils could be selected for viewing or further analysis.

3. Modeling procedures, or functions for the prediction of what data might be at a different time and place. Predictions could be made about which soils would be highly vulnerable to erosion in high winds or during flooding, or the type of soil present in an unmapped area.

The way of analyzing the data in GIS is principally depends on the scope of the study, available data, financial and time constraints. The decision method can be developed by the analyst considering the most feasible alternative among all others. The key point should be based on performing the most effective and useful analysis within the pre-specified borders of study.

4.5.4. Data Output

The data output component of GIS provides a way to represent the data or information in the form of maps, tables, diagrams, etc. The selection of the form of data output depends on cost constraints, user needs, audience to whom the results are presented and the available output facilities. While a local authority may produce simple tables, graphs and maps, professional map makers may prefer to produce detailed plots for publication. It is also possible to transfer the digital data into another computer-based system for statistical analysis, desktop publishing and further analysis (43, 44). In general, maps are the most standard output format, which provides to the user and audiences an easy and clear interpretation on the overall picture of the study.

"A picture is worth more than a thousand words."

4.6. GIS Analysis Functions

As emphasized in the previous sections, the most valuable characteristics of GIS rather than other information systems is its ability to combine and integrate data using a set of analytical functions. These analyses functions integrate the spatial and attribute data in a single database in order to find the answers of the questions related with the real world (49).

A GIS analysis function can be fundamentally defined as the investigation of the relationship between the objects having geographical references with themselves and with the other objects located in the database. These analyses functions may only include querying processes to create clearly visualized maps but also they may be used for the integration of several layers for complex studies. In this manner, analyses functions can be mainly classified into *querying*, *(re)classification*, *overlay*, *connectivity* and *neighborhood* operations.

4.6.1. Querying

Performing queries on a GIS database, which is used for retrieving data, forms the basis of the GIS analysis functions. Queries principally offer a method of data retrieval that involves the selective search, manipulation and output of data without any requirement to modify the geographic location of the features involved in the query. Asking queries can be performed on the main GIS database or on new data produced as a result of data analysis. The answers can be used for finding out what exists at a particular location; for the retrieval of spatial data (points, lines, polygons or mapping units in a raster map) and for the retrieval of information using conditional, logical and arithmetical operators (37, 43). For instance, a query may be used to map the most dangerous criminal locations of a city or to obtain the number of traffic accidents occurred on a specific road for the specified time period.

4.6.2. (Re)classification

(Re)classification operations transform the attribute data associated with a single map layer (50, 51). It involves grouping of objects into classes according to the new attribute values assigned to the objects of the input data. (44). For example, traffic accidents occurred in a certain location can be reclassified in 3 classes (human, road, vehicle factors) based on their causes. Reclassification would result in a new map without making change on the main database.

4.6.3. Overlay

Map overlaying is the integration of two or more different input layers in order to generate a new output layer. It can be performed with raster and vector data layers. If the input data is vector structured, the output is a new set of polygons, formed by the intersection of all boundaries of the input data layers. Overlay operations may include any combination of points, lines and polygons (Figure 4.14). The important point in overlaying of the vector-based layers is to construct a common coordinate system for each of the layers. Vector overlaying is based on the Boolean operators; *intersection-and*, *union-or*, *exclusion-xor*, *negation-not* (44) as explained in Figure 4.15 (37).

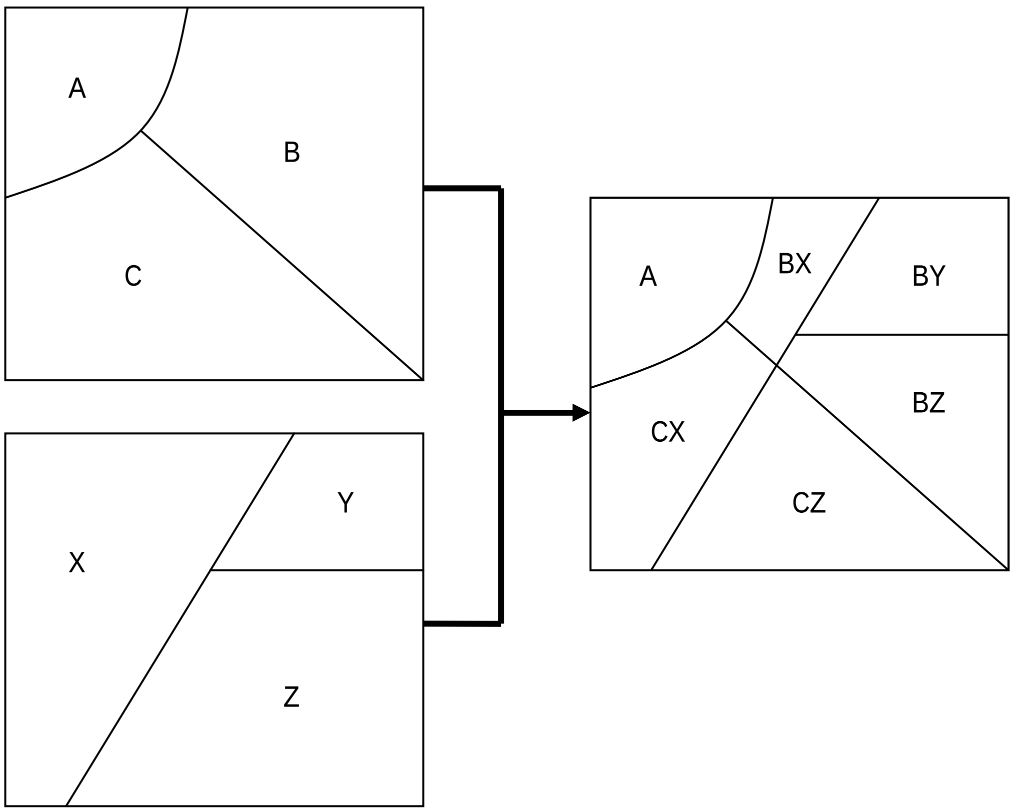


Figure 4.14. Processing of overlay procedure for polygon data layers (44)

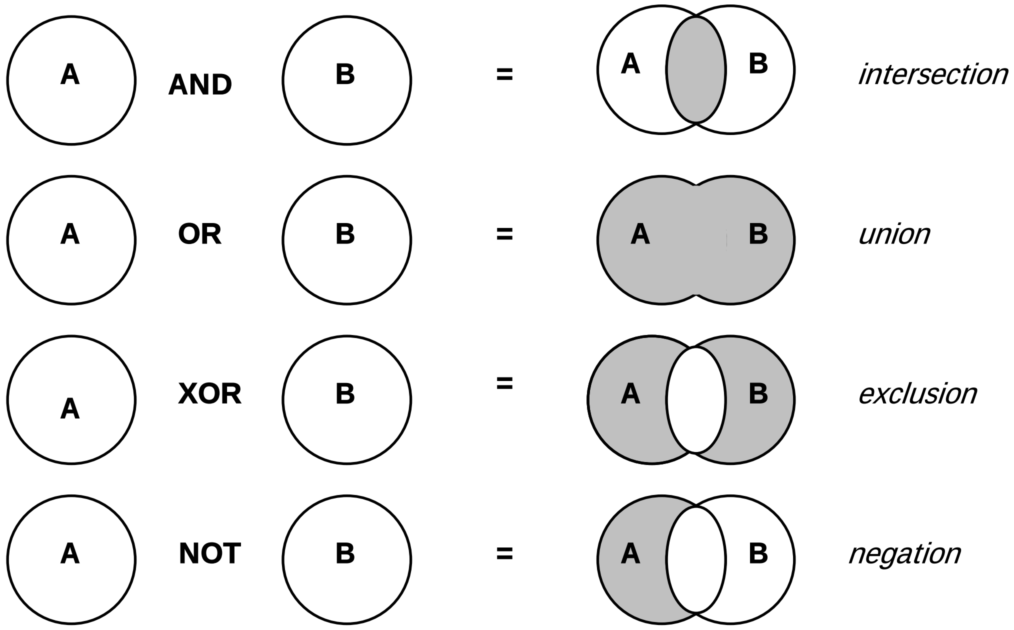


Figure 4.15. Boolean Operators (37)

On the other hand, if the input data is in raster format, the output data is obtained as a raster data layer containing a value in each cell, which is formed by the combination of the overlay operations of the input data values (Figures 4.16 and 4.17). The overlay procedure requires a common grid reference for the raster data layer as in the vector one (44).

The overlay operation tools (raster-based) can be categorized into 4 groups:

1. Arithmetic operations (+ , - , * , /)
2. Relational operations (< , > , =)
3. Logical operations (and, or, xor, not)
4. Conditional operations (if, then, else)

Input Layer A

5	8	2	7
3	4	6	1
2	5	3	5
1	5	7	9

Input Layer B

2	1	7	2
5	4	3	6
7	2	5	3
8	4	1	0

A + B = C →

Output Layer C

7	9	9	9
8	8	9	7
9	7	8	8
9	9	8	9

Figure 4.16. Arithmetic overlay operations (44)

Input Layer A

X	X	X	
X	X		
	X	X	
X	X	X	X

Input Layer B

6	8	2	7
3	4	6	1
2	7	3	4
1	8	7	9

C = (A=”X”) and (B<5) →

0 = false; 1 = true

Output Layer C

0	0	1	0
1	1	0	0
0	0	1	0
1	0	0	0

Figure 4.17. Relational and logical operations (44)

4.6.4. Connectivity

Connectivity operation characterizes spatial units that are connected with each other. It has different concepts for vector and raster structured data. For vector data, it is used to describe the linkage of points or polygons to each other. Alternatively, it describes the existence and type of connection between two pixels for raster data. Connectivity functions particularly accumulate the values over the area being traversed (48).

Proximity analysis is the mostly used operation of connectivity functions. Proximity analysis can be defined as the measurement of distances between features (distance in units of length, travel time, etc.). For instance, the question of "which parcels are within 60 m. of a main road?" can be answered by using proximity analysis (Figure 4.18 (37)). Determination of *buffer* zones; i.e. creating boundaries around objects at an equal proximity in all directions is the most typical proximity function (44).

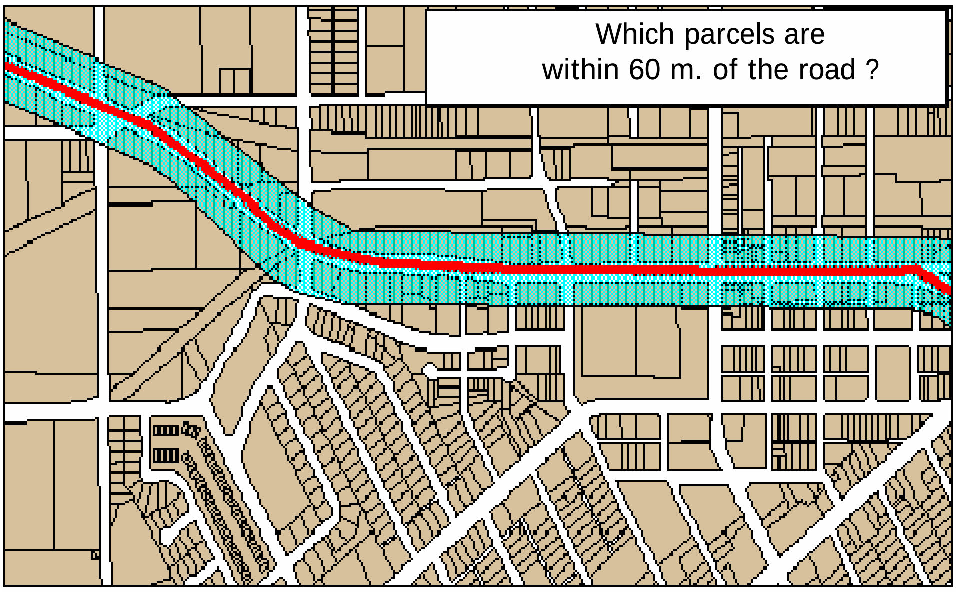


Figure 4.18. An example for proximity analysis (37)

4.6.5. Neighborhood

In the neighborhood analysis, the values of individual cells are altered on the basis of adjacency (43). Explicitly, the value assigned to each location on the output layer is a function of independent values in the neighborhood of that location on the input layer. Search function is the most important one, which evaluates the characteristics of the specified location with specified parameters (44).

The major search operations can be counted as follows (37):

- **average:** the average of the values in the neighborhood
- **diversity:** a measure of diversity of the values in the neighborhood, such as variance or standard deviation
- **majority:** the number of occurrences for each value in the neighborhood is determined; the value occurring most frequently is the calculated result
- **maximum/minimum:** the maximum / minimum value in the neighborhood is returned
- **total:** the summed total of the values in the neighborhood

In order to perform search functions, neighborhood membership is referred as a window. After defining the target locations and neighborhood, algebraic or statistical operations are applied to the locations within the window. Finally, the resulting value is assigned to each location in the neighborhood (44).

4.7. GIS Application Areas

GIS plays a serious and crucial role for understanding and overcoming the vital problems of our world caused by global heating, acid rains, exhausted natural resources, dense traffic volume, rapid urbanization, etc.

GIS is now extensively used in government, business and research for a wide range of applications. These applications mainly include (41, 45):

- Natural Resource Management
    - Wildlife habitat
    - Recreation resources
    - Floodplains and wetlands
    - Agricultural lands and aquifers
    - Forests and woodcutting
- Facilities Management
    - Transportation systems
    - Balancing loads in electrical networks
    - Locating underground pipes and cables
    - Tracking energy use
- Land Management
    - Zoning and subdivision planning
    - Land acquisition

    - Cultural heritage
    - Environmental impact policy
    - Water quality management
    - Maintenance of ownership
- Location Analysis
    - Commercial and industrial location
    - Spatial marketing (stores, warehouses, competitors)
    - Location analysis or site selection for services
    - Development of evacuation plans
- Public Services
    - Emergencies
        - Vehicle routing (police, fire ambulance)
        - Location of medical facilities
    - Spatial crime statistics
    - Location of water sources
    - Location of chemical and hazard materials
- Education
    - Research
    - Teaching tool
    - Administration

4.8. Importance of GIS in Transportation

Transportation engineering and planning require vast and variable amount of information. The collection, management and use of such kind of data require expensive, labor-intensive and time consuming tasks that need the cooperation of many organizations and individuals. In this respect, GIS, a computer-based technology, is particularly adopted to the management and sharing of transportation-related data (46).

In addition to the necessity of efficient and sharable information system, the application of GIS is significantly related with transportation due to the spatially distributed nature of transportation-related data and the need for various types of spatial analysis (52).

Since transportation systems are complicated and have continuously changing nature, they are required to be studied with an enhanced tool. In this manner, GIS has an important role for gaining better understanding of the transportation systems and better perception of the problem, for providing an integrated information system platform and for performing comprehensive and valuable spatial analysis, which enable better and well informed decisions (10).

GIS can be described as an *intelligent* mapping tool because of its ability to represent objects geographically by constructing the relationships between the objects and also associating thematic information to these objects (46). The ability of accessing and analyzing spatially distributed data regarding the topological relationships between objects and also attribute data linked to the spatial data along with the final presentation of results on a digital map and flexibility for the linkage with external procedures and softwares constitute the main advantages of GIS and distinguish it from other database management systems (52).

4.9. Specific Applications of GIS in Transportation

GIS is an emerging technology, which has been used by a variety of institutions such as governmental organizations and universities and also by companies and multinational corporations. It also becomes a valuable asset for civil engineers that mostly deal with geographic data.

The transportation related studies can be defined in a field featured as space extension (linear feature) and always data hungry (7). Moreover, their necessity to integrate various kinds of geographic data into the same system proceeds GIS to be a technology with considerable potential for achieving precious gains in efficiency and productivity for multitude of transportation applications. The widespread use of GIS for transportation applications can involve data retrieval, data integration or data analysis (52).

There are several studies in the literature basically describing how GIS helps the integration of many transportation elements. Information systems management, pavement management systems that work with road segment, highway maintenance management, safety management, intelligent transportation systems, bridge management, transportation systems management, traffic modeling, accident analysis, demographic analysis for funding justification, route planning, environmental assessment of road schemes and the option of displaying any form of tabular data that has a spatial component are some of the typical examples of GIS applications within the concept of transportation.

GIS can be used as a decision aid tool for evaluating the route alignments for roadways and highways by relating topographic, urban and environmental features to geometric design controls. An early transport-related application of GIS was developed by Kastelic and Zura (1992) in order to select the least risky route in the transportation of hazardous materials. This model determines an optimal path based on existing road network geometry and characteristics (road width, radius and slope) and the class of the hazardous material to be transported (4). Similarly, Bridgehouse (1993) used GIS to determine optimal routes for emergency vehicles. With this system, which matches the addresses obtained from incoming 911 calls to a road network, it is possible to find the quickest route. (53). Furthermore, a combination of GIS tools was used to locate a feasible road corridor through the southern Appalachian Mountains of northern Georgia (54).

Aifandopoulous (1995) et al. studied on a tool, which was based on GIS technology demand forecasting, by testing different alternative transport infrastructure development schemes and by evaluating their impact on traffic and environmental conditions for the city of Budapest (4). Illinois Department of Transportation has used GIS for displaying critical environmental themes relative to the proposed roadway alignments (55).

Chisalita and Shahmehri (2002) analyzed the benefits of introducing GIS tools to increase the awareness of drivers against low visibility, meteorological conditions, unsafe road topography or traffic flow with a collision avoidance systems located in vehicles. The developed system integrates the relevant data for traffic safety and takes into consideration the specific requirements of the vehicular domain (56).

GIS has been successfully used for the transportation safety problems. The capability of GIS software in order to display historical accident information pertaining to a particular site helps to the planner and/or designer for studying the impacts of a particular roadway design within consideration of safety. Gharaibeh et al. (1994) used GIS for obtaining spatial and statistical analyses of roadway characteristics such as safety, congestion level and pavement conditions. This study is also important for recognizing various application areas of transportation (23).

Analysis of crash data is one of the most important application areas that is extensively performed by GIS for taking effective countermeasures. From this respect, a highway safety information system for North Carolina for the area of Wake County, created by Federal Highway Administration (FHWA), can be given as a good example. The system includes a user interface by which data is entered, edited, analyzed and exported to other applications. By using the database including crash, roadway, traffic signal, pavement management/ condition, railroad grade crossing inventories and average daily traffic file, five different analyses can be performed: Spot/Intersection Analysis, Strip Analysis, Cluster Analysis, Sliding-Scale Analysis and Corridor Analysis (57).

Spot/Intersection Analysis (Figure 4.19) is used to evaluate crashes at a user-defined spot or intersection within a given search radius. The end result of this analysis is a report listing the number of crashes by injury severity, crash cost and other designated variables. In order to study crashes along a designated length of roadway, Strip Analysis (Figure 4.20) is applied. Cluster Analysis enables to study crashes around a given roadway feature such as a bridge, railroad crossing or traffic signals. Sliding-Scale Analysis is used to identify roadway segments with a high crash occurrence. Lastly, in order to locate high crash concentration within a corridor, Corridor Analysis is performed (57).

Austin et al. (1997) applied GIS for identifying the roads with a high exposure to risk and for investigating exposure of school children to risk during their travels to and from school. They investigated the accidents occurred along specific routes and examined the travel patterns of school children. GIS also enabled them to generate disaggregated travel data and then a statistical package was integrated to examine the relationships between selected variables (53).

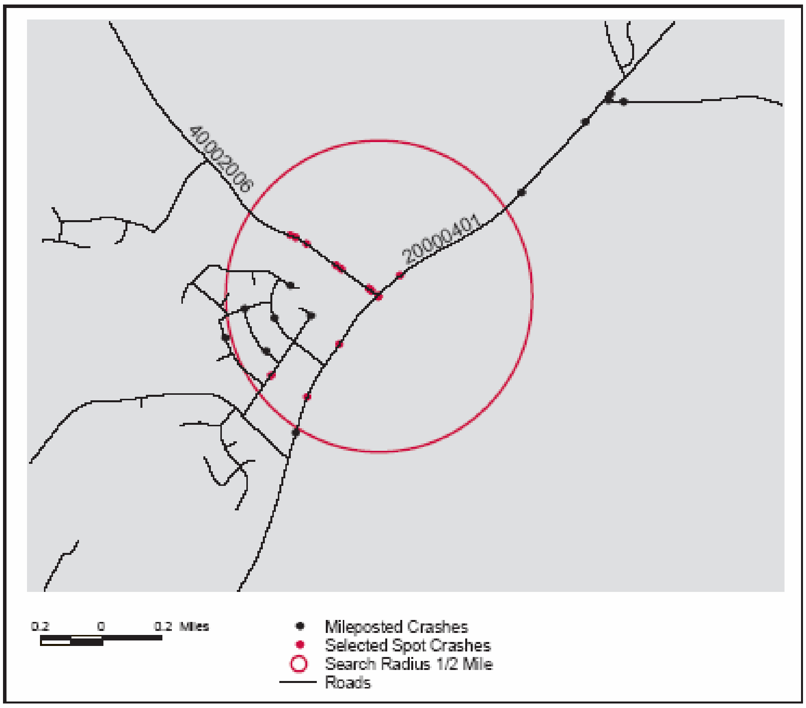


Figure 4.19. Spot-Intersection analysis (57)

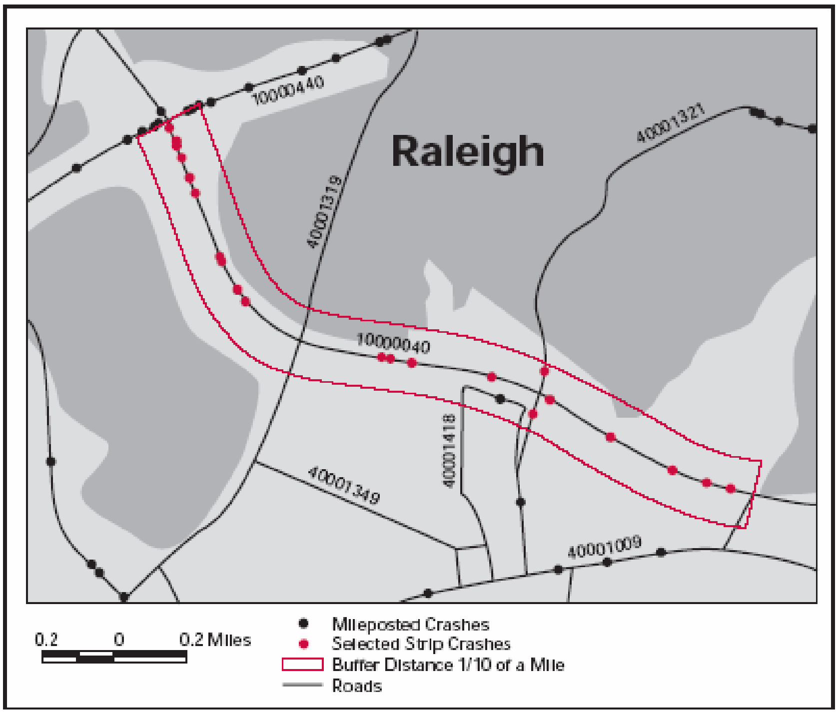


Figure 4.20. Strip analysis (57)

In addition to being an analysis tool for the spatially distributed roadway data, GIS can be applied in accident management systems. A GIS-based traffic accident information system for Kent Country, Delaware was developed by Faghri and Raman (1995). This system, which includes information about the occurrences of crashes, such as conditions of accident site and frequency of accidents at any given location (mile-point) on a roadway, is extremely useful for identifying potentially dangerous (accident prone) locations in planning new infrastructure development and also correctly evaluating and choosing among different design alternatives (23).

Accident Data Management Systems also include the following GIS-related applications (23):

- Providing quick and easy access to roadway segments' information
- Evaluating and monitoring the magnitude of accidents
- Identifying safety improvement needs
- Analyzing alternative solutions to safety problem and assessing their effectiveness in solving them
- Measuring the effectiveness of the implemented actions
- Retrieving accident information that satisfies the query criteria established by the user
- Obtaining a diverse and user-friendly delivery format

Center for Transportation Research and Education (CTRE) is developing a GIS-based accident location and analysis system for the state of Iowa, which facilitates spatial analyses of crash incidences. This system includes a comprehensive location-based database covering 10 years of traffic crashes on all road systems (58).

CHAPTER 5

STATISTICAL AND MULTICRITERIA DECISION ANALYSES OF TRAFFIC ACCIDENTS

5.1. Traffic Accident Data Used

In this study, accident data of the year 2003 collected by police officers in Turkey -taken from Ministry of Interior-General Directorate of Security of Turkey- are used as the basis. Actually, as emphasized before, three year period of data are generally required for a more reliable traffic accident analysis. However, since the coordinates of accident locations have been recorded for the last two years and data of the year 2004 are not ready to use, only one year data were analyzed during this study. On the other hand, this does not create a major problem because of the reason that the primary aim of this thesis study is to show the promising role of GIS, integrated with statistical and Multicriteria Decision Analyses in traffic accident analysis.

Although location-based traffic accident data have gained an incredibly valuable place and have been studied for a longer time in Europe and USA, Turkey -as like most of the developing countries- is quite behind the time, considering GIS-based accident studies. Since Turkey has recently started to collect and store accident coordinate data, the system has not yet had a concrete background.

Therefore, the traffic accident data collected in 2003 in Turkey have many missing and erroneous values. The most essential problem, indeed, comes from the inadequacy of the training given to the police officers. In view of the fact that collecting coordinate data is a new thing for our country, police officers should be well-informed about the usage of devices and method of true recording. Another problem lies in the importance given to the accidents. In Turkey, mostly, accidents without fatality or severe injury are not taken into consideration seriously and mostly not reported. Furthermore, traffic accidents are evaluated only at the scene; i.e. the change in the health conditions of people injured in the accidents are not followed and this causes lack of knowledge on the actual results of the accidents. Thus, accident reports cannot be considered as reflecting completely valid and true information. In order to develop more meaningful and useful safety measures, the problem of incomplete and inaccurate accident data recording should be solved with education, encouragement and enforcement legislations as soon as possible.

The data set used in this study was derived from the fatal or injury accidents occurring in 2003 in Turkey occurred on the state highways or motorways. The set includes 20162 traffic accidents and each of these accidents has its own characteristics, which are defined according to the traffic accident report form (as shown in Figure 5.1) filled for all accidents by police officers.

For the statistical analyses, the total number of accidents; i.e. 20162 traffic accident data, were used; however, the spatial analyses were performed only with 13184 traffic accidents. The rest, unfortunately, were eliminated because of the incompleteness or inaccuracy of the coordinate data. During the error checking and organization of spatial data, many of the actual data were tried to be corrected and so saved. For instance, some of the accidents not existing within the border of Turkey were placed to their accurate location with the help of their attribute data; or the confusion occurring due to reversely recording the x and y coordinates were adjusted systematically.

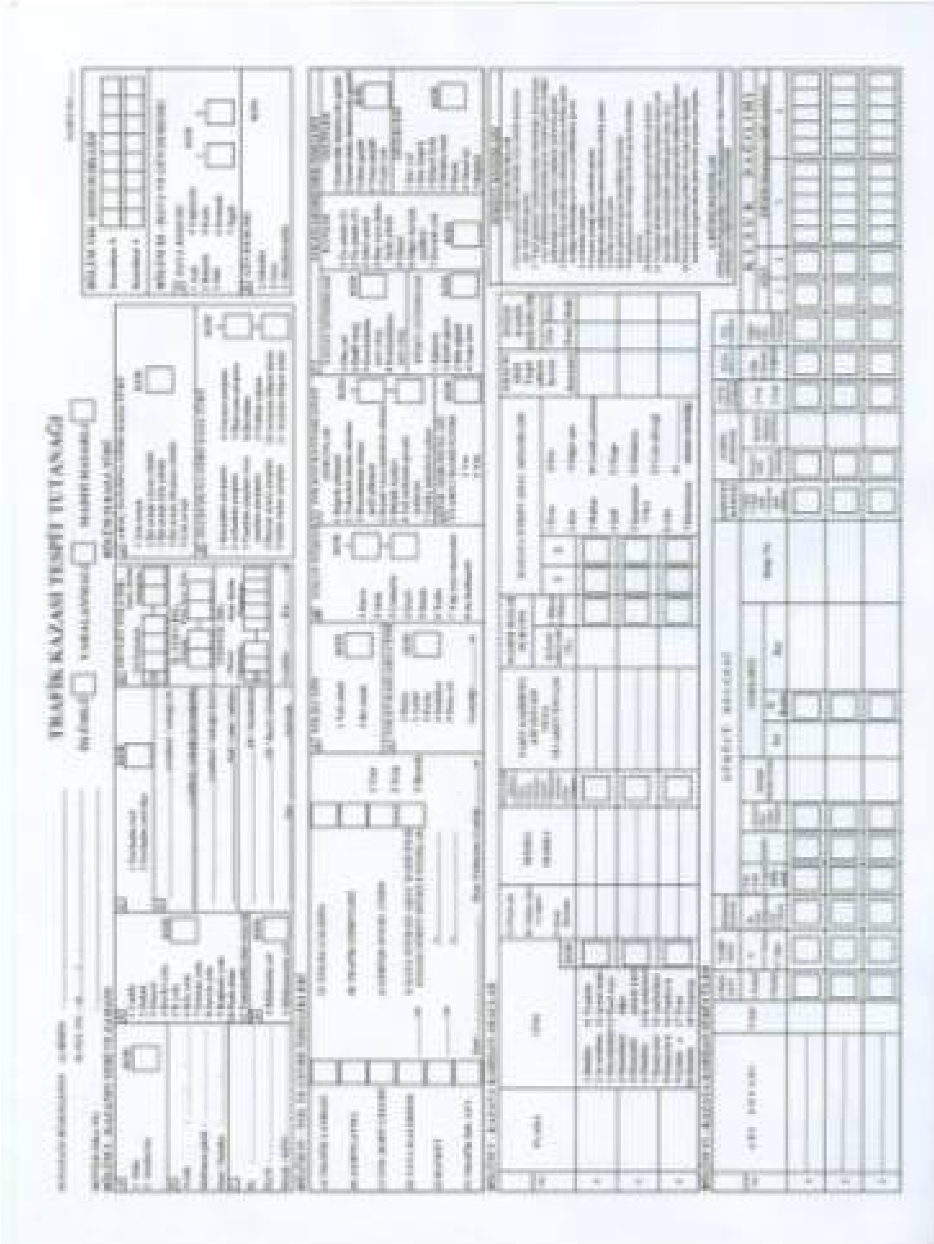

TRAFİK KAZASI TESPİT TUTANAĞI

Figure 5.1. Traffic Accident Report Form used in Turkey

Figure 5.1. Traffic Accident Report Form used in Turkey (continued)

5.2. Development of a Statistical Model of Traffic Accidents

Identifying the contribution of each variable to the accident severity is a special concern in order to examine the potential causal factors to traffic accidents. In this manner, there are many statistical studies in the literature, which investigate the factors reducing safety and efficiency of transportation systems and increasing number of traffic accidents. From this respect, regression methods constitute essential components of the traffic accident analysis by which it is possible to construct the relationship between the response variables (dependent variables) and the model variables (independent variables).

In most of the earlier researches, multiple linear regression based methods were used; however these methods are unsuitable for determining the distribution of accident events. Because of the fact that linear regression model assumes a continuous assumption on data and homogeneous assumption on error terms, the coefficients found from this hypothesis would fail to understand the significance of the parameters (59).

Consequently, the models such as Poisson, Negative Binomial and Logistic regression have been used for the analysis of traffic accident events. Because of the 'rare' and 'random' nature of the traffic accidents, the drawbacks associated with the application of the conventional multiple linear regression models in accident analysis can be overcome by the nonlinear regression models. However, the application of Poisson regression may be erroneous in certain situations where the mean deviates from the variance significantly (if variance is larger than the mean, overdispersion exists) (59).

5.2.1. Application of Logistic Regression

In this study, traffic accident data was studied with logistic regression, which better fits with the nature of the accident data. The reason of choosing logistic regression lies in structuring the outcome (response variable). If the response (dependent) variable is binary or dichotomous, then the dependent variable can take the value *1* with a probability of success π, or the value *0* with probability of failure $1 - \pi$. Hence, logistic regression can be successfully applied to the accident-related data. The logic behind the logistic model is basically designed to describe a probability, which is always some number between 0 and 1 (60). In traffic accident analysis, for instance, such a probability gives the risk of an accident resulting in fatality if the response variable is organized whether the accident is fatal or not.

In any of the analysis, the key quantity is the mean value of the response variable under the given values of the independent (predictor) variable;

$$E(Y/x) = \beta_0 + \beta_1 x$$

where Y denotes the response variable, x denotes a value of independent variable, and β values denote the model parameters or coefficients. The quantity is called the conditional mean or the expected value of Y given the value of x (61).

The two main differences between the linear and logistic regression is the relationship between the response and independent variables and the conditional distribution of the response variable. In logistic regression, conditional distribution of the response variable follows a binomial distribution with probability given by the conditional mean '$\pi(x)$' and conditional mean of the regression equation should be bounded between 0 and 1 (61).

The specific form of logistic regression model is;

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \qquad \text{where} \quad \pi(x) = E(Y/x)$$

The transformation of $\pi(x)$ logistic function is known as the logit transformation;

$$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 x$$

Briefly, if Y is coded as 0 or 1 (binary variable), the expression $\pi(x)$ provides the conditional probability that Y is equal to 1 given x, denoted as P(Y = 1 / x); and so the quantity 1- $\pi(x)$ gives the conditional probability that Y is equal to 0 given x, denoted as P(Y = 0 / x).

5.2.2. Description of the Statistical Model

Throughout the study, the risk factors were determined with the calculation of odds ratios (OR) within a confidence interval of 95 % (CI 95 %) through logistic regression. The OR is one of the statistics tools used to asses the risk of a particular outcome if a certain factor is present.

OR can be also defined as the *relative measure* of risk that is found as the ratio of the odds that the independent variable will be present, to the odds that it will not present. Odds, on the other hand, are simply the way of presenting probabilities. The odds of an event happening is the probability that the event will happen divided by the probability that the event will not happen.

The dependent variable in this study was taken as the result of traffic accident; and the associated independent variables (Table 5.1) were selected from the parameters related with location and constant characteristics of roadways and variable environmental parameters in order to emphasize the importance of improving the roadway factors, which would reduce the rate and severity of the traffic accidents.

As indicated in Table 5.1, the selected parameters are presented in three main levels for developing a hierarchical configuration with respect to the contributing independent variables, which also eases the construction of the logistic model.

In this manner, logistic regression was firstly applied to the third level (core level) and then the relationship within the upper levels (first and second levels) -both within themselves- were obtained with an assumption regarding the connection between the number of traffic accidents that are categorized according to their results and the obtained coefficient factors for the core level from logistic regression, which will be explained in the following section.

For logistic regression analysis, each accident in the model was firstly categorized with a severity index. An additional study was also performed by only changing the response variable as fatal/non-fatal categorization in order to have a different perspective. For overcoming the misleading random nature of the data, the severity index is constructed with a predetermined comparative factor. In Turkey, while one injury accident equals to '3 PDO' (property damage only accident), one fatality accident equals to '9 PDO' (62).

Table 5.1. Structure of the selected road and environmental parameters for the statistical analysis

	Level 1	Level 2	Level 3 (core level)
Traffic Accident Risk	Location Characteristics	Road Type	Motorway
			State Highway
		Urbanization	Rural
			Urban
	Road Characteristics	Divided/Undivided	Divided
			Undivided
		Road Coating	Surface Treatment
			Asphaltic Concrete
		Horizontal Curve	No curve
			Larger Curve
			Sharp Curve
		Vertical Curve	No Gradient
			Moderate Gradient
			High Gradient
		Intersection	None
			3-Leg
			4-Leg/5-Leg
			Rounded/Others
	Environmental Parameters	Road Surface	Dry
			Wet (wet-muddy-dusty-oily-watery)
			Snowy (snowy-icy)
		Weather Condition	Good (open)
			Medium (misty-cloudy-rainy)
			Tough (snowy-stormy-blizzard)
		Time of Day	Day
			Night
			Twilight

In this respect, the outcome of each accident was categorized according to a constant *importance* index of '9 to 3' (fatality/injury); simply '3/1'. For instance, if an accident results in two deaths and three injuries, the outcome was calculated as "2 * 3 + 3 *1 = 9". Then, outcomes were re-categorized to transform them into binary form by taking '3' as the critical threshold. This means that, if the calculated outcome values were equal to or greater than '3', the outcome was accepted as '1', otherwise '0'. This selected threshold represents an accident including at least one death or three injuries, and so it is defined as respectively more severe.

The response variable, namely accident result, has two criteria: for the severity index categorization, the response variable takes the value of *1* if the accident index is calculated as equal to or more than three, and *0* otherwise; for the fatal/non-fatal categorization, the response variable is assigned to *1* if accident is fatal and *0* if not.

As can be seen from Table 5.1, independent variables in the core level have several sub-levels, which should be coded according to the logic of the software used through the analysis. In this study, MINITAB (63) was used as the statistical software to find the related coefficients or OR's in the logistic model.

Defining 'k-1' design variables for 'k' levels (named as sub-levels in this study) of that variable is one possibility of coding the independent variables in the model (64). Table 5.2 represents an example for the accident time variable, which has three sub-levels (day, night, twilight) and therefore, two design variables (D_1 and D_2). According to the coding, one of these levels is identified as reference/base level, such as *day*, and all of the belonging values are set to zero if accident occurs in day time. If it occurs at *night*, D_1 takes the value of 1 and D_2 takes 0 and lastly, if accident occurs at *twilight*, D_1 takes the value of 0 and D_2 takes 1. This coding was applied to entire independent variables used in this study in order to find their effects on the accident severity or fatality.

Table 5.2. The design variables for accident time

time of day	D1	D2
Day	0	0
Night	1	0
Twilight	0	1

During the statistical analysis, the number of sub-levels was reduced rather than using all sub-levels of the independent variables. Some of the variables were neglected or at least merged if the data were not available, have statistically insignificant proportion, or if the accident data taken from the traffic police records were not coincident with the inventory data taken from the General Directorate of Highways. The latter one is due to the fact that roadway GIS layer was created according to the data obtained from the General Directorate of Highways. In fact, the aim of visualizing the likelihood of the more severe accidents with their spatial and attribute characteristics necessitates comparable data. Reduction of the variable levels also simplifies the model and makes the interpretation easier and more reliable. As a result, Table 5.1 shows the final (reduced) form of the selected independent variables according to their corresponding categories, extracted from the first four parts of the traffic accident report form (Figure 5.1) filled by police officers in Turkey.

5.2.3. Interpretation of the Statistical Model

When the coded data are analyzed, the statistical software gives the estimated coefficients for the independent variables in the model, which represent the slope or rate of change in the dependent variable per unit change in the independent variable. Hence, interpretation of the model involves both determining the functional relationship between the dependent variable and the independent variable and approximately defining the unit change for the independent variable (64).

In the logistic regression model, as stated previously, the slope coefficient in the logit transformation represents the change in the logit for a change of one unit in the independent variable x. In order to make an accurate interpretation of the coefficients in a logistic regression model, it is required to place a meaning on the difference between two logits. The exponent of this difference in logits gives the odds ratio (64).

In this study, OR greater than 1 shows the increase in the likelihood that the accident will result in at least one death or three injuries for the severity index categorization. On the other hand, for the fatal/non-fatal categorization, OR greater than 1 shows the increase in the likelihood that the accident will result in fatality.

Because of the reason that while some of the independent variables have two sub-levels (single design variable), the others have more than two sub-levels and so at least two design variables, one should be careful about finding the accurate odds ratio. While the value of OR's is directly equal to the exponential of the corresponding parameter estimated for the two sub-leveled variables, the logit difference is not necessarily equal to the parameter estimates of a polytomous variable with more than two sub-leveled ones.

For two sub-leveled independent variables:

Suppose that an accident is investigated according to whether it occurs at an intersection or not (logit function (y) = $\beta_0 + \beta_1$ x). The system is coded as '0' if it occurs at an intersection and '1' if not. The response variable has two values: '1' for fatal, '0' for non-fatal.

<u>independent variable (x):</u>	intersection = 0 ; non-intersection = 1
<u>dependent variable (y):</u>	fatal = 1 ; non-fatal = 0

logit (fatal accident/non-intersection) = $\beta_0 + \beta_1$

logit (fatal accident/intersection) = β_0

logit difference = $\beta_0 + \beta_1 - (\beta_0) = \beta_1$

odds ratio $= e^{\beta 1}$ (which means that the odds ratio of being in a fatal accident occurring at a non-intersection location is "β_1" times higher than those occurring at an intersection)

For more than two sub-leveled independent variables:

Suppose that, this time, an accident is investigated according to some probable causes; speed, run red light, wrong way, follow too close (logit function (y) = $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$). The system is coded as shown in the table below and the response variable has two values: '1' for fatal, '0' for non-fatal.

Table 5.3. The design variables for probable accident causes

accident cause	D1	D2	D3
speed	0	0	0
run red light	1	0	0
wrong way	0	1	0
follow too close	0	0	1

In this table speed refers to the reference/base level; run red light refers to x_1; wrong way refers to x_2 and follow too close refers to x_3.

logit (fatal accident/run red light) = $\beta_0 + \beta_1$

for any other cause but not running red light;

logit (fatal accident/not run red light) = $\beta_0 + \beta_2 + \beta_3$

logit difference = $\beta_0 + \beta_1 - (\beta_0 + \beta_2 + \beta_3) = \beta_1 - \beta_2 - \beta_3$

odds ratio $= e^{\beta 1 - \beta 2 - \beta 3}$ (which means that the odds ratio of being in a fatal accident because of running red light is "$\beta_1 - \beta_2 - \beta_3$" times higher than not running red light related accident)

5.2.4. Results of the Statistical Analysis

Regarding the basics of logistic regression, traffic accident data of Turkey for the year 2003 were processed with the software MINITAB (63) and the following results were obtained for two cases: categorization of the response variable using a severity index and separating it directly as fatal/non-fatal. In the logit model used in this study with the logit function (y) = $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_{17} x_{17}$, 18 coefficients existed for each of the response categorization (Table 5.4) and OR of each variable was calculated through these coefficients, as explained in the previous section. Thus, impacts of the independent variables -relative to each other- are presented in Table 5.5 and Table 5.6, respectively.

These tables also represent the relative risk factors of the variables, which can be reduced with effective location-based safety measures. For instance, the expression of 1.048 in the 3*3 'horizontal curve' matrix in Table 5.5 -representing the pairwise risk ratios between the types of horizontal curves- explains that the probability that an accident will be more severe according to the predefined severity index (i.e. at least one fatal or three injuries exist) if it occurs within a larger horizontal curve rather than flat road. In the same way, the expression of 1.472 in the 2*2 'road type' matrix in Table 5.6 -representing the pairwise risk ratios between the motorway and state highway- indicates that the probability that an accident will be fatal when it occurs at motorway is 1.472 times higher than if it occurs at state highway.

Table 5.4. Coefficients of the logit model used according to the response variables

Level 2	***Level 3 (core level)***	***Number of design variables***	***Coefficients according to the severity index (β_0 = -1.2202)***		***Coefficients according to the fatality condition (β_0 = -2.8812)***	
Road Type	Motorway State Highway	1	β_1	-0.15655	β_1	0.38691
Urbanization	Rural Urban	1	β_2	0.49705	β_2	0.28937
Divided / Undivided	Divided Undivided	1	β_3	-0.33493	β_3	-0.37458
Road Coating	Surface Treatment Asphaltic Concrete	1	β_4	0.14020	β_4	0.13870
Horizontal Curve	No curve Larger Curve Sharp Curve	2	β_5 β_6	0.04700 -0.00530	β_5 β_6	-0.13344 -0.20070
Vertical Curve	No Gradient Moderate Gradient High Gradient	2	β_7 β_8	0.00909 -0.14901	β_7 β_8	0.08965 -0.06020
Intersection	None 3-Leg 4-Leg/5-Leg Rounded/Others	3	β_9 β_{10} β_{11}	0.02806 0.05610 0.03616	β_9 β_{10} β_{11}	-0.50900 -0.61240 -0.33070
Road Surface	Dry Wet Snowy	2	β_{12} β_{13}	0.00669 0.21880	β_{12} β_{13}	-0.19313 -0.71040
Weather Condition	Good Medium Tough	2	β_{14} β_{15}	-0.05806 -0.11230	β_{14} β_{15}	0.05802 0.20660
Time of Day	Day Night Twilight	2	β_{16} β_{17}	0.06587 0.16161	β_{16} β_{17}	0.46411 0.48900

Table 5.5. Odds Ratios for the selected independent variables (level 3) affecting traffic accidents according to the predefined severity index

Location Characteristics

road type	**motorway**	**state highway**	
motorway	1.000	0.855	
state highway	1.169	1.000	
urbanization	**rural**	**urban**	
rural	1.000	1.644	
urban	0.608	1.000	

Road Characteristics

divided/ undivided	**divided**	**undivided**	
divided	1.000	0.715	
undivided	1.398	1.000	
road coating	**surface treatment**	**asphaltic concrete**	
surface treatment	1.000	1.151	
asphaltic concrete	0.869	1.000	
horizontal curve	**no curve**	**larger curve**	**sharp curve**
no curve	1.000	0.954	1.005
larger curve	1.048	1.000	1.054
sharp curve	0.995	0.949	1.000
vertical curve	**no gradient**	**moderate gradient**	**high gradient**
no gradient	1.000	0.991	1.161
moderate gradient	1.009	1.000	1.171
high gradient	0.862	0.854	1.000

Table 5.5. Odds Ratios for the selected independent variables (level 3) affecting traffic accidents according to the predefined severity index (continued)

intersection	**none**	**3-leg**	**4-leg / 5-leg**	**rounded / etc.**
none	1.000	0.972	0.945	0.964
3-leg	1.028	1.000	0.972	0.992
4-leg / 5-leg	1.058	1.028	1.000	1.020
rounded / etc.	1.037	1.008	0.980	1.000

Environmental Parameters

road surface	**dry**	**wet**	**snowy**
dry	1.000	0.993	0.803
wet	1.007	1.000	0.809
snowy	1.245	1.236	1.000
weather condition	**good**	**medium**	**tough**
good	1.000	1.060	1.119
medium	0.944	1.000	1.056
tough	0.894	0.947	1.000
time of day	**day**	**night**	**twilight**
day	1.000	0.936	0.851
night	1.068	1.000	0.909
twilight	1.175	1.100	1.000

Table 5.6. Odds Ratios for the selected independent variables (level 3) affecting traffic accidents according to the fatal/non-fatal categorization

Location Characteristics

road type	**motorway**	**state highway**
motorway	1.000	1.472
state highway	0.679	1.000

urbanization	**rural**	**urban**
rural	1.000	1.336
urban	0.749	1.000

Road Characteristics

divided/ undivided	**divided**	**undivided**
divided	1.000	0.688
undivided	1.454	1.000

road coating	**surface treatment**	**asphaltic concrete**
surface treatment	1.000	1.149
asphaltic concrete	0.870	1.000

horizontal curve	**no curve**	**larger curve**	**sharp curve**
no curve	1.000	1.143	1.222
larger curve	0.875	1.000	1.070
sharp curve	0.818	0.935	1.000

vertical curve	**no gradient**	**moderate gradient**	**high gradient**
no gradient	1.000	0.914	1.062
moderate gradient	1.094	1.000	1.162
high gradient	0.942	0.861	1.000

Table 5.6. Odds Ratios for the selected independent variables (level 3) affecting traffic accidents according to the fatal/non-fatal categorization (continued)

intersection	**none**	**3-leg**	**4-leg / 5-leg**	**rounded / etc.**
none	1.000	1.664	1.845	1.392
3-leg	0.601	1.000	1.109	0.837
4-leg / 5-leg	0.542	0.902	1.000	0.755
rounded / etc.	0.718	1.195	1.325	1.000

Environmental Parameters

road surface	**dry**	**wet**	**snowy**
dry	1.000	1.213	2.035
wet	0.824	1.000	1.677
snowy	0.491	0.596	1.000

weather condition	**good**	**medium**	**tough**
good	1.000	0.944	0.813
medium	1.060	1.000	0.862
tough	1.229	1.160	1.000

time of day	**day**	**night**	**twilight**
day	1.000	0.629	0.613
night	1.591	1.000	0.975
twilight	1.631	1.025	1.000

Since the independent variables (Table 5.1) were arranged in an expressive leveling order, each level was evaluated in itself. As stated before, logistic regression was directly applied to the core level (level 3); on the other hand, an assumption based on the relationship between the coefficients of the core level and the accident results used to find these coefficients, was applied in finding the relative factors for the first and second levels. Actually, when the variables in the first and second levels are investigated, it is seen that, constructing a meaningful and objective relationship like in the third level, is too hard. Therefore, the variables in the first and second levels can be assigned to be independent from each other and factorized as 1.

Although further studies of the analysis of the study were also performed with this independency approach, a different view (assumption) was proposed in order to use and benefit from the factors found from the logistic regression. At that point, it should be clarified that, the results attained in this study were extracted from the traffic accidents occurred in 2003 in Turkey; and so the entire risk values or risk maps in this study belong to that year. Hence, all of the computed factors were derived from the 2003 year accident data. If more data can be collected with a systematic manner in the future, more accurate and representative results can be reached, which will enable researchers to implement invaluable safety measures to the deficient locations.

In order to find the factors of the second level, first of all, the third (core) level factors derived from logistic regression was converted to a common scale. For this reason, the factor of one of the variables in each sub-level was assigned to '1' and then the others were found accordingly since ORs give the relative risk values in pairs. After that, these values were normalized and then, they were multiplied with the number of accidents of the corresponding sub-level, which is found according to the parameter selected as the response variable (i.e. both for predefined severity index and for fatal/non-fatal condition). Then, the second level factors were obtained by dividing the sum of all multiplied values by the total number of accidents that result in the selected response.

For instance, suppose that the probability that an accident will be fatal on undivided roads is 1.454 times higher than the ones on divided roads (OR, found by concerning fatality as response variable). When this ratio is normalized, divided road takes 0.407; undivided road takes the value of 0.593.

The total number of accidents with fatality is 1439; 588 of them occurred on divided roads and the rest 851 on undivided roads. Therefore, when the sum '0.407*588+0.593*851' is divided by the total number of accidents with fatality '1439', the value of 0.517 is obtained as a factor in the case of divided-undivided condition. The factors of the first level were calculated with similar logic; in that case, the normalized risk ratio values in each sub-level was also multiplied with their corresponding factors, which were found in the second level; and then, the average of the sum was divided by the total number of accidents according to the selected response variable. Figure 5.2 briefly summarizes the processes of these assumptions (equations used for each level).

Odd Ratio	A1	A2
A1	1	x
A2	y	1

x = 1 / y

t = 1 / z

Odd Ratio	B1	B2
B1	1	t
B2	z	1

level 1	**level 2**	**level 3**	**normalized risk ratio**	**total # of accidents (TA)**	**equation for level 2**	**equation for level 1**
AB	A	A1	**k=**1/(1+y)	ε	**v1=** (k*ε + m*μ) / TA	**v3=** [(k*ε + m* μ)*v1 + (n*δ + p*θ) *v2] / (2*TA)
		A2	**m=**y/(1+y)	μ		
	B	B1	**n=**1/(1+z)	δ	**v2=** (n*δ + p*θ) / TA	
		B2	**p=**z/(1+z)	θ		

Figure 5.2. The equations of assumption-based calculated factors for levels 1&2

Tables 5.7 and 5.9 show the required values used in the methodology and assumption-based calculated factors for the first and second levels according to the response variables; and Tables 5.8 and 5.10 show the relative risk values (risk ratios) obtained from the values in Tables 5.7 and 5.9, respectively.

Table 5.7. First and second level importance factors concerning the pre-determined severity index as the response variable

		normalized third level factors found from logistic regression	# of accidents according to the severity index used in core level (5953)	% (in severe)	total # of accidents (fatal+injured) (20162)	% (in total)	assumption-based calculated factors for the second level	assumption-based calculated factors for the first level
location characteristics	**road type**							0.301
	motorway	0.461	522	8.77	2093	10.38	0.532	
	state highway	0.539	5431	91.23	18069	89.62		
	urbanization							
	rural	0.622	4546	76.36	13744	68.17	0.564	
	urban	0.378	1407	23.64	6418	31.83		
road characteristics	**divided/undivided**							0.168
	divided	0.417	2413	40.53	9839	48.80	0.516	
	undivided	0.583	3540	59.47	10323	51.20		
	road coating							
	surface treatment	0.535	5882	98.81	19901	98.71	0.534	
	asphaltic concrete	0.465	71	1.19	261	1.29		
	horizontal curve							
	no curve	0.329	4632	77.81	16132	80.01	0.331	
	larger curve	0.344	971	16.31	2970	14.73		
	sharp curve	0.327	350	5.88	1060	5.26		

Table 5.7. First and second level importance factors concerning predetermined severity index as the response variable (continued)

	vertical curve							
	no gradient	0.348	4073	68.42	14170	70.28	0.347	
	moderate gradient	0.352	1582	26.57	4990	24.75		
	high gradient	0.300	298	5.01	1002	4.97		
	intersection							
	none	0.243	5076	85.27	16762	83.14	0.244	
	3-leg	0.249	283	4.75	1050	5.21		
	4-leg/5-leg	0.257	316	5.31	1278	6.34		
	rounded/others	0.251	278	4.67	1072	5.32		
environmental parameters	**road surface**							0.106
	dry	0.308	4521	75.94	15463	76.69	0.312	
	wet	0.310	1101	18.49	3766	18.68		
	snowy	0.383	331	5.56	933	4.63		
	weather condition							
	good	0.352	4172	70.08	14163	70.25	0.346	
	medium	0.333	1503	25.25	5164	25.61		
	tough	0.315	278	4.67	835	4.14		
	time of day							
	day	0.308	3547	59.58	12249	60.75	0.318	
	night	0.329	2131	35.80	7072	35.08		
	twilight	0.362	275	4.62	841	4.17		

Table 5.8. The relative risk factors for the first and second levels concerning the predetermined severity index as the response variable **

level 2					
location characteristics					
	road type	**urbanization**			
road type	1.000	0.943			
urbanization	1.060	1.000			
road characteristics					
	divided / undivided	**road coating**	**horizontal curve**	**vertical curve**	**intersection**
divided / undivided	1.000	0.965	1.557	1.487	2.113
road coating	1.036	1.000	1.613	1.540	2.189
horizontal curve	0.642	0.620	1.000	0.955	1.357
vertical curve	0.672	0.649	1.047	1.000	1.421
intersection	0.473	0.457	0.737	0.704	1.000
environmental parameters					
	road surface	**weather condition**	**time of day**		
road surface	1.000	0.903	0.981		
weather condition	1.107	1.000	1.086		
time of day	1.020	0.921	1.000		

level 1			
	location	**road**	**environmental**
location	1.000	1.789	2.835
road	0.559	1.000	1.585
environmental	0.353	0.631	1.000

** The values were all found from the transformation of the assumption-based calculated factors in Table 5.7 into a pairwise format.

Table 5.9. First and second level importance factors concerning fatality as the response variable

		normalized third level factors found from logistic regression	# of accidents with fatality (1439)	% (in fatal)	total # of accidents (fatal+injured) (20162)	% (in total)	assumption-based calculated factors for the second level	assumption-based calculated factors for the first level
location characteristics	**road type**							0.238
	motorway	0.596	189	13.13	2093	10.38	0.430	
	state highway	0.404	1250	86.87	18069	89.62		
	urbanization							
	rural	0.572	1111	77.21	13744	68.17	0.539	
	urban	0.428	328	22.79	6418	31.83		
road characteristics	**divided/undivided**							0.182
	divided	0.407	588	40.86	9839	48.80	0.517	
	undivided	0.593	851	59.14	10323	51.20		
	road coating							
	surface treatment	0.535	1422	98.82	19901	98.71	0.534	
	asphaltic concrete	0.465	17	1.18	261	1.29		
	horizontal curve							
	no curve	0.371	1152	80.06	16132	80.01	0.361	
	larger curve	0.325	213	14.80	2970	14.73		
	sharp curve	0.304	74	5.14	1060	5.26		

Table 5.9. First and second level importance factors concerning fatality as the response variable (continued)

	vertical curve							
	no gradient	0.329	980	68.10	14170	70.28	0.337	
	moderate gradient	0.360	388	26.96	4990	24.75		
	high gradient	0.310	71	4.93	1002	4.97		
	intersection							
	none	0.349	1305	90.69	16762	83.14	0.337	
	3-leg	0.210	43	2.99	1050	5.21		
	4-leg/5-leg	0.189	42	2.92	1278	6.34		
	rounded/others	0.251	49	3.41	1072	5.32		
environmental parameters	**road surface**							0.120
	dry	0.432	1142	79.36	15463	76.69	0.411	
	wet	0.356	249	17.30	3766	18.68		
	snowy	0.212	48	3.34	933	4.63		
	weather condition							
	good	0.304	1024	71.16	14163	70.25	0.311	
	medium	0.322	362	25.16	5164	25.61		
	tough	0.374	53	3.68	835	4.14		
	time of day							
	day	0.237	722	50.17	12249	60.75	0.307	
	night	0.377	638	44.34	7072	35.08		
	twilight	0.386	79	5.49	841	4.17		

Table 5.10. The relative risk factors for the first and second levels concerning fatality as the response variable**

level 2					
location characteristics					
	road type	**urbanization**			
road type	1.000	0.797			
urbanization	1.255	1.000			
road characteristics					
	divided / undivided	**road coating**	**horizontal curve**	**vertical curve**	**intersection**
divided / undivided	1.000	0.968	1.432	1.535	1.533
road coating	1.033	1.000	1.479	1.585	1.583
horizontal curve	0.698	0.676	1.000	1.072	1.070
vertical curve	0.652	0.631	0.933	1.000	0.999
intersection	0.652	0.632	0.934	1.001	1.000
environmental parameters					
	road surface	**weather condition**	**time of day**		
road surface	1.000	1.322	1.339		
weather condition	0.756	1.000	1.013		
time of day	0.747	0.987	1.000		

level 1			
	location	**road**	**environmental**
location	1.000	1.306	1.978
road	0.766	1.000	1.515
environmental	0.506	0.660	1.000

** The values were all found from the transformation of the assumption-based calculated factors in Table 5.9 into a pairwise format.

5.3. Multicriteria Decision Analysis of the Traffic Accident Data

The values computed in the previous sections for the first, second and third levels are the pairwise (relative) factors within their levels. In order to find the direct risk factor of each independent variable; i.e. weight of the potential causal factors contributing to the occurrence of traffic accidents, it is required to integrate all of these values into a single model. This model can be achieved by Multicriteria Decision Analysis (MCDA) methods.

Decision analysis is a set of systematic procedures in order to analyze complex decision problems. It is based on the division of the decision problem into small understandable parts. Then, these parts are analyzed and integrated in a logical manner for a meaningful solution. Analytic Hierarchy Process (AHP), which was developed by Saaty (1980), is one of the most important decision aiding method that makes a pairwise comparison in order to create a ratio matrix. AHP, principally, takes the pairwise comparisons as an input and produces the corresponding weights (importance factors) as output. The comparisons are performed subjectively on scale from 1 to 9. Finally, AHP normalizes the eigenvector associated with the maximum eigenvalue of each ratio matrix -reciprocal- (44).

In addition to its aim of quantifying relative priorities for a given set of alternatives, AHP also judges the consistency of the comparison of alternatives in the decision making process (65). The strength of this approach comes from its ability to organize tangible and intangible factors in a systematic way, and to provide a structured yet relatively simple solution to the decision making problems. (66).

In this step of the study, 14 matrices ("10" of them for the third (core) level, "3" for the second level and "1" for the first level) were constructed for each of the accepted response variables; i.e. according to severity index and fatality condition. Then, required eigenvalues were obtained and the eigenvectors fitting to the maximum eigenvalue were determined through a mathematical software, MathCAD (67). The importance weights of each variable were reached after normalizing the lastly found eigenvectors. Figure 5.3 exemplifies this procedure for one of the categories existing in the second level by taking the severity index as the response variable. The final values in that figure represent the importance weights (risk factors in the occurrence of accidents) obtained for the corresponding independent variables, respectively.

All of these operations were carried out with the logic used in AHP, with a difference in the method of comparing the variables; i.e. in giving the pairwise risk ratios to the variables. Although AHP makes the comparison in a subjective manner, the variables in this study were directly found from statistical analysis based on real traffic accident data that significantly increase the consistency of the risk factors. With this innovative model, contributing weights of the potential causes of traffic accidents, which increase the risk with respect to both predetermined severity index and fatality condition, were found. Besides, through accepting the assumption of independency for all of the pairwise risk values for the second and first levels, these factors were re-determined with the same procedure.

Tables 5.11 and 5.12 indicate AHP factors of the independent / predictor variables and final form of these risk factors (in percentage) with respect to the response variables for each level, respectively. In addition, Tables 5.13 and 5.14 represent the same results with the independency approach; i.e. equal importance between the parameters within the first and second levels.

Lastly, Figure 5.4 demonstrates the flowchart of all of these steps (including statistical and Multicriteria Decision Analyses) for an easy interpretation of the process followed during this study.

Eigenvalues:

$$\text{eigenvals}\left(\begin{pmatrix} 1 & 0.903 & 0.981 \\ 1.107 & 1 & 1.086 \\ 1.020 & 0.921 & 1 \end{pmatrix}\right) = \begin{pmatrix} 3.0001 \\ -0.0004 \\ 0.0002 \end{pmatrix}$$

Eigenvector corresponding to the maximum eigenvalue:

$$\text{eigenvec}\left[\begin{pmatrix} 1 & 0.903 & 0.981 \\ 1.107 & 1 & 1.086 \\ 1.020 & 0.921 & 1 \end{pmatrix}, 3.0001\right] = \begin{pmatrix} 0.5534 \\ 0.6127 \\ 0.5643 \end{pmatrix}$$

Normalization of the values taken from the corresponding eigenvector:

Road Surface:
$$\frac{0.5534}{0.5534 + 0.6127 + 0.5643} = 0.3198$$

Weather Condition:
$$\frac{0.6127}{0.5534 + 0.6127 + 0.5643} = 0.3541$$

Time of Day:
$$\frac{0.5643}{0.5534 + 0.6127 + 0.5643} = 0.3261$$

Figure 5.3. An example of finding the importance weights of the variables with AHP (for the second level; according to the severity index response)

At that point, GIS gains a significant place in order to visualize and interpret the whole picture and so to apply the most suitable and effective countermeasure to the high risked location. In fact, GIS offers a unique role for managing and analyzing a variety of spatial data organized with statistical and Multicriteria Decision Analyses.

Table 5.11. AHP factors of the predictor variables with respect to the response variables for each level

	Level 1	severity index	fatality condition	Level 2	severity index	fatality condition	Level 3	severity index	fatality condition
Results of AHP *according to response variables of 'severity index' and 'fatality condition'*	Location Characteristics	0.5231	0.4402	Road Type	0.4854	0.4435	Motorway	0.461	0.5955
							State Highway	0.539	0.4045
				Urbanity	0.5146	0.5565	Rural	0.6218	0.5718
							Urban	0.3782	0.4282
	Road Characteristics	0.2924	0.3372	Divided/ Undivided	0.2615	0.2478	Divided	0.4170	0.4075
							Undivided	0.5830	0.5925
				Road Coating	0.2709	0.2560	Surface Treatment	0.5351	0.5347
							Asphaltic Concrete	0.4649	0.4653
				Horizontal Curve	0.1680	0.1730	No Curve	0.3286	0.3713
							Larger Curve	0.3445	0.3249
							Sharp Curve	0.3269	0.3038
				Vertical Curve	0.1758	0.1615	No Gradient	0.3484	0.3294
							Moderate Gradient	0.3514	0.3604
							High Gradient	0.3002	0.3102
				Intersection	0.1238	0.1616	None	0.2425	0.3495
							3-Leg	0.2494	0.2101
							4-Leg/5-Leg	0.2566	0.1894
							Rounded/Others	0.2515	0.2510
	Environmental Parameters	0.1845	0.2226	Road Surface	0.3198	0.3995	Dry	0.3075	0.4319
							Wet	0.3097	0.3559
							Snowy	0.3828	0.2122
				Weather Condition	0.3541	0.3022	Good	0.3524	0.3040
							Medium	0.3326	0.3222
							Tough	0.3150	0.3738
				Time of Day	0.3261	0.2983	Day	0.3083	0.2369
							Night	0.3293	0.3768
							Twilight	0.3623	0.3863

Table 5.12. The final form of risk factors (in percentage) with respect to the response variables in each level

	Level 1	severity index	fatality condition	*Level 2*	severity index	fatality condition	*Level 3*	severity index	fatality condition
Traffic Accident Risk (accepted as 1 unit) *according to response variables of 'severity index' and 'fatality condition'*	Location Characteristics	52.31	44.02	Road Type	25.391	19.523	Motorway	11.705	11.626
							State Highway	13.686	7.897
				Urbanity	26.919	24.497	Rural	16.738	14.007
							Urban	10.181	10.490
	Road Characteristics	29.24	33.72	Divided/ Undivided	7.646	8.356	Divided	3.188	3.405
							Undivided	4.458	4.951
				Road Coating	7.921	8.632	Surface Treatment	4.239	4.616
							Asphaltic Concrete	3.683	4.017
				Horizontal Curve	4.912	5.834	No Curve	1.614	2.166
							Larger Curve	1.692	1.895
							Sharp Curve	1.606	1.772
				Vertical Curve	5.140	5.446	No Gradient	1.791	1.794
							Moderate Gradient	1.806	1.963
							High Gradient	1.543	1.689
				Intersection	3.620	5.449	None	0.878	1.904
							3-Leg	0.903	1.145
							4-Leg/5-Leg	0.929	1.032
							Rounded/Others	0.910	1.368
	Environmental Parameters	18.45	22.26	Road Surface	5.900	8.893	Dry	1.814	3.841
							Wet	1.827	3.165
							Snowy	2.259	1.887
				Weather Condition	6.533	6.727	Good	2.302	2.045
							Medium	2.173	2.167
							Tough	2.058	2.515
				Time of Day	6.017	6.640	Day	1.855	1.573
							Night	1.981	2.502
							Twilight	2.180	2.565

Table 5.13. AHP factors of the predictor variables with respect to the response variables for each level (level 1&2 independency)

	Level 1	severity index	fatality condition	*Level 2*	severity index	fatality condition	*Level 3*	severity index	fatality condition
Results of AHP **"with the independency approach ; (equal importance) between the parameters within level 1 and level 2"** *according to response variables of 'severity index' and 'fatality condition'*	Location Characteristics	0.333	0.333	Road Type	0.5	0.5	Motorway	0.4610	0.5955
							State Highway	0.5390	0.4045
				Urbanity	0.5	0.5	Rural	0.6218	0.5718
							Urban	0.3782	0.4282
	Road Characteristics	0.333	0.333	Divided/ Undivided	0.2	0.2	Divided	0.4170	0.4075
							Undivided	0.5830	0.5925
				Road Coating	0.2	0.2	Surface Treatment	0.5351	0.5347
							Asphaltic Concrete	0.4649	0.4653
				Horizontal Curve	0.2	0.2	No Curve	0.3286	0.3713
							Larger Curve	0.3445	0.3249
							Sharp Curve	0.3269	0.3038
				Vertical Curve	0.2	0.2	No Gradient	0.3484	0.3294
							Moderate Gradient	0.3514	0.3604
							High Gradient	0.3002	0.3102
				Intersection	0.2	0.2	None	0.2425	0.3495
							3-Leg	0.2494	0.2101
							4-Leg/5-Leg	0.2566	0.1894
							Rounded/Others	0.2515	0.2510
	Environmental Parameters	0.333	0.333	Road Surface	0.333	0.333	Dry	0.3075	0.4319
							Wet	0.3097	0.3559
							Snowy	0.3828	0.2122
				Weather Condition	0.333	0.333	Good	0.3524	0.3040
							Medium	0.3326	0.3222
							Tough	0.315	0.3738
				Time of Day	0.333	0.333	Day	0.3083	0.2369
							Night	0.3293	0.3768
							Twilight	0.3623	0.3863

Table 5.14. The final form of risk factors (in percentage) with respect to response variables in each level (level 1&2 independency)

	Level 1	severity index	fatality condition	*Level 2*	severity index	fatality condition	*Level 3*	severity index	fatality condition
Traffic Accident Risk (accepted as 1 unit)	Location Characteristics	33.3	33.3	Road Type	16.650	16.650	Motorway	7.676	9.915
"with the independency approach ; (equal importance) between the parameters within level 1 and level 2"							State Highway	8.974	6.735
according to response variables of 'severity index' and 'fatality condition'				Urbanity	16.650	16.650	Rural	10.353	9.520
							Urban	6.297	7.130
	Road Characteristics	33.3	33.3	Divided/ Undivided	6.660	6.660	Divided	2.777	2.714
							Undivided	3.883	3.946
				Road Coating	6.660	6.660	Surface Treatment	3.564	3.561
							Asphaltic Concrete	3.096	3.099
				Horizontal Curve	6.660	6.660	No Curve	2.188	2.473
							Larger Curve	2.294	2.164
							Sharp Curve	2.177	2.023
				Vertical Curve	6.660	6.660	No Gradient	2.320	2.194
							Moderate Gradient	2.340	2.400
							High Gradient	1.999	2.066
				Intersection	6.660	6.660	None	1.615	2.328
							3-Leg	1.661	1.399
							4-Leg/5-Leg	1.709	1.261
							Rounded/Others	1.675	1.672
	Environmental Parameters	33.3	33.3	Road Surface	11.089	11.089	Dry	3.410	4.789
							Wet	3.434	3.947
							Snowy	4.245	2.353
				Weather Condition	11.089	11.089	Good	3.908	3.371
							Medium	3.688	3.573
							Tough	3.493	4.145
				Time of Day	11.089	11.089	Day	3.419	2.627
							Night	3.652	4.178
							Twilight	4.018	4.284

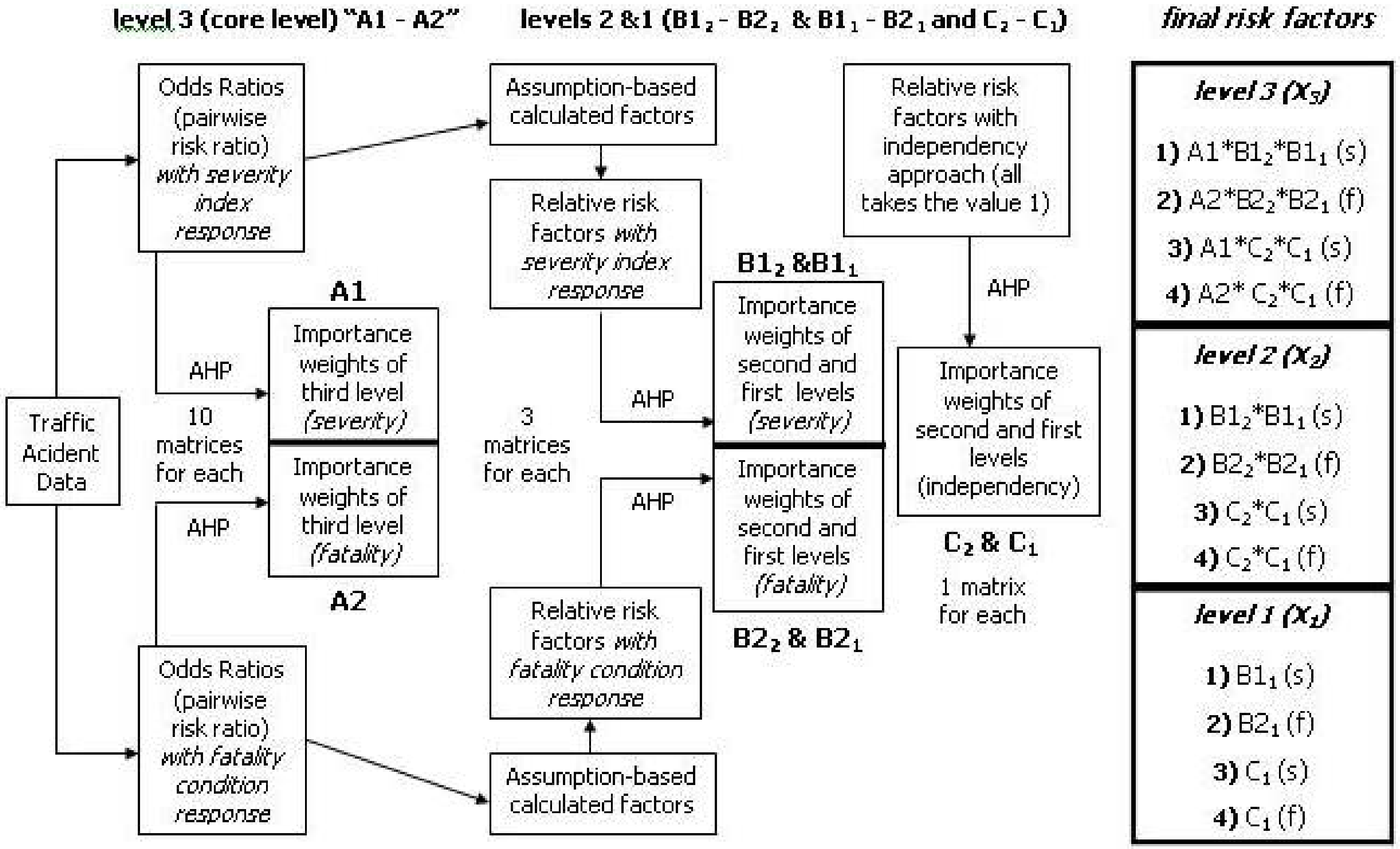


Figure 5.4. Flowchart of the processes followed during the statistical and Multicriteria Decision Analyses of traffic accident data

CHAPTER 6

INTEGRATION OF SPATIAL AND STATISTICAL ANALYSES FOR TRAFFIC ACCIDENT STUDIES

6.1. Importance of Constructing a Hybrid between Statistical and Spatial Analyses

Generally, statistical methods are utilized for the analysis of traffic accidents. However, characteristics of traffic accident data are highly suitable for spatial analyses and it is easier and more efficient to comprehend, interpret and analyze spatially distributed data. From this respect, GIS is required for geographical representation of the accident data (as a mapping tool) and for performing effective spatial analyses, which integrate the information from different sources in a single pot, especially by using the overlaying advantage of GIS.

The aim of this study is to develop a model to analyze the nature of traffic accidents using both statistical and spatial analyses methods in order to determine the characteristics of the accidents by which it is possible to propose effective remedies and to perform necessary future research. Multicriteria Decision Analysis methods were also performed in order to organize the results of statistical analysis in a logical manner.

In fact, constructing a hybrid of statistical analysis, Multicriteria Decision Analysis and spatial analysis will be helpful to integrate the statistical studies with the geographic information systems, and this ensures the development of location-based safety measures. Therefore, this study has a special goal of introducing the advantages of GIS into the area of traffic accident analysis and so showing/emphasizing the importance of studying on spatially distributed data incorporated with an accurate and logical statistical distribution.

6.2. Spatial Analysis of the Traffic Accident Data

The algorithm developed to estimate the spatial distribution of traffic accidents has a significant importance for assessing the risk zones of traffic accidents. While statistical analyses only give a general idea about the distribution of traffic accidents (mostly because of numeric representation), spatial analyses support this with the location data, which permit to obtain zonal distribution.

The mapping and displaying capabilities of GIS ensure visualization and better interpretation of high accident areas. In addition, the ability of simultaneously displaying two or more base maps enables the user to realize the risky zones regarding the severity of accidents. It would be possible to extract any required information by constructing queries. These location-based queries would be helpful to perceive and interpret the accelerating trend and patterns of the accidents within the consideration of high crash locations. In fact, queries can be enhanced to any level if more detailed data are available.

Within the borders of this study, firstly, some queries were developed by using the 13184 traffic accident data of the year 2003 in Turkey. All of these accidents only cover the ones that occurred at **motorways** or at **state highways** of our country with injury and/or fatality. Figure 6.1 shows the distribution of all these accidents as a GIS layer. All of the spatial analyses in this study were performed with ArcGIS (68) software.

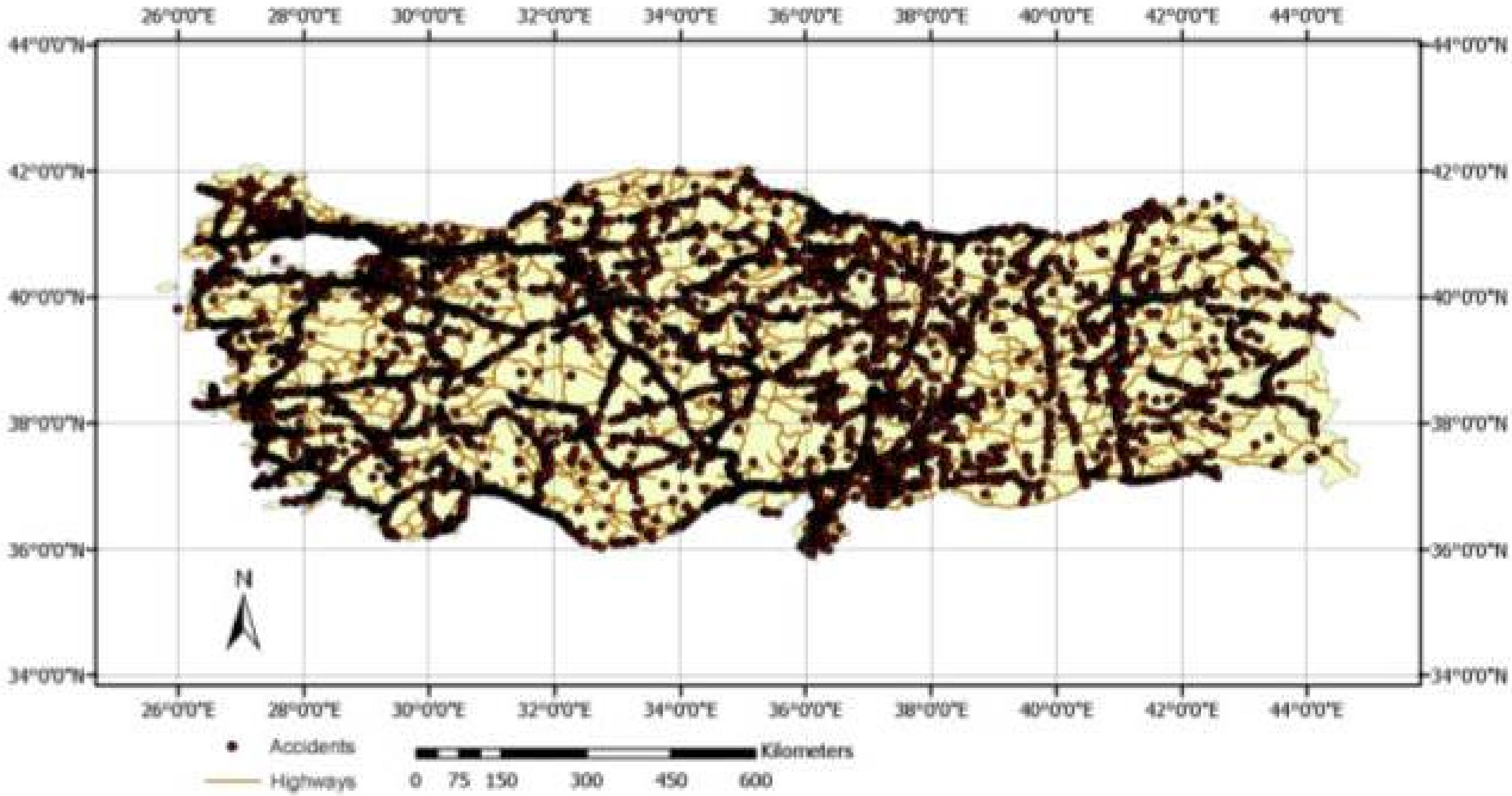


Figure 6.1. Distribution of traffic accidents at motorways and state highways of Turkey in 2003

In each query, the whole accident data were either clustered and mapped or just mapped according to a selected attribute. Some of these query results are presented in Figures 6.2 to 6.9; each representing the distribution of the selected data within the country. In addition to single attribute queries, multi-attribute queries were also constructed to give a better picture of GIS and its promising use in accident analysis. Results of these queries were found to be interesting and are also mapped in Figures 6.10 to 6.14.

6.3. An Innovative Model Integrating Spatial and Statistical Analyses

It is for sure that two distinct areas of research can benefit from each other if they are integrated in an effective manner and thus, invaluable solutions can be achieved. Considering the traffic accident problem, it is possible to implement the most appropriate location-based safety measures with the integration of spatial and statistical analyses supported by Multicriteria Decision Analysis.

The primary aim of safety experts is to make roadways safer by improving engineering measurements or by traffic monitoring. However, it is not always efficient or feasible to implement safety improvements through the whole city/country mostly due to random occurrence of traffic accidents and variable characteristics of roadways. Therefore, it is more meaningful to apply the safety improvement programs in those areas/zones where accident occurrence is high within a considerably small area (8). In this respect, high risk zones of traffic accidents are tried to be determined within the concept of this thesis study.

In order to realize the distribution of accidents and the accident risk, GIS is used in locating traffic accidents on a digital map. While regression analysis is mainly applied to interpret the risk factors (weights) of probable potential causes giving rise to the occurrence of traffic accidents, spatial analysis performed with GIS specifies the locations of the risky areas according to these weight factors.

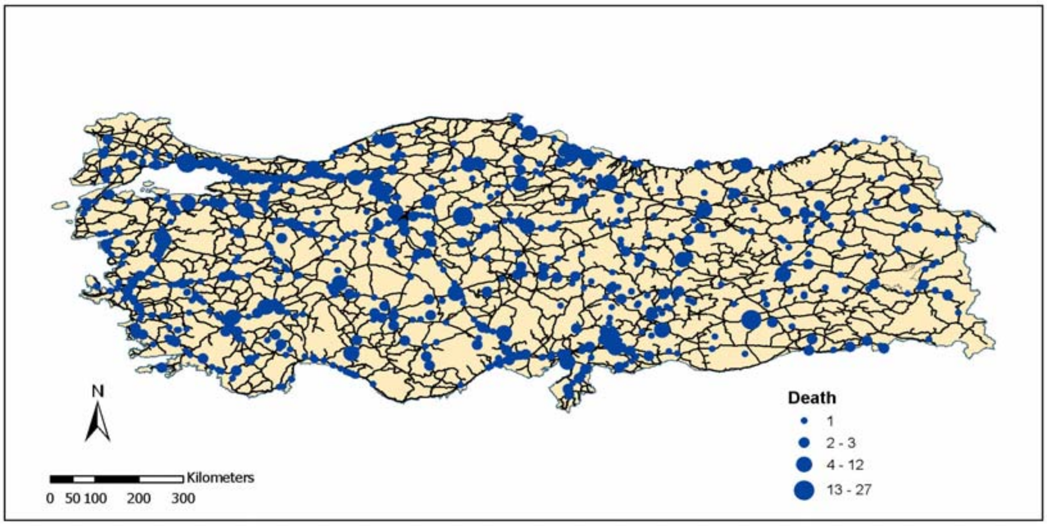


Figure 6.2. Distribution of traffic accidents with fatality in 2003 in Turkey

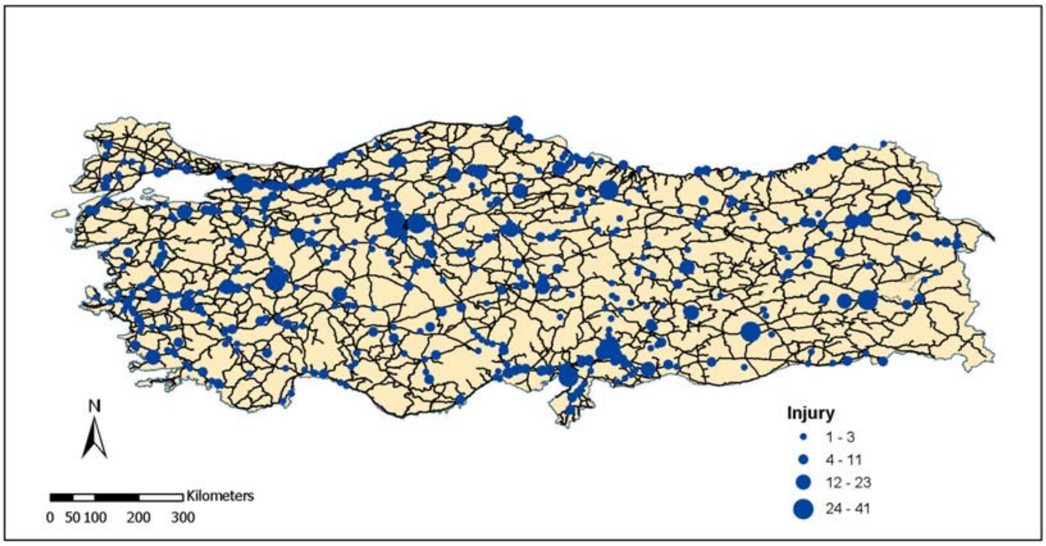


Figure 6.3. Distribution of traffic accidents with injury in 2003 in Turkey

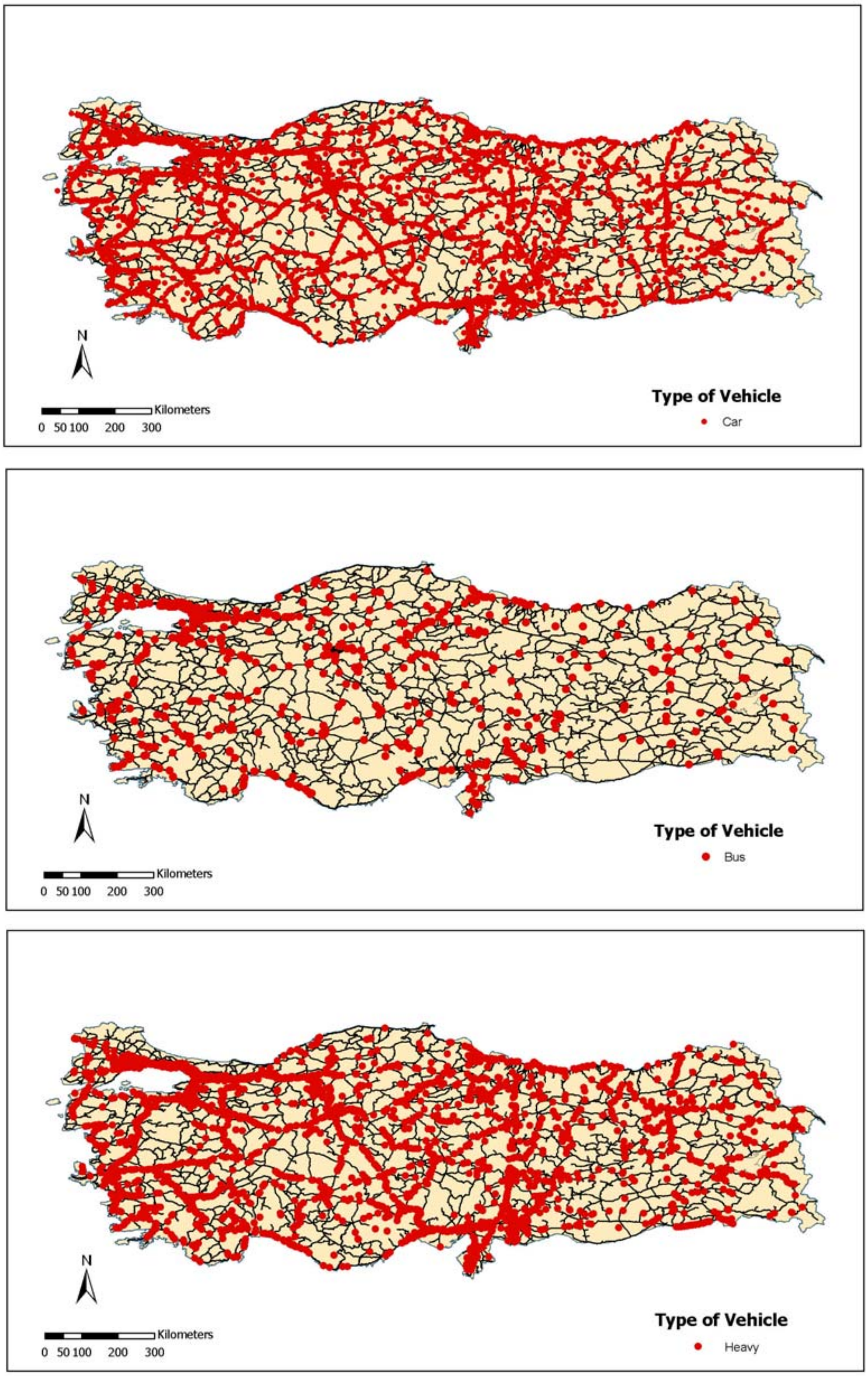


Figure 6.4. Distribution of vehicular traffic accidents in 2003 in Turkey

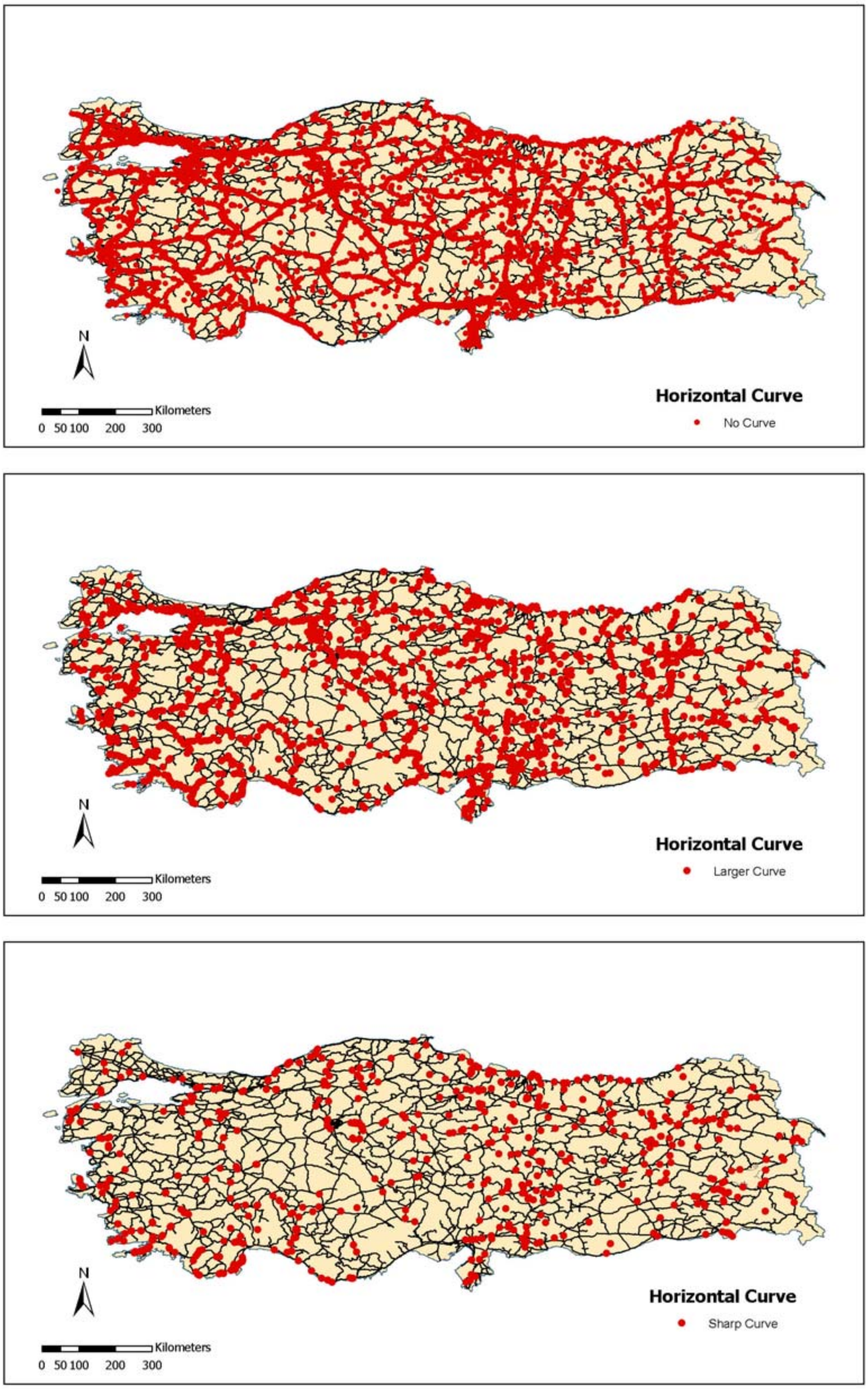


Figure 6.5. Distribution of accidents among horizontal curves in 2003 in Turkey

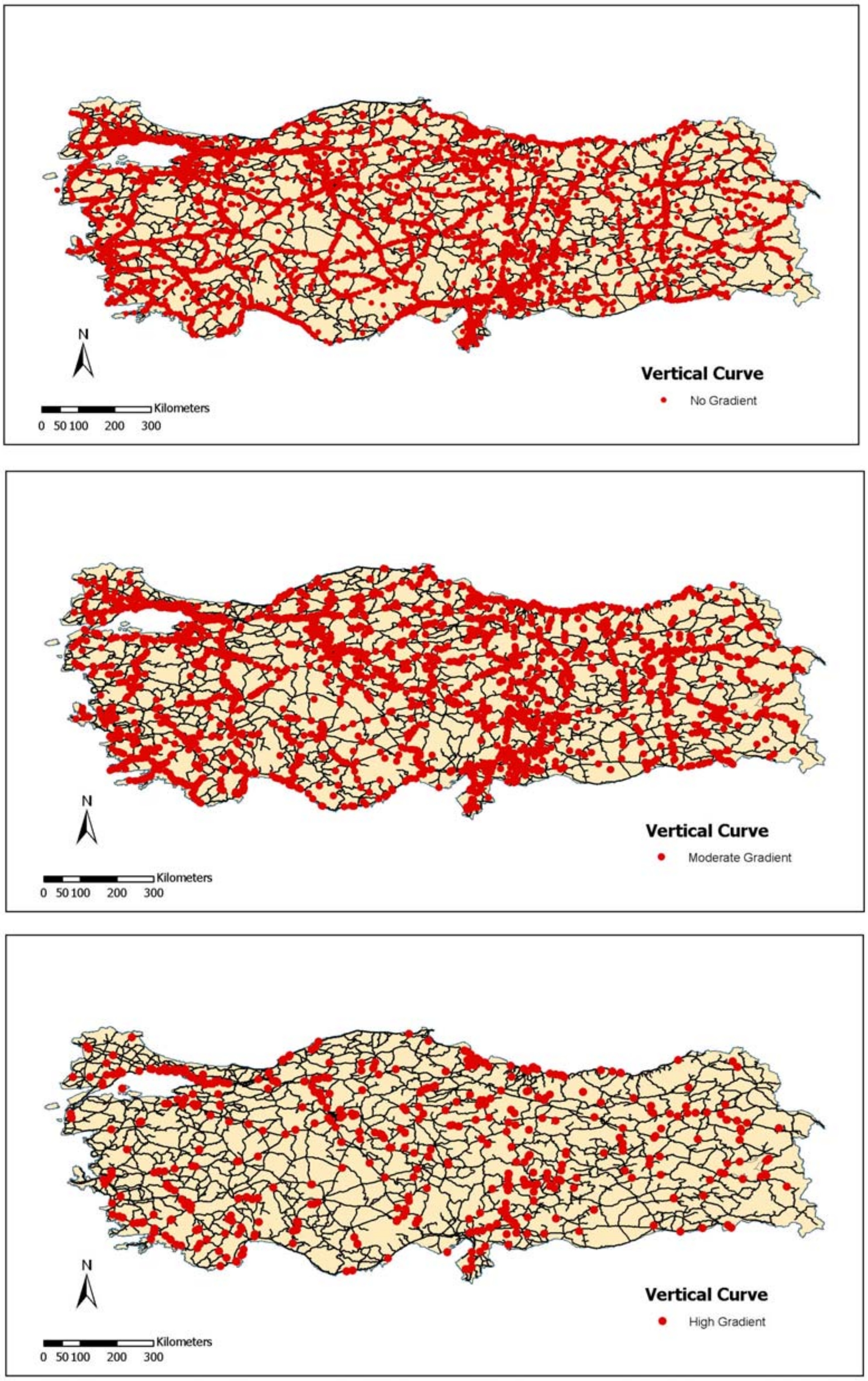


Figure 6.6. Distribution of accidents among vertical curves in 2003 in Turkey

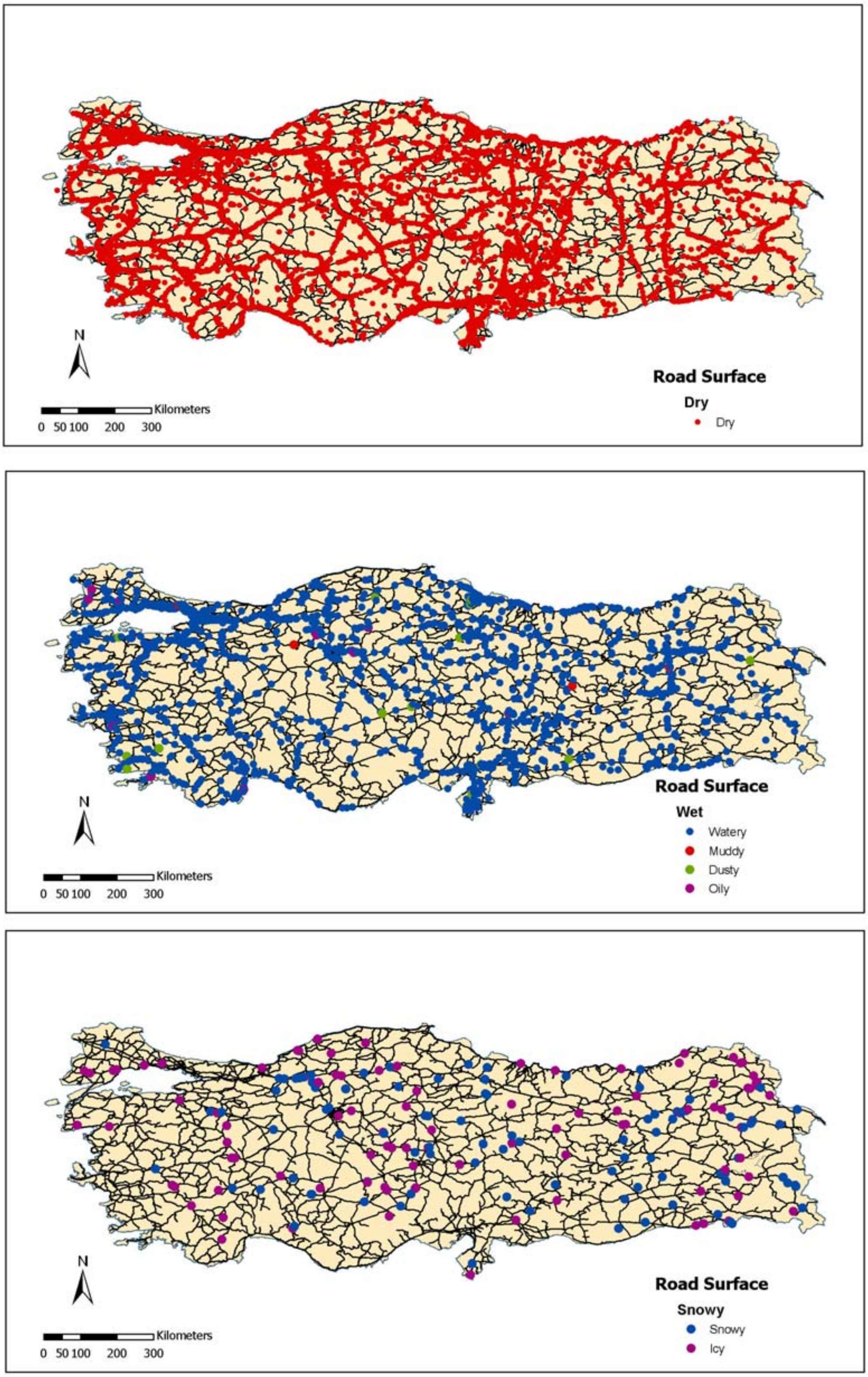


Figure 6.7. Distribution of accidents among road surface in 2003 in Turkey

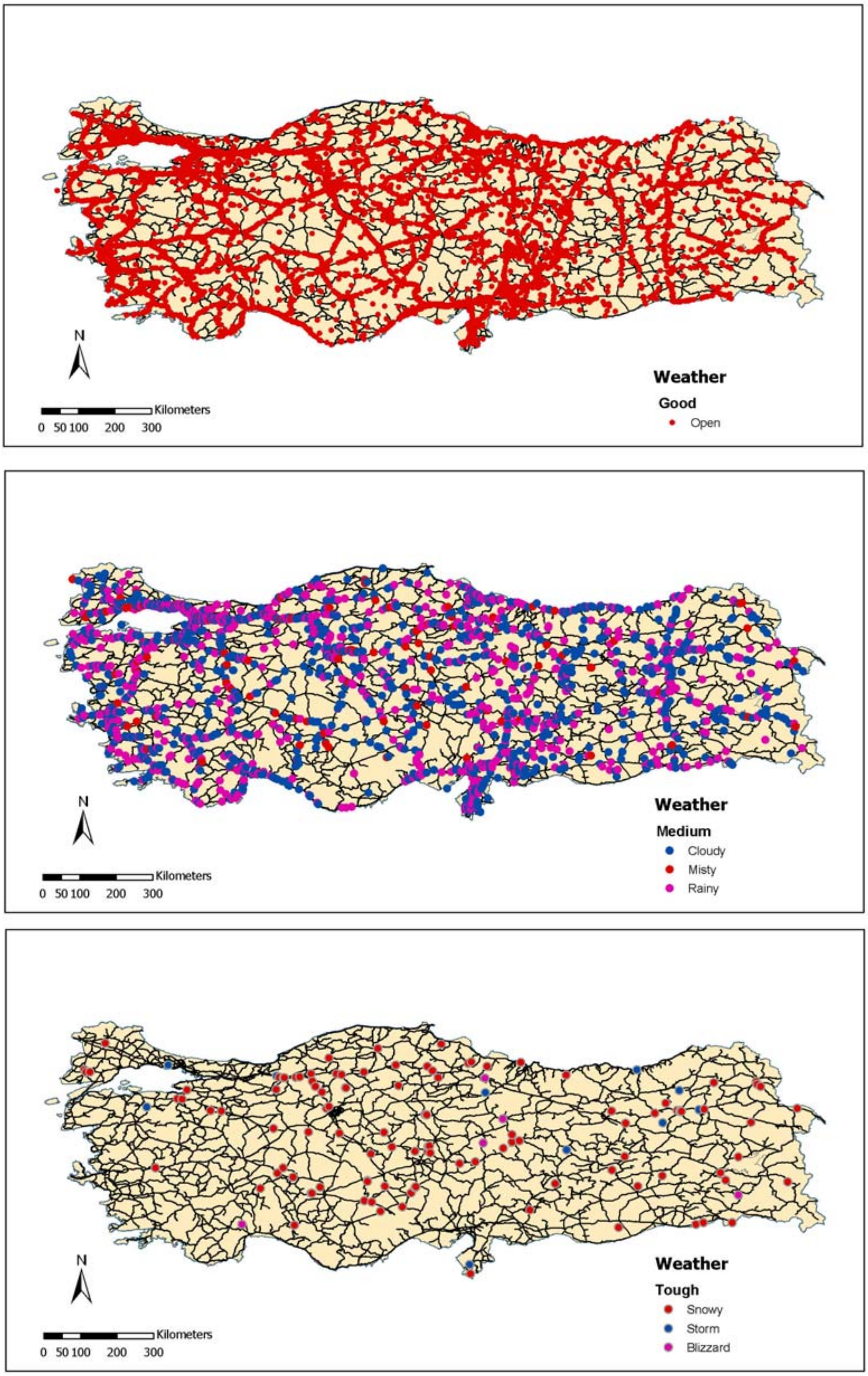


Figure 6.8. Distribution of accidents among weather condition in 2003 in Turkey

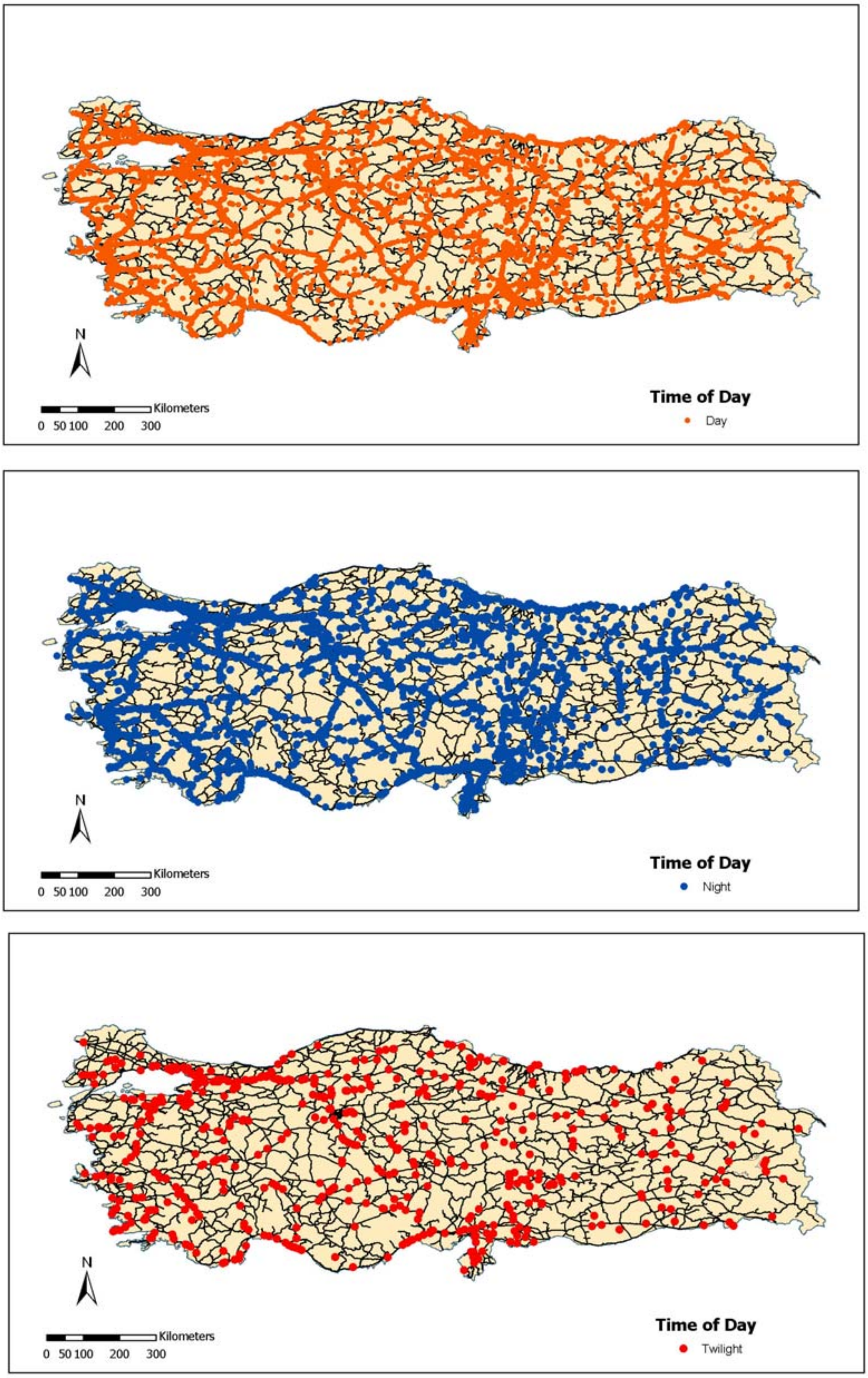


Figure 6.9. Distribution of accidents among time of day in 2003 in Turkey

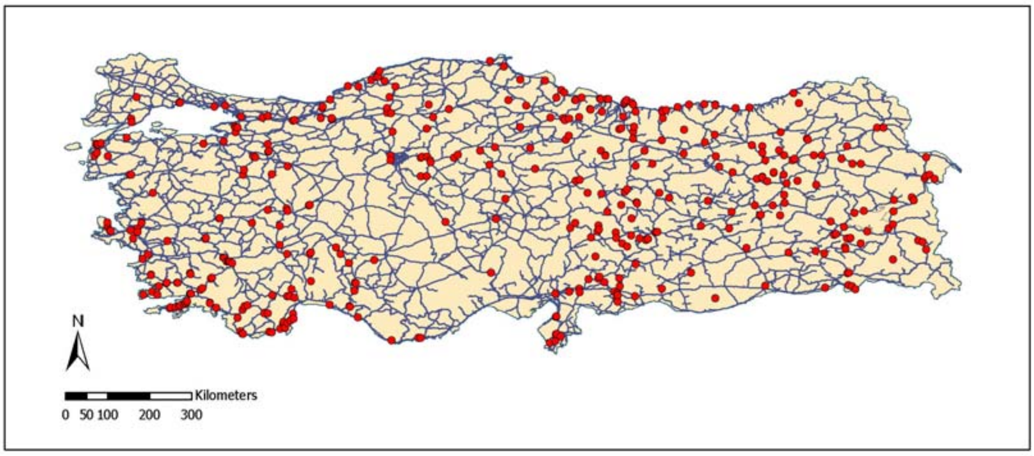


Figure 6.10. Distribution of head-on, sideswipe and rear-end accidents at sharp curves in 2003 in Turkey

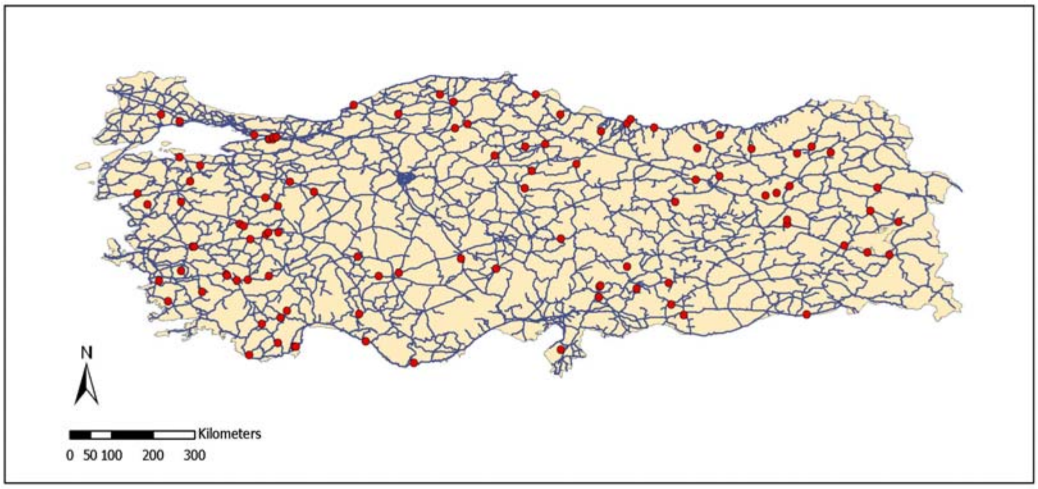


Figure 6.11. Distribution of head-on collisions on undivided roads and in the case of non-existent lane marking in 2003 in Turkey

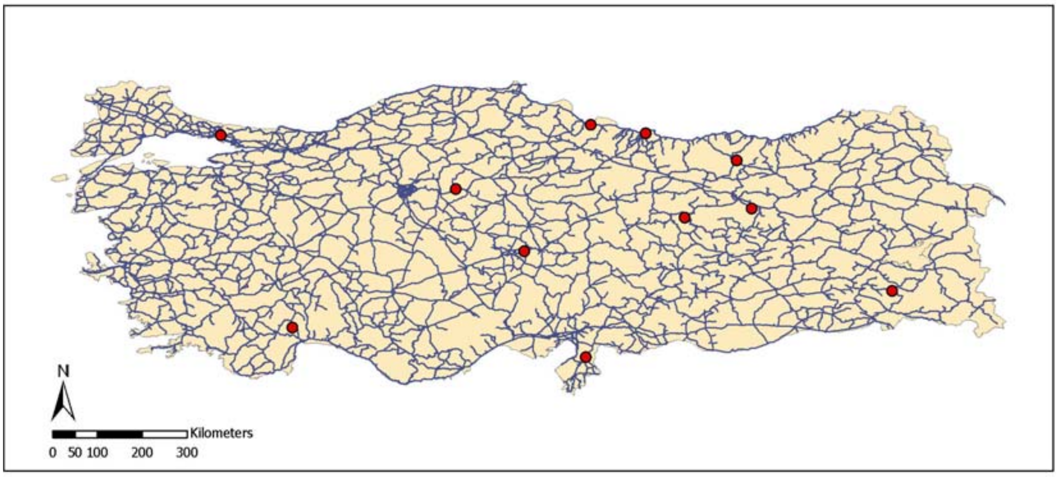


Figure 6.12. Distribution of overturning and lose of control accidents in the existence of sight obstruction in 2003 in Turkey

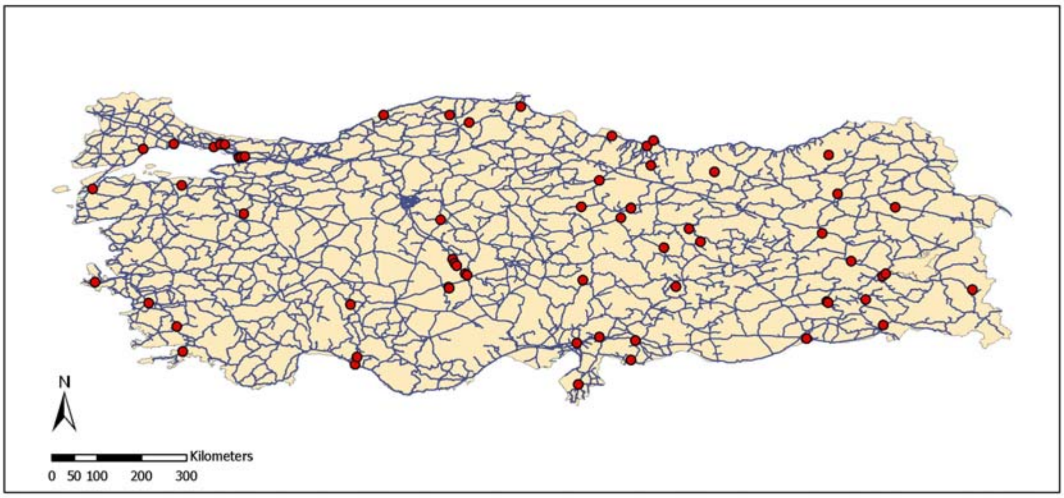


Figure 6.13. Distribution of accidents in the existence of road deficiency and inadequate warning sign in 2003 in Turkey

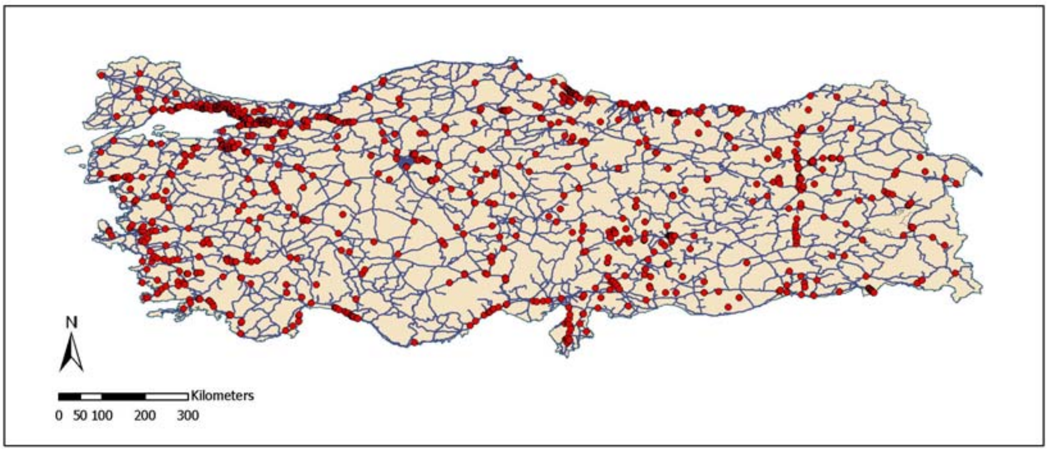


Figure 6.14. Distribution of accidents resulting from pedestrian-vehicle conflict in the case of non-existent pedestrian cross-walk in 2003 in Turkey

In the last decade, it is discovered that GIS is a powerful tool for analysis of geographical and environmental issues. In general, it holds data in several layers. For example, sensitive environmental areas forming one layer, soil conditions as the second layer, topographical information as the third layer and the fourth layer that consists of the existing highway network may be superimposed for the identification of possible network extensions, which meet environmental and geotechnical requirements (Figure 6.15). For the non-spatial (thematic) attributes of the features included in GIS; such as the population of a traffic analysis zone, the length, capacity and free-flow speed of a highway link or the severity and property damage associated with a traffic accident, relational database is generally used (46).

In this study, an innovative model, which can be effectively used for investigating the potential causal factors that have considerable influence on the occurrence of traffic accidents, is developed. The model also aims to build traffic accident risk map through our country regarding constant physical road parameters integrated with some varying environmental parameters. The required constant and variable parameters used for this model are the ones that are determined from the statistical analysis and Multicriteria Decision Analysis.

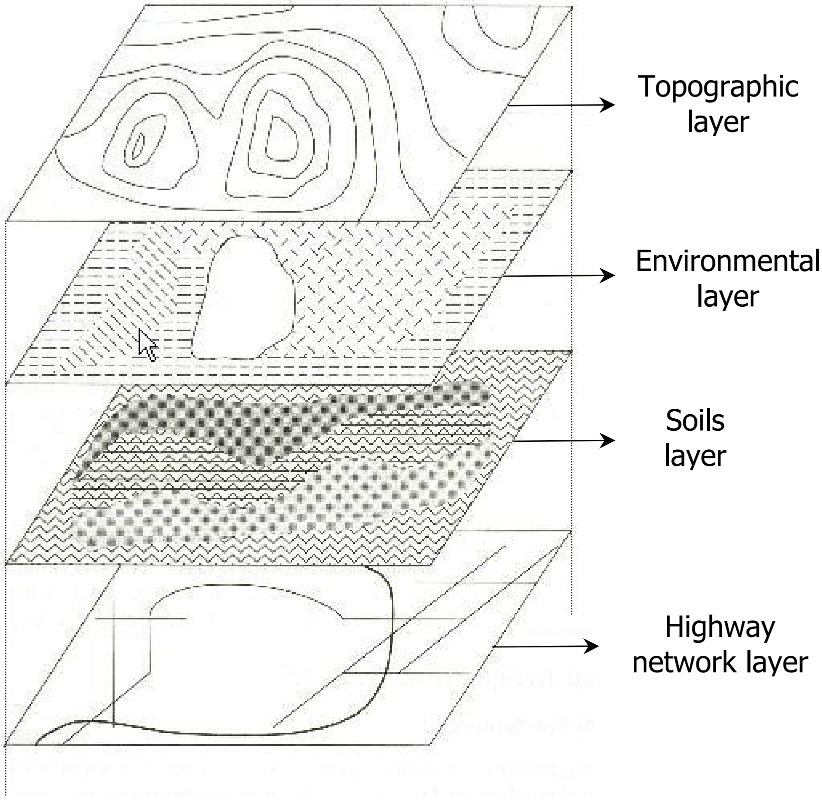


Figure 6.15. Typical multi-layer structure of GIS (46)

In order to represent the road characteristics (such as horizontal curves, road type at the desired location; i.e. location and road parameters shown in Table 5.1) geographically on a map, it is required to construct the road characteristics as a GIS layer. Similar to the traffic accident data maps, several queries can be constructed for better visualization of the road network.

For this thesis study, the digitized form of the road map of state highways and county roads was taken from Ministry of Interior-General Directorate of Security of Turkey, whereas the digitized road map of the motorways was taken from General Directorate of Highways.

On the other hand, all of the attribute data of the roads, which are named as inventory data (namely the location and road characteristics of the parameters defined in Table 5.1, i.e. road type, urbanization, divided/undivided, road coating, horizontal curve, vertical curve and intersection), were obtained from General Directorate of Highways. Except the motorways, the inventory data of all roads were obtained in hardcopy format.

Indeed, integrating all of the attribute data into the road map for the entire country is not possible. Even, a project for building up detailed road maps of the whole country in GIS has been recently put out to tender. The approximate period even for gathering the entire information to construct the GIS database of Turkish roadways is approximated as 18 months. Therefore, the road network of Ankara County was taken into consideration for this study and all of the spatial analyses were performed for the state highways and motorways of the county. In addition to the Ankara Ring Road and Ankara-Gerede Motorway, all of the state highways of Ankara County (Table 6.1) were included within this study. The road network of Ankara County is also mapped in Figure 6.16.

Since characteristics of roads do not exist in a spatial form, the first required process is to integrate the attribute data into the existing digitized road map. This was done after separating the roads into 50 m. intervals, which was selected after a careful judgment and found to be sufficient for a sensitive study and also appropriate according to the inventory data taken. When the values of the inventory data were integrated into the related road segments (Figure 6.17), it is possible to link the road characteristics with the corresponding location, which gives the ability to implement effective location-based studies. Thus, some integrated road inventory data of Ankara County are also mapped in Figures 6.18 to 6.20.

Briefly, since the model aims to find the risk maps with respect to the constant road parameters with the addition of variable environmental parameters; i.e. under different scenarios, the weight factor of each parameter evaluated from AHP that has a base of logistic regression, were firstly utilized. Then, for each parameter, a raster layer was constructed by matching the road characteristics of that parameter with the associated weight factors.

Table 6.1. Motorways and state highways of Ankara County

Road type	Name / ID	Length (km)
motorway	Ankara Ring	116.0
	Ankara-Gerede	68.1
state highway	140-03	69.1
	140-04	40.8
	140-05	30.9
	140-06	59.0
	170-02	25.6
	180-07	6.6
	200-11	23.4
	200-12	52.7
	200-13	38.5
	260-04	41.5
	260-05	55.0
	260-06	54.5
	655-03	10.7
	695-01	77.7
	750-05	33.4
	750-06	53.8
	750-07	14.9
	750-08	67.9
	750-08	67.9
	750-09	47.5
	757-03	29.0

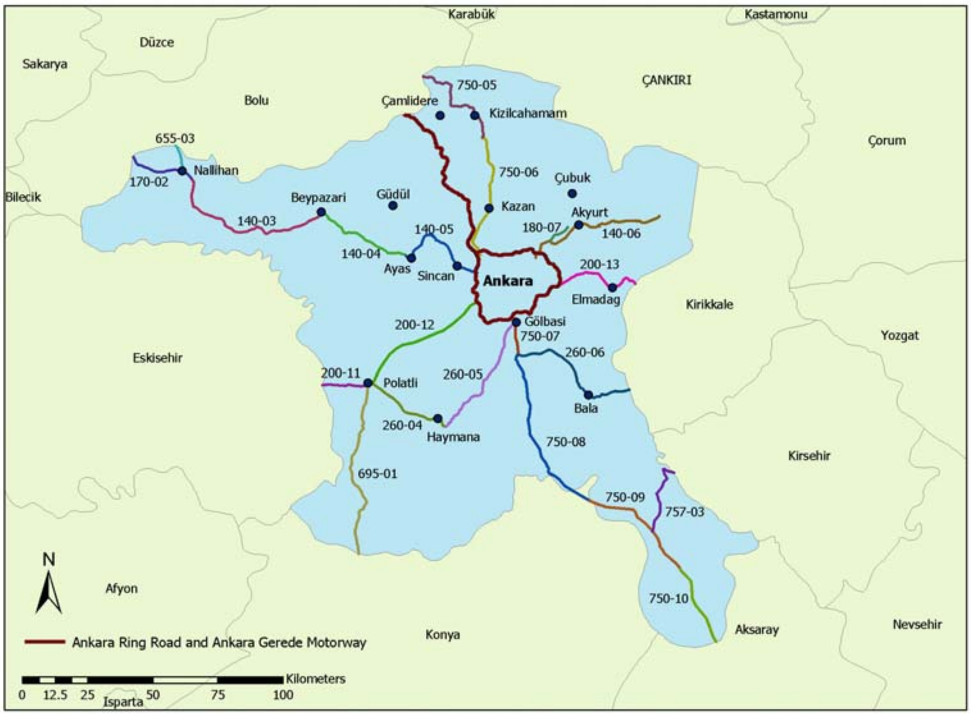
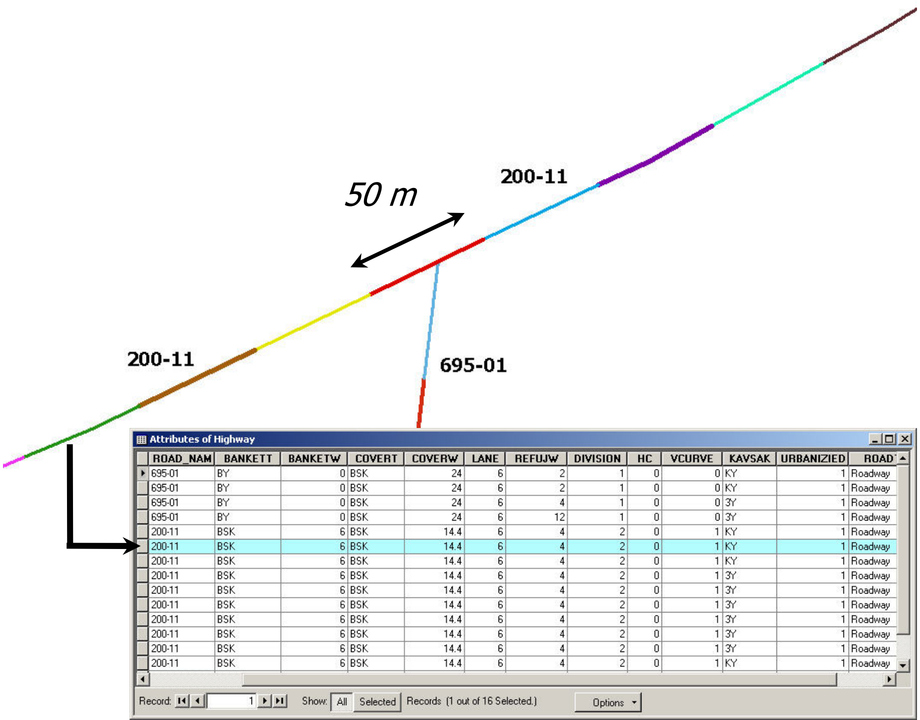


Figure 6.16. Motorways and State Highways of Ankara County



ROAD_NAM	BANKETT	BANKETW	COVERT	COVERW	LANE	REFUJW	DIVISION	HC	VCURVE	KAVSAK	URBANIZIED	ROAD
695-01	BY	0	BSK	24	6	2	1	0	0	KY	1	Roadway
695-01	BY	0	BSK	24	6	2	1	0	0	KY	1	Roadway
695-01	BY	0	BSK	24	6	4	1	0	0	3Y	1	Roadway
695-01	BY	0	BSK	24	6	12	1	0	0	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	KY	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	KY	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	KY	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	3Y	1	Roadway
200-11	BSK	6	BSK	14.4	6	4	2	0	1	KY	1	Roadway

Figure 6.17. Integrating attribute/inventory data with the road segments

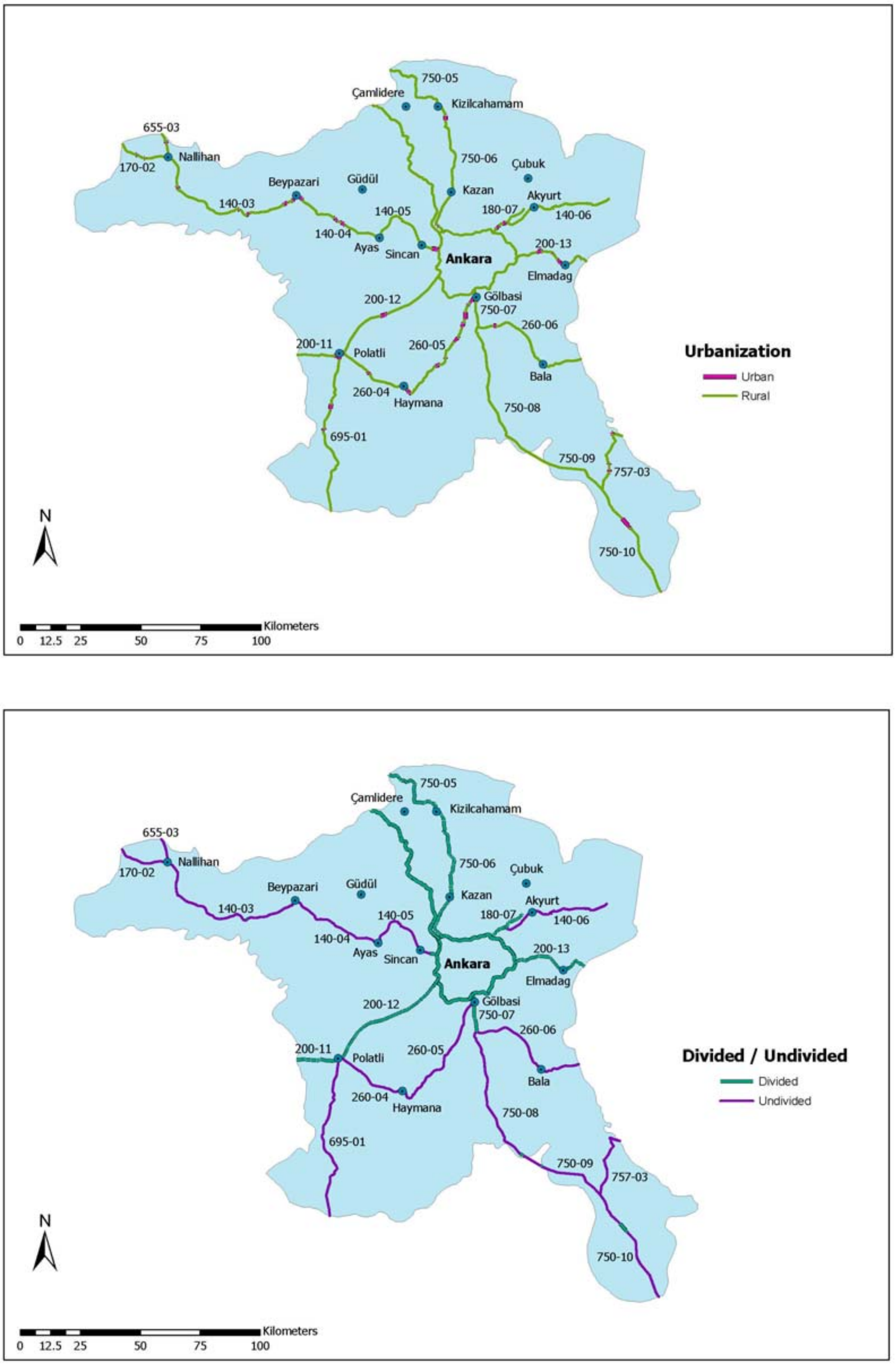
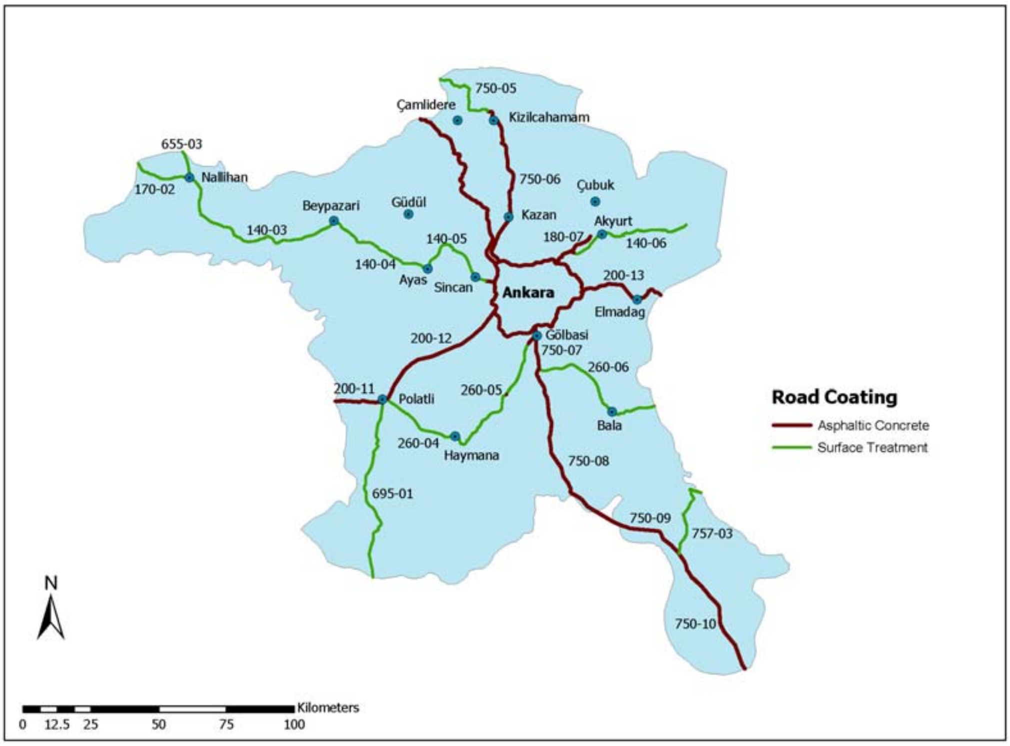


Figure 6.18. Inventory data of Ankara County integrated into road segments-1



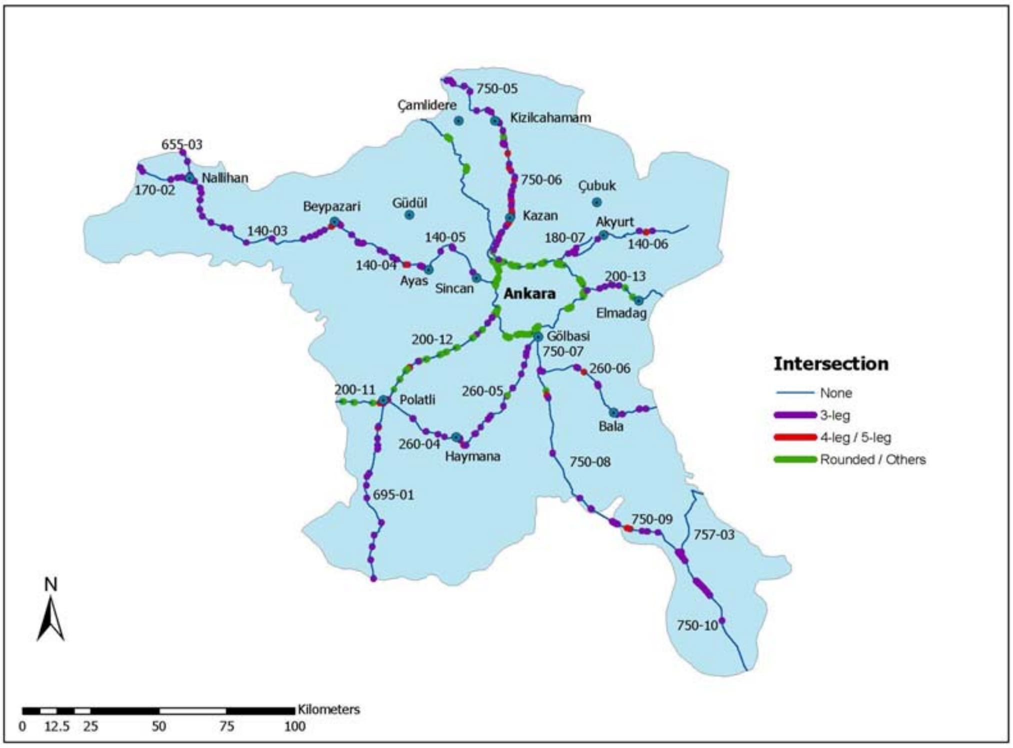


Figure 6.19. Inventory data of Ankara County integrated into road segments-2

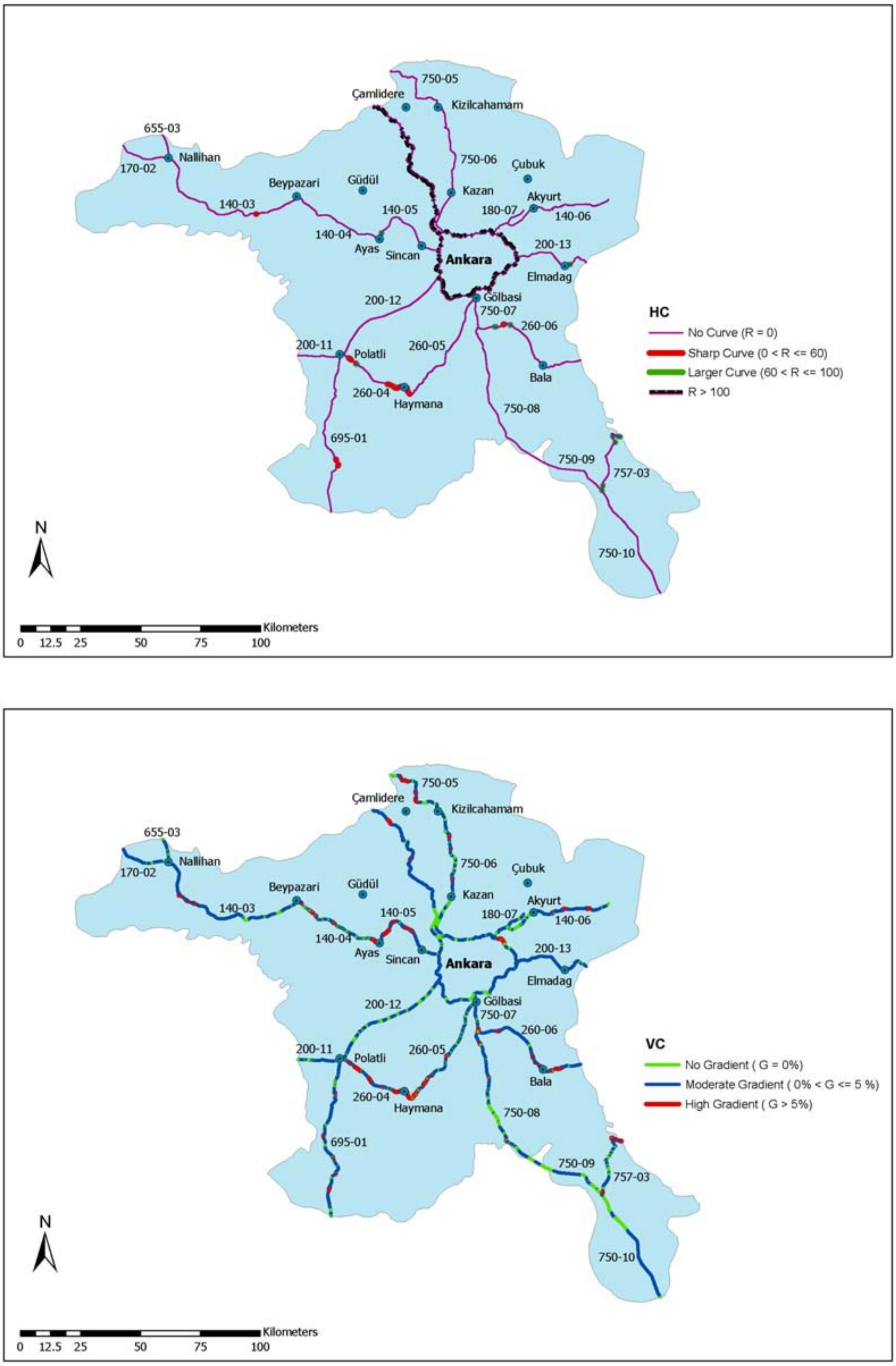


Figure 6.20. Inventory data of Ankara County integrated into road segments-3

However, before realizing the matching process, it is desired to create buffer zones around the road network for better visualization of the accident data. In order to decrease the overlaps and so to minimize the mixed information within the overlapping sections, rectangular buffering method -rather than the round buffering- is used by which data loss at edges was prevented due to non-existent buffer at line edges (Figure 6.21). In addition, due to incorrect recording of the accident data, the accidents may not definitely lie on the road's centerline but within a corridor around the centerline. Hence, buffering process for the entire road network of Ankara County ensures to better include the accident data within the analysis results. In this study, 100 m. rectangular buffer were used around both of the left and right side of the road centerline.

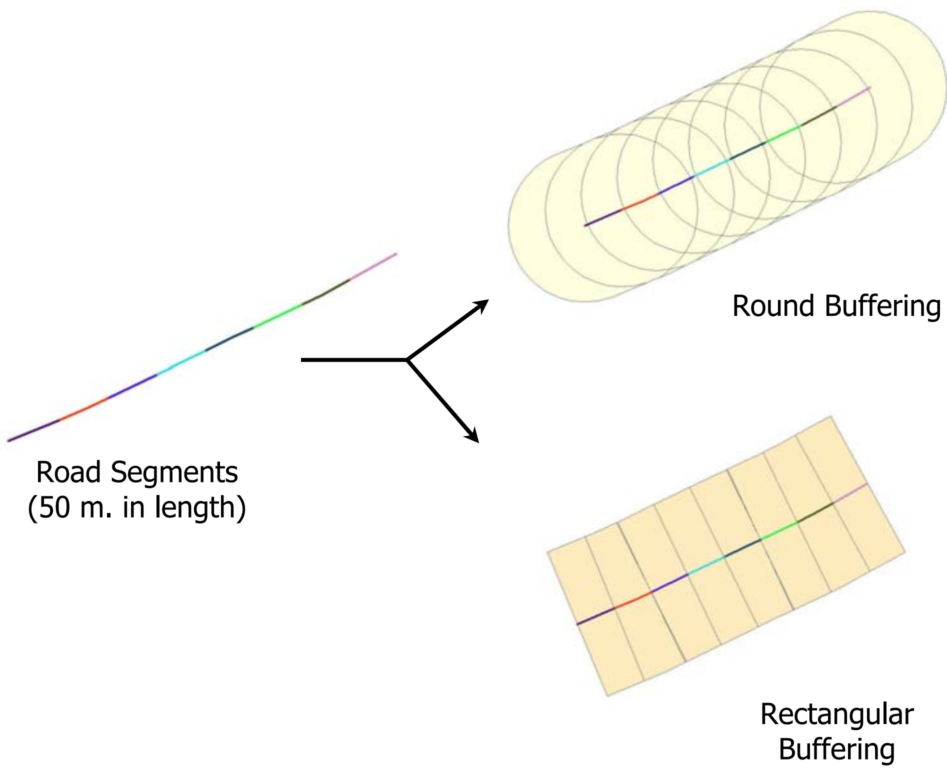


Figure 6.21. Round and rectangular buffering methods

For the matching process, firstly, the vector type road map is converted to a raster layer for taking the advantage of overlaying different raster layers including various road and environmental characteristics. Resolution (cell size) of the raster road layer was selected as 50 m. that is found to be supporting sufficient detail and appropriate for the 50 m. road segments. As a result of the converting process from vector to raster, totally 7 constant road parameter raster layers were obtained. These are namely road type, urbanization, divided/undivided, road coating, horizontal curve, vertical curve and intersection raster layers, i.e., location and road characteristics as indicated in Table 5.1.

The next thing to be done is to write a code into the GIS software. The code principally ensures a system that checks the road attribute of each cell and assigns appropriate weights found from the beforehand statistical analyses. In other words, new raster layers were created by reclassifying the values of the constant road parameter raster layers according to the risk factors of the corresponding parameter. Finally, a final raster -named as the weighted constant road parameters raster- was obtained by overlaying all individual weighted physical road parameter raster layers. In Figure 6.22, this method is exemplified by using random numbers.

During the reclassification analysis (matching process), it was required to coincide the inventory road data of General Directorate of Highways with the selected road parameters taken from the traffic accident report forms of Ministry of Interior-General Directorate of Security of Turkey. This mainly includes identifying, for instance, the meaning of the sharp curve with respect to the values determined by General Directorate of Highways. Therefore, two (vertical curve and horizontal curve) of the seven road parameters were needed to be systematically re-grouped as indicated in Table 6.2.

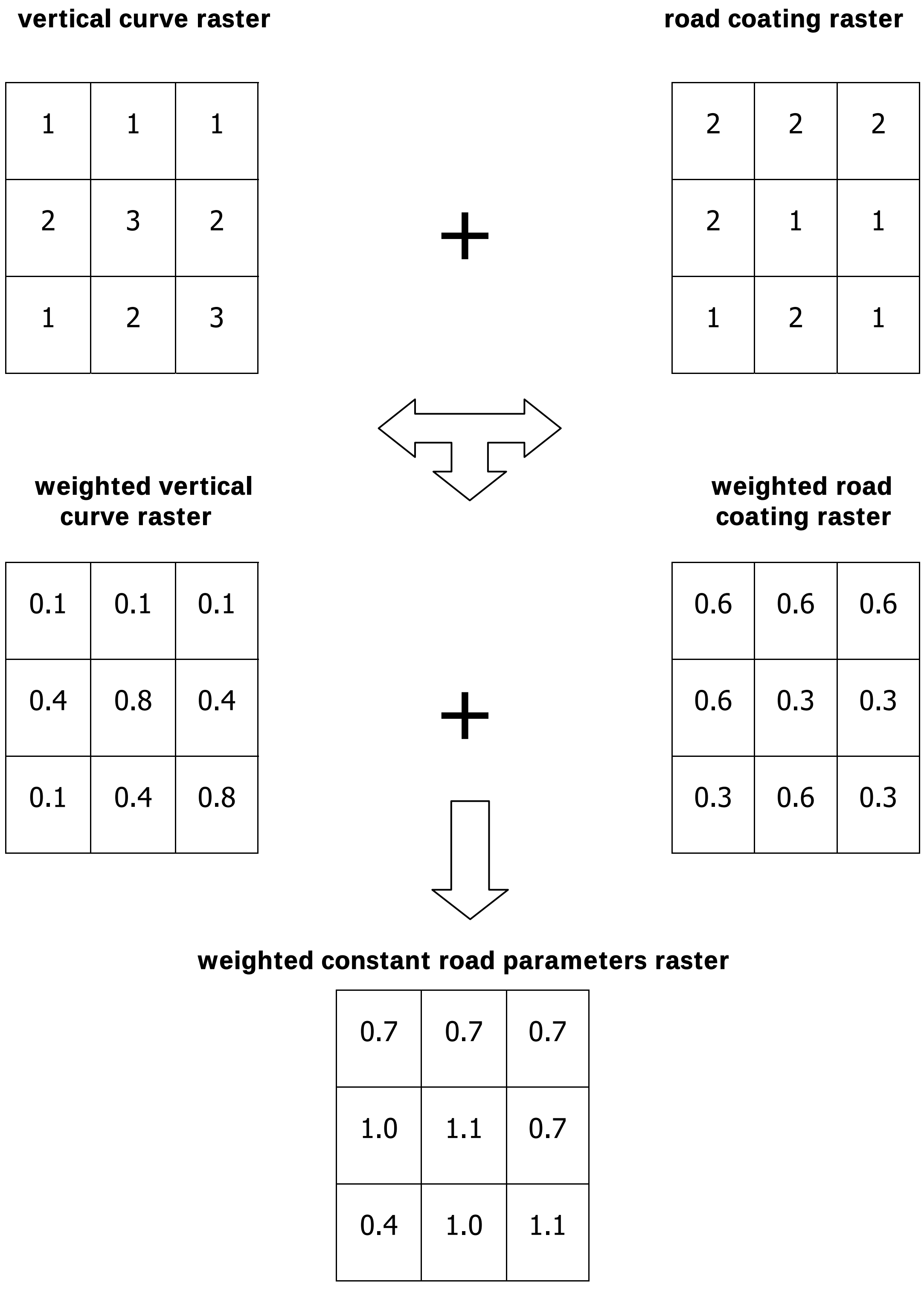


Figure 6.22. An example for arithmetic operations between constant road parameter raster layers

Table 6.2. Re-grouping the vertical curve and horizontal curve parameters according to General Directorate of Highways and matching them with the ones on traffic accident report form

	Description of the parameter (in traffic accident form filled by police officers)	**Classification of intervals according to General Directorate of Highways**
Vertical Curve (% gradient)	no gradient	gradient = 0
	moderate gradient	0 < gradient ≤ 5
	high gradient	gradient > 5
Horizontal Curve (curve radius in meter)	no curve	curve radius = 0
	sharp curve	0 < curve radius ≤ 60
	larger curve	60 < curve radius ≤100
	no curve	curve radius > 100

Since four different sets of risk factors were found for this study (for four response variables according to **1)** severity index/assumption-based for the first and second levels; **2)** severity index/independency approach for the first and second levels; **3)** fatality condition/assumption-based for the first and second levels; **4)** fatality condition/independency approach for the first and second levels), four different weighted constant road parameters raster layers were created.

Figures 6.23 to 6.26 show the distribution of these risk factors and thus the risk corridors according to the road parameters and with respect to different response variables.

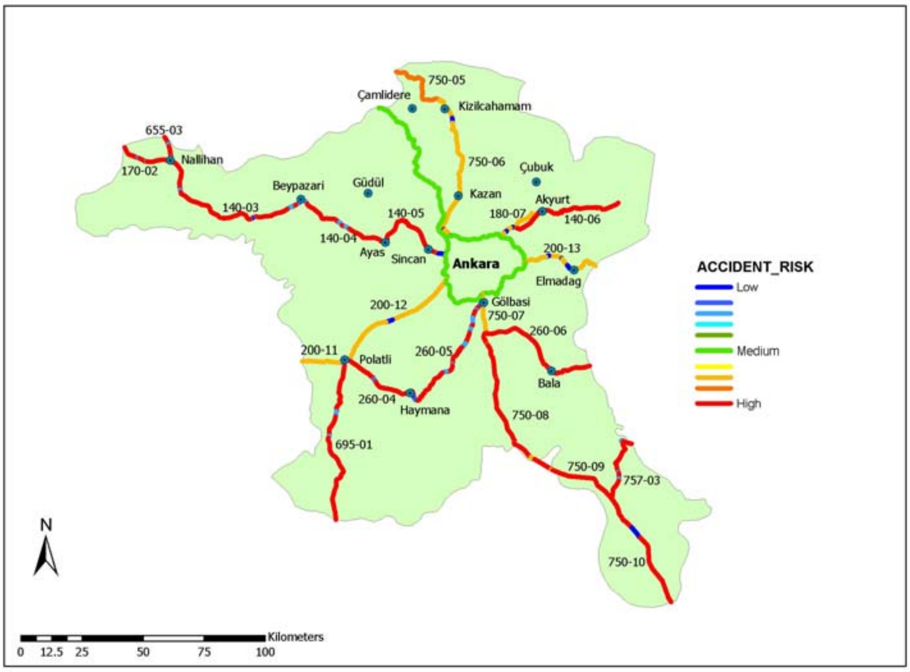


Figure 6.23. Accident Risk of Ankara County according to the weighted constant road parameters (severity index response variable)

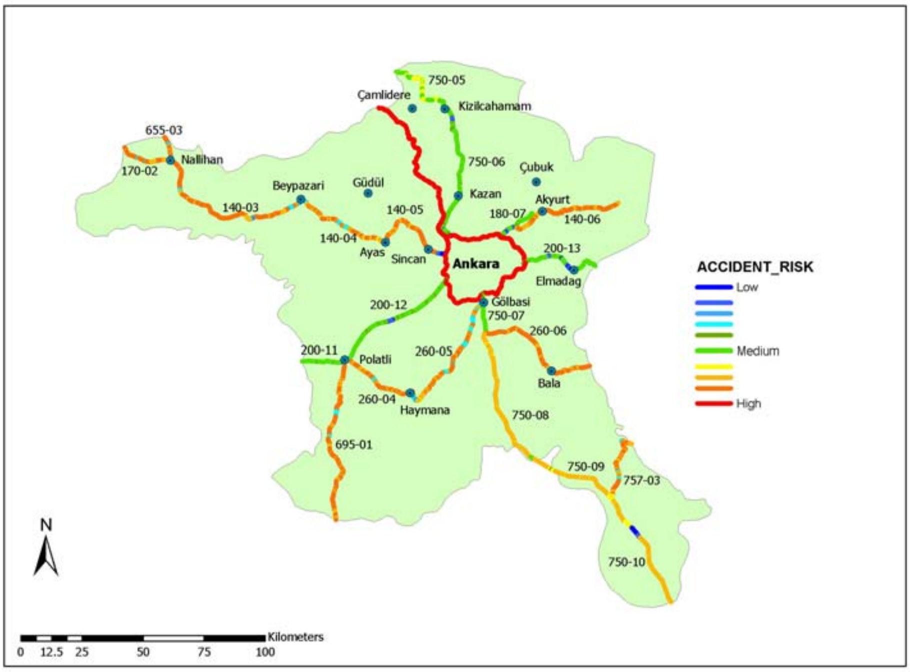


Figure 6.24. Accident Risk of Ankara County according to the weighted constant road parameters (fatality condition response variable)

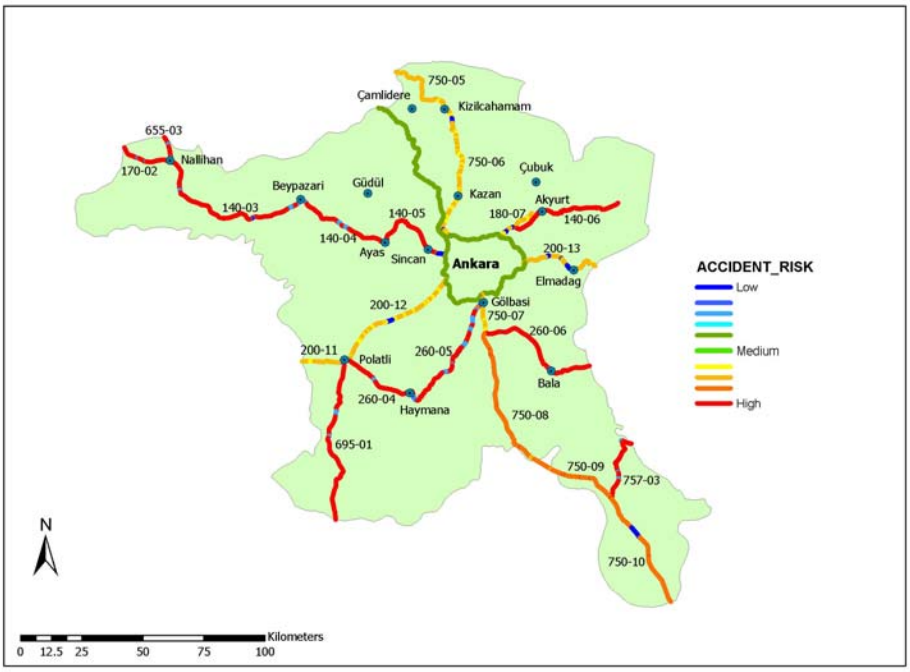


Figure 6.25. Accident Risk of Ankara County according to the weighted constant road parameters (severity index response variable/independency)

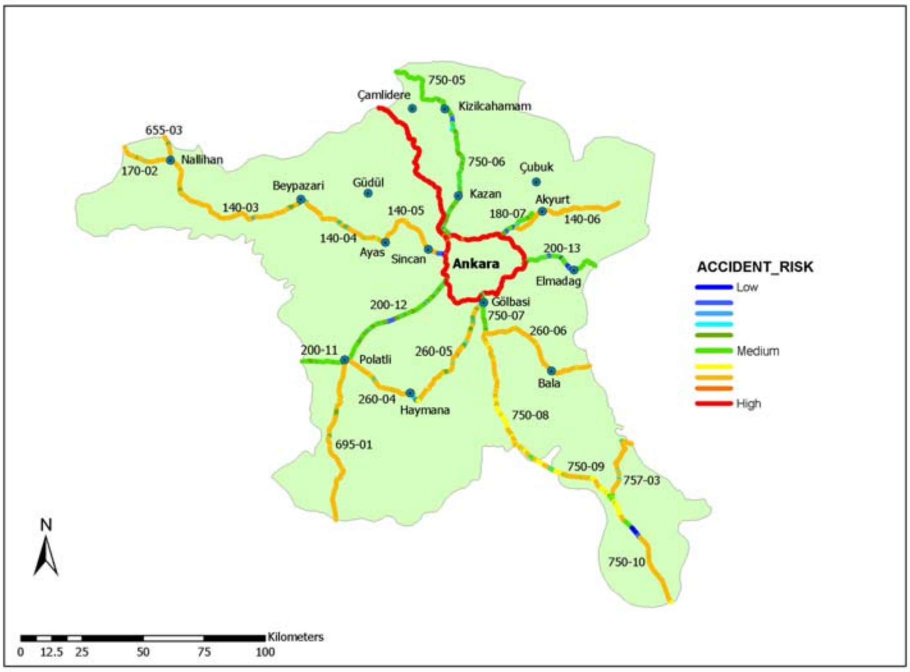


Figure 6.26. Accident Risk of Ankara County according to the weighted constant road parameters (fatality condition response variable/independency)

Afterwards, the weighted constant road parameters raster layers were aggregated with the weighted environmental parameters (risk factors of environmental parameters found in the previous chapter (Table 6.3)). For this reason, 27 scenarios having different combination of variable parameters (road surface, weather condition, time of day; each has three subgroups) were created for each of the four different response variables. Each one of these scenarios represents a single condition of a variable environmental parameter, which means that the scenario has a single total environmental risk factor obtained with the addition of risk factors of the selected conditions taken from each of the sub-groups of the environmental parameters, as represented in Table 6.4. Moreover, Table 6.5 indicates the maximum and minimum values within these scenarios according to different response variables.

Table 6.3. Risk factors of environmental parameters for each of the response variable; obtained from statistical and Multicriteria Decision Analyses

environmental parameters		**assumption-based risk factors**		**independency approach risk factors**	
		severity index	**fatality condition**	**severity index**	**fatality condition**
road surface	dry (d)	1.814	3.841	3.410	4.789
	wet (w)	1.827	3.165	3.434	3.947
	snowy (s)	2.259	1.887	4.245	2.353
weather condition	good (g)	2.302	2.045	3.908	3.371
	medium (m)	2.173	2.167	3.688	3.573
	tough (t)	2.058	2.515	3.493	4.145
time of day	day (dy)	1.855	1.573	3.419	2.627
	night (n)	1.981	2.502	3.652	4.178
	twilight (tw)	2.180	2.565	4.018	4.284

Table 6.4. Total environmental risk factors according to different scenarios

scenario		total assumption-based risk factors		total independency approach risk factors	
		severity index	fatality condition	severity index	fatality condition
s1	**d-g-dy**	5.971	7.459	10.737	10.787
s2	**w-g-dy**	5.984	6.783	10.761	9.945
s3	**s-g-dy**	6.416	5.505	11.572	8.351
s4	**d-m-n**	5.968	8.510	10.750	12.540
s5	**w-m-n**	5.981	7.834	10.774	11.698
s6	**s-m-n**	6.413	6.556	11.585	10.104
s7	**d-t-tw**	6.052	8.921	10.921	13.218
s8	**w-t-tw**	6.065	8.245	10.945	12.376
s9	**s-t-tw**	6.497	6.967	11.756	10.782
s10	**d-g-n**	6.097	8.388	10.970	12.338
s11	**d-g-tw**	6.296	8.451	11.336	12.444
s12	**w-m-da**	5.855	6.905	10.541	10.147
s13	**w-m-tw**	6.180	7.897	11.140	11.804
s14	**s-g-n**	6.542	6.434	11.805	9.902
s15	**s-g-tw**	6.741	6.497	12.171	10.008
s16	**d-m-da**	5.842	7.581	10.517	10.989
s17	**d-m-tw**	6.167	8.573	11.116	12.646
s18	**d-t-dy**	5.727	7.929	10.322	11.561
s19	**d-t-n**	5.853	8.858	10.555	13.112
s20	**w-g-n**	6.110	7.712	10.994	11.496
s21	**w-g-tw**	6.309	7.775	11.360	11.602
s22	**w-t-dy**	5.740	7.253	10.346	10.719
s23	**w-t-n**	5.866	8.182	10.579	12.270
s24	**s-m-da**	6.287	5.627	11.352	8.553
s25	**s-m-tw**	6.612	6.619	11.951	10.210
s26	**s-t-dy**	6.172	5.975	11.157	9.125
s27	**s-t-n**	6.298	6.904	11.390	10.676

Table 6.5. Extreme values in scenarios

Extreme values in scenarios	**severity index**		**fatality condition**	
	max	min	max	min
	s-g-tw (s15)	d-t-dy (s18)	d-t-tw (s7)	s-g-dy (s3)
assumption-based	6.741	5.727	8.921	5.505
independency approach	12.171	10.322	13.218	8.351

As can be seen from Table 6.5, maximum and minimum risk values showed difference according to the response variables; but, it was observed that independency approach did not alter the scenarios in which these extreme values occurred and they were also nearly similar with respect to their importance sequence with assumption-based approach.

In addition, an interesting remark is made from Table 6.5 that while the most risky condition occurs under dry road surface and tough weather condition for fatality condition response variable, these parameters change to snowy road surface and good weather condition for severity index response variable. Whereas, the parameters for the least risky situations were snowy road surface & good weather condition and dry road surface & tough weather condition, respectively. Actually, when the logic of the response variable is changed, the effects of environmental parameters significantly change. In this respect, it can be understood that before identifying the reasonable response variable, the problematic locations should be well-discussed and the results of the accidents should be evaluated correctly.

After determining the risk factors of the environmental parameters, these risk factors were added to the total weight of constant road parameters raster obtained previously. Thus, different scenarios were created for different variable conditions. As a result, for each of these scenarios, an accident risk raster (totally 27*4=108 raster layers), including constant and variable parameters, was modeled. The methodology is also exemplified in Figure 6.27. In addition, Figures 6.28 and 6.29 represent the maximum and minimum scenario results - aggregated onto the constant road parameters- for severity index and fatality condition response variables, which are presented on a common scale within themselves for better understanding of the scenarios.

In brief, the model ensures to obtain a final integrated raster under different environmental conditions, which gives the chance of investigating the most risky locations with respect to their reasons (Figure 6.30 displays the most risky zones among the state highways and motorways of Ankara County). This also enables the researchers to implement the solution more effectively and quickly. Especially for the urgent conditions, one of the most important things is being aware of the actual cause of the existing problem as soon as possible. Also, taking the varying conditions into consideration will improve the decision making process since the problematic location can be analyzed geographically within the consideration of its changeable nature, which is composed of related combination of environmental parameters.

With the help of this methodology, one can also observe the possible future changes or the results of planned improvements since the system allows the user to manipulate with the weights. For instance, it is possible to investigate for the engineering measures proposed to be implemented on the selected location and surely to make a comparison by changing different parameters systematically.

Moreover, the probability of the future accidents under different conditions can be found, which would be also helpful for better land developments. As a result, in addition to its preventing and minimizing role, integration of statistical and spatial models will be helpful to enhance the future developments, which are surely very essential from the planning point of view.

Total weight of variable parameters for different scenarios

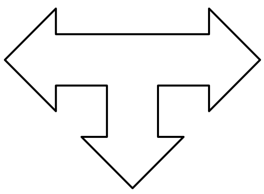
	Road Surface	***Weather Condition***	***Time of day***	***Total weight***
Scenario 1	Dry (wr1)	Good (ww1)	Night (wt1)	wr1+ww1+wt1
Scenario 2	Wet (wr2)	Medium (ww2)	Day (wt2)	wr2+ww2+wt2

---

weighted constant
road parameters raster

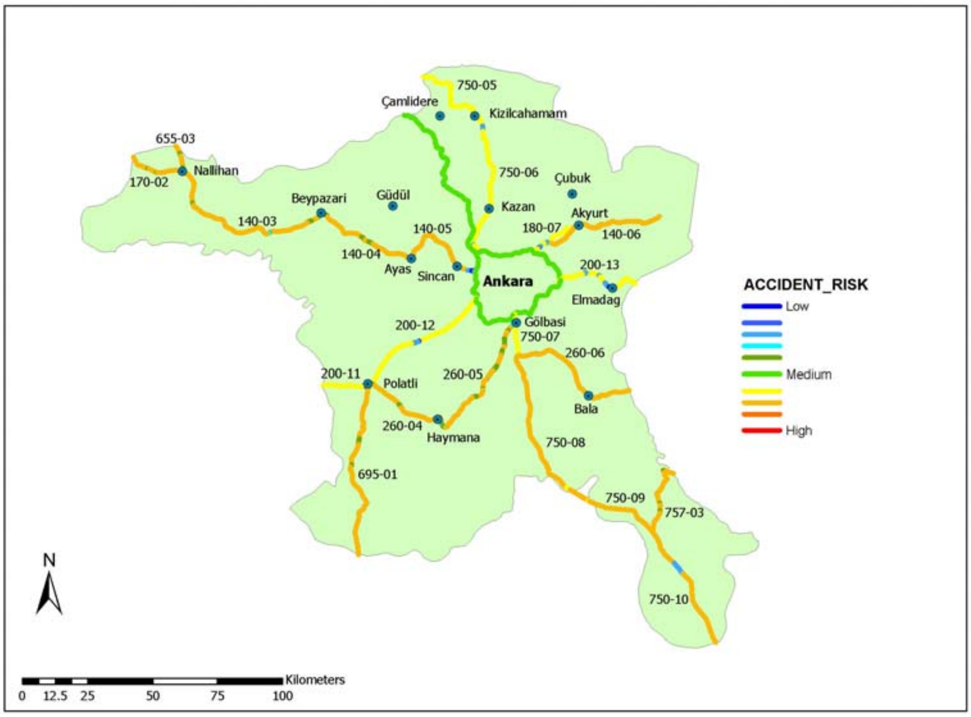
0.7	0.7	0.7
1.0	1.1	0.7
0.4	1.0	1.1

\+ **"total weights coming from scenarios"**

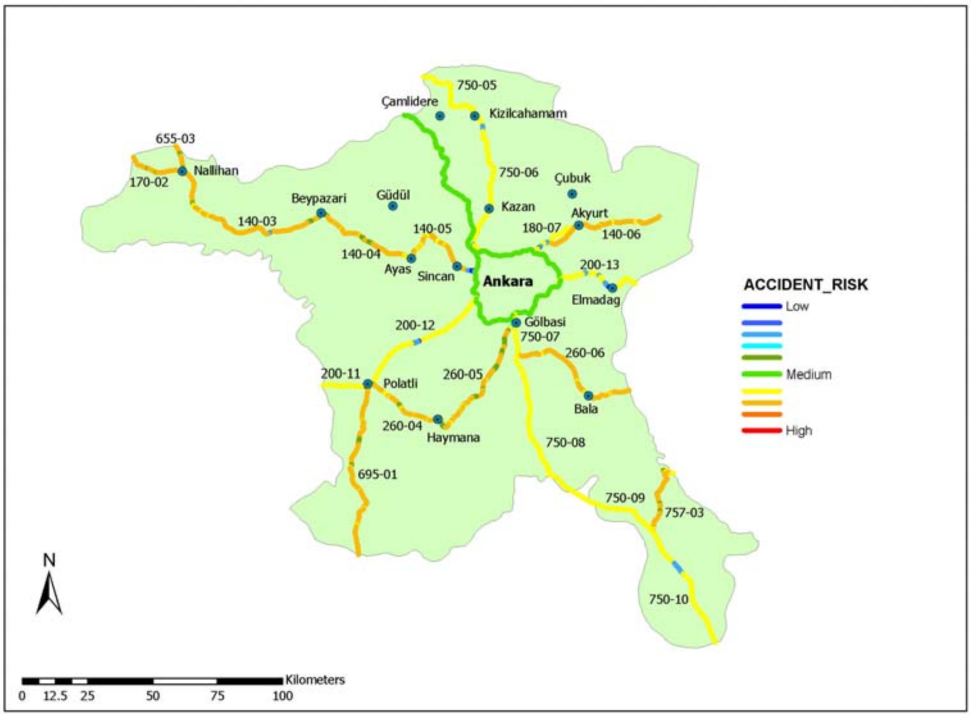


Accident Risk Raster
(including constant and variable parameters)

Figure 6.27. Creation of different scenarios by combining the constant road and variable environmental parameters

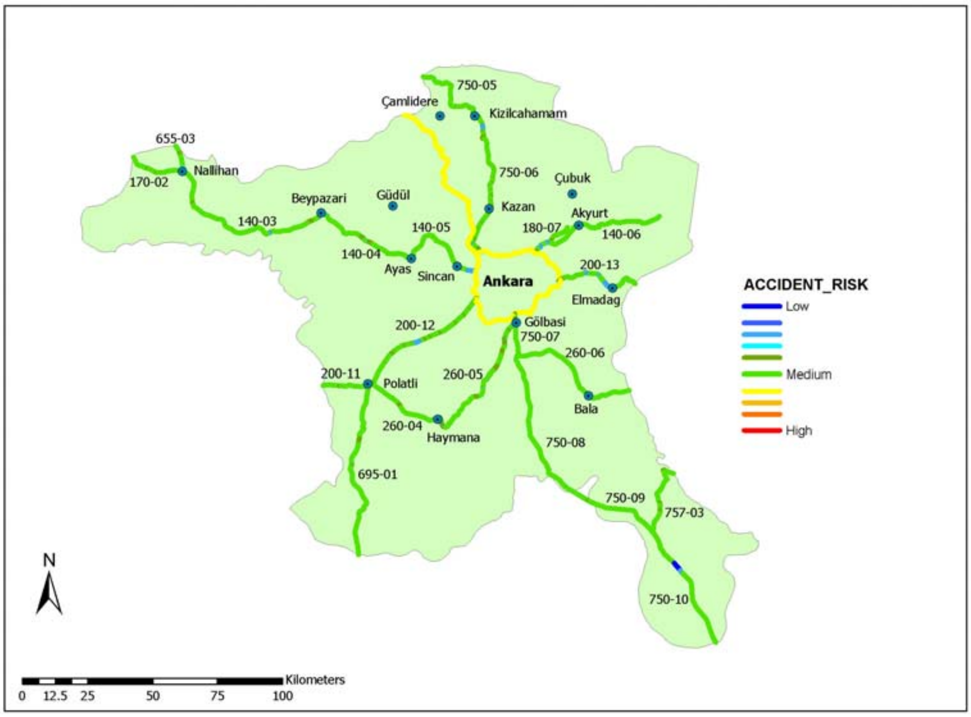


(a)

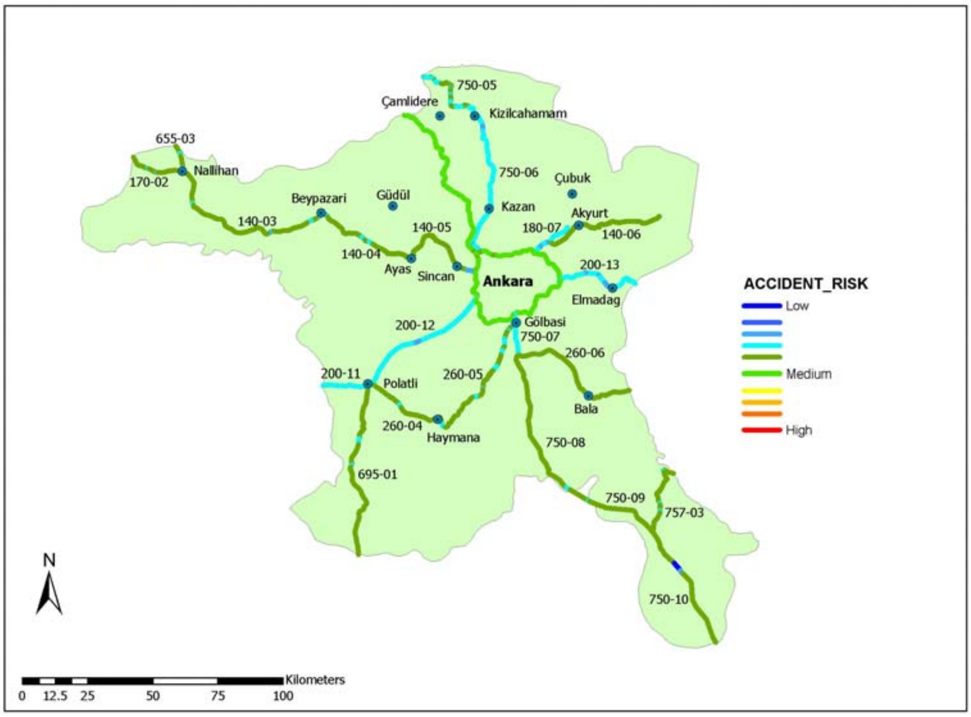


(b)

Figure 6.28. Maximum (a) & minimum (b) risky scenario results-severity index

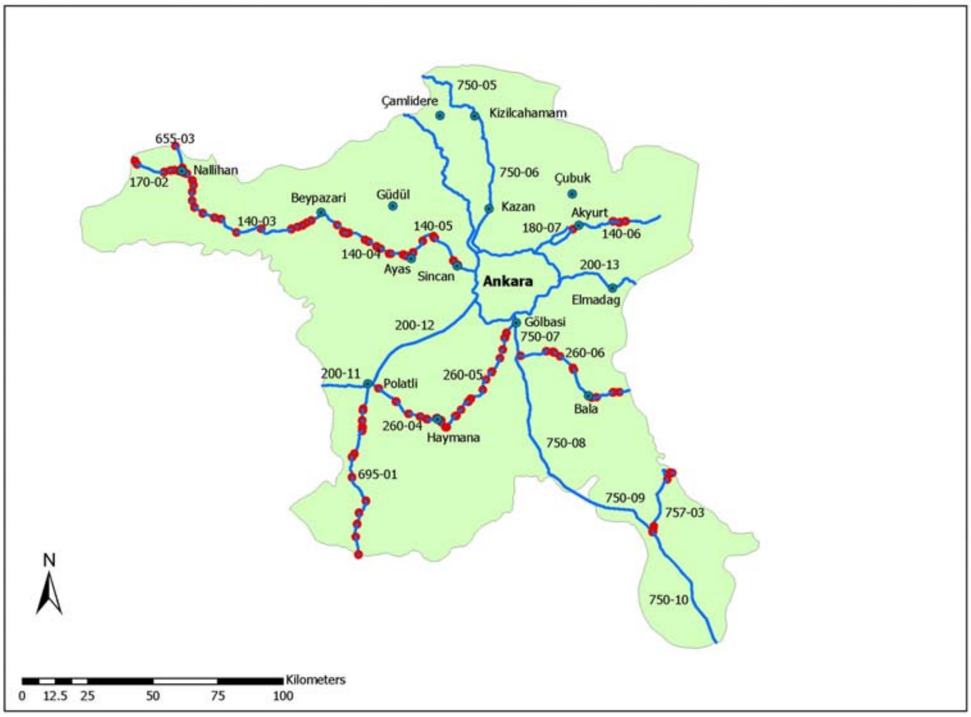


(a)

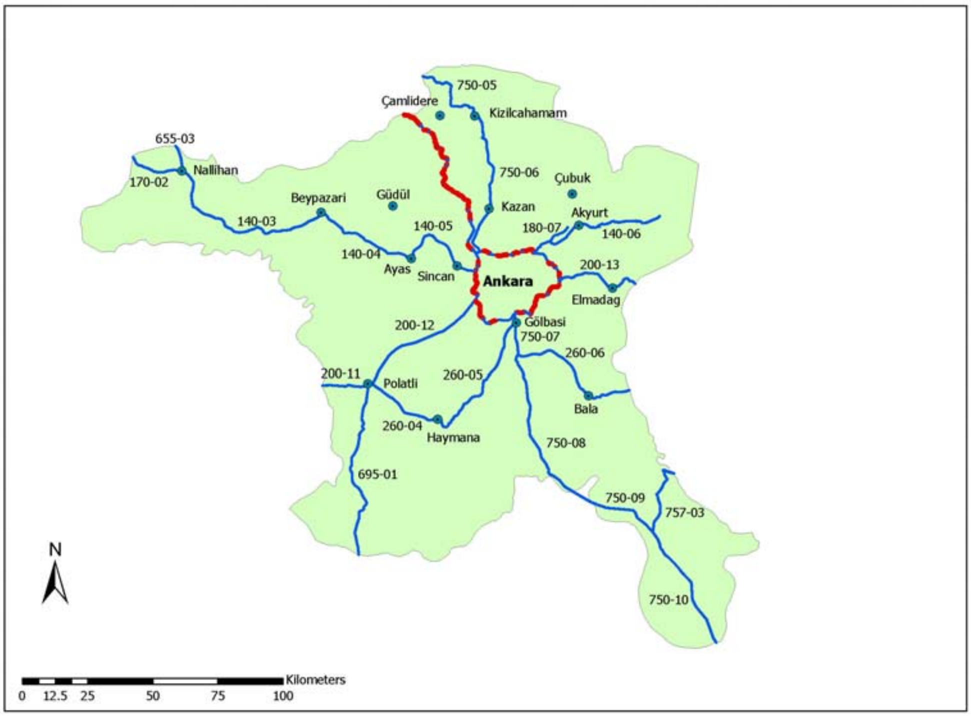


(b)

Figure 6.29. Maximum (a) & minimum (b) risky scenario results-fatality condition



(a)



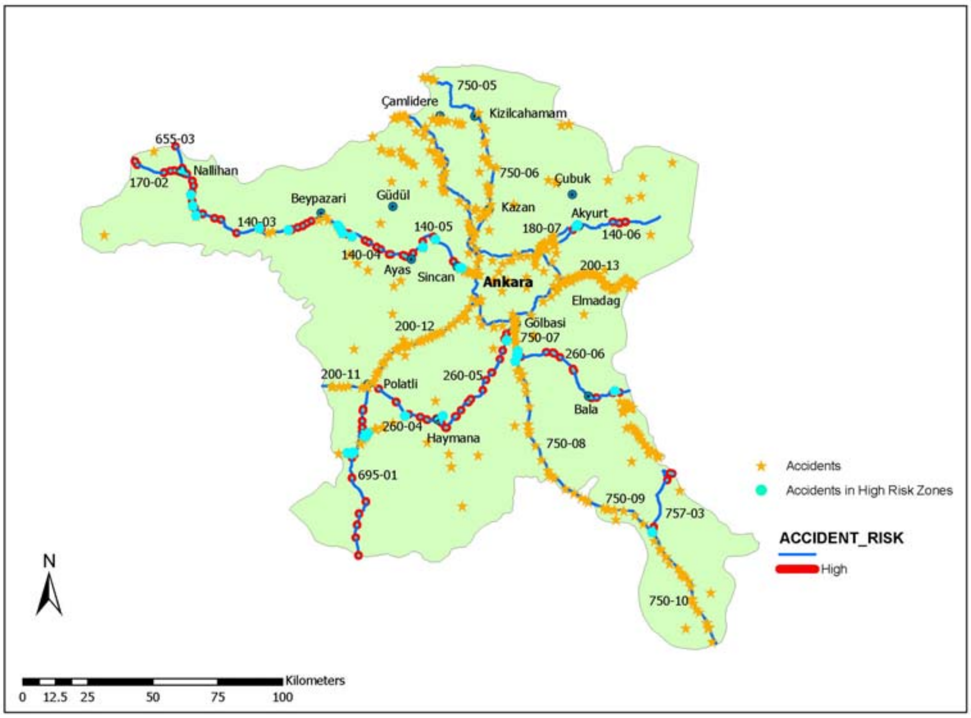
(b)

Figure 6.30. The most risky corridors of the state highways and motorways of Ankara County with respect to (a) severity index (b) fatality condition

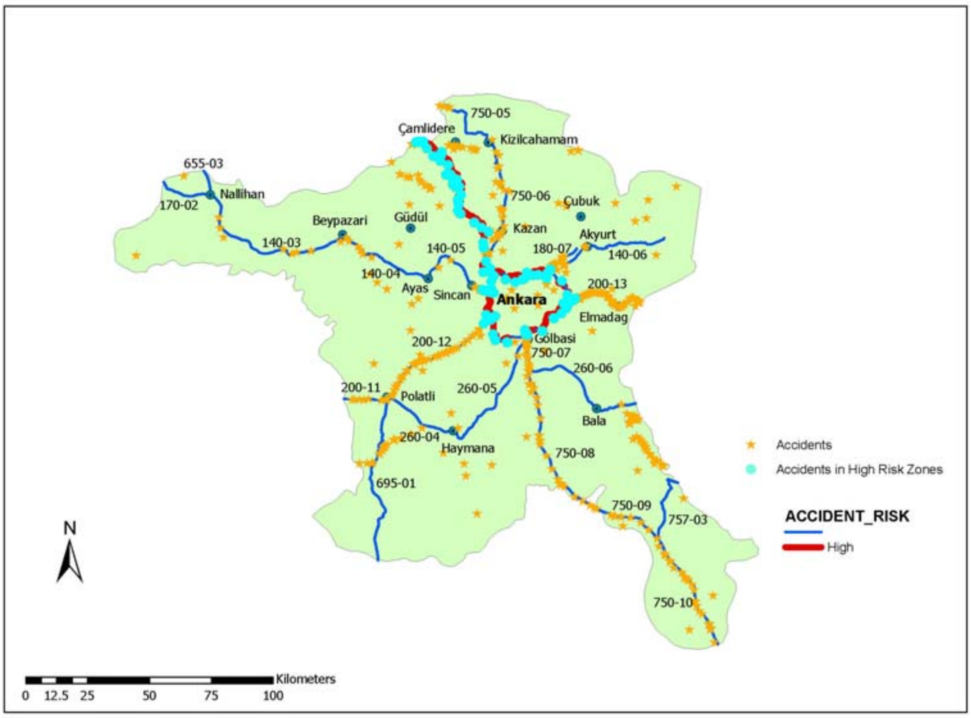
One last thing within the concept of this thesis is the comparison of the accident risk maps with the real traffic accident cases occurred on the state highways and motorways of Ankara County for the year 2003. For this reason, the highest risk zones were determined and then the same regions were overlaid with the actual accident distribution map, which was prepared at the beginning of the study for the year 2003 (Figure 6.1); and the accidents within these regions were extracted by location-based GIS queries. This procedure was carried out for the accident risk maps having maximum risk ratios according to the severity index and fatality condition response variables. The query results are presented in Figure 6.31.

In this manner, the accidents located in the most risky corridors and areas were extracted with respect to both of the response variables and also the actual traffic accidents occurred within those regions were investigated according to their attribute characteristics. Then, the attribute data of the actual accidents were compared with the attribute data of the roadway. This comparison was also performed for a definite case; i.e. for an accident located just on the risk corridor, as an example (Figure 6.32) and the comparison of the related numeric values are given in Table 6.6.

When the results of this comparison (Table 6.6) were evaluated, it was observed that while some of the attribute data matched with each other, some did not. The differences most probably results from the erroneous and incomplete structure of the data collecting and recording system. For instance, when the road type parameter of the accidents is investigated, it is seen that 47.71 % of accidents, occurred on motorways, were recorded as they occur in state highways.



(a)



(b)

Figure 6.31. Distribution of actual accidents in high risk zones of Ankara County with respect to (a) severity index (b) fatality condition

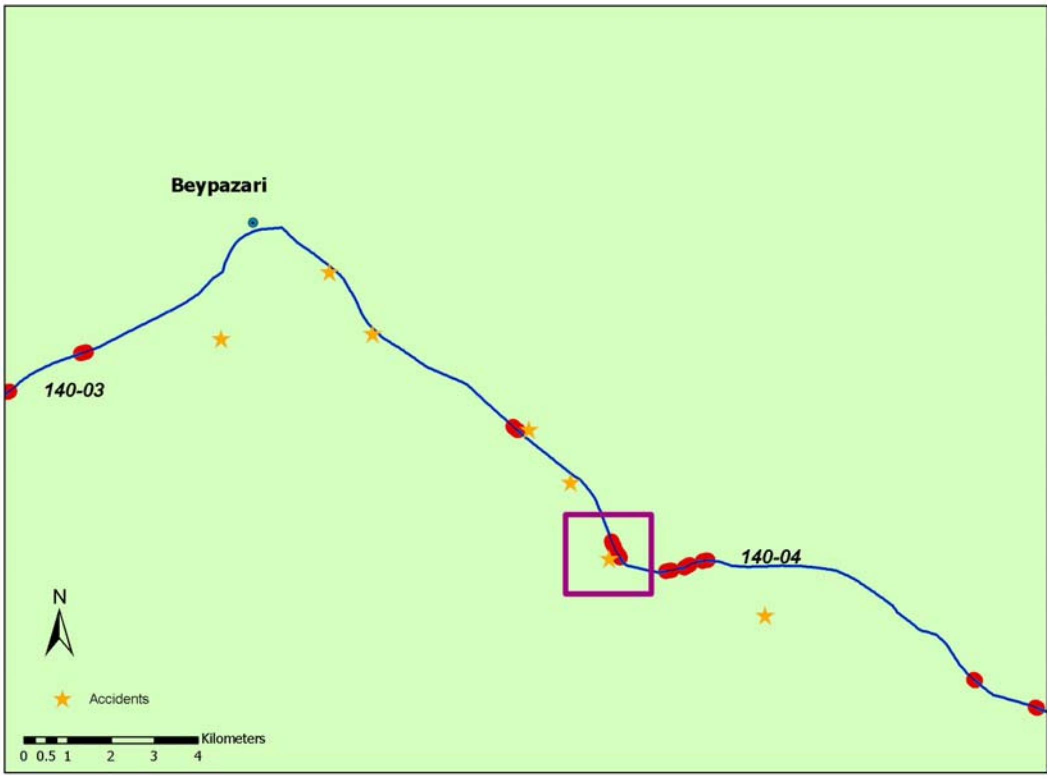


Figure 6.32. Location of the selected case on 140-04 state highway

If more reliable data can be collected, more meaningful results can be obtained in comparing the attribute tables of the accidents located within those regions. Conclusively, the significant causal factors of the accidents can be investigated for the specific locations that would be particularly helpful to understand the causes of the accidents with respect to their locations and hence to identify the locations, which require more urgent improvements.

Table 6.6. Comparison of the attribute data between the maximum risk zones and the actual traffic accident data occured on state highways and motorways of Ankara County in 2003 in Turkey

Percentage of the parameters according to the response variables		Severity Index		Fatality Condition		Case (severity index)	
		Accident	Road	Accident	Road	Accident	Road
Road Type	Motorway	5.41	0	52.29	100	0	0
	State Highway	94.59	100	47.71	0	100	100
Urbanization	Rural	75.68	100	84.40	100	100	100
	Urban	24.32	0	15.60	0	0	0
Divided / Undivided	Divided	83.78	0	96.33	100	100	0
	Undivided	16.22	100	3.67	0	0	100
Road Coating	Surface Treatment	100	99.19	99.08	0	100	100
	Asphaltic Concrete	0	0.81	0.92	100	0	0
Horizontal Curve	No Curve	86.49	96.73	77.06	100	100	100
	Larger Curve	10.81	3.27	18.35	0	0	0
	Sharp Curve	2.70	0	4.59	0	0	0
Vertical Curve	No Gradient	62.16	14.59	53.21	0	100	100
	Moderate Gradient	32.43	85.41	36.70	100	0	0
	High Gradient	5.41	0	10.09	0	0	0
Intersection	None	78.38	3.78	82.57	100	100	0
	3-Leg	2.70	92.43	7.34	0	0	100
	4-Leg / 5-Leg	16.22	3.78	4.59	0	0	0
	Rounded / Others	2.70	0	5.50	0	0	0
Road Surface	Dry	83.78	0	85.32	100	100	0
	Wet	16.22	0	12.84	0	0	0
	Snowy	0	100	1.84	0	0	0
Weather Condition	Good	83.78	100	82.57	0	100	100
	Medium	13.51	0	14.68	0	0	0
	Tough	2.70	0	2.75	100	0	0
Time of Day	Day	75.68	0	66.97	0	0	0
	Night	24.32	0	26.61	0	0	0
	Twilight	0	100	6.42	100	100	100

CHAPTER 7

CONCLUSION

7.1. In General

Traffic accidents are one of the most serious problems of Turkey as it is almost for the whole world. The economic and social losses of road accidents can be evaluated as the negative impacts of the developing technology, which seriously destroy our lives. These losses absolutely show that effective and immediate solutions are required for the development of road safety.

As emphasized before, high usage of highways in addition to the increasing number of vehicles per km gives rise to the occurrence of traffic accidents. In the last decade, many people have died or injured and more will die or injure in the next years because of traffic accidents all over the world. Especially, statistical analyses indicate that Turkey seems risky in terms of traffic accidents.

In order to decrease this accelerating rate of traffic accidents, it is required to investigate the potential causal factors on the occurrence of accidents. In Turkey, accident analysis studies are generally performed by using statistical methods. Traffic accident data used for these analyses are obtained from the traffic accident report forms, which are filled by security soon after an accident.

However, an effective solution to the traffic accidents can actually be obtained by determining the hazardous locations. Otherwise, the results and related countermeasures would not be realistic since it is not possible to see and interpret the whole picture clearly. From this respect, the traffic accident records in Turkey have also been started to be kept with their geographic coordinates for the last two years.

Integrating spatial data into the traffic accident analyses directly helps for developing more effective location-based precautions. Therefore, GIS technology has a significant role in accident analysis methods in order to decrease the rate and severity of traffic accidents and so to enhance the road safety.

Particularly, integration of statistical analysis with spatial analysis would be the best methodology for the accuracy of the accident based studies and for the applicability of the solutions to the correlated problem. Additionally, increasing the visualization through GIS improves understanding and interpretation of the picture easily, quickly and completely, which also greatly enhances the success of the study.

Ensuring safer roadways through the improved engineering, enforcement and education measurements is the main aim in traffic safety. Among these three basic rules of thumb, engineering measurements can be regarded as the most effective ones especially for the urgent conditions. For more feasible studies, it is required to identify the high risk zones and apply the precautions to these bordered areas rather than the whole county or country. Otherwise, such applications will become time and money loss instead of being effective safety measures.

In this study, it is aimed to develop a model, which establishes and visualizes the distribution of the occurrence of traffic accidents by means of probable potential causal factors, identifies the significant influence of road factors on accidents and besides, evaluates high risky accident zones with respect to their potential causes.

7.2. Summary and Concluding Remarks

At the end of this study, there exist several inferences. The most important ones that are drawn from this study are summarized item by item as follows:

- Transportation is one of the most essential parts of our social and economic life; however, it is just an obligatory for the movement of people and goods and in fact, it is not a necessary event for itself. Having an unwanted structure; therefore, results in serious social and economic costs. Deaths and injuries in traffic accidents, unbalanced land use, pollution and loss of productivity are some of the negative impacts of transportation.

- Accelerated traffic accident problem is unfortunately a natural outcome of increasing traffic and can be regarded as the most important transportation related problem because of its terrible and destructive results. Especially, when the value of human life is considered, it is seen that it becomes meaningless to discuss the greatness and seriousness of the traffic accident problem.

- There are three main factors affecting the occurrence of traffic accidents: human, road and vehicle factors. The statistical studies carried out all over the world indicate that the most dominating factor in accidents is the human factor. In Turkey, this figure increases approximately to a percent of 99. However, detailed studies show that the contributions of roadway and vehicle factors are not as small as they are indicated and both require special attention within the concept of traffic accident analysis studies.

- It is actually very difficult to control the human factor, which necessitates a well-developed education program and enforcement of legislation. In short run, engineering improvements will be more effective and save the lives when effective countermeasures are applied to the accurate locations. These measures will also be extremely helpful for decreasing the severity and number of traffic accidents in long term if they are implemented within the planning of future developments. In addition, human behavior may even be influenced by means of engineering solutions if the effect of road user comes from the deficiency or inadequacy of the road characteristic and road environment.

- Similar to the other developing countries, Turkey has been facing with the traffic accident problem with quite high rates. When the dominant use of roads (since highway is the major transportation mode of our country) is combined with the increasing number of vehicles, comparatively much more little increase in the length of total road network and insufficient improvements in the standards, it is not too surprising to have such a tremendous number of traffic accidents. Statistics indicate that Turkey, unhappily, places at one of the highest levels with its number of traffic accidents, fatality and injury rates.

- The success of traffic safety improvement programs and so corresponding analyses, which are performed for preventing accidents and/or reducing their negative impacts, are directly related with accurate, reliable, updated, easily accessible and affordable traffic accident data. But, the quality of data necessary for the analysis is not always sufficient. Much of the accident information available in police records is often incomplete and erroneous (69). The situation in Turkey is also not so different from this figure. Since traffic accident report forms are not recorded accurately and completed sufficiently, accident studies can not be conducted with high level of accuracy.

- Traffic accident studies are generally adopted with statistical analysis, which gives a general numeric view about the occurrence of accidents. It is possible to find the probable causal factors and their importance weights by using the tools of statistics. However, in order to develop effective safety measures for the problematic areas, it is necessary to integrate the statistical information with its related location and also with the characteristics of that location. This means that without the knowledge of spatial data and without incorporating the attribute and spatial data, the existing, expected or probable high risk zones cannot be accurately determined. Hence, building up effective and comprehensive countermeasures is required in order to decrease the severity of traffic accidents and surely to prevent them as far as possible.

- Briefly, integration of statistical and spatial analyses can be stated as the most comprehensive, effective and immediate solution for the development of road safety. Location-based studies open the way of interpreting the situation easily, accurately and with high efficiency.

- Regression analysis is used for the prediction or estimation of the value of a response variable based on the values of several predictor (independent) variables. There is generally a single response variable of interest that depends on a number of related quantities. Regression procedures, intended for quantitative response variables, are among the most widely used statistical techniques (70).

- Since traffic accident data are rare and random, multiple linear regression models are mostly unsuitable and invalid for determining the number of accident events. A continuous assumption and homogeneity on error terms give rise to a wrong distribution and lead to erroneous results (59). Therefore, Poisson, negative binomial or logistic regression models, which assume a non-linear distribution, are appropriate for the characteristics of traffic accident data.

- Logistic regression provides an alternative method for categorical responses. The model of logistic regression is basically constructed to predict the probability of a certain event. The data used are based on binary response variables where, for each observation, the response variable is recorded as “1” or “0” denoting that the event of interest did occur or did not occur, respectively (70). In this study, in order to examine the relative risk factor (or importance weight) of each potential factor -including constant road parameters and variable environmental parameters- to the traffic accidents, logistic regression issues were applied to the traffic accident related data extracted from traffic police records.

- Analytic Hierarchy Process (AHP) is one of the most vital Multicriteria Decision Making Analysis Methods that has a very critical role for this study to construct the link between the statistical and spatial analyses. The relative risk factors found from the logistic regression model was generalized and the direct importance weights (risk factors) of the variables were obtained through the logic of AHP. Fundamentally, comparison matrices were constructed in AHP in order to quantify and unify the pairwise risk factors regarding the consistency between the values. AHP also strengthens and simplifies the interpretation of the study because of its systematic organization and hierarchical structure.

- It is possible to display database information geographically with the help of Geographic Information Systems (GIS). The ability of associating spatial objects; *such as street name, route number*, with attribute information; *such as accident cause*, can be stated as the most useful aspect of GIS as a management tool (23).

- Many approaches are developed in order to analyze crash data by the tools of GIS. The variability of data; however, creates a complexity that can be simplified by the development of a topology to structure an assessment of the best approach for the required situation (58). In a GIS database, spatial relationships between objects are explicitly represented by the term of topology. It is basically used to describe the connection of linear objects, to define areas and to identify the relationships between adjacent areas (23).

- GIS technology provides the core framework for an integrated transport information system. The topological structure available in GIS database opens the way of analyzing transportation related data for different purposes in addition to querying capabilities. GIS can be thought as the most efficient tool for the integration of all types of data used in the area of transportation. Huge amount of data can be put together, organized, modeled and continuously updated with new data items effectively, easily, quickly and cheaply. Various alternatives can be assessed; zonal areas or themes can be detailed in order to analyze the characteristics of spatial data in the concept of the study. This, indeed, has a considerable importance for an effective, useful and comprehensive utilization in planning, design, construction, maintenance and management of transport system (43, 52).

- GIS provides a powerful tool for geographical and environmental issues. As stated before, it is possible to link or merge qualitative and quantitative data through spatial relationships, which is mostly done by overlaying function of GIS. Overlaying permits the GIS user to visually or topologically combine multiple maps (42). In general, the spatial data is stored in a number of layers in GIS.

- In addition to its ability of integrating several different data, the most appealing feature of GIS is its capability to present analysis results visually, as a map. Visualization of data using displaying features of GIS makes the studies more understandable and clear even for the inexperienced people. This ensures a better communication between the experts from various disciplines, which increases the efficiency of the study and its usage areas.

This study mainly presents a model for an effective traffic accident analysis, which integrates statistical and spatial tools. The logic of Multicriteria Decision Analysis is also utilized as an adhesive tool. In other words, the bridge between the statistical and spatial models is constructed with the help of a methodology based on Multicriteria Decision Making Systems.

The primary objective of this study is to show the importance, efficiency and valid applicability of a system that introduces GIS into a well-developed statistical traffic accident analysis. Furthermore, it is aimed to investigate the significance of studying the effects of roadway factors, which can be improved directly by engineering measures. Thus, an appropriate regression model -logistic regression- is applied to find the relative probable weights of the selected several predictor variables. Then, the general independent effect (i.e. risk factor or importance weight) of each variable is determined by using the logic of Analytic Hierarchy Process.

Afterwards, these risk factors are attached to the GIS layer, which is digitized according to the characteristics of roadways. Since the effects of the predictor variables are definite and introduced into GIS, high accident areas with respect to their road characteristics or road environment are visualized as clusters through the displaying and mapping capabilities of GIS.

In order to see the ability and efficiency of the model, several queries are made for extracting the required/desired information from GIS. These location-based queries ensure to comprehend the high crash locations and to discuss their probable causes. Thus, developing the most appropriate and effective safety measure can be easily achieved.

Moreover, since it is possible to change the weight factors, several future cases can be investigated by just changing the value of the factor. For instance, if the results show that the probability of having fatality in an accident at a definite location increases because of the high risk value of any road attributes, one can see the effect of any improvement by changing the value of that factor accordingly.

The high risky regions under different scenarios are also investigated and compared with the actual distribution of accidents in order to see the differences and similarities. This comparison gives the chance of evaluating the correctness of the estimated (probable) causes of accidents at that location for the construction stages of the model. As well, different scenarios show the variation of the same location under different environmental conditions, which can be beneficial to take effective remedies in the case of those circumstances.

The obtained results show that the proposed model supports the accident risk estimation based on traffic accident records. As a conclusion, the innovative model integrating statistical and spatial analyses is certified to be effective, appropriate, qualified and so comprehensive for identifying high accident risk areas, discussing their corresponding potential causal factors and developing appropriate location-based safety measures; in general for the analysis of traffic accidents. This directly improves safety levels and decreases the greatness and severity of the traffic accident problem in time.

7.3. Recommendations for Further Studies

As stated before, this study primarily emphasizes the invaluable role of using Geographic Information Systems coupled with statistical tools for traffic accident analysis. For this reason, a respected statistical study is combined with spatial analysis with one year traffic accident data.

Since the geographic coordinates of traffic accidents in Turkey are started to be recorded lately, it was not possible to use the most reliable time period -generally 3 years- for determining the importance weights of predictor variables. Moreover, the incomplete and erroneous structure of the collected data also obstacle to reach more accurate and valid coefficients.

For more sophisticated studies that would be carried out in the future with the basis of this study, more reliable and wide-range data should be gathered, which would ensure to reach better results and helps to develop better location-based safety measures. In addition, statistical analysis can be performed by changing the structure of the response variable according to the condition of accidents. Selection of the response variable requires a special care since it considerably affects the importance of the risk factors. Hence, accident studies should be performed after a detailed and careful decision-making process.

The spatial part of the proposed model is, unfortunately, limited with Ankara region since characteristics of roadway network of Turkey is not available as a GIS layer. Because of the time consideration, it is also not possible to digitize and assign attributes to all road network of the country. Therefore, the algorithm could also be arranged in a more detailed way for a more comprehensive study after constructing a complete GIS inventory layer of Turkey.

Finally, the model could be converted in a systematic form in order to increase the efficiency and easiness of the analysis and decrease the time spent for the application of the procedure required to reach a qualified solution. Thus, a computer program could be created, which uses and organizes the spatial and statistical data and realizes all of the processes with a simple management.

REFERENCES

1. U.S Department of Transportation, Federal Highway Administration Offices of Research and Development Implementation Division, 1980, *Safety Design and Operational Practices for Streets and Highways*, Washington D.C., USA

2. Commission of the European Communities, 2001, *White Paper-European Transport Policy for 2010: Time to Decide*, Brussels

3. Almqvist, S. and Hyden, C., 1994, *Methods for Assessing Traffic Safety in Developing Countries*, Building Issues, 6 (1)

4. Sadek, S., Bedran, M. and Kaysi, I., 1999, *GIS Platform for Multicriteria Evaluation of Route Alignments*, Journal of Transportation Engineering, ASCE, 125 (2), 144-151

5. Anderson, T., 2003, *Review of Current Practices in Recording Road Traffic Incident Data: With Specific Reference to Spatial Analysis and Road Policing Policy*, Centre for Advanced Spatial Analysis, CASA, University College London, London, UK

6. Roberts H. J., 1971, *The Causes, Ecology and Prevention of Traffic Accidents*, Charles Thomas Publisher, Springfield, Illinois, USA

7. Gao, H., *GIS Applications in Traffic Accidents*, Retrieved, March 14, 2005, from http://www.geog.ubc.ca/courses/geog516/presentations/hongtao2.htm

8. Saxena, A. and Babu, G., *GIS as an Aid to Identify Accident Patterns*, Retrieved, March 14, 2005, from http://www.gisdevelopment.net/application/Utility/transport/utilitytr0017.htm

9. Ekonomi Forumu - Wirtschaftsforum, 1998, *Karayollarında Trafik Güvenliği*, Friedrich Ebert Stiftung, İstanbul, Türkiye

10. Han, K., 2001, *Development of an Interoperable Geographic Information System Platform for Transportation Applications*, M.S. Dissertation, University of Manitoba, Canada

11. Karlaftis, M. G. and Tarko, A. P., 1998, *Heterogeneity Considerations in Accident Modeling,* Accident Analysis and Prevention, Elsevier Science, 30 (4), 425-433

12. Herrstedt, L., 1999, *Road Safety Introductory Report*, PIARC

13. İnal, A., 1995, *A Methodology of Accident Analysis for Road Safety Programs*, International Journal of Environmental Studies, Overseas Publishers Association (OPA), 47, 55-61

14. The World Bank, Retrieved, February 17, 2005, from http://www. worldbank.org/

15. Northwestern University Transportation Institute, 1946, *Accident Investigation Manual*, USA

16. XVIIIth World Road Congress, 1987, *Road Safety and Infrastructure*, Permanent International Association of Road Congress, Brussels

17. Grime, G., 1987, *Handbook of Road Safety Research*, Butterworth & Co Ltd., UK

18. *Belediyelere Ulaşım ve Trafik Hizmetlerinin Devri*, Retrieved, February 4, 2005, from http://www.icisleri.gov.tr/strateji/kitap/trafik/B7.doc

19. Wahlberg, A. E., 2002, *Characteristics of Low Speed Accidents with Busses in Public Transport,* Accident Analysis and Prevention, Elsevier Science, 34 (5), 637-647

20. O'Flaherty, C. A., 1997, *Transport Planning and Traffic Engineering*, John Wiley & Sons, Inc., New York, USA

21. The Institution of Highways and Transportation, 1986, *Highway Safety: Guidelines for Accident Reduction and Prevention 2nd Edition*, London, UK

22. Roess, R. G., Pragrassas, E. S. and McShane, W. R., 2004, *Traffic Engineering 3rd Edition,* Pearson Education, Inc., New Jersey, USA

23. Martin, P. T., Perrin, J., Hansen, B. and Barrios, A., 2000, *Accident Data Availability*, Department of Transportation, University Transportation Center Program, University of Utah, USA

24. Pignataro, L. J., 1973, *Traffic Engineering-Theory and Practice*, Prentice Hall, Inc., New Jersey, USA

25. DeSILVA, H. R., 1942, *Why We Have Automobile Accidents*, John Wiley & Sons, Inc., New York, USA

26. Gattis, J. L., Alguire, M. S. and Narla, S. R. K., 1996, *Guardrail End-Types, Vehicle Weights and Accident Severities*, Journal of Transportation Engineering, 122 (3), 210-214

27. Muhlrad, N., 1987, *Traffic Safety Research for Developing Countries*, Institut National De Recherche Sur Les Transports Et Leur Securite, Joinville, Arcueil

28. Khisty, C. J. and Lall, B. K., 1998, *Transportation Engineering – An Introduction 2nd Edition*, Prentice Hall, Inc., New Jersey, USA

29. International Road Traffic and Accident Database, IRTAD, Retrieved, March 2, 2005, from http://www.bast.de/htdocs/fachthemen/irtad/

30. Lee, J. and Mannering, F., 1999, *Analysis of Roadside Accident Frequency and Severity and Roadside Safety Management*, Washington State Transportation Center (TRAC), Washington, USA

31. Broyles R. W., Clarke, S. R., Narine, L. and Baker, D. R., 2001, *Factors Contributing to the Amount of Vehicular Damage Resulting from Collisions Between Four-Wheel Drive Vehicles and Passenger Cars*, Accident Analysis and Prevention, Elsevier Science, 33 (5), 673-678

32. Al-Ghamdi, A. S., 2003, *Analysis of Traffic Accidents at Urban Intersections in Riyadh*, Analysis and Prevention, Elsevier Science, 35 (5), 717-724

33. SWEROAD, 1998-2002, *Black Spot Manual*, The Turkish Traffic Safety Project General Directorate of Highways

34. Coppock, J. T. and Rhind, D. W., 1991, *The History of GIS.* In: Maguire, D.J., Goodchild, M. F. and Rhind D. W. (eds) , *Geographical Information Systems: Principles and Applications,* Volume 1, 21-43, Longman, London, UK

35. Longley P. A., Goodchild M. F., Maguire, D. J. and Rhind, D. W., 2001, *Geographic Information Systems and Science*, John Wiley & Sons, Ltd., England

36. *GIS History*, Retrieved, March 16, 2005, from http://www.geography.wisc.edu/sco/gis/history.html#60

37. Arcasoy, A., 2004, *SA 514 Geographical Information Systems in Settlement Archaeology Class Notes*, Settlement Archeology Graduate Program, Middle East Technical University, Ankara, Turkey

38. ESRI, 1996, *GIS Glossary*, Environmental Systems Research Institute, Redlands, California, USA

39. Osegueda R., Garcia-Diaz, A., Ashur, S., Melchor, O., Chang, S., Carrasco, C. and Kuyumcu, A. , 1999, *GIS Based Network Routing Procedures for Overweight and Oversized Vehicles*, Journal of Transportation Engineering, ASCE, 125 (4), 324-331

40. *Geographic Information Systems-An Introduction*, Retrieved, March 16, 2005, from http://visc.sis.pitt.edu/resources/tutorials/GIS_intro.pdf

41. *Geographic Information Systems as an Integrating Technology: Context, Concepts, and Definitions*, Retrieved, March 16, 2005, from http://www.colorado.edu/geography/gcraft/notes/intro/intro.html

42. Miles, S. B. and Ho, C. L., 1999, *Applications and Issues of GIS as Tool for Civil Engineering Modeling*, Journal of Computing in Civil Engineering, ASCE, 13 (3), 144-152

43. Heywood, I., Cornelius, S. and Carver, S., 1998, *An Introduction to Geographical Information Systems*, Longman Limited, New York, USA

44. Malczewski, J., 1999, *GIS and Multicriteria Decision Analysis*, John Wiley & Sons, Inc., Canada

45. Rodrigue, J. P., *Introduction to Geographical Information* Systems, Retrieved, March 16, 2005, from http://people.hofstra.edu/faculty/Jean-paul_Rodrigue/downloads/ GEOG 60 Topic 1a.ppt#16

46. Papacostas, C. S. and Prevedourous, P. D., 2001, *Transportation Engineering and Planning 3rd Edition*, Prentice Hall, New Jersey, USA

47. Chrisman, N., 1997, *Exploring Geographic Information Systems*, John Wiley & Sons, Inc., Canada

48. Aranoff, S., 1989, *Geographic Information Systems: A Management Perspective*, WDL Publications, Canada

49. Aranoff, S., 1989, *An Introduction to Geographic Information Technologies*, WDL Publications, Canada

50. Davis, B. E., 1996, *GIS: A Visual Approach*, On World Press, Santa Fe, New Mexico, USA

51. DeMers, M. N, 1997, *Fundamentals of Geographic Information Systems*, Wiley, New York; USA

52. Gupta, P., Jain, N., Sikdar, P. K. and Kumar, K., *Geographical Information System in Transportation Planning*, Retrieved, March 16, 2005, from http://www.gisdevelopment.net/technology/gis/ma03242.htm

53. Andrey, J., *Risk Assessment*, Retrieved, March 16, 2005, from http://www.people.hofstra.edu/geotrans/eng/ch7en/meth7en/ch7m3en.html

54. Gilbrook, M., 1998, *Finding the Appalachian Scenic Corridor*, Proc., 18th Annual (ESRI) User Conference, Environmental Systems Research Institute, Redlands, California, USA

55. Hall, J. P., Paulis, M. A., Wright, J. A., Lindquist, R. C. and Romanelli, A. J., 1997, *Implementation of a GIS at the Illinois Department of Transportation*, Proc., 17th Annual (ESRI) User Conference, Environmental Systems Research Institute, Redlands, California, USA

56. Chisalita, I. and Shahmehri, N., 2002, *An In-Vehicle Approach for Improving Traffic Safety through GIS Utilization*, IEEE International Conference on Systems, Man and Cybernetics, Hammamet, Tunisia

57. Federal Highway Administration US Department of Transportation, 1999, *GIS-Based Crash Referencing and Analysis System*, Turner-Fairbank Highway Research Center, Virginia, USA

58. Pawlovich, M. D., Souleyrette, R. R. and Strauss, T., 1998, *A Methodology for Studying Crash Dependence on Demographic and Socioeconomic Data*, 1998 Transportation Conference Proceedings, Center for Transportation Research and Education, Iowa, USA

59. Ng., K., Hung, W. and Wong, W., 2002, *An Algorithm for Assessing the Risk of Traffic Accident*, Journal of Safety Research, Elsevier Science, 33 (3), 387-410

60. Kleinbaum, D. G., 1994, *Logistic Regression: A Self-Learning Text*, Springer-Verlag New York, Inc., USA

61. Hosmer, D. W. and Lemeshow, S., 2000, *Applied Logistic Regression 2nd Edition*, John Wiley & Sons, Inc., Canada

62. Acar, O. S., 2002, *CE 452 Traffic Safety and Accident Investigation Class Notes*, Middle East Technical University, Ankara, Turkey

63. MINITAB Statistical Software, *Meet Minitab*, Minitab Inc.

64. Al-Ghamdi, A. S., 2002, *Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity*, Accident Analysis and Prevention, Elsevier Science, 34 (6), 729-741

65. Saaty, T. L., 1980, *The Analytical Hierarchical Process*, McGraw Hill, New York, USA

66. Skibniewski, M. J. and Chao, L., 1992, *Evaluation of Advanced Construction Technology with AHP Method*, Journal of Construction Engineering and Management, ASCE, 118 (3), 577-593

67. MathCAD Release 11, *Mathsoft-MathCAD.11 User's Guide*, Mathsoft Engineering & Education, Inc.

68. ArcGIS, *Using ArcMap*, Environmental Systems Research Institute

69. Tiglaco, N. C., *Development of Traffic Accident Information System using Geographic Information System (GIS)*, Retrieved, March 17, 2005, from http://www.gisdevelopment.net/aars/acrs/1998/ts11/ts11002.shtml

70. Conerly, M., Gray, B., Busby, K. and Mansfield, E., *Data Mining and Visualization of the Alabama Accident Database*, Retrieved, April 6, 2005, from http://utca.eng.ua.edu/ projects/final_reports/99113report.pdf