USING GEOGRAPHIC INFORMATION SYSTEMS IN ANALYSING THE PATTERN OF CRIME INCIDENTS AND THE RELATIONSHIP BETWEEN LANDUSE AND INCIDENTS

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

BY

EBRU AKPINAR

IN PARTIAL FULLFILMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
GEODETIC AND GEOGRAPHIC INFORMATION TECHNOLOGIES

FEBRUARY 2005

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: Ebru Akpınar

Signature:

ABSTRACT

USING GEOGRAPHIC INFORMATION SYSTEMS IN ANALYSING THE PATTERN OF CRIME INCIDENTS AND THE RELATIONSHIP BETWEEN LANDUSE AND INCIDENTS

Akpınar, Ebru
M.Sc., Department of Geodetic and Geographic Information Technologies
Supervisor: Assoc. Prof. Dr. Melih Pınarcıoğlu

February 2005, 138 pages

The location where crime events occur is an important factor in the analysis of crime. It is not only important where a crime incident takes place but also the characteristics of those places and the environment in which incident occurs. Thus, examination of spatial data such as street networks, parcel information, school locations, commercial and residential zoning, is imperative for effective crime analysis, crime prevention and police activities.

Using variables, relating to five incident types as burglary, auto, pick pocket, usurp and murder, which are recorded by the police in 2003 and including the information of where and when crime incident occurs, this thesis attempts to analyze the high-incident areas and the relationship between these incidents and land uses for two important police precincts of Çankaya district of Ankara; Centre of Çankaya Police Station Zone and Bahçelievler Police Station Zone.

The aim of the study is to improve a methodology for the determination of the effects of land use variables on the distribution of crime incidents. The study applies some methods related to the spatial data analysis which can be also integrated with

Geographic Information Systems. In the study, in order to identify the current pattern of the incidents, kernel estimation method is found as the best method while Nearest Neighbor Hierarchical Clustering method is preferred to analyze what land uses are prone to incidents. In order to determine the properties of the relationships between land use and incidents correlation calculations are performed. At the end, the potential crime incident areas are determined by referring these relationships with GIS functions.

Keywords: crime incidents, spatial data analysis, Geographic Information Systems (GIS), land use, correlation

ÖZ

SUÇ OLAYLARININ DAĞILIMI VE ARAZİ KULLANIMI İLE OLAYLAR ARASINDAKİ İLİŞKİNİN ANALİZ EDİLMESİNDE COĞRAFİ BİLGİ SİSTEMİNİN KULLANILMASI

Akpınar, Ebru
Yüksek Lisans, Joedezi ve Coğrafi Bilgi Teknolojileri Bölümü
Tez Yöneticisi: Doç. Dr. Melih Pınarcıoğlu

Şubat 2005, 138 sayfa

Suç olaylarının gerçekleştiği yer suç analizinde önemli bir faktördür. Sadece suçun nerede gerçekleştiği değil, aynı zamanda suç olaylarının meydana geldiği bu alanların ve çevrenin özellikleri de önemlidir. Bu nedenle, yol ağı, parsel bilgisi, okul yerleri, ticari ve konut bölgeleri gibi mekansal verilerin incelenmesi, etkili suç analizi, suç önleme ve polis aktiviteleri için gereklidir.

Bu tez, Ankara’nın Çankaya ilçesi’ nin iki önemli polis mıntıkası; Çankaya Merkez Karakolu ve Bahçelievler Karakolu; için 2003 yılında polis tarafından kaydedilen ve suç olayının nerede ve ne zaman meydana geldiği bilgisini içeren hırsızlık, oto, yankesicilik, gasp ve cinayet olaylarına ait değişkenleri kullanarak, yoğun olay alanlarını ve bu olaylar ile arazi kullanımı arasındaki ilşkiyi analiz eder.

Bu tezin amacı, arazi kullanım değişkenlerinin suç olayları üzerindeki etkilerini belirlemek için bir metodoloji geliştirmekdir. Çalışma, Coğrafi Bilgi Teknolojileri ile bütünleşebilen mekansal veri analizi metotlarını uygular. Çalışmada, olayların mevcuttaki dağılımını tespit etmek için kernel metodu en iyi metot olarak bulunurken, Nearest Neighbor Hierarchical Clustering metotu ne tür arazi kullanımlarınının olaylara yatkın olduğunu analiz etmekte tercih edilir. Arazi kullanımı ve olaylar arasındaki ilişkilerin özelliklerini bulmak için korelasyon hesaplamaları yapılır. Sonunda, bu ilişkilere dayanarak CBS fonksiyonlarının yardımı ile potansiyel suç olayları alanları belirlenir.

Anahtar Kelimeler: suç olayları, mekansal veri analizi, Coğrafi Bilgi Sistemleri (CBS), arazi kullanımı, korelasyon

To my parents Nilgün and Ahmet Akpınar
and to my sister Pınar Akpınar

Thank you for your love and support

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Assoc. Prof. Dr. Melih Pınarcıoğlu for his guidance and encouragement throughout the development of this thesis. His assistance made this thesis valuable.

Thanks go to my teachers Assoc. Prof. Dr. Nurunnisa Usul, Assist. Prof. Dr. Zuhal Akyürek and Assist. Prof. Dr. Şebnem Düzgün who are the people that bring me together with GIS and share their invaluable knowledge. This research would have been impossible without them.

I also appreciate the support of Ankara Security Directorate and General Directorate of Security. Special thanks to Ercüment Yılmaz who is the Director of Ankara Security for his support in providing data. In addition assistances of Hasan Yön and Mesut Demirbilek are gratefully acknowledged.

I would like to thank my colleagues M. Emin Halitligil, Evren Mercimek, Ayhan Saraçoğlu, Köksal Sanlı and Cemal Ertaş, in INTA SpaceTurk, for their supports and understandings.

I would like to express my deepest thanks to Murat Kaya for his morale support, comments and suggestions. In addition, thanks go to my friends Ebru Yıldırım, Seçkin Ak, Ayça Turhan, Aslı Eski, Özgün Balkanay, Birsen Coşkun, Balkan Uraz, Ezgi Ulusoy, Nilhan Çiftçi, Nilay Köse, Ece Aksoy and Mehmet Tunçer for their keen friendship and support at my desperate times. It is great to know that I have such good friends throughout my life.

Finally, I would like to thank my family for their encouragement and patience during my study.

TABLE OF CONTENTS

PLAGIARISM iii
ABSTRACT iv
ÖZ vi
DEDICATION viii
ACKNOWLEDGMENTS ix
TABLE OF CONTENTS x
LIST OF TABLES xii
LIST OF MAPS xiii
LIST OF FIGURES xv
CHAPTER
1. INTRODUCTION 1
2. THEORETICAL AND PRACTICAL VIEW ON CRIME INCIDENTS 6
2.1. Main Concept of Crime and the Causes of Crime Incidents 6
2.2. Environmental Criminology 11
2.3. Geographic Information Systems in Crime Incident Analysis 14
2.3.1. Hot Spot Analysis 15
2.3.2. Spatial Data Analysis 19
2.3.2.1. Quadrat Method 21
2.3.2.2. Kernel Estimation 21
2.3.2.3. Nearest Neighbor Distances 23
2.3.2.4. K-function 24
2.3.2.5. Nearest Neighbor Hierarchical Clustering (Nnh) 25
2.3.2.6. Spatial and Temporal Analysis of Crime (STAC) 26
2.3.2.7. K-Means Clustering 27
2.3.2.8. Correlation Analysis of Geographic Data 27
3. CURRENT STATUS OF INCIDENTS IN THE STUDY AREA 30
3.1. Definition of the Study Area 30
3.2. Data Used in the Study 34

3.3. Descriptive Analysis for Incident Rates 42
4. GENERATION OF CRIME INCIDENT ANALYSIS IN THE STUDY AREA 51
4.1. Application of Centrographic Techniques 51
4.2. Exploring Spatial Point Pattern Analysis 54
4.2.1. Application of Quadrat Method 54
4.2.2. Application of Kernel Estimation 60
4.2.3. Application of Nearest Neighbor Distances 71
4.2.4. Application of K-function 73
4.2.5. Application of Nearest Neighbor Hierarchical Clustering 76
4.2.6. Application of Spatial and Temporal Analysis of Crime 78
4.2.7. Application of K-Means Clustering 81
4.2.8. Evaluation of the Outcomes 83
5. DETERMINATION OF THE RELATIONSHIP OF INCIDENTS BETWEEN LANDUSE 86
5.1. The Inputs of the Correlation Calculations 86
5.2. Correlation Between Land Use and Incidents 95
5.2.1. Determining the Relationship Between Burglary Incidents and Land Use 95
5.2.2. Determining the Relationship Between Auto Incidents and Land Use 99
5.2.3. Determining the Relationship Between Pick Pocket Incidents and Land Use 102
5.2.4. Determining the Relationship Between Usurp Incidents and Land Use 105
5.2.5. Determining the Relationship Between Murder Incidents and Land Use 109
5.3. Determining the Risky Areas according to the Correlation Results 112
5.4. Evaluation of the Outputs 125
6. CONCLUSION AND RECOMMENDATIONS 128
REFERENCES 134
ONLINE REFERENCES.. 137
INTERVIEWS 138

LIST OF TABLES

TABLES

2.1: Crime Prevention Approaches 9

3.1: Data used in the Study Area 37

5.1: The Size and Rate of the Land Use Types inside each Neighborhood 88

5.2: The Amount and Proportion of the Incident Types inside Each Neighborhood 89

5.3: Incidents Inside Buffer Rings of Schools 92

5.4: Incidents Inside Buffer Rings of Police Stations 92

5.5: Incidents Inside Buffer Rings of Hospital 93

5.6: Incidents Inside Buffer Rings of Subway Stops 93

5.7: Incidents Inside Buffer Rings of Carparks 94

5.8: Correlation Coefficient Results for Burglary Incidents 97

5.9: Correlation Coefficient Results for Auto Incidents 101

5.10: Correlation Coefficient Results for Pick Pocket Incidents 104

5.11: Correlation Coefficient Results for Usurp Incidents 107

5.12: Correlation Coefficient Results for Murder Incidents 111

5.13: The Directions of the Correlations 113

5.14: Correlation coefficient Values and Weights 121

LIST OF FIGURES

FIGURES

2.1: Distribution Types for Points 20
3.1: Incidents in Ankara 33
3.2: Incidents in Çankaya 33
3.3: Land Use Distribution of the Neighborhoods 40
3.4: Distribution of Iincident Types 43
3.5: The Neighborhoods and the Places of the Incidents 44
3.6: Temporal Distribution of the Burglary Incidents 48
3.7: Temporal Distribution of the Auto Incidents 48
3.8: Temporal Distribution of the Pick Pocket Incidents 49
3.9: Temporal Distribution of the Usurp Incidents 49
3.10: Temporal Distribution of the Murder Incidents 50
4.1: Spatial Summary Statistics for the Incidents 53
4.2: The Nearest Neighbor Distances for the Incidents 72
4.3: K-function for the Incidents 74
5.1: Scatter Graphs for Burglary Between Land Use and Landmarks 96
5.2: Scatter Graphs for Auto Between Land Use and Landmarks 100
5.3: Scatter Graphs for Pick Pocket Between Land Use and Landmarks 103
5.4: Scatter Graphs for Usurp Between Land Use and Landmarks 106
5.5: Scatter Graphs for Murder Between Land Use and Landmarks 109
5.6: Raster Calculation for Burglary Incidents 122

LIST OF MAPS

MAPS

2.1: A Map Showing Robbery Hot Spots, using STAC, West Bronx, New York ….. 18
3.1: The Study Area …………………………………………………………………… 31
3.2: Crime Rate in Turkey,2000 ……………………………………………………… 32
3.3: Distribution of incidents occurred in 2003 in the Study Area …………………… 35
3.4: Land Use and Landmark data in Study Area ………………………………… 36
3.5: Point Pattern of the Incidents in the Study Area ……………………………… 37
4.1: Quadrat analysis for Burglary Incidents ……………………………………… 55
4.2: Quadrat Analysis for Auto Incidents ………………………………………… 56
4.3: Quadrat Analysis for Pick Pocket Incidents ………………………………… 57
4.4: Quadrat Analysis for Usurp Incidents ……………………………………… 58
4.5: Quadrat Analysis for Murder Incidents ……………………………………… 59
4.6: Kernel Estimation for Burglary Incidents …………………………………… 61
4.7: Kernel Estimation for Auto Incidents ……………………………………… 63
4.8: Kernel Estimation for Pick Pocket Incidents ………………………………… 65
4.9: Kernel Estimation for Usurp Incidents ……………………………………… 67
4.10: Kernel Estimation for Murder incidents …………………………………… 69
4.11: Nearest Neighbor Hierarchical Clustering for the Incidents ………………… 76
4.12: Spatial and Temporal Analysis of Crime (STAC) for the Incidents ………… 79
4.13: K-means Clustering for the Incidents ……………………………………… 81
5.1: Buffer Zones for Landmark Points ………………………………………… 90
5.2: Reclassification of Residential Areas ……………………………………… 114
5.3: Reclassification of Commercial Areas ……………………………………… 114
5.4: Reclassification of Park Areas …………………………………………… 115
5.5: Reclassification of Roads ……………………………………………… 116
5.6: Straight line for Point Features ………………………………………… 117
5.7: Reclassification of the Distances to Schools ………………………………… 118
5.8: Reclassification of the Distances to Police Stations ………………………… 119
5.9: Reclassification of the Distances to Car Parks ……………………………… 119

5.10: Reclassification of the Distances to Hospitals 120
5.11: Reclassification of the Distances to Subway Stops 120
5.12: Comparison of Potential Crime Areas with Hotspot Areas 123

CHAPTER I

INTRODUCTION

The rate of crime incidents is increasing in all developing countries due to change of technology and materialistic way of life and also due to poor social, political, economic and environmental conditions. Distribution of the incidents across the landscape is not geographically random since incidents are human phenomena. For incidents to occur, offenders and their targets - the victims and/or property - be required to exist at the same location for a period of time. Several factors, including the lure of potential targets and simple geographic convenience for an offender, influence where people choose to break the law.

Brantingham and Brantingham (1991) describe crime as a complex event and outline four things that have effects on the occurrence of an incident: a law, an offender, a target, and a place. A crime occurs when these four dimensions are in concurrence. Without a law there is no crime. Without an offender, someone who breaks the law, there is no crime. Without some object, target, or victim, there is no crime. Without a place in time and space where the other three come together, there is no crime.

The first dimension of crime, the legal dimension, is studied by lawyers and sociologists and is practiced in the legislatures and in the courts. The law defines some specific behavior, prohibits it, and provides for the punishment of those who continue to do it. Social and life scientists study the offender dimension as the second dimension which is practiced in courts, in prisons, and in social work agencies. Offender-centered criminologists ask questions about the specific behaviors of criminal offenders and about the motives for those behaviors. The third dimension of crime, the target dimension, is studied by social scientists and practiced by police and private security firms, by social agencies, and by individuals. The target or object or victim of criminal events is investigated in third dimension.

When the target of crime is a person, criminologists talk about their psychology, social situation or economic circumstance, or politics.

Environmental criminology is the study of the fourth dimension of crime which is place, discrete location in time and space at which the other three dimensions intersect and a criminal event occurs. Environmental criminologists begin their study of crime by asking questions about where and when crimes occur. They ask about the physical and social characteristics of crime sites; the movements that bring the offender and target together at the crime site; the perceptual processes that lead to the selection of crime sites and the social processes of ecological labeling; the spatial patterning in laws and the ways in which legal rules create crime sites, spatial distributions of targets and offenders in urban, suburban and rural settings.

Socio-demographic information which refers to the characteristics of individuals and groups such as sex, race, income, age and education is also equally important with place, but the thesis deals with only the fourth dimension of crime instead of the other dimensions including law, offender and target.

Many authors like Lab (2000), Weisburd and McEwen (1998), Bowers and Hirschfield (1999) mentioned the relationship between land use, time and incidents in their studies. They argue that particular areas may be devoted to different types of land use (residential development, retailing, industry, leisure, open space) and based on its land use type the activities and population profile of an area may vary considerably according to the day of the week or time of day.

Neighborhood bars, for example, experience more aggression and violence than lounges in nice restaurants. Similarly, assault is more prevalent in the evenings than during mid-morning. Knowing even these two basic facts; time and place; shows that many problems cluster in certain time and location.

Naturally crime does not disappear on its own. Police departments are on the duty of protecting the citizens' safety and taking precautions to minimize the risk of incidents. It has long been common practice for the police to identify locations and times that are more prone to criminal activity (Lab, 2000). To reduce or completely eliminate the incident, some actions such as crime prevention methods should be taken. Crime prevention can be signified as a set of ideas for combating incident and includes the activities taken by individuals and groups, both public and private.

The first step of crime prevention is to analyze the current status of incidents such as determining the density or pattern of the incidents. Lab (2000) claims that the findings that incident often concentrates in certain locations or at certain times, suggests that the targeting of hot spots: areas of the highest incident concentration (Ratcliffe and McCullagh, 1999) should be the starting point for crime prevention. Thus determining the hotspot areas is the prompt analysis to uncover what factors make a location a good spot for crime in the thesis.

The usage of Geographic Information Systems (GIS) in data storage, manipulation and display makes incident prevention process more manageable, more realistic and case specific. CrimeStat, which is the most widely used software for determining the pattern of incidents in American police agencies (web1) is used to show hotspot areas in the study area. The results of integrated CrimeStat into GIS give an idea about the current status of incident pattern. Due to its spatial operation capability, GIS helps police and also other people who are interested in incidents, to visualize and analyze the spatial relationships between different data layers such as incidents and land use, to forecast and take precautions for future incidents.

The knowledge derived through GIS provides a simple geographic inventory (e.g., indicate the location of all events that occurred in a specific location) or display different spatial patterns (e.g., the locations of burglaries may vary from the locations of pickpockets) or presents the results of more complex analyses (Wieczorek and Hanson, 1997).

With improving a methodology this thesis aimed to answer the questions such as whether or not the events within a certain area on the map are densely clustered and what land use types are prone to incident and how land use types are used in order to detect the risky incident areas. In addition considering five different incident types (burglary, auto, pick pocket, usurp and murder) the thesis examines if the spatial pattern of each incident are similar.

In order to answer these questions two basic steps are performed. Firstly the locations of the highest concentrations of recorded incidents and changes in the locations of these areas are identified based on incident types by hot spot analysis, within the software of CrimeStat and ArcGIS 8.2. Since there are numerous methods that can be used to analyze hotspot areas and which provide different results, a few of them are generated in order to find the best visualization. Secondly the relationship between the incident locations and the land use types is tried to be established with the help of SPSS and ArcGIS.

For finding out the land use parameters that are supposed to be more prone for incidents the land use types within the hotspot areas are investigated. After selecting land use parameters, the relationship between these land use and the incidents are studied by the calculation of correlation coefficients. Thus the strength and the direction of these relationships are determined. By referring to the properties of these correlations some values which indicate the importance on the occurrence of the incidents, are given to each land use parameter. Then, potential risky areas for the occurrence of the incidents are determined without considering the information of incidents and the results are compared with the hotspot areas in order to test the accuracy of these risky areas.

The lack of systematic records and the geographic database with regard to both land use and crime incidents, make it difficult to reach more detailed and correct results. Especially, it is the most problematic step to obtain the incident data from the police who would not like to publish the information of crime, occurred in their precincts. On the other hand, some land use and incident data that were obtained for this study, may not be concerned to the same period. Because, it is impossible to get the required data from one source and for the same time in Turkey. Thus many organizations and companies had to be visited while collecting the necessary data in this thesis. However, within the framework of the data that can be reached, following steps are carried in the chapters below.

In the second chapter of the thesis, definitions and causes of the crime and incidents are identified. Since the thesis deal with the location of the events, it emphasizes the effects of the land uses instead of the social, economic or demographic factors. Crime prevention approaches are explained briefly in this chapter and the importance of the crime area analysis in crime prevention process is introduced. Than, the concept of environmental criminology is explained and the role and effective use of GIS in this context and the methods of spatial data analysis are described.

In chapter three, the characteristics of the study area are dealt with. The chapter sets the current status of incidents in the study area. In order to do so, the data related to incidents are presented. In addition the characteristics of the land use of the study area are given in this chapter. The data used in this study include the digital maps of the study area with the necessary layers of information for land use and official incident records.

In the fourth chapter the spatial patterns of incidents are determined by CrimeStat within a GIS software. In order to find the best presentation of hotspot areas which mean the highest concentration of the incidents, some GIS based statistical methods are applied to the study area in this chapter. The methods which are developed for spatial point analysis give different results and different visualizations.

The relationship between incidents and land use will be tried to be determined based on some specific variables in the fifth chapter. The strength and the direction between these variables are determined by the correlation coefficient calculations. Regarding to these calculations, their weights which are the indicators of their effects on the crime events are determined for each land use based on the properties of the correlation. Then potential areas for the occurrence of the incidents are determined and compared with the hotspot areas.

In the conclusion chapter, an evaluation is maid by the aim, objectives of the study and the analyses carried to satisfy these objectives. Significant findings of the thesis are assessed and opinions are stated. One last consideration is given to the outcomes to guide future studies.

CHAPTER II

THEORETICAL AND PRACTICAL VIEW ON CRIME INCIDENTS

There are numerous references in varied disciplines that include the investigations about the causes and origins of crime incidents. Various approaches and techniques are being developed to assist crime mappers to aggregate spatial data both to make their analysis more reasonable and protect identification of the locations of victims. In this framework the chapter gives brief definitions for crime incidents and crime analysis. Than the approaches and techniques for incident prevention are investigated.

2.1. Main Concept of Crime and the Causes of Crime Incidents

Crime is a comprehensive concept that can be defined in legal and non-legal sense (Ahmadi, 2003). From a legal point of view, it refers to breaches of the criminal laws that govern particular geographic areas (jurisdictions) and are aimed at protecting the lives, property and rights of citizens within those jurisdictions. A crime is a criminal act against a person (for example, murder and sexual assault), or his/her property (for example, theft and property damage) and regulation (for example, traffic violations).

Non-legal point of view would define crime as acts that violate socially accepted rules of human ethical or moral behavior. As the moral principles that underpin the notion of crime are subject to gradual change over time, the types of behavior defined by the legal system as criminal may also change. Examples of behaviors that have been de-criminalized in some jurisdictions include prostitution, abortion, attempted suicide and homosexual intercourse. Other behaviors, such as tax evasion or credit card fraud, have been criminalized over time.

Ratcliffe and McCullagh (1999) made a distinction between incidents and recorded crime, where an incident is a request for police assistance but may not involve criminal activity whereas recorded crime may lead to criminal prosecution.

At the interview with Atasoy (2004), she outlined three components of the occurrence of the incidents: offender, victim and place. Some precautions could be taken for victims, for instance, people are warned to place an alarm into their cars, not to carry their bags full of valuable things. However, apart from arresting the offender and social and economical developments there are not too many ways to decline the rate of the offenders. In this case the best and the easiest way to reduce the rate of the crime incidents, is to take some precautions by changing the usage or the structure of the place.

Investigating the causes or origins of crime requires many researches by various disciplines such as sociology, psychology, criminology or economy. Economic status, unemployment, age, household size, female headed households, education, race, religion, nationality, country of origin, disability, gender or sexual orientation, history of receiving public assistance are only few examples of the possible affects of crime.

In addition, the geographic pattern of the incidents indicates that incident does not occur in random or at unpredictable locations. Rather, incidents occur in observable structures that are influenced by the landscape in which they occur. It is this maxim that makes geographic profiling a powerful investigative tool.

Many authors like Block (1998) mentioned the relationship between incident and land use. Below issues have an influence on the location, time, type and rate of the incidents:

- Land use data sets containing information on each parcel of land in the city (for example, vacant or not, abandoned or not, residential or commercial, state of repair, specific function such as tavern or convenience store, and so on)
- Public transit data sets (train or bus stops and routes)
- Schools (grammar, high schools, private)

- Community organizations (block clubs, religious centers, social service agencies)
- Parks and other open areas (with park roads, field houses, lagoons)
- Emergency locations (hospitals, fire houses, police stations)
- Places holding liquor licences (by type of establishment and licence)

In this case, in order to identify incident pattern and areas which are suitable for the occurrence of incidents, some analysis may be performed. Below, few of them are given (Canter, 1998; Block 1998):

- identify areas that may likely be targeted by an offender;
- determine whether common attributes exits among a group of reported cases
- explore relationships between incidents and other geographic features such as land use and the built environment
- study the movement of offenders to predict the location of future targets to establish interdiction points along escape routes
- detect whether the incident locations are clustered
- determine if incidents tend to be located close to a specific location such as taverns or gang territories or the periphery of a county or in the center

With the purpose of abating incidents, preventing loss and saving lives, some precautions should be taken. At this point crime prevention is important for everyone from the individual to the public. Lab (2000) defines crime prevention that it entails any action designed to reduce the actual level of crime and/or the perceived fear of crime. Brantingham and Faust (1976) divide crime prevention into three approaches: primary, secondary and tertiary (Table 2.1).

Table 2.1: Crime Prevention Approaches (Lab, 2000)

Primary Prevention	Secondary Prevention	Tertiary Prevention
Environmental Design • Architectural Design • Lighting • Access control • Property identification	**Identification and Prediction** • Early ID of problem individuals • Crime area analysis	**Specific Deterrence**
Neighborhood Watch • Surveillance • Citizen patrols	**Situational Crime Prevention** • Problem identification	**Incapacitation**
General Deterrence • Arrest and conviction • Sentencing methods	**Community Policing**	**Rehabilitation and Treatment**
Public Education • Levels of incident • Fear • Self-help	**Substance Abuse** • Prevention and treatment	
Social Incident Prevention • Unemployment • Poverty • Employment/Job training	**Schools and Crime Prevention**	
Private Security		

According to Brantingham and Faust (1976) primary prevention identifies conditions of the physical and social environment that provide opportunities for criminal acts. Secondary prevention engages in early identification of potential offenders and seeks to intervene prior to commission of illegal activity. Implicit in secondary prevention is the ability to correctly identify and predict problem people and situations. These solutions may involve physical design changes, altering social behaviors, improving surveillance, or any number of other activities. Tertiary prevention deals with actual offenders and involves intervention in such a fashion that they will not commit further offenses. The majority of tertiary prevention rests within the workings of the criminal justice system.

The distinction between primary and secondary prevention efforts rests on whether the programs are aimed more at keeping problems that lead to criminal activity from arising (primary prevention) or if the efforts are focused on factors that already exist and are fostering deviant behavior (secondary prevention). For example schools can play an important role in secondary prevention both in terms of identifying problem youths and in providing a forum for interventions.

The results obtained from the first analyses in the thesis are the parts of secondary incident prevention in which crime area analysis and problems that have effects on incidents are identified. The further analyses are the parts of primary prevention in which physical conditions that may precipitate incidents are investigated.

Crime analysis involves the collection and analysis of data pertaining to a criminal incident, offender and target (Canter, 2000). One of the most important purposes of crime analysis is to identify and generate the information needed to assist in decisions regarding the deployment of police resources to prevent and suppress criminal activity. Additionally, it is possible to use crime analysis to evaluate the effectiveness of programs such as community policing and crime prevention, develop policy through research, justify budget requests, and help define a problem.

Boba (2001) divided Crime Analysis into five types: intelligence, criminal investigative, tactical, strategic and administrative. Intelligence analysis is the study of "organized" criminal activity, whether or not it is reported to law enforcement, to assist investigative personnel in linking people, events, and property. Criminal investigative analysis is the study of serial criminals, victims, and/or crime scenes as well as physical, socio-demographic, psychological, and geographic characteristics to develop patterns that will assist in linking together and solving current serial criminal activity.

Tactical crime analysis is the study of recent criminal incidents and potential criminal activity by examining characteristics such as how, when and where the activity has occurred to assist in problem solving by developing patterns and trends, identifying investigative leads/suspects, and clearing cases. Strategic crime analysis is the study of crime and law enforcement information integrated with socio-demographic and spatial factors to determine long term "patterns" of activity, to assist in problem solving, as well as to research and evaluate responses and procedures.

Finally, administrative crime analysis is the presentation of interesting findings of crime research and analysis based on legal, political, and practical concerns to inform audiences within law enforcement administration, city government/council, and citizens. At the first glance strategic and tactical crime analysis look like same. In both these two types, incident patterns with the information of how, when and where the incident has occurred are studied. The main difference between strategic and tactical crime analysis is the timeliness of the data. Strategic crime analysis usually involves data covering at least a year-long period, same as the data used in the thesis, whereas tactical crime analysis uses data collected during several days (Canter, 2000).

2.2. Environmental Criminology

Mapping itself has a long history, but crime mapping specifically can be traced back to early 1800s in France when social theorists began to create maps to illustrate their theories and research about crime. In relation to crime and policing, maps initially were used to examine issues like poverty or demographic characteristics and crime (Boba, 2001; Weisburd and McEwen, 1998).

In 1829, Adriano Balbi and Andre-Michel Guerry created the first maps of crime. Using criminal statistics for the years 1825 to 1827 and demographic data from France's latest census, they developed maps of crimes against property, crimes against human, and levels of education. Comparing these maps, they found that the northeastern portion of France (from Orleans to the FrancheComte) was better educated, that area with high levels of crimes against property had low incidences of attacks on people, and that the areas with more property crime were populated by people with higher levels of education (Weisburd and McEwen, 1998).

In 1831 and 1832 the Belgian astronomer and statistician Lambert-Adolphe Quetelet independently published three maps dealing with the same themes but spreading across larger areas. He saw a correlation between crime and several variables including transportation routes, education levels, and ethnic and cultural variations.

One of the first police departments to use mapping was New York City in the 1900s. The maps consisted of simple wall maps in which "push pins" were used to indicate crimes that had occurred (Boba, 2001). In order to examine the spatial distribution of incident locations, police places pushpins in wall maps. In the maps which contain all of the streets for an area of interest such as a police precinct or a municipality, incident locations are usually represented by a pin. Thus the relationship between a particular point location such as street robbery to other geographic features such as bus stop or shopping center could be determined (Canter, 1998).

More sophisticated maps were developed by a group of scholars associated with the University of Chicago in the 1920s and 1930s. These urban sociologists, led by Robert Park, looked to characteristics of the urban environment to explain the crime problem in American cities (Boba, 2001).

However because of the increase in the number of incidents the amount of effort for maintaining pin maps manually, become difficult and problematic. The limitations of the manual pin mapping compromise the geographic accuracy of the incident locations. Since the geographic location of an incident on a pin map is not quantifiable, analysts are not able to test hypotheses about the spatial distribution of incident. The maps need to be periodically updated for new roads or other geographic features.

For this reason, in the late 1960s the idea of automated crime mapping was created (Canter, 1998). At the same time that advances in computer and information systems have largely turned over the technological barriers to mapping of crime, innovations in crime prevention theory have pushed the concept of place to the center of research and practice in controlling crime.

In the 1980s the focus of crime prevention began to shift from their focus on people and involvement in criminality. In its broadest terms, this new perspective sought to develop a greater understanding of crime and more effective crime prevention strategies through concern with the physical, organizational, and social environments that make crime possible. From the outset, the concept of place became a central concern of scholars in this area (Weisburd and McEwen, 1998).

This shift emerged *environmental criminology*, and academics began to examine the spatial characteristics of crime as well as how location characteristics might contribute to criminal activity at particular locations over others.

Since the location information can provide clues for identifying the suspects, assist in the design of prevention or apprehension strategies, aid in the evaluation of programs, and help gain a better understanding of environmental factors that may affect the incident, the location of an incident and any other geographic features associated with a criminal event become important attribute features. Many researchers argued that the modification of specific features of urban design would reduce crime with their works. This idea attracted criminologists, planners, geographers, environmental psychologists and architects to the study of the environment in which incidents occur (Brantingham and Brantingham, 1991).

Environmental Criminology emphasized the roles of space and place in determining the time, location and character of crimes (Brantingham and Brantingham, 1991). Environmental Criminology argues that full crime analysis has four dimensions: a legal dimension, an offender dimension, a victim or target dimension, and a spatio-temporal or locational dimension.

Brantingham and Brantingham (1991) states that victims are victimized near places where they spend most of their time – home, work, school, shopping, entertainment – and along the major pathways in between. They imply that crime can be understood and predicted through analysis of a city's land use patterns, its street network, and its transportation system. It has become clear that particular types of land uses such as fast food restaurants, bars or high schools are criminogenic because of the volume of people they attract and, sometimes, because of the nature of their juxtaposition with other land uses can affect the crime rates of entire neighborhoods. Street networks along with traffic and transit patterns strongly affect the distribution of crimes. Indeed, incident rates vary quite substantially by land use type, and it is possible to build predictive models of the incident potential of any given point within a city.

Environmental criminology is currently studied at three separate levels of analysis (Brantingham and Brantingham, 1991). The first one is macro-analysis which involves studies at the highest levels of spatial aggregation. It involves studies of distribution of crime between countries, between the states or provinces or cities within a particular country, or between the countries or cities within a state.

The second one is meso-analysis which involves the study of crime within the subareas of a city or metropolis. The range of such studies includes very large subunits such as the constituent cities of a metropolitan area, or smaller subunits such as planning areas or police precincts or census tracts, or very small subunits such as faces of individual city blocks. Research at this level of analysis which is also the level of the thesis, examines the distribution of crimes in relation of the distribution of targets; of routine daily activities such as work, school, shopping and recreation locations; of zoned land uses.

The third one is micro-analysis which involves the study of specific crime sites. At this level of analysis the focus is on building type and its placement on a lot, landscaping and lightning, interior form, and security hardware.

2.3. Geographic Information Systems in Crime Incident Analysis

With the improvement in Geographic Information Systems, police became able to produce more versatile electronic maps by combining their databases of reported incident locations with digitized maps of the areas they serve. GIS opens new opportunities for the use of digital mapping in incident control and prevention programs. GIS allows police personnel to make plans effectively for emergency response, determine mitigation priorities, analyze historical events, and predict future events and it helps the related officers to determine potential incident sites and facilitates to explore the relationship between incident and land use.

GIS has been used to examine crime incidents in relation to the distribution of residential properties, community facilities, administrative boundaries and street network. The analysis of relations between land use and crime and the targeting of crime prevention strategies can be facilitated greatly by information systems capable of handling spatially – referenced incident data and cross-referencing them contextual information on land use (Bowers and Hirschfield, 1999). This is where GIS-linked applications have the most to offer.

According to Canter (1998) GIS have three broad applications. One of them is the forwarded data mapping which is used to map attributes contained in the database files that are linked to a geographic location by a GIS. Thus it is possible to determine when and where to deploy police resources in response to a particular incident problem.

The backward data mapping allows an analyst to display the information based on the relationship between an incident's location and geography. It is the first step in gaining a better understanding of the relationship between incident and geography. It is used for spatial analysis such as cluster analysis, quadrat analysis, or nearest neighbor analysis. The analyzing ability is enhanced in two ways: first, by associating a map feature such as a road, transit station, or school to an incident location and second, to relate incident locations to a geographic component, such as longitude and latitude, size or distance.

And the third one is interactive data modeling which involves using a GIS to predict or simulate some phenomenon, in other words develop some scenarios. Because testing and evaluating different simulations, such as estimating response times, have a significant affect on police operations and policy.

The ability of GIS to relate and synthesize data from a variety of sources enables analysts to examine various aspects of crime analysis, including the determination of the problematic areas with the term of hotspot areas and the identification the spatial factors. There are several ways to determine the hotspot areas and also the factors that cause these hotspot areas by utilizing GIS.

2.3.1. Hot Spot Analysis

Police analysts monitor the geographic distribution of incidents in order to identify high-incident areas or "hot-spots". The definition of Ratcliffe and McCullagh (1999) says that "Hotspots are aggregations of the raw crime data, designed to identify the sites of highest incident concentration". Canter (1998) determines three criteria for the definition of high incident area: incident frequency, geography and time. These areas contain at least two criminal incidents of the same crime type. A high incident area is usually identified as such based on the number of offenses reported over the most recent one-to two week period. Once a high-incident area is identified, the analyst will continue to monitor the pattern over time until it abates (Canter, 1998).

Block and Block (1995) examine incident data for three Chicago communities. Using mapping techniques, they report that hot spots often surround elevated transit stops and major intersections. These areas are locations where potential victims can be located and potential offenders have options for escape.

Another example of hot spot analysis is the work of Rengert (1997) looking at auto theft in Philadelphia. While his results show that hot spots clearly exist, the analysis provides an important qualifier concerning the influence of time of day. Rengert (1997) shows that, the locations of hot spots change according to different times of the day and night. Tourist attractions and educational institutions may be hot spots for auto theft during the day, while entertainment venues, bars and other adult night spots become greater target areas in the evenings and at night.

Clearly it can be said that hot spots can be anywhere – businesses, schools, abandoned buildings, vacant lots, housing complexes or intersections – or anytime – evenings, late night, weekends, holiday, or vacation months. For instance, the study Nelson et al. (2001) shows primary clusters at night in the pub/club leisure zones; and secondary clusters during the shopping day in major retail streets. Disorder data also reveal subsidiary afternoon clusters near licensed premises, and a late-night confluence flashpoint at a node of pedestrian activity.

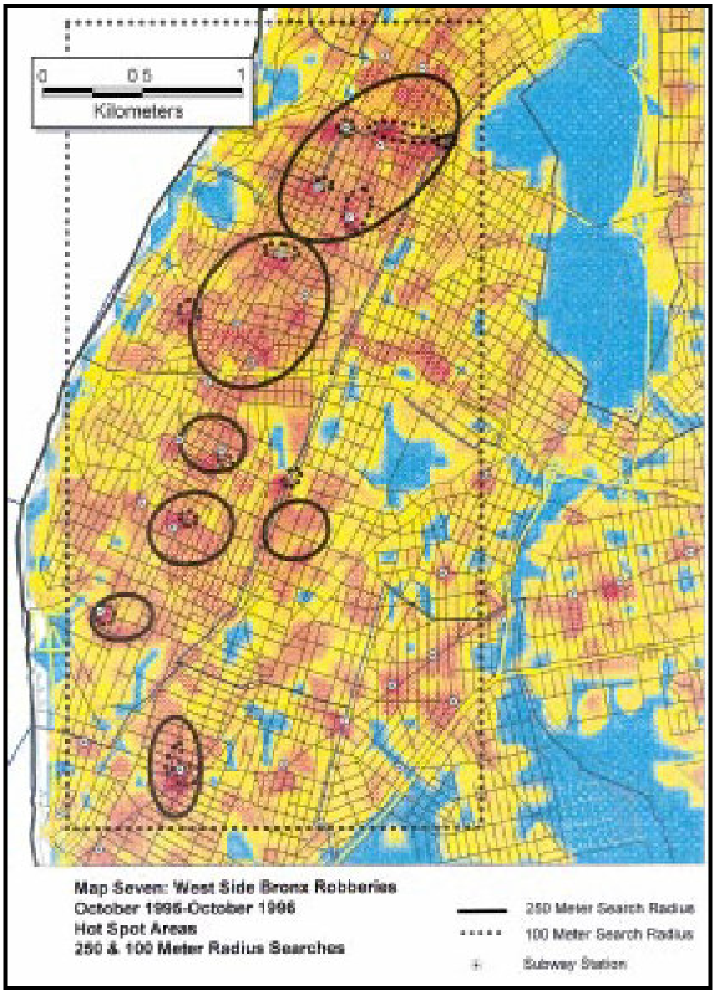
There are advantages and disadvantages of determining hotspot areas. On one hand, it has the potential to cause falling house prices, increasing insurance premiums or business abandonment, conflicts may exist between providing a public service and protecting the individual (Ratcliffe, 2001). It may prevent citizens to feel themselves in a safe and it causes disorder. Good employees may be unwilling to work in high incident regions and this may impact on the quality of education available to schoolchildren in the region.

Highlighting high incident areas have an adverse affect on a number of socio-economic indicators such as real estate value. In addition personal privacy could be damaged. For example, victims of a burglary would not want their address to be in the public domain, advertising that their property was vulnerable to burglary.

On the other hand, attempts to identify hot spots are useful in pointing out what types of incidents and geographic locations coincide. Hot spot research represents innovative approaches to narrowing the individuals or situations that will be targeted by incident prevention activities. It can be used for informing the public about relative risk. Increased community knowledge about incident may increase community co-operation, public awareness of neighborhood problems and police department accountability.

At the interview with Chainey (2004) he stated that hotspot analysis and hotspot maps are tools used to assist in the allocation and targeting of policing and crime reduction resources, and have application in areas for immediate operational patrolling responses to the direction of strategic crime prevention resources. They provide that first step that helps to identify problem areas which can then be analysed in further detail to reveal the specific problems that need addressing.

Most of the police analysts in developed countries identify a high-incident area – hotspots by the method of standard deviational ellipse. The standard deviational ellipse is computed using the Illinois Criminal Justice Authority's Spatial and Temporal Analysis of Crime (STAC) (Map 2.1). STAC is a spatial statistics program for the analysis of incident locations developed by Ned Levine and Associates under grants from National Institute of Justice, has new version called Crime Stat II (Levine, 2002).



Map 2.1. A Map Showing Robbery Hot Spots, using STAC, West Bronx, New York (Harries, 1999)

CrimeStat is a spatial statistics package (web2) that can analyze crime incident location data and it is able to link to a GIS (Levine, 2002). In addition to standard ellipses many point pattern analysis could be generated by Crime Stat.

Generating hot spot areas is possible with an input that includes x, y coordinate pairs for each incident location, a search radius, and a set of parameters used to define the search area. Hot Spot Areas turn point data into areas that reflect the actual scatter of points over the map. Hot Spot Areas do not lose the detailed information of point patterns scatter. It is an objective, quick, database-driven method for finding dense areas (Block and Block, 1995).

Once hot spots which are geographical areas are defined, it is possible to describe the characteristics of crimes located within each hot spot area, including the number of crimes, their locations, and the dates and times of their occurrence. Also, analysts can zoom in to get more detailed view of the locations of crimes within hot spots and to assess the links between crime occurrence and the locations of streets, buildings, parks, and other facilities. Hot spot areas are great importance for crime control and prevention (McLafferty et al., 2000).

2.3.2. Spatial Data Analysis

Spatial data analysis deals with the situation where observational data are available on some process operating in space and methods are sought to describe or explain the behavior of this process and its possible relationship to other spatial phenomena (Bailey and Gatrell, 1995). Spatial data analysis involves the accurate description of data relating to a process operating in space, the exploration of patterns and relationships in such data and the search for explanations of such patterns and relationships.

Spatial data analysis deals with spatial statistics of which have two branches (Walford, 1994):

- Centrographic techniques produce a quantitative measure describing the central tendency (mode, mean and median) or dispersion (range, interquartile range, variance and standard deviation) of location of a class of entities.

- Pattern analysis helps to determine the distribution type of spatial entities which are points, lines or areas. A point, by which crime incident locations are symbolized in the thesis, is the most fundamental means of representing a geographical phenomenon and requires a single pair of coordinates. The three main types of spatial distribution which refers to density; clustered, regular and random are illustrated in Figure 2.1.

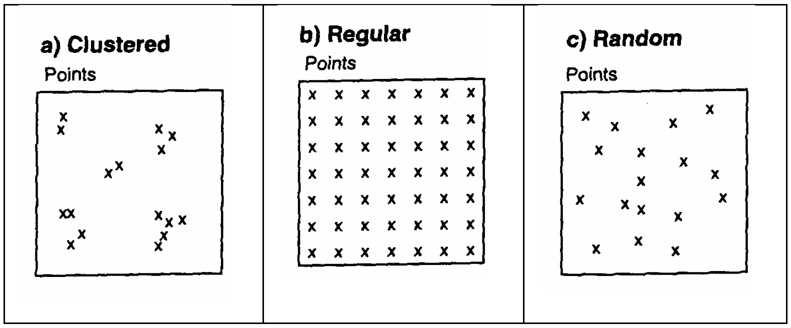


Figure 2.1: Distribution Types for Points (Walford, 1994)

The spatial behavior of the point is often studied within two concepts; first order effects and second order effects. **First order effects** are described in terms of the intensity, of the process, which is the mean number of events per unit area at the point in space – a global or large scale trend. **Second order effects,** or spatial dependence, involve the relationship between numbers of events in pairs of areas in the study – local or small scale effects (Bailey and Gatrell, 1995).

First-order properties are global because they represent the dominant pattern of distribution – where it is centered, how far it spreads out, and whether there is any orientation or direction to its dispersion. Second-order or local properties, on the other hand, refer to sub-regional patterns or neighborhood patterns within the overall distribution. If there are distinct 'hot spots' where many crime incidents cluster together, their distribution is spatially related not so much to the overall global pattern as to something unique in the sub-region or neighborhood (Levine, 2002).

There are so many techniques for summarizing the pattern in the locations of the events in a given area. In the thesis different methods are used in obtaining the pattern of the incidents in order to complement each other, compare if they provide different results and test if the risky areas which are determined in the end of the thesis present similar pattern. These methods are quadrat analysis, kernel estimation, Nearest Neighbor Hierarchical Clustering (NNH), STAC and K-Means Clustering which are generated for first order effects; nearest neighbor distances and the K function which are generated for second order effects. In order to perform these methods Crime Stat II, Golden Software Surfer 8 (web3) Excel and ArcGIS 8.2 softwares are used.

2.3.2.1. Quadrat analysis

Quadrat analysis involves overlaying the study area with a regular square grid and counting the number of points falling in each of squares and converting this to an intensity measure dividing by the area of each of the squares. The result will give some indication of whether and how the intensity of the process is changing over the study area.

The formulation of finding the length of a grid's size is much simpler:

$$l = \sqrt{\frac{2 * A}{N}} \qquad \text{Eq. (2.1)}$$

In which l is the length of a grid's size, A is the size of the search area and N is the number of incidents (Canter, 1998).

However results may depend on the size and orientation. Since it is based primarily on the density of points, and not their arrangement in relation to one another, it does not give the real pattern.

In addition, the problem is that although Quadrat methods give a global idea of sub-regions with high or low intensity they throw away much of the spatial detail in the observed pattern. If the quadrats are smaller to retain more spatial information, variability in quadrat counts will be high and this degenerates into a mosaic with many empty quadrats (Bailey and Gatrell, 1995).

2.3.2.2. Kernel Estimation

A way round the problem of Quadrat method is through the use of counts per unit area in a 'moving window' (Bailey and Gatrell, 1995). One defines a suitable size 'window' which is then moved over a fine grid of locations in the study area and the intensity at each grid point is estimated from the event count per unit area within the 'window' centred on that point. This produces a more spatially 'smooth' estimate of the way in which intensity is varying. This technique is known as **Kernel Estimation**.

A smaller value of the bandwidth would be more appropriate to avoid smoothing out too much detail. In other words setting a bandwidth which would be appropriate for sparse areas unfortunately obscures detail in dense areas.

Crime Stat includes five kernel estimators: normal, uniform, quartic, triangular and negative exponential (Levine, 2002). The main difference between these is that the normal includes all points in the pattern, whereas the others have a distance cut-off beyond which no points are included in the kernel estimation (web4).

The normal distribution weighs all points in the study area, though near points are weighted more highly than distant points. The other techniques use a circumscribed circle around the grid cell. The uniform distribution weighs all points within the circle equally. The quartic function weighs near points more than far points, but the fall of is gradual. The triangular function is weighs near points more than far points within the circle, but the fall off is more rapid. The negative exponential weighs near point much more highly than far points within the circle (Levine, 2002). The use of any of one of these depends on how much the user wants to weigh near points relative far points.

The normal distribution produces an estimate over the entire region whereas the other four produce estimates only for the circumscribed bandwidth radius. Within the four circumscribed functions, the uniform and quartic tend to smooth the data more whereas the triangular and negative exponential tend to emphasize 'peaks' and 'valleys'.

Each of the kernels allows to specify a kernel bandwidth as a fixed interval or to choose an adaptive interval. With a fixed bandwidth, the user must specify the interval to be used and the units of the measurement. An adaptive bandwidth adjusts the bandwidth interval so that the number of points is found.

2.3.2.3. Nearest-neighbor Distances

Quadrat counting and kernel estimation are basically concerned with exploring the first order properties of a spatial point pattern, in other words with estimating the way in which the intensity varies in the study region. However Bailey and Gatrell (1995) claim that, there is a close relationship between the distribution of inter-event distances and second order properties. **Nearest-neighbor analysis** concentrates on the spacing between points rather than their overall pattern. Using distances between observed events in the study area, it is possible to investigate second order properties.

In nearest neighbor analysis the nearest neighbour event-event distance w is taken as being the distance between randomly chosen event and the nearest neighbouring event. In order to perform nearest neighbour analysis the empirical cumulative probability distribution or distribution function G(w) of w is estimated by below formula in which n is the number of events w is the event-event nearest neighbour distances (Bailey and Gatrell, 1995).

G(w)=number of (wi≤w) / n Eq. (2.2)

Nearest Neighbour analysis involves calculating the ratio between the observed and expected mean minimum distance between the points to produce the quantity nearest neighbour index (NNI), which can lie either side of unity (1) with the range 0.0 to 2.1491 (Walford, 1994).

NNI=0 ⇨ Perfect clustered pattern; all points lie on top of each other
NNI=1.0 ⇨ Perfect random pattern
NNI=2.1491 ⇨ Perfect regular pattern; forms a triangular lattice

The nearest-neighbor index is defined by below equations (Walford, 1994):

$$NNI = \frac{r_o}{r_e} \qquad r_e = 0.5\sqrt{\frac{A}{N}} \qquad r_o = \frac{\sum_{i=1}^{n} d_i}{n} \qquad \text{Eq. (2.3)}$$

Where; ro = Observed mean minimum distance

re = Expected density

A = The area

N = The number of points

d = Minimum distance

2.3.2.4. K function

One of the problems of the nearest neighbour distance method is that they use distances only to the closest events and hence only consider the smallest scales of pattern. Information on larger scales of pattern is ignored. An alternative approach is to use an estimate of the reduced second moment measure or **K function** of the observed process, which provides a more effective summary of spatial dependence over wider range of scales (Bailey and Gatrell, 1995).

Bailey and Gatrell (1995) defines the K function:

$\lambda K(h) = E$ (#(events within distance h of an arbitrary event)) Eq. (2.4)

Where; # = Number of

E () = Expectation operator

λ = Intensity (mean number of events / unit area)

The practical value of *K (h)* as a summary measure of second order effects is that it is feasible to obtain a direct estimate of it, ($\hat{K}(h)$) from an observed point pattern. If *R* is the area of the study area, then the expected number of events in the study area is λR. The expected number of pairs of events a distance at most h apart is $\lambda^2 RK(h)$. If d_{ij} is the distance between i_{th} and j_{th} observed events in the study area and $Ih(d_{ij})$ is an indicator function which is 1 *if* $d_{ij} \leq h$ and 0 otherwise, then the observed number of pairs is $\sum\sum_{i \neq j} I_h(d_{ij})$. Therefore a suitable estimate of *K(h)* is :

$$\hat{K}(h) = \frac{1}{\lambda^2 R} \sum\sum_{i \neq j} I_h(d_{ij})$$ Eq. (2.5)

The random occurrences of the events imply that an event at any point in *R* is independent of other events and equally likely over the whole of *R*. Hence for a random process the expected number of events within a distance of h of a randomly chosen event would be $\pi\lambda h^2$.

The K function for a random event should be: $K(h) = \lambda\pi h2/\lambda = \pi h2$
If the point pattern has regularity then $K(h) < \pi h2$
If the point pattern has clustering $K(h) > \pi h2$

For the observed data (Bailey and Gatrell, 1995), the estimated $\hat{K}(h)$ is compared with πh^2. One way of doing this can be achieved by plotting *L(h)* against *h*, where;

$$\hat{L}(h) = \sqrt{\frac{\hat{K}(h)}{\pi}} - h \qquad \text{Eq. (2.6)}$$

2.3.2.5. Nearest Neighbor Hierarchical Clustering (Nnh)

The **nearest neighbor hierarchical clustering (Nnh)** routine in CrimeStat identifies groups of incidents that are spatially close. It is a hierarchical clustering routine that clusters points together on the basis of a criteria and proceeds to group the clusters together. The clustering is repeated until either all points are grouped into a single cluster or else the clustering criteria fails (Levine, 2002).

The CrimeStat Nnh routine uses a nearest neighbor method that defines a threshold distance and compares the threshold to the distances for all pairs of points. Only points that are closer to one or more other points than the threshold distance are selected for clustering. Only points that fit both criteria - closer than the threshold and belonging to a group having the minimum number of points, are clustered at the first level (first-order clusters).

The routine then conducts subsequent clustering to produce a hierarchy of clusters. The first-order clusters are themselves clustered into second-order clusters. Again, only clusters that are spatially closer than a threshold distance (calculated a new for the second level) are included. This re-clustering process is continued until no more clustering is possible, either all clusters converge into a single cluster or, more likely, the clustering criteria fail.

2.3.2.6. Spatial and Temporal Analysis of Crime (STAC)

The **STAC** Hot Spot Area routine in CrimeStat searches for and identifies the densest clusters of incidents (hotspot areas) based on the scatter of points on the study area. STAC is a scan-type clustering algorithm in which a circle is repeatedly laid over a grid and the number of points within the circle is counted. It, thus shares with those other scan routines the property of multiple tests, but it differs in that the overlapping clusters are combined into larger cluster until there are no longer any overlapping circles.

STAC in CrimeStat program implements a search algorithm, looking for Hot Spot Areas. Firstly it lays out a 20 x 20 grid structure on the plane defined by the area boundary. Than it places a circle on every node of the grid, with a radius equal to 1.414 (the square root of 2) times the specified search radius. Thus, the circles overlap.

After that, it counts the number of points falling within each circle, and ranks the circles in descending order. For a maximum of 25 circles, STAC records all circles with at least two data points along with the number of points within each circle. The X and Y coordinates of any node with at least two incidents within the search radius are recorded, along with the number of data points found for each node.

These circles are then ranked according to the number of points and the top 25 search areas are selected. If a point belongs to two different circles, the points within the circles are combined. This process is repeated until there are no overlapping circles. The result is called Hot Clusters and represented by the best-fitting standard deviational ellipses (Levine, 2002).

STAC is a useful approach for visual pattern identification or to obtain a quick picture from a large quantity of data, but it is not suitable for more rigorous analysis or model building (Chakravorty and Pelfrey, 2000).

2.3.2.7. K-Means Clustering

The K-means clustering routine is a partitioning procedure where the data are grouped into K groups defined by the user. The routine tries to find the best positioning of the K centers and then assigns each point to the center that is nearest. Like the Nnh routine, the K-means assigns points to one, and only one, cluster.

However, unlike the Nnh procedure, all points are assigned to clusters. Thus, there is no hierarchy in the routine, that is there are no second- and higher-order clusters. This technique is useful when a user want to control the grouping (Levine, 2002).

2.4.2.8. Correlation Analysis of Geographic Data

Correlation analysis is used to determine whether a relationship exists between two variables (Walford, 1994). Correlation coefficient (r) is a numerical value between -1 and +1 ($-1<r<1$) which measures the extent and nature of any relationship. The magnitude of r quantifies the strength of the relationship while the sign of r denotes the direction of the relationship.

r = 1 ⇨ Perfect positive correlation
r = -1⇨ Perfect negative correlation
r = 0 ⇨ No correlation.

There are 5 ways of determining a correlation coefficient depending on measurement scale of the data:

- Pearson's product moment correlation coefficient
- Spearman's rank correlation coefficient
- Kendall's tau correlation coefficient
- Phi correlation coefficient
- Biserial correlation coefficient.

Pearson's Moment Correlation Coefficient is based on the notion of covariance, which is the development of the concept of variance, and describes the correspondence or co-variation of two variables together. It is a parametric measure of the strength and direction of relationship between two variables, measured on an interval scale.

The formula for the coefficient where the two variables are represented by x and y, n denotes the number of items, ẋ and ẏ indicates the means, is (Walford, 1994):

$$r = \frac{\frac{\sum (x-\overline{x})(y-\overline{y})}{n-1}}{\sqrt{\frac{\sum (x-\overline{x})^2}{n-1}}\sqrt{\frac{\sum (y-\overline{y})^2}{n-1}}} \qquad \text{Eq. (2.7)}$$

In Boggs's study (1966), the occurrence of crime is related to environmental opportunity rather than to a population base or to some indicator of seriousness. Boggs suggests that "using these crime-specific occurrence rates, one can determine whether crime targets in certain areas are exploited at higher rates than targets in other neighborhoods". She further notes the disparity between high crime occurrence in urban central business districts contrasting with high offender rates in "lower class, non-white, anomic neighborhoods."

Boggs (1996) calculates correlations between crime-specific and standard crime occurrence rates in order to evaluate the concordance between the two types of measurement. Rank order correlations are lowest for business-related offenses and highest for residential burglary, rape, and homicide/aggravated assault in her study.

In the study of Phillips (1973) an example of risk-related crime rates for Minneapolis, Minnesota is developed. He suggests that the business-to-residential land use ratio as a base for business-oriented crimes was somewhat misleading in the sense that business and residential land uses vary substantially in intensity within the city. He combines business robbery and burglary owing to the similarity of targets and high degree of correlation in their patterns of occurrence. The population-based pattern and the pattern based on employment are approximately inversely correlated (rank-order r=-0.38). While the population-based rates are highly centralized, employment-based rates are lowest in the central business district.

For armed robberies, an additional rate is calculated based on commercial structures; for residential burglaries, a rate is calculated on a base of dwelling units in the study of Pyle (1974) in Akron, Ohio. Correlation analysis produces an r of 0.098 between residential burglary calculated on the bases of population and dwelling units.

The use of correlation analysis helps to find out the relationship between various variables and incidents. They can give more concrete information about the relationships. Furthermore the results of these correlation calculations bring about new researches for the study area. The following parts of the thesis try to demonstrate these possible analyses by using the results of the correlations.

CHAPTER III

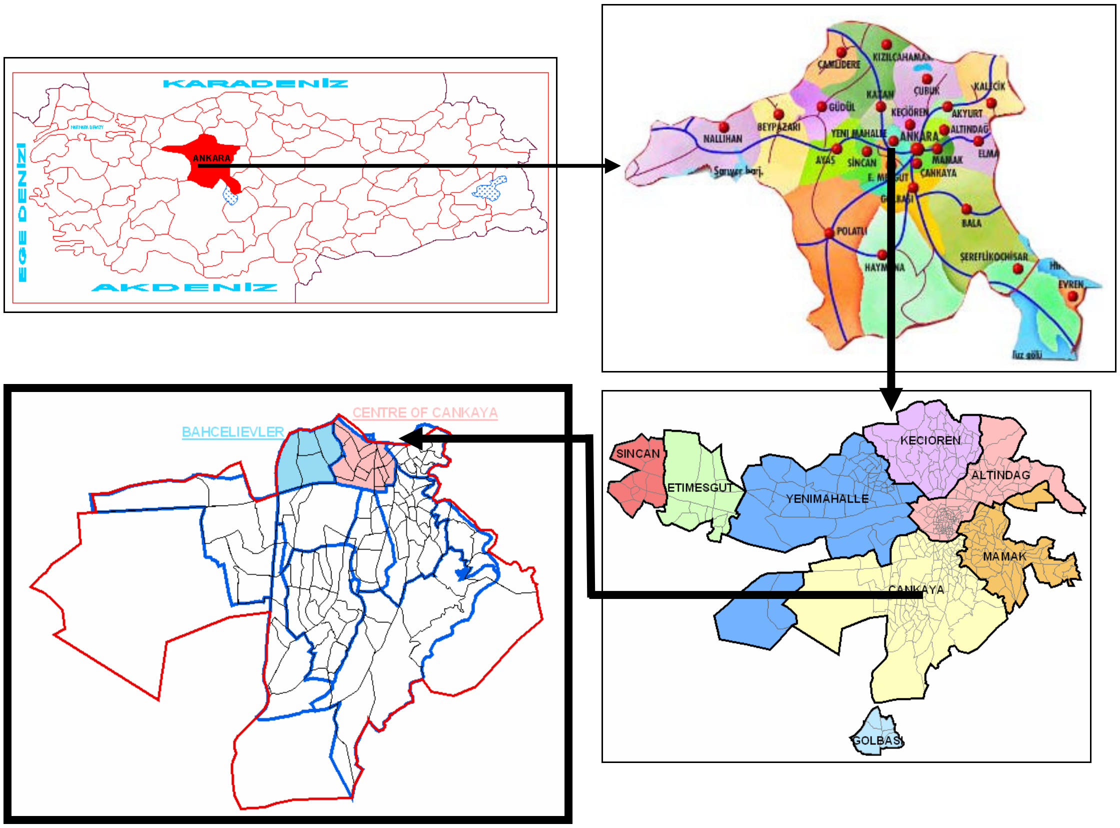
CURRENT STATUS OF INCIDENTS IN THE STUDY AREA

In this chapter, the case study area and the incidents which have occurred in 2003 in the study area are described. In order to fulfill the explanation of the current status in the area, descriptive analyses are presented which involves the investigation of the study area in terms of incident distribution based on time, type and place information and land use characteristics by using GIS.

3.1. Definition of the Study Area

The study area is located in Çankaya which is the most developed district of Ankara, the capital city of Turkey. The district of Çankaya is controlled by 10 police precincts; these are Bahcelievler, Dikmen, Centre, Cebeci, On Nisan, Esat, Kavaklıdere, Yıldızevler, Şehit Mustafa Düzgün and Ellinci Yıl.

The study area is comprised of two police precincts - Centre of Çankaya Police Station Zone and Bahcelievler Police Station Zone - located in Çankaya district. Centre of Çankaya Police Station Zone includes 15 neighborhoods - Kızılay, Meşrutiyet, Yücetepe, Anıttepe, Eti, Maltepe, Korkut Reis, Kavaklidere, Sağlık, Fidanlık, Namık Kemal, Cumhuriyet, Kültür, Kocatepe and Devlet. The second zone - Bahçelievler Police Station Zone - includes 4 neighborhoods - Emek, Bahçelievler, Yukarı Bahçelievler and Mebusevleri. With these two precincts the study area is approximately 8 sq.km. (Map 3.1).

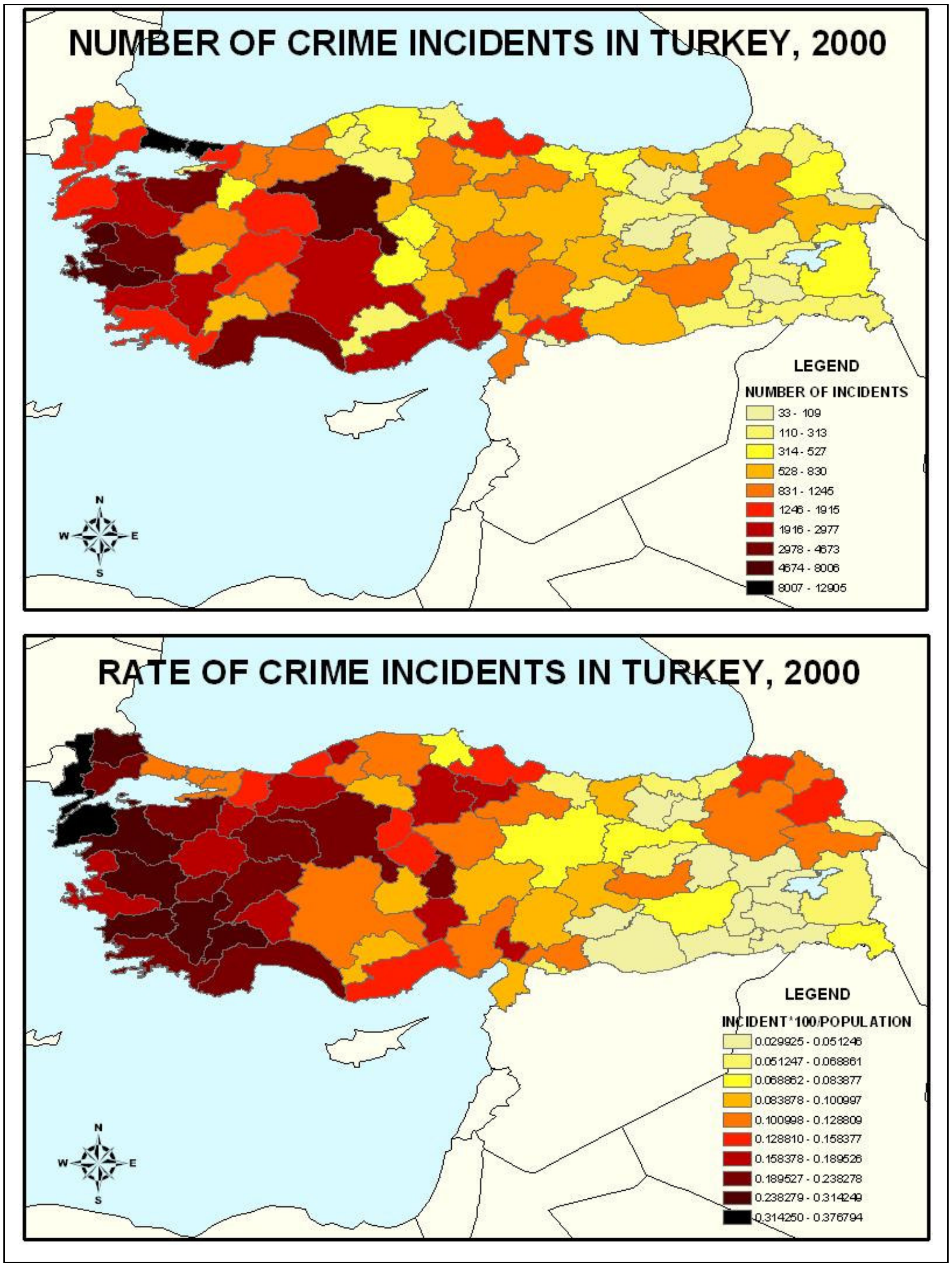


Map 3.1 The Study Area

The urban population of Ankara is 3.540.522 in the year of 2000 and the most crowded district is Çankaya with the population of 758.490 (DİE, 2000). In view of the fact that 80% of the cultural facilities in Ankara like cinema, theatre, culture and convention centre, museum and art galleries are located in Çankaya and this district is the centre of culture for Ankara (Çankaya Belediyesi, 2002). In Ankara, also in Çankaya, the economy mostly concentrates in service sector since governmental top positions are generally located in Ankara, especially in Çankaya. Hence public associations, embassies and military functions cover large areas in Çankaya.

The commercial areas in Çankaya cover 4% while residential areas cover 75% of the total surface of Çankaya which is 203 sq.km. The mixed usages including both residential and commercial cover 16% of the total. The rest covers 5% which includes education, culture, health, social, sports, public association and military (Çankaya Belediyesi, 2001).

Map 3.2 indicates the number and the rate of crime incidents, happened in 2000 in Turkey. As seen in the upper side of the Map, the most incidents were recorded in Istanbul than Ankara and Izmir. On the other hand when the rate of the incidents to the population in 2000 is considered, under side of the Map is obtained which explains the percentage of the incidents based on the population.



Map 3.2. Crime Incidents in Turkey, 2000 (web5)

According to the statistical analysis of crime rate in Ankara which is done by Ankara Police Directorate, the results show that Çankaya with its crime rate in December 2002 takes part in the first line comparing to the other districts of Ankara (Figure 3.1). As for the police precincts in Çankaya the same analysis indicates that while Centre of Çankaya Police Station Zone is in the first line, Bahçelievler sets in the fifth line (Figure 3.2).

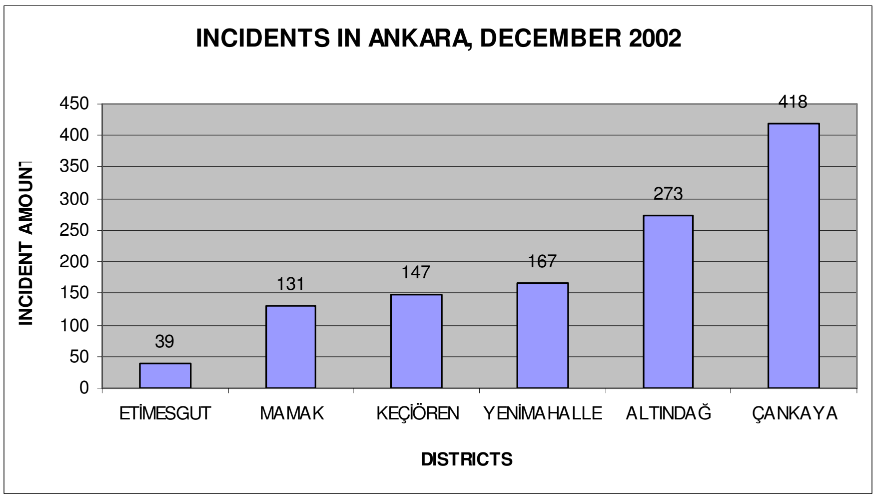


Figure 3.1: Incidents in Ankara (Ankara Emniyet Müdürlüğü, 2002)

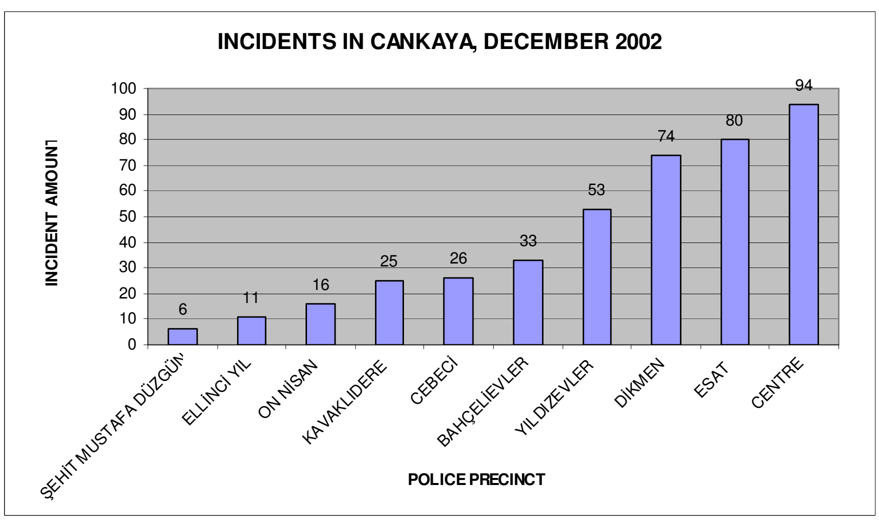


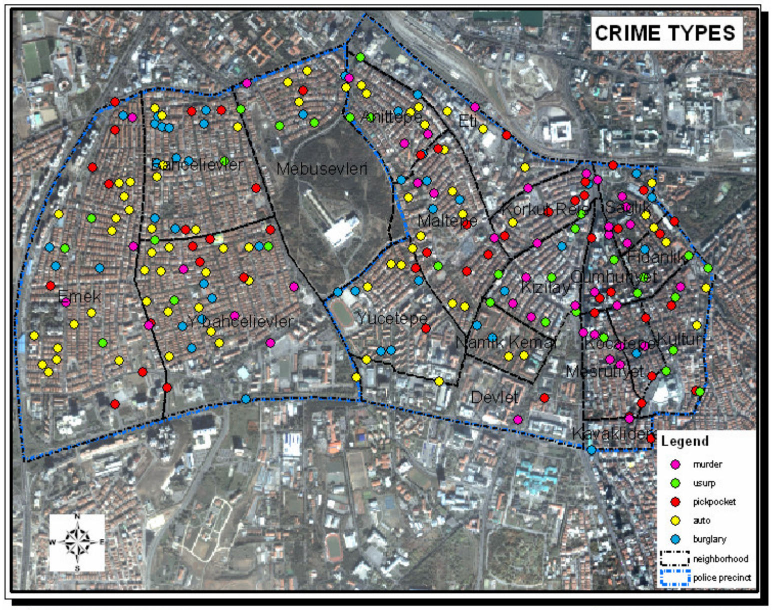
Figure 3.2: Incidents in Çankaya (Ankara Emniyet Müdürlüğü, 2002)

In Çankaya, mostly incidents of burglary occurred with a higher amount when compared to the other incident types, in December 2002. The amount of burglary incidents was recorded as 236 in December 2002. Then pick pocket with 48 incidents follows auto with 93 incidents. On the other hand, the amount of usurp incidents is 16 while the amount of murder is 16 (Ankara Emniyet Müdürlüğü, 2002).

In addition to the economic factors, district of Cankaya has also significant concentrations of both day, evening and late-night leisure areas. Especially Kızılay which is the city centre of Ankara and surrounding neighborhoods have many entertainment functions and shopping opportunities. In addition mixed land use type and income level make Bahçelievler neighborhood an attractive area for the offenders. These properties of the district influence the rate and location of incidents. Therefore these two police control zones are chosen as a case study area.

3.2. Data used in the Study

Data used in the study are complied from governmental organizations and private companies. The thesis concentrates on five incident types: murder, burglary, auto, pickpocket and usurp which are represented by point feature (Map 3.3). Spatial and temporal information regarding to these incidents were obtained from Ankara Police Directorate. The data of the incidents, occurred in 2003, include the information of the number, address, occurrence time, location and type. These data were recorded by the two police stations – Centre of Çankaya Police Station and Bahçelievler Police Station - and forwarded to Ankara Police Directorate.

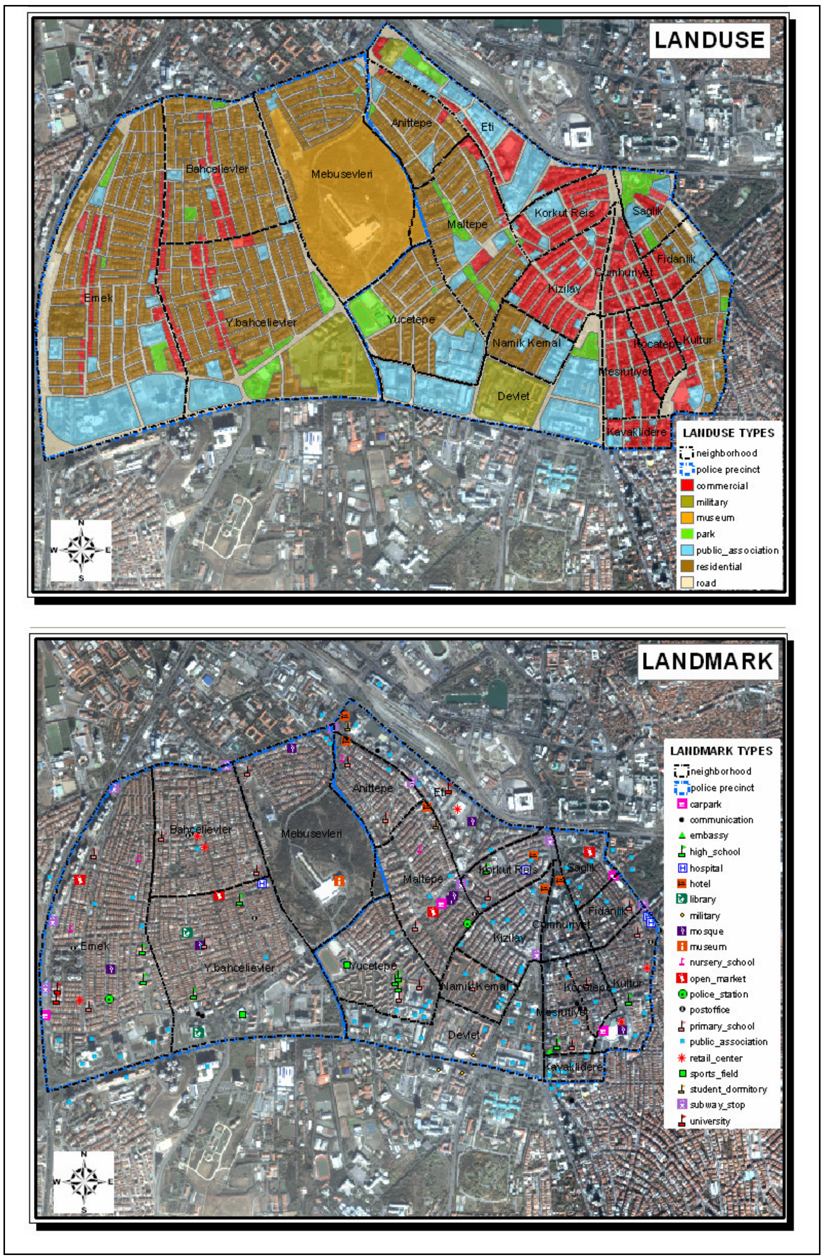


Map 3.3: Distribution of Incidents occurred in 2003 in the Study Area

However some problems related to the reliability of the data are realized. Polices are often referred to as the gatekeepers of the criminal justice system. This means that they control, to a large extent, the numbers and types of problems handled by latter stages of the system. In Ankara, technical staff in each police station enters the information of incidents and also the offenders and victims into a database which is connected to the centre office - Ankara Police Directorate - by a network system. However untrustworthiness or indifference of the data may occur.

Besides the decision of police, many victims and witnesses decide not to call the police in the aftermath of offenses. The reasons for this non-reporting are many: people believe that the criminal justice system will not / cannot help, victims fear act of vengeance, offender is unsuccessful, insurance would not cover, or other solutions to the problem are found. Among the many problems inherent in the surveys is the lack of knowledge of what constitutes various incidents (Lab, 2000). Thus the information which is obtained from the police and entered arbitrary into system may not reflect the real incident status.

The other data that include the land use types were obtained by digitizing the satellite imagery that was acquired in 2003 (Map 3.4). Digitizing 1m resolution IKONOS satellite imagery, land use data set is obtained which contains the information of each parcel such as residential, commercial, park, public association and military; landmark types such as school, police stations, subway stops, hospitals, carparks; and roads. The resources of the obtained information are summarized in the below table (Table 3.1).

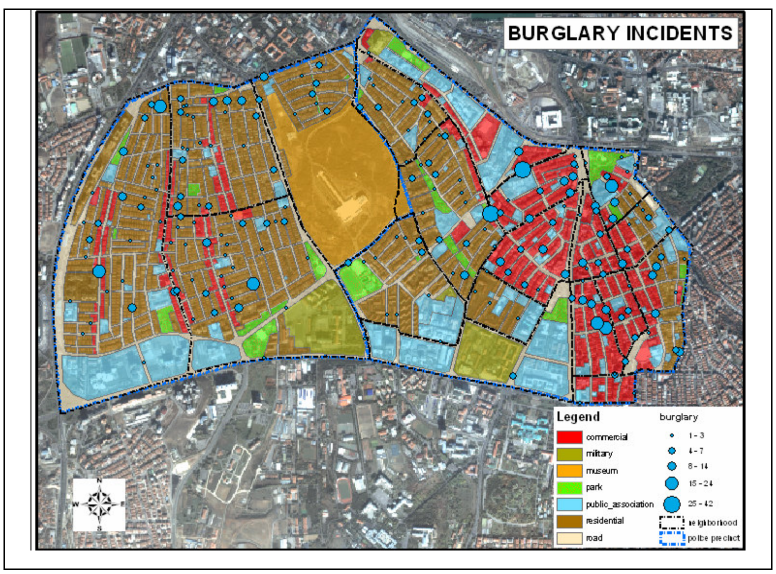


Map 3.4 : Land Use and Landmark data in Study Area

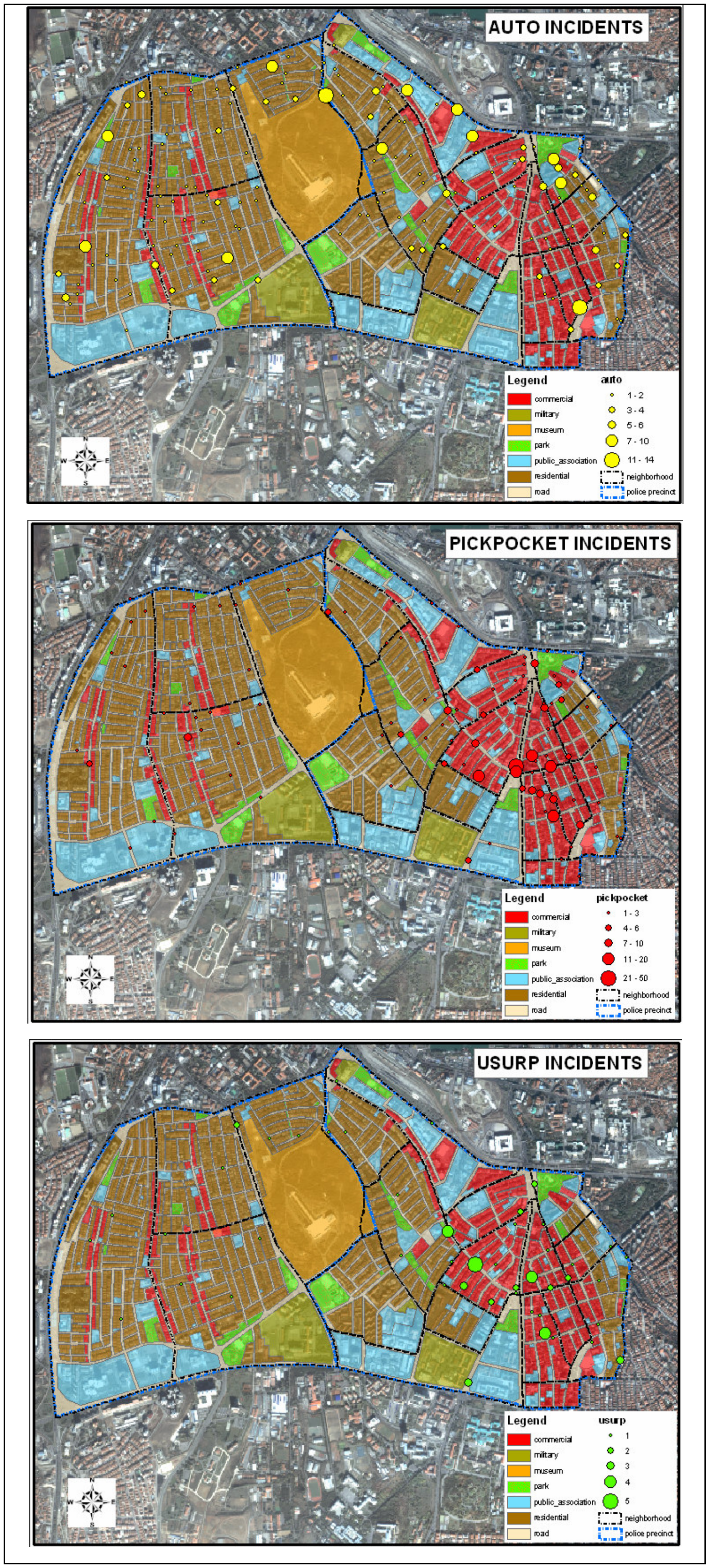
Table 3.1. Data used in the Study Area

DATA	INFORMATION	RESOURCE
Incident Data	Incident type	- Ankara Police Directorate
	Address of the incident location	
	Responsible police station zone	
	Place of the incident	
	Date of the incident	
	Time of the incident	
	X,Y Coordinates of the incident	
Land use Data	Land use type	- INTA Space Turk, Ankara - Site Survey
	Size of the land use area	
	Landmark type	
	Road class	

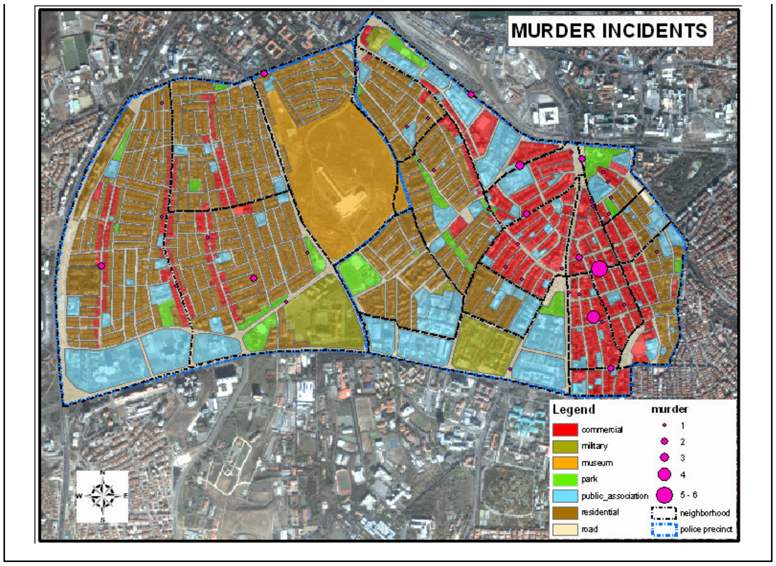
The easy way to visualize a spatial point pattern is to plot the data in the form of a dot map (Bailey and Gatrell, 1995) (Map 3.5). This will give an initial idea of the shape of the study area and the distribution of the events. However when the data set is large and there are multiple occurrences of events at the same or nearly the same location it will be hard to come to definite conclusions.



Map 3.5: Point Pattern of the Incidents in the Study Area



Map 3.5: Point Pattern of the Incidents in the Study Area (cont'd)



Map 3.5: Point Pattern of the Incidents in the Study Area (cont'd)

As seen in the below figures (Figure 3.3) neighborhoods of the study area include different types and size of land use areas. Commercial areas are generally located in Kızılay, because of it is the trade center of Ankara. Kızılay and around neighborhoods include various commercial usages from entertainment functions like bars, cafes and restaurants to shopping and service places such as shops, retail centers, manufacture units and banks. On the other hand Bahcelievler and around neighborhoods are generally covered by residential parcels. However Bahcelievler have also important commercial functions in which mostly cafes and shops could be found. Parks are generally found where the residential areas are located. This variety in land use is the reason to choose these precincts as a study area.

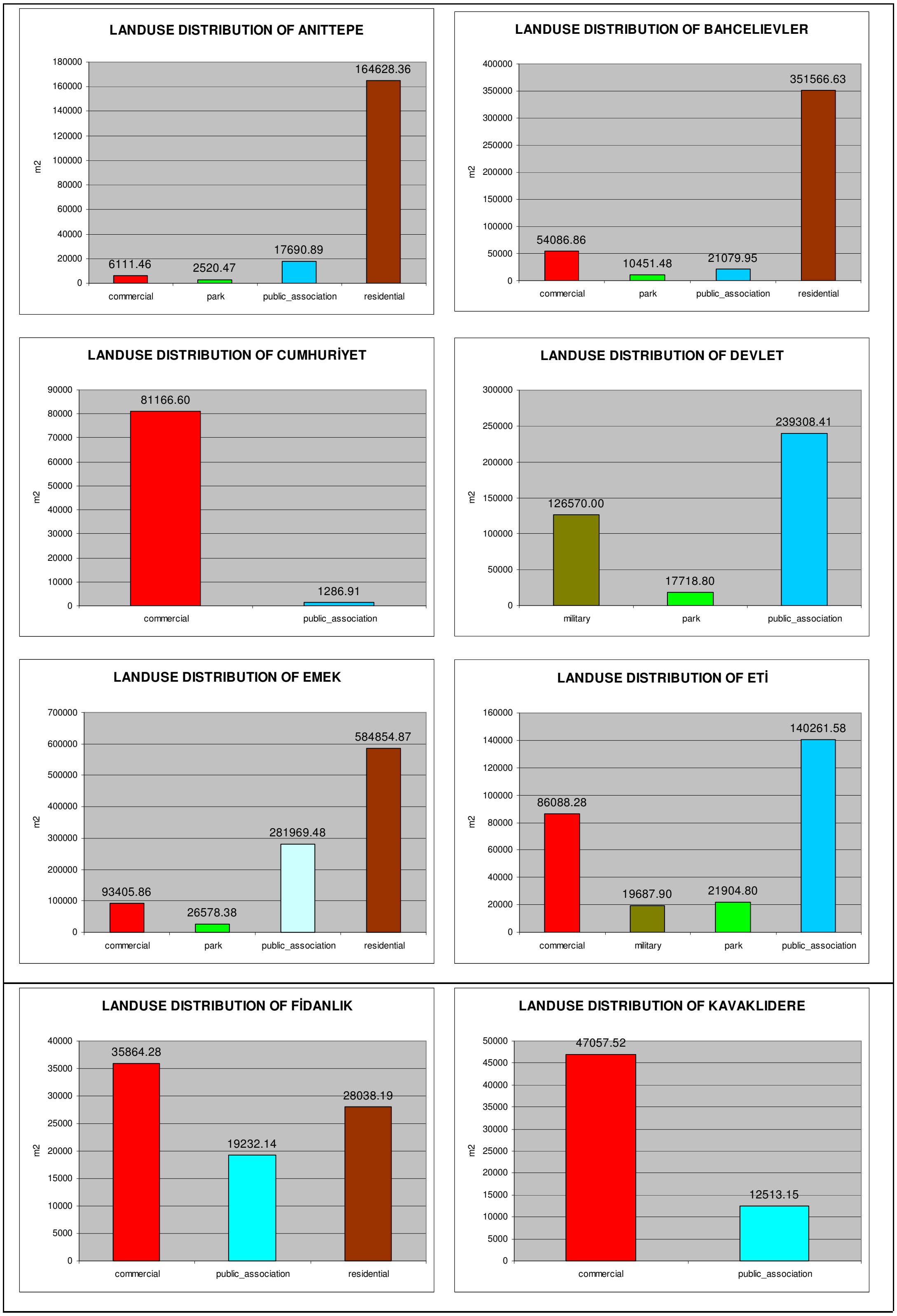


Figure 3.3: Land Use Distribution of the Neighborhoods

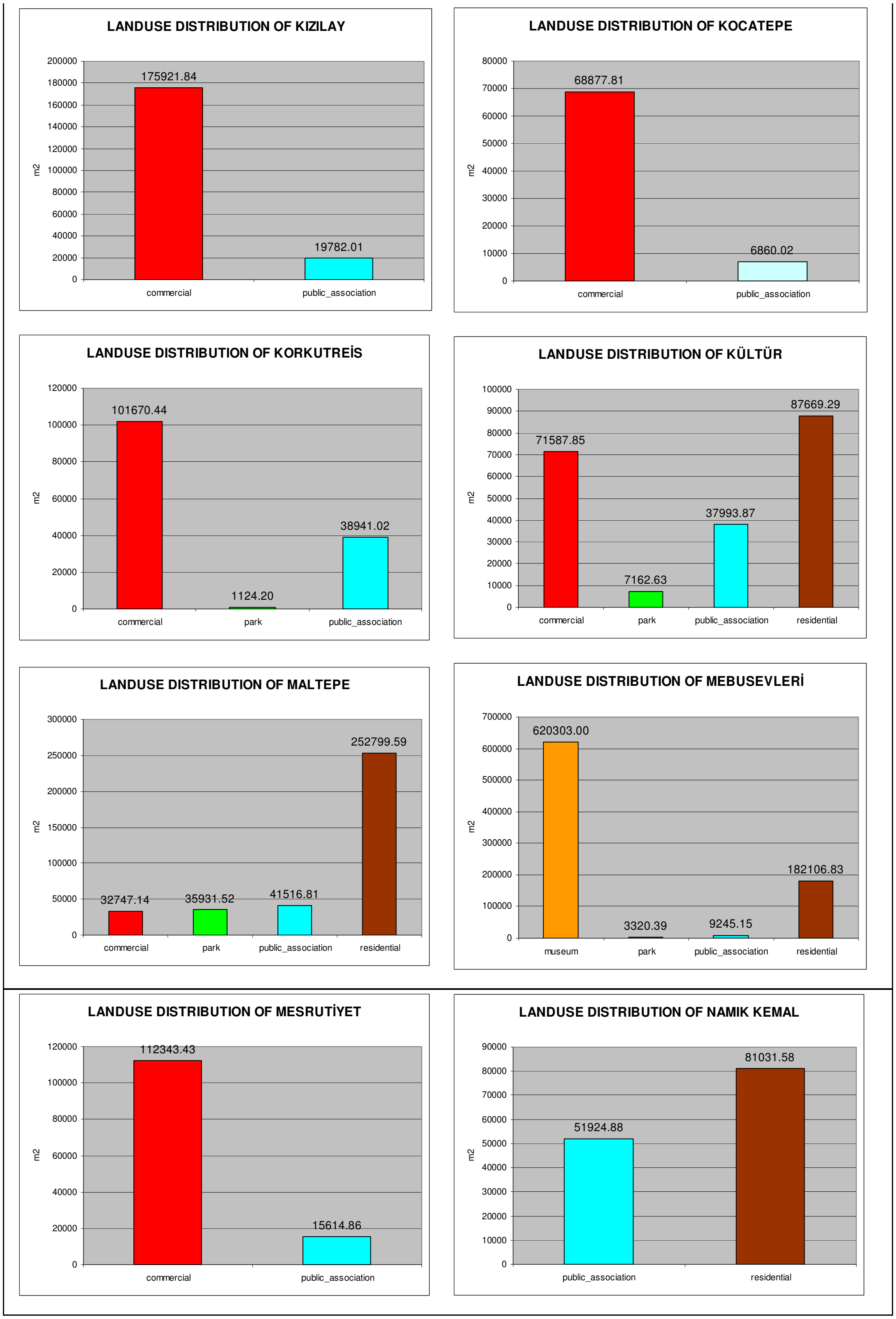


Figure 3.3: Land Use Distribution of the Neighborhoods (cont'd)

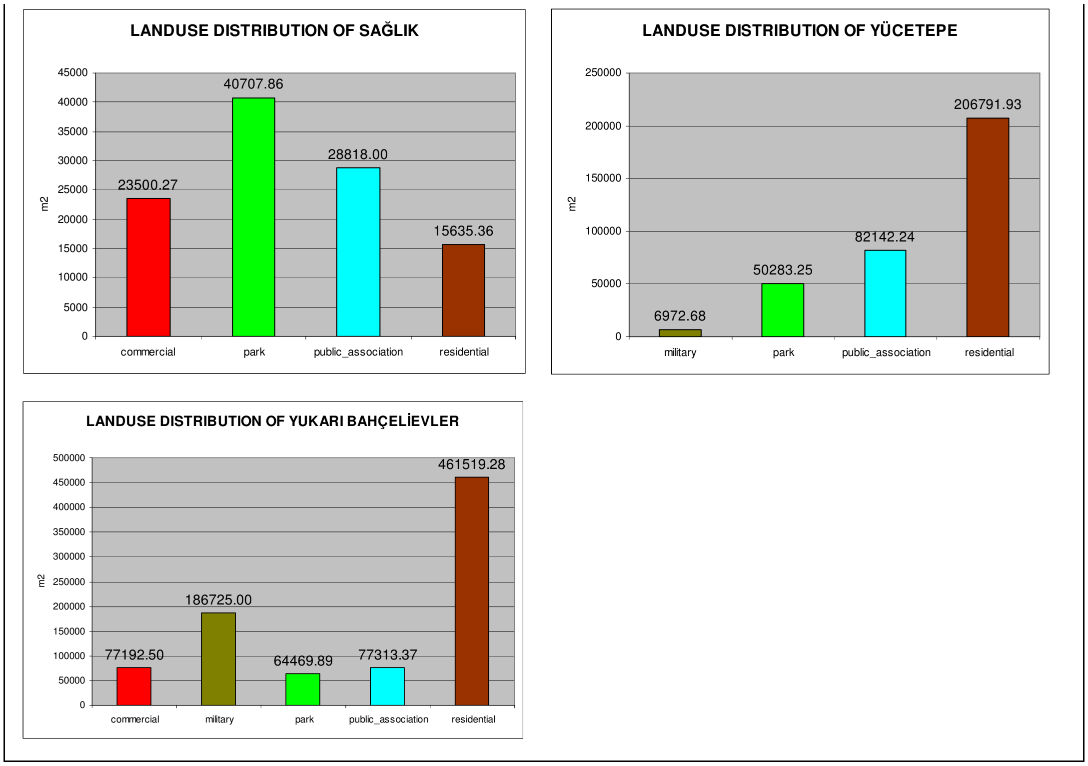


Figure 3.3: Land Use Distribution of the Neighborhoods (cont'd)

3.3. Descriptive Analysis for Incident Rates

In the study area, there are 1910 incidents were recorded in 2003. Mostly burglary incidents which include the actions of theft from home, theft from office, theft from hotel room, bicycle theft and theft from warehouse occurred with a high rate when compared to the other incidents (Figure 3.4). Then pickpocket incidents with the events of swindling, purse snatching, pocket picking and breach of trust follows auto incidents which contain auto theft and theft from auto. The incidents of usurp that include the acts of beating, molestation, assault, grab, resist the police, shooting a business place, arson, bribe, threat, kidnap, using counterfeit identity and the murder incidents with the events of suicide, killing and wound is coming after usurp have lower rates. One of the reasons that mostly incidents against property happen may be the major punishments to the offenders who realize an incident against human (interview with Demirbilek, 2004).

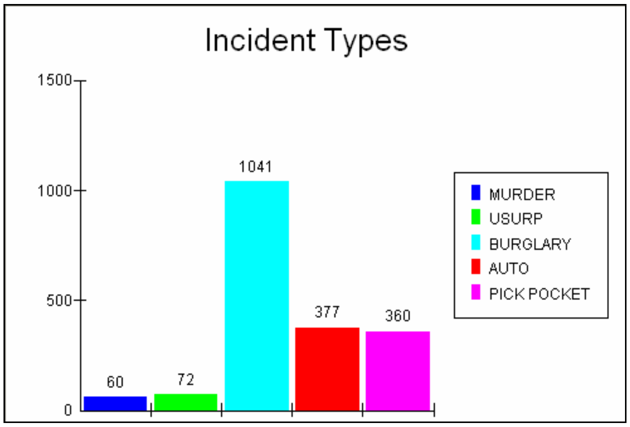


Figure 3.4: Distribution of Incident Types

On the other hand the classification of each incident which is done by police may not be true since some irrelevant classes are noticed. For instance, at the interview with Atasoy (2004), she says that suicide events could not be examined under murder incidents or resist to police could not be the class of usurp. For example, Vellani and Nahoun (2001) mention that in Uniform Crime Report of the Federal Bureau of Investigation in the United States, the crimes are grouped into two classes, of which the first one is violent crimes including murder, rape, robbery and aggravated assault incidents, the second one is property crimes including burglary, larceny theft, motor vehicle theft and arson incidents.

Furthermore, Atasoy also emphasizes that there may be some duplicated events. For example a criminal may swindle by using counterfeit identity. However this event may be recorded as both a usurp event and a pickpocket event. Thus such problems may affect the result of the analyses.

These incidents have distinctively different targets and different patterns. For instance; burglary, auto and pickpocket incidents are offense against property; murder is human; and usurp is mixed offense against both property and human. Similarly land use types demonstrate differences based on their intentions. Therefore the places and naturally the neighborhoods in which the incidents take place changes (Figure 3.5).

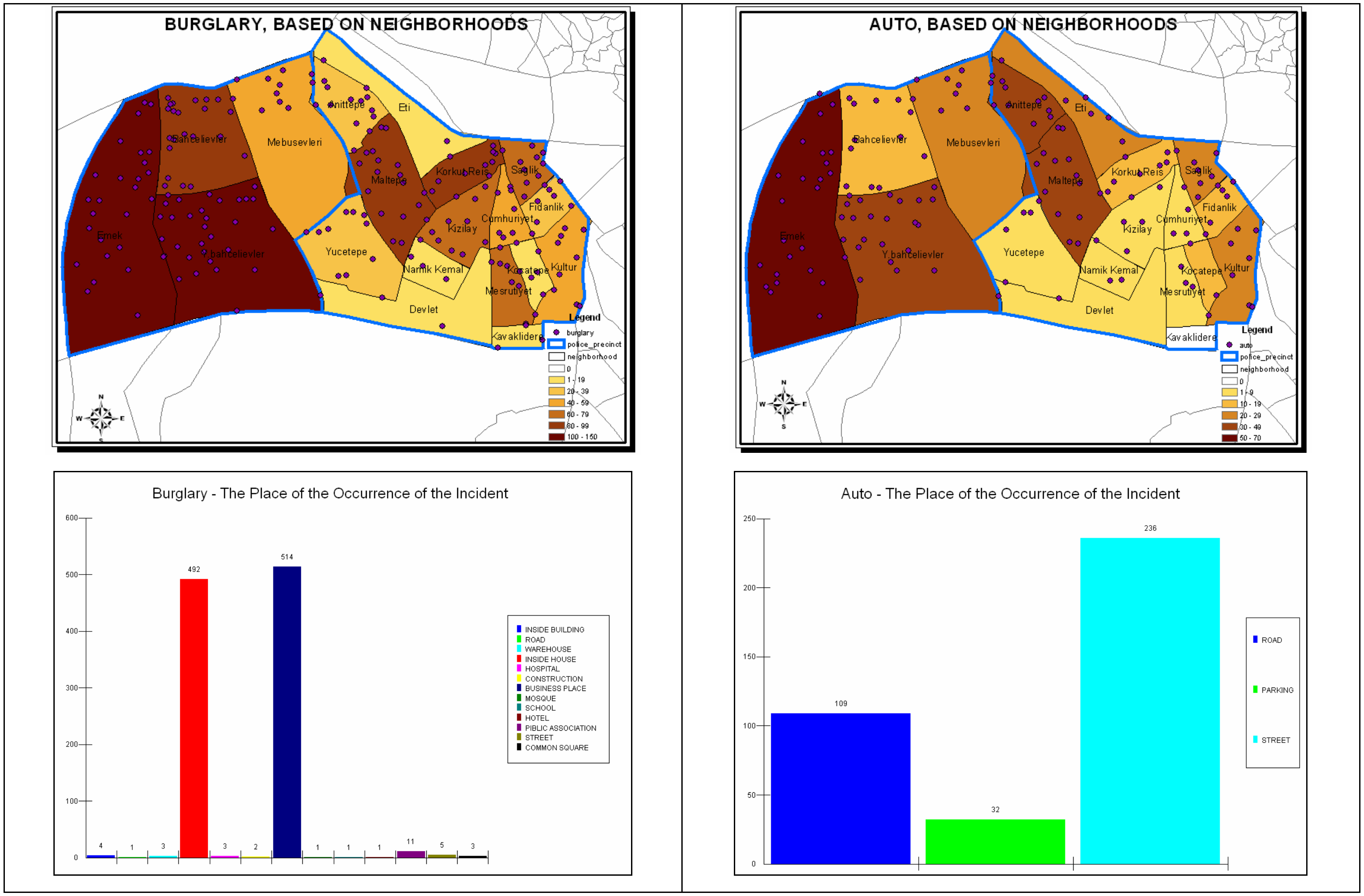


Figure3.5: The Neighborhoods and the Places of the Incidents

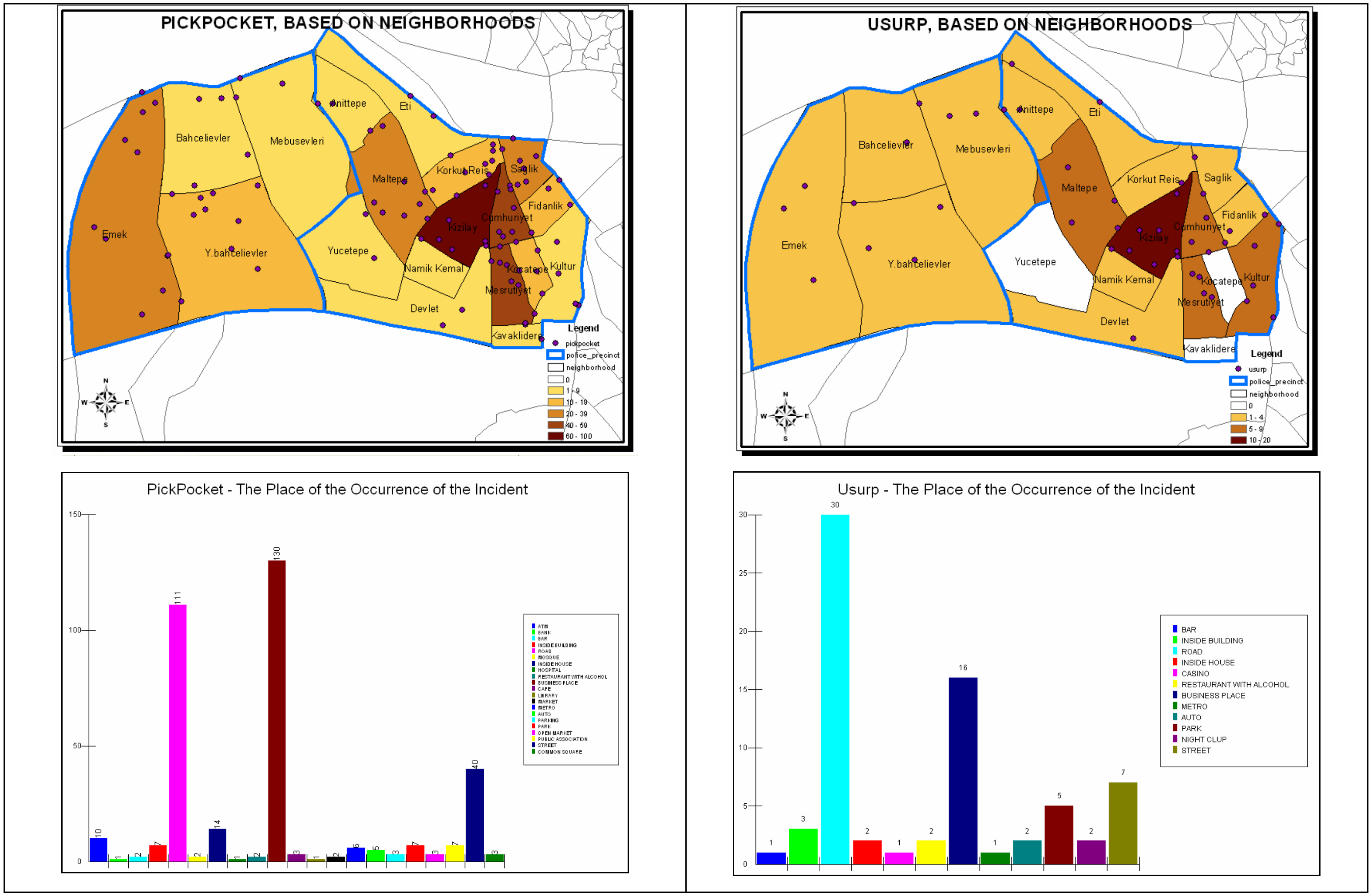


Figure3.5: The Neighborhoods and the Places of the Incidents (cont'd)

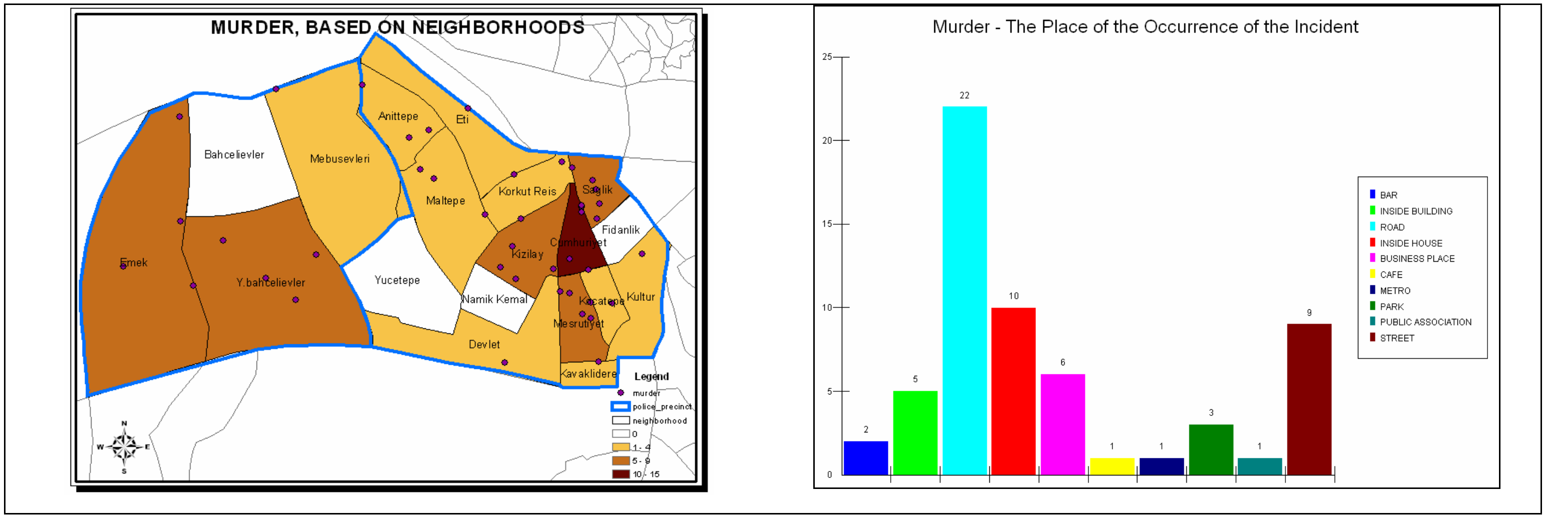


Figure3.5: The Neighborhoods and the Places of the Incidents (cont'd)

Burglary events are mostly occurred inside business buildings and houses. So burglary incidents take place in the neighborhoods like Emek, Yukari Bahçelievler, Bahçelievler, Maltepe and Korkutreis where commercial and residential usages are mostly seen together. As seen in the graphics theft from auto or auto thefts come into existence in roads and streets but near to the residential areas which are mostly seen in the neighborhoods of Emek, Yukarı Bahçelievler, Anıttepe and Maltepe.

It is the excepted result that pickpocket events are generally happened in the neighborhoods which have shopping areas. Pocket pickers who are mostly seen in the streets or roads prefer the places in which the movement of the pedestrians is dense. Like pick pocket incidents above figures show that usurp events are mostly seen in Kızılay neighborhood. As the least seen incident in the study area, murder events occur in roads, streets and inside houses and business buildings.

However when the effect of the size of the boundary on the number of incidents is considered, this way is not appropriate for analyzing the highest concentration areas. Thus, instead of determining the hot spot areas by counting the number of incidents inside the neighborhoods, the methods based on the analysis of spatial point pattern which may give more precise results should be taken.

Besides the spatial distribution of the incidents, the temporal pattern of the incidents is considered in terms of the time of the day, day of the week and month of the year as shown in figures below. This kind of information reveals that some specific days, times and months when the levels of incidents are at their greatest.

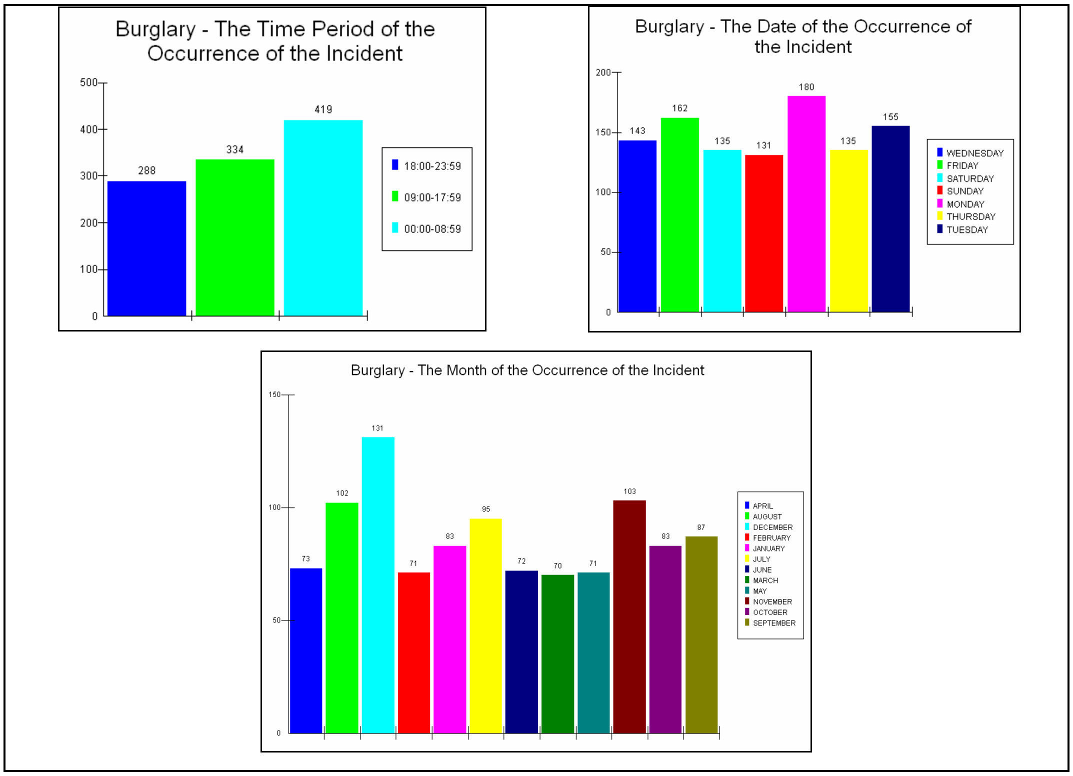


Figure 3.6: Temporal Distribution of the Burglary Incidents

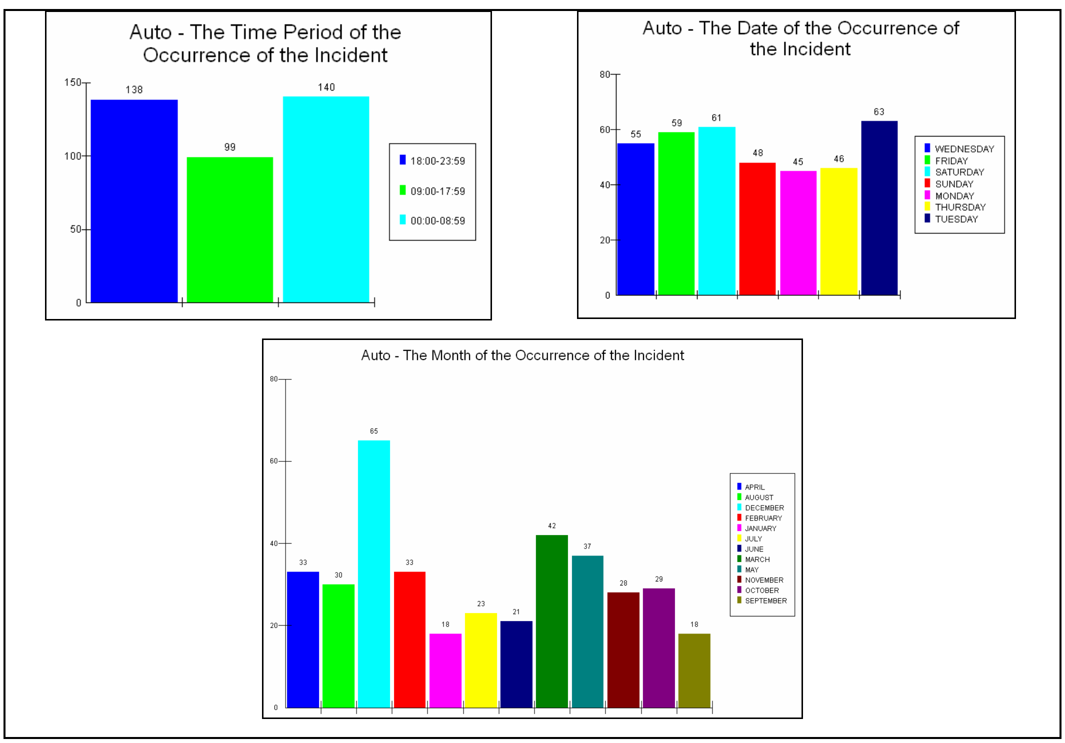


Figure 3.7: Temporal Distribution of the Auto Incidents

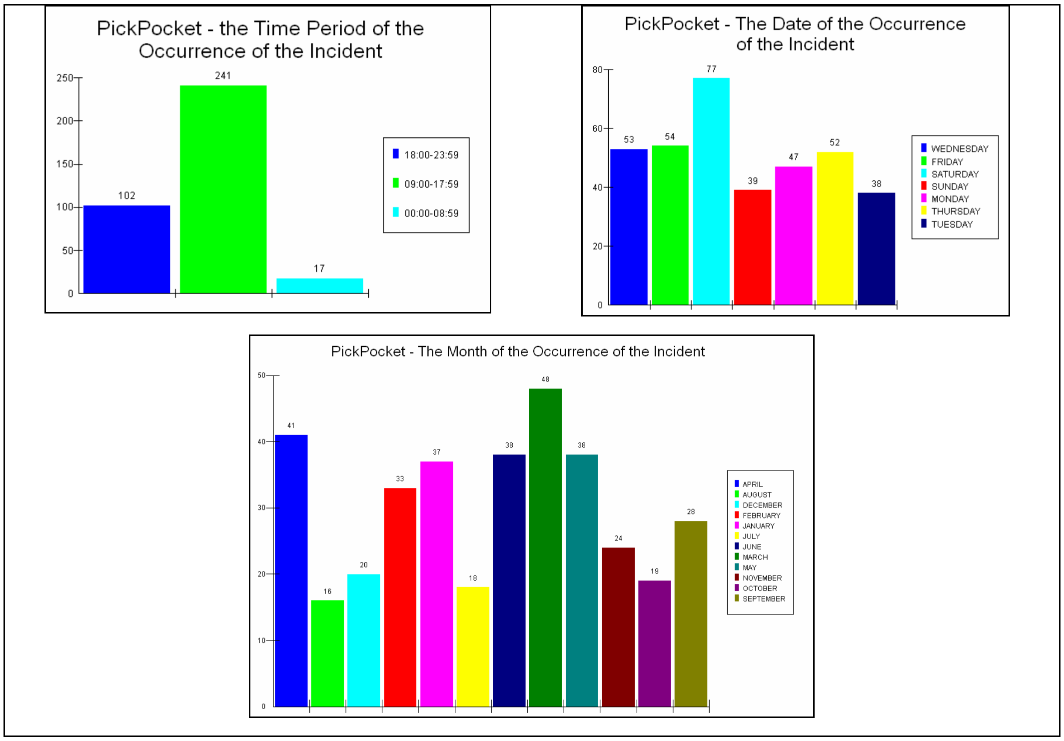


Figure 3.8: Temporal Distribution of the Pick Pocket Incidents

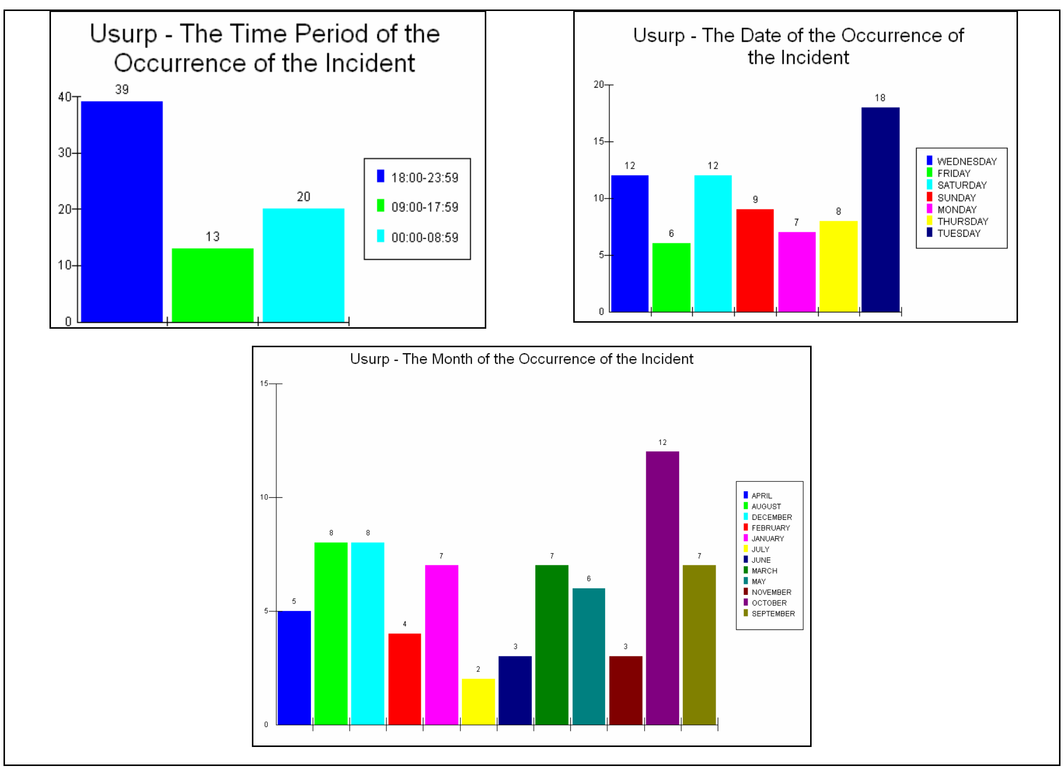


Figure 3.9: Temporal Distribution of the Usurp Incidents

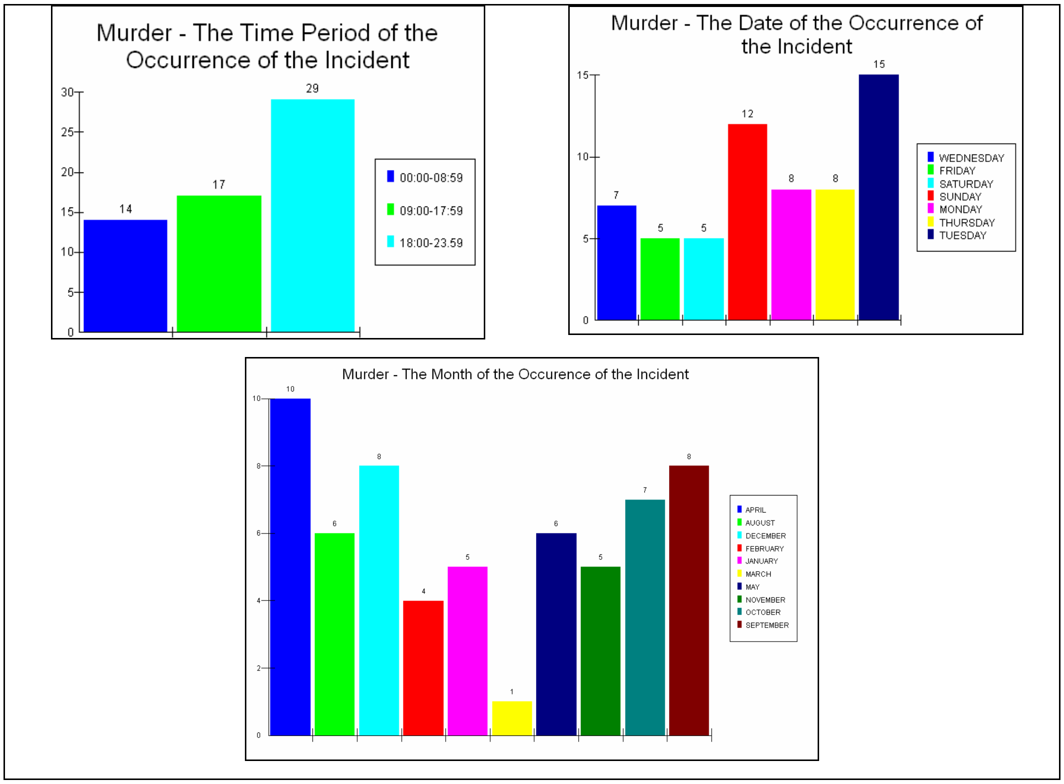


Figure 3.10: Temporal Distribution of the Murder Incidents

As seen in the figures the time of the day has a significant effect on the occurrence of the incidents. Burglary and auto offenders prefer midnight hours, since it is the time for victims to sleep. Pocket pickers have a preference on daytime, since people are outside for work, school or shop at these times. So the dense movement of the victims is higher outside on day times. Murder and usurp events for which desolation and darkness are the required factors, are naturally detected on the evenings.

Different from time of the day, neither month nor day have ocular effect on the rate of the incidents in the study area. Whereas, it is expected that burglary events happen on the summer months instead of winter, since people are generally on vacation and far away from their houses or offices or they hold open their balconies or windows because of the hot weather. In addition except pick pocket events which arise on Saturdays, since it is the day for people to go outside and make shopping or walk around, other incidents do not occur on a day which has a logical explanation. These results may be clarified with social reasons by sociologists or criminologists.

CHAPTER IV

GENERATION OF CRIME INCIDENT ANALYSIS IN THE STUDY AREA

In this chapter, it is considered to investigate methods for analysis of a set of point locations, which is often referred as "point pattern". The methods of analysis where various summary statistics or plots are derived from the observed distribution of events are taken into consideration. The results are used to investigate hypothesis of interest or to suggest possible models. Some of these methods are more concerned with investigating first order effects in the process; others address the possibility of spatial dependence or second order effects. In general main objective is to ascertain if there is a tendency for points to exhibit a systematic pattern, for example regularity or clustering and to investigate how intensity of points varies across the study area.

The spatial point pattern analyses are generated under two frameworks. One of them includes centrographic techniques which describe the central tendency or dispersion of the location of the points. The other one includes pattern analysis which helps to identify the distribution of these points and contains such methods: quadrat analysis, kernel estimation, nearest neighbour distances, K-function, Nearest Neighbor Hierarchical Clustering (Nnh), Spatial and Temporal Analysis of Crime (STAC) and K-Means Clustering. ArcGIS 8.2, Crime Stat II and Golden Software Surfer 8 are employed to carry out these analyses. In some stages of these methods Microsoft Excel is used.

4.1. Application of Centrographic Techniques

Basic descriptive statistics which are performed in the thesis are mean, median, center of minimum distance, standard deviation of X and Y coordinates standard distance deviation and standard deviational ellipse. The mean center summarizes the central tendency of a spatial distribution of points or of aggregate frequencies. It lies at the intersection of the means of the X and Y coordinates.

The spatial equivalent of the median can also be calculated to summarize the central tendency of a point distribution. It divides a single set of numbers in half and is the value of the middle observation or the value of mid-way between the observations either side of the middle, whether or not this actually occurs within the set.

The center of minimum distance is the unique statistic in that it defines the point at which the sum of the distance to all other points is the smallest. The location of the center of minimum distance for the incidents indicates that there are slightly more incidents.

Besides the mean center, median center and center of minimum distance, CrimeStat will calculate various measures of spatial distribution. One of them is standard deviation of the X and Y coordinates which describe the degree of dispersion, orientation and shape of the distribution of a variable. A measure of standard distance deviation (Sd) is the spatial equivalent of standard deviation (Walford, 1994). Standard distance deviation measures dispersion around the mean centre, with larger values signifying that the points are more spread out. There is another statistic which gives dispersion in two dimensions called the standard deviational ellipse (Levine, 2002). These statistics are generated for the five incident types in the study area with the help of Crime Stat II and the results are presented in ArcGIS 8.2 (Figure 4.1).

Auto and burglary incidents have similar centrality measures than usurp, murder and pick-pocket incidents of which the centers of gravity are more alike. The location of the center of minimum distance for each incidents show that burglary and auto events are again similar in their location of the minimum distance at which there are more occurred incidents. This information guides police who could station the patrol car at the center of minimum distance to allow it to respond quickly to calls for service. Auto events are more dispersed and rectangular in shape, than the events of burglary. On the other hand pick-pocket and usurp are more concentric, square-like and clustered in distribution than murder events.

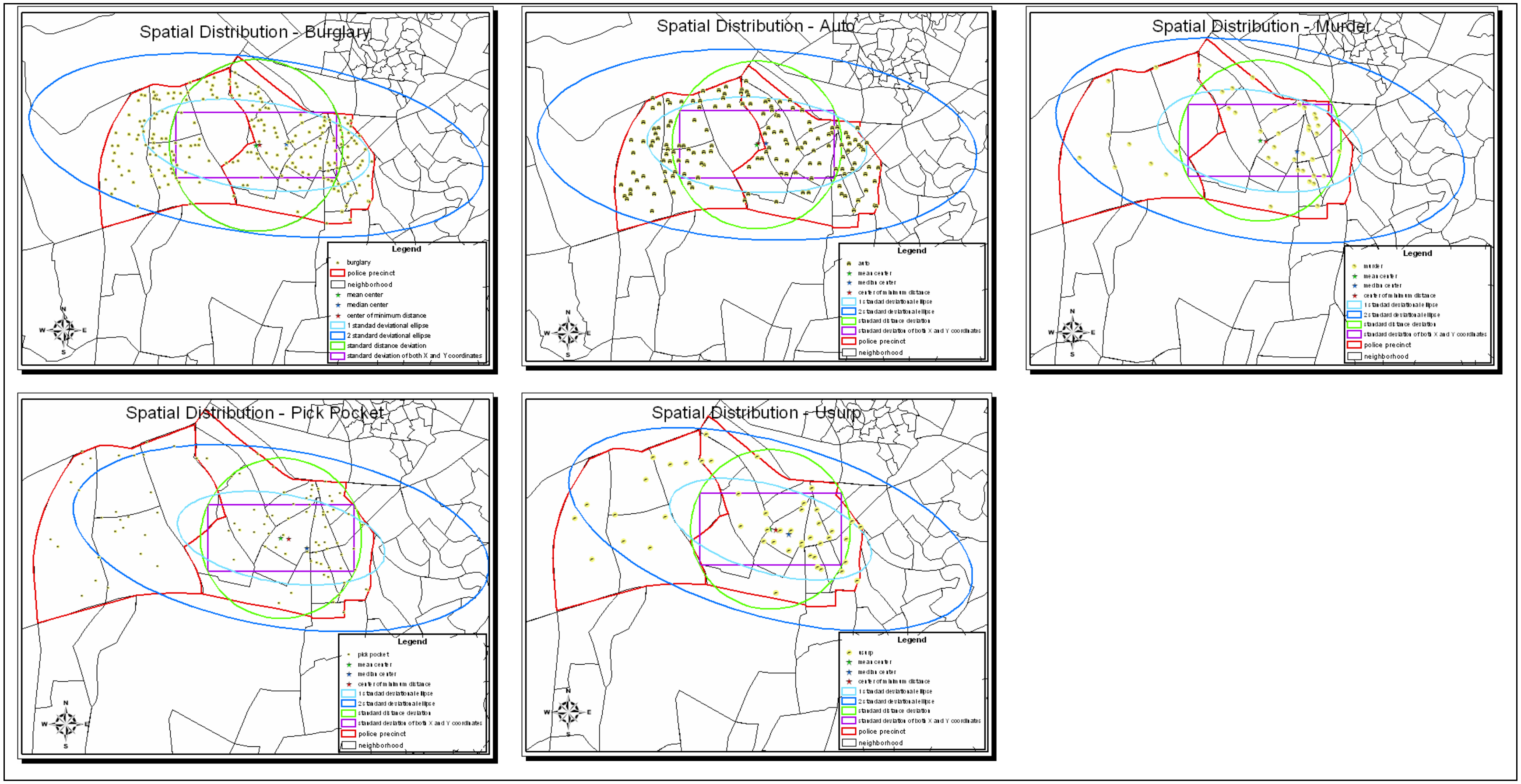


Figure 4.1: Spatial Summary Statistics for the Incident

4.2. Exploring Spatial Point Pattern Analysis

In order to analyze point pattern a number of techniques such as quadrat analysis, kernel estimation, nearest neighbor distances, K function, Nearest Neighbor Hierarchical Clustering (NNH), STAC and K-Means Clustering which are explained in the previous chapter, are generated for the study area for each incident types. For this purpose Crime Stat II, ArcGIS 8.2 and Ms Excel are used for the computing and creating related outputs and ArcGIS and Ms Excel are used for the visualization.

4.2.1. Application of Quadrat Analysis

The first technique for summarizing the pattern of the incidents is Quadrat analysis. In order to apply the Quadrat method Crime Stat II and Arc GIS 8.2 are used. To form the uniform grid and compute the frequencies of each point the extensions of Crime Stat are used, then the obtained grid and the points are joined to each other by ArcGIS. In this way the number of events in each of the quadrat is obtained to visualize the intensity over the study area. It is based primarily on the intensity of points and not arrangement in relation to one another.

To find the best grid size for each incident calculations given in Eq 4.1 are performed but also different grid sizes are tried for the analysis in order to find the finest visualization. The results are presented in Maps 4.1, 4.2, 4.3, 4.4 and 4.5.

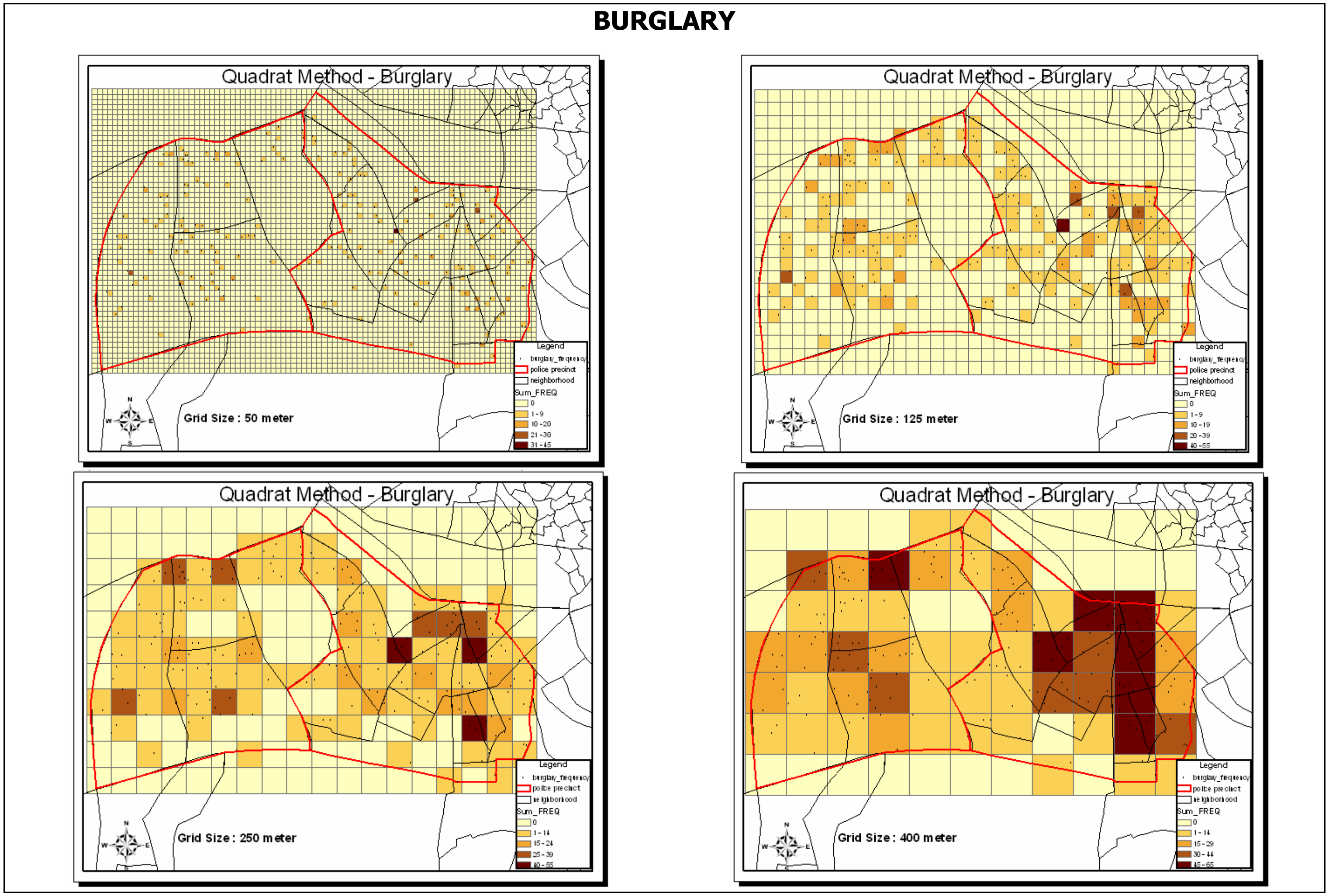
Burglary: $l = \sqrt{\frac{2*8}{1041}} = 0.124km \cong 125m$

Auto: $l = \sqrt{\frac{2*8}{377}} = 0.2060km \cong 205m$

Pick Pocket: $l = \sqrt{\frac{2*8}{360}} = 0.2108km \cong 210m$

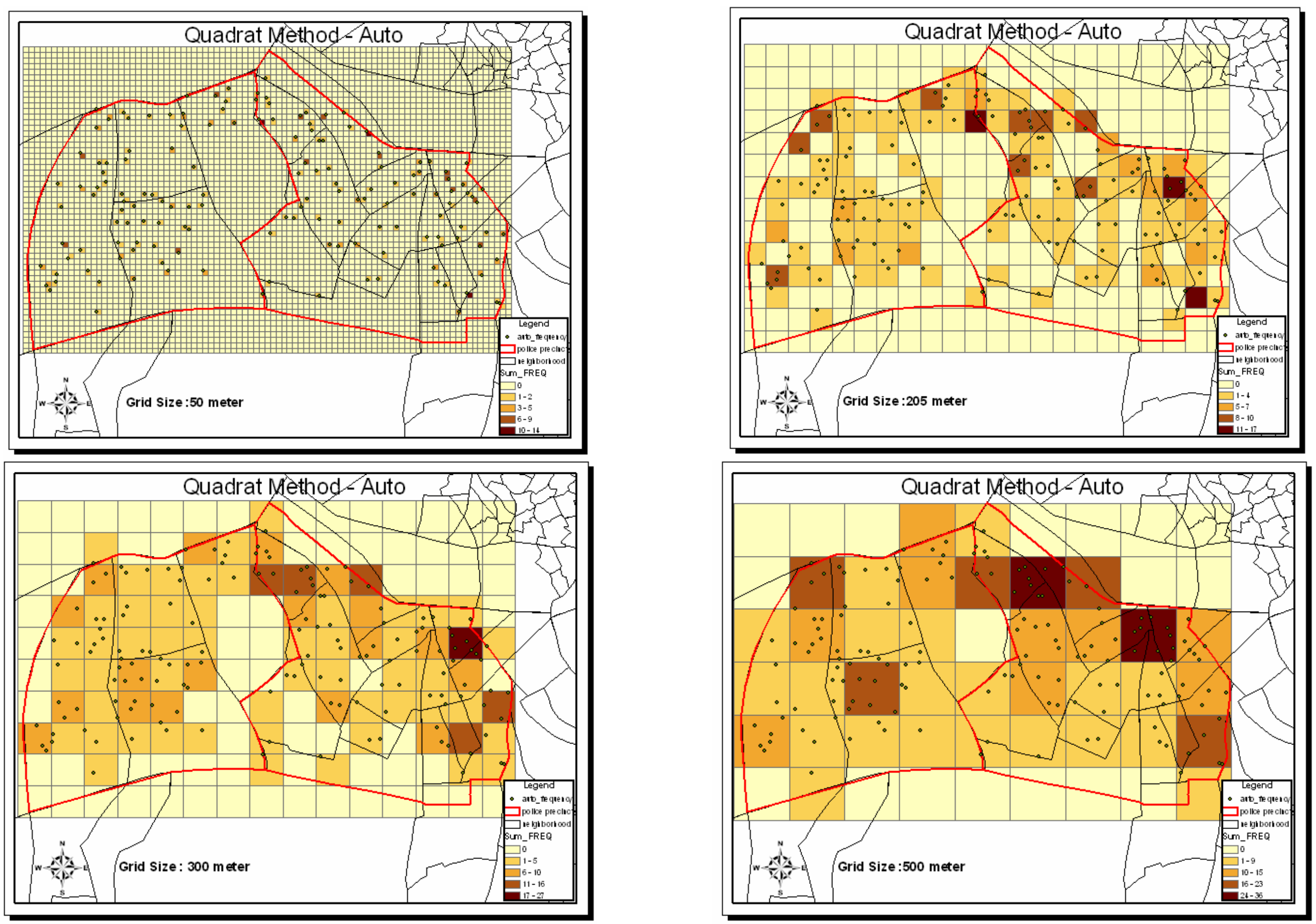
Usurp: $l = \sqrt{\frac{2*8}{72}} = 0.2222km \cong 220m$

Murder: $l = \sqrt{\frac{2*8}{60}} = 0.5164km \cong 515m$ Eq. 4.1



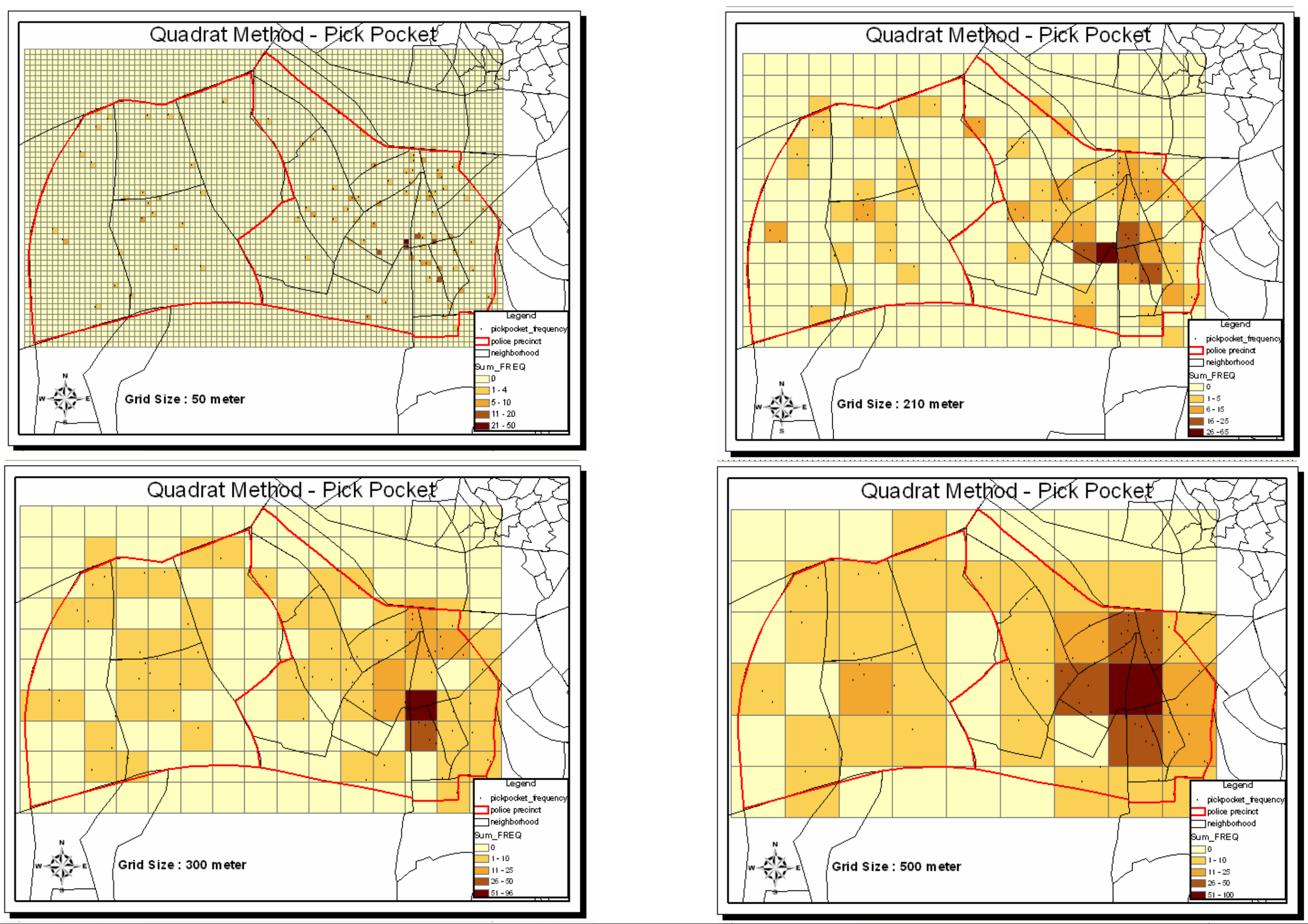
Map 4.1: Quadrat Analysis for Burglary Incidents

AUTO

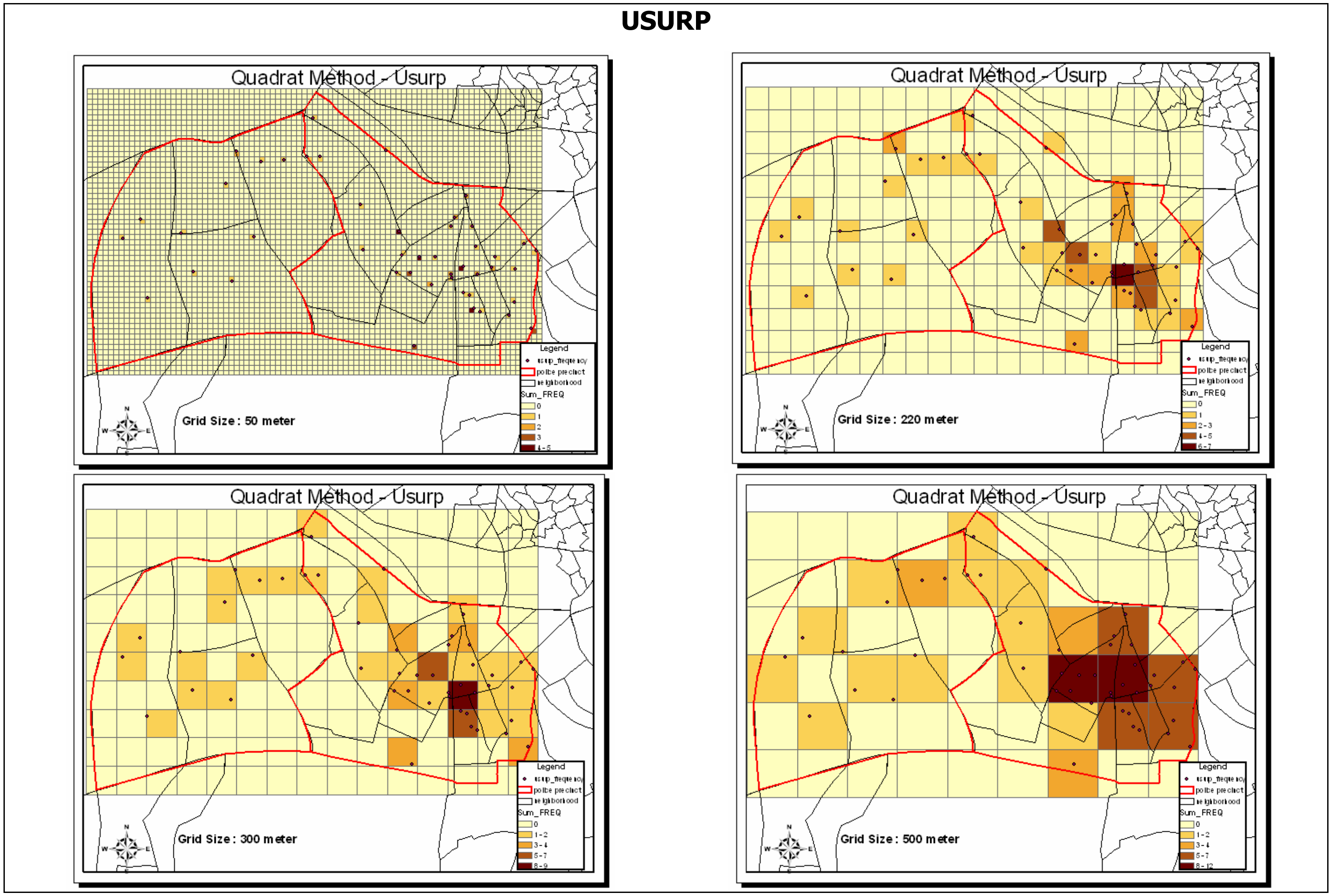


Map 4.2: Quadrat Analysis for Auto Incidents

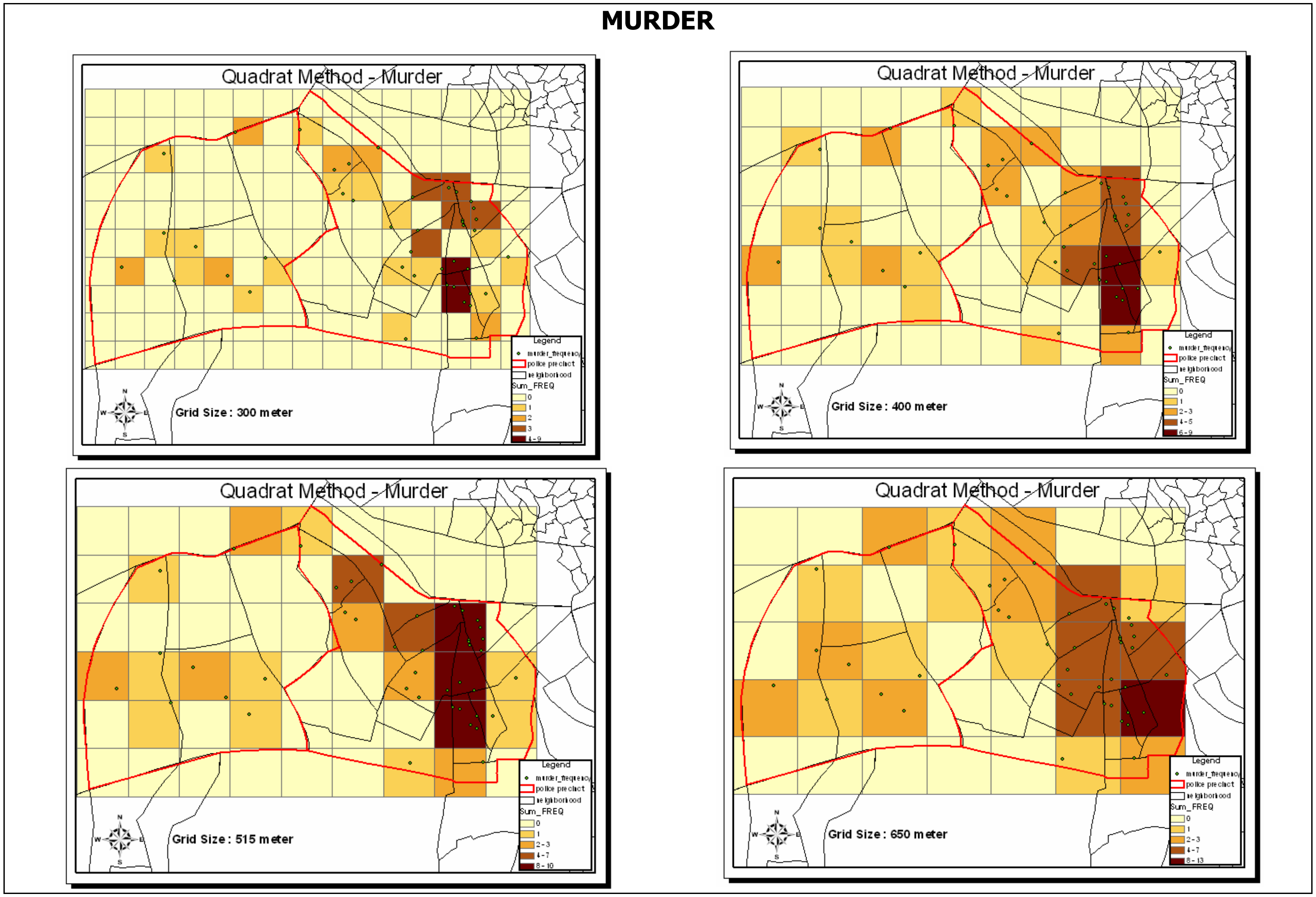
PICK POCKET



Map 4.3: Quadrat Analysis for Pick Pocket Incidents



Map 4.4: Quadrat Analysis for Usurp Incidents

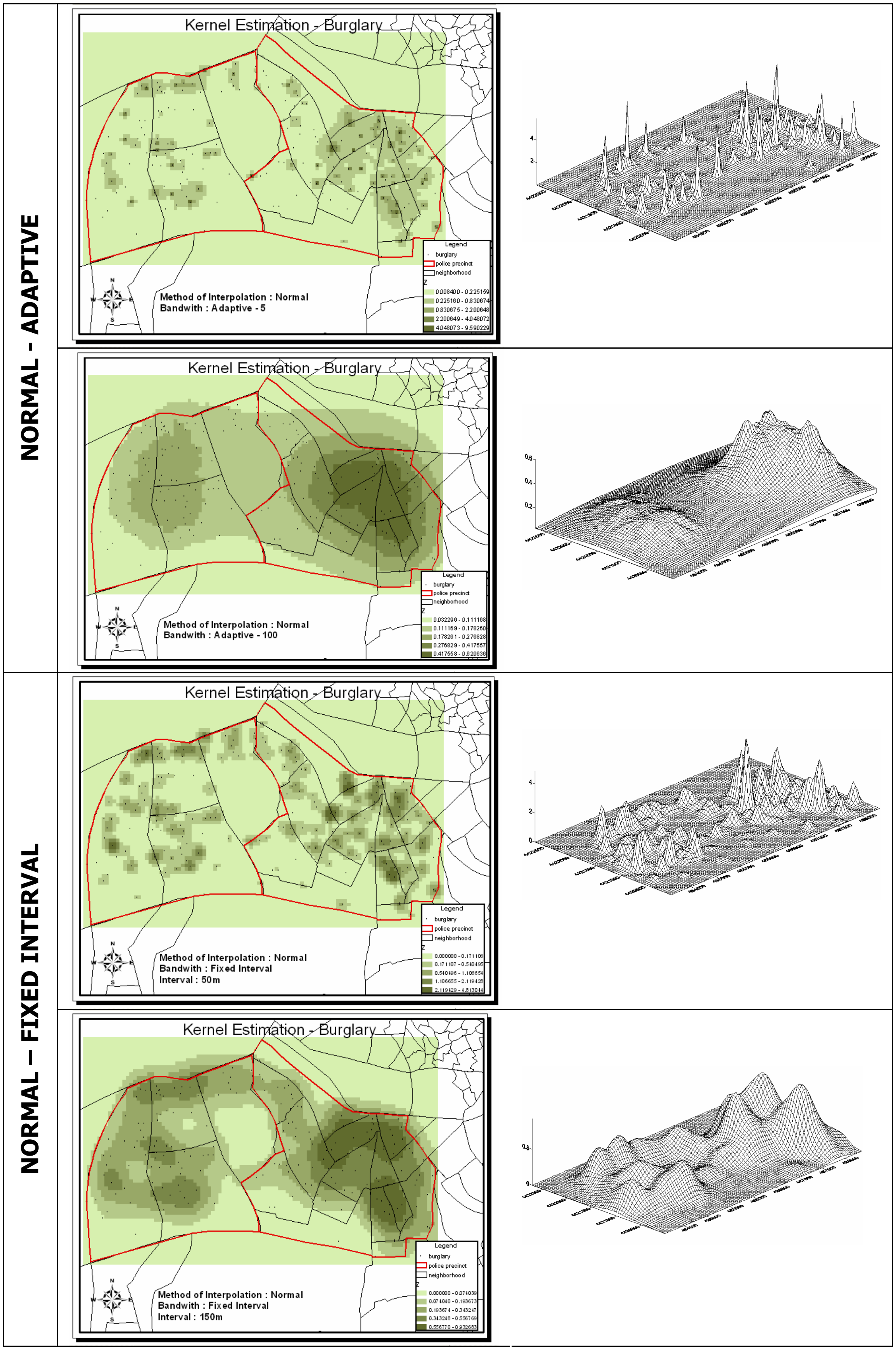


Map 4.5: Quadrat Analysis for Murder Incidents

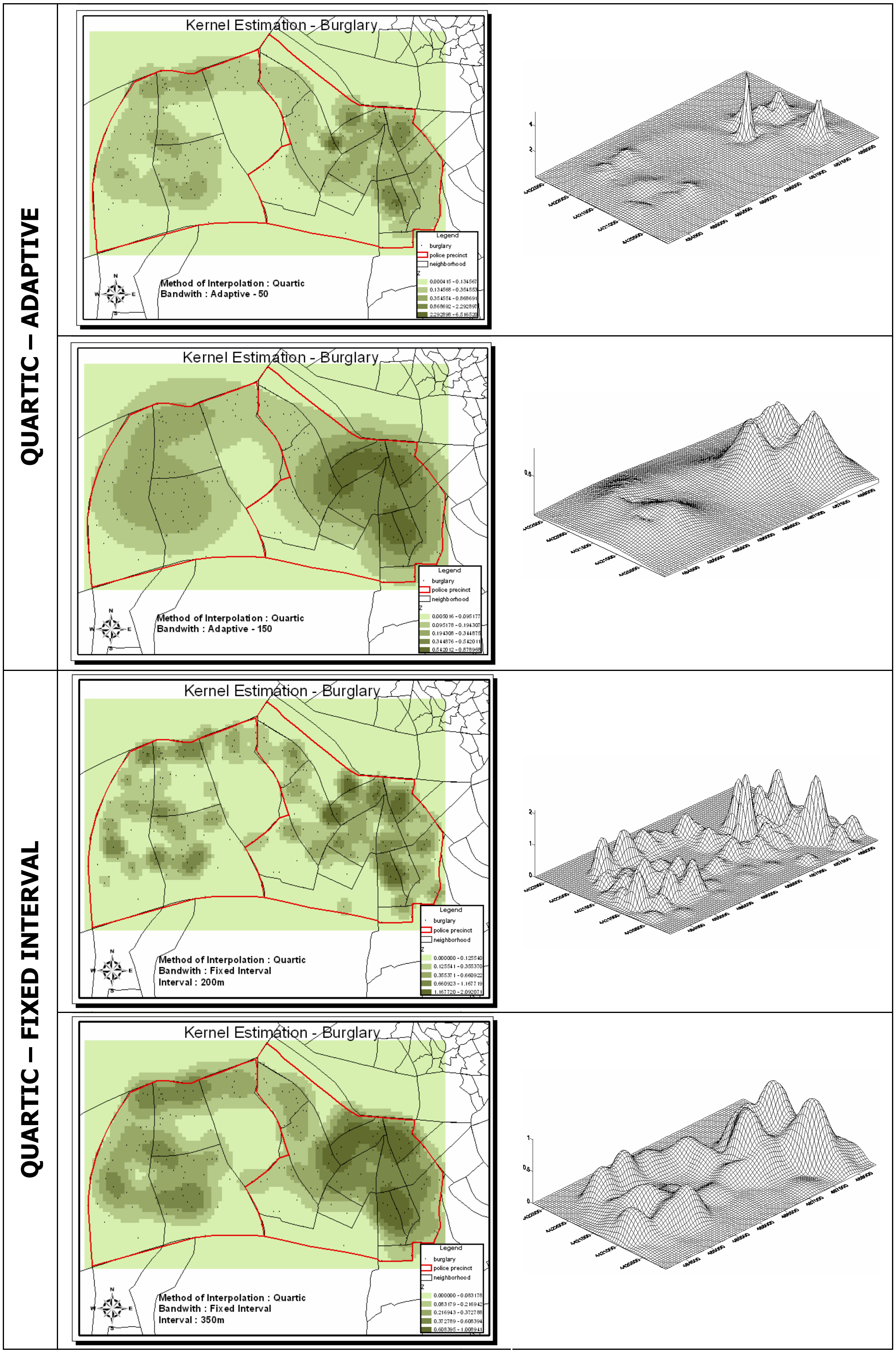
Although the quadrat method is a simple way for the visualization, the results are too sharpened and much of the spatial details are thrown away as noticed in the above figures. When the quadrats are too small, they may contain only a couple of points. But if they are too large, they may contain too many points and the distribution of the points becomes more clustered. So the results are completely based on the size of the grids and do not indicate the real pattern.

4.2.2. Application of Kernel Estimation

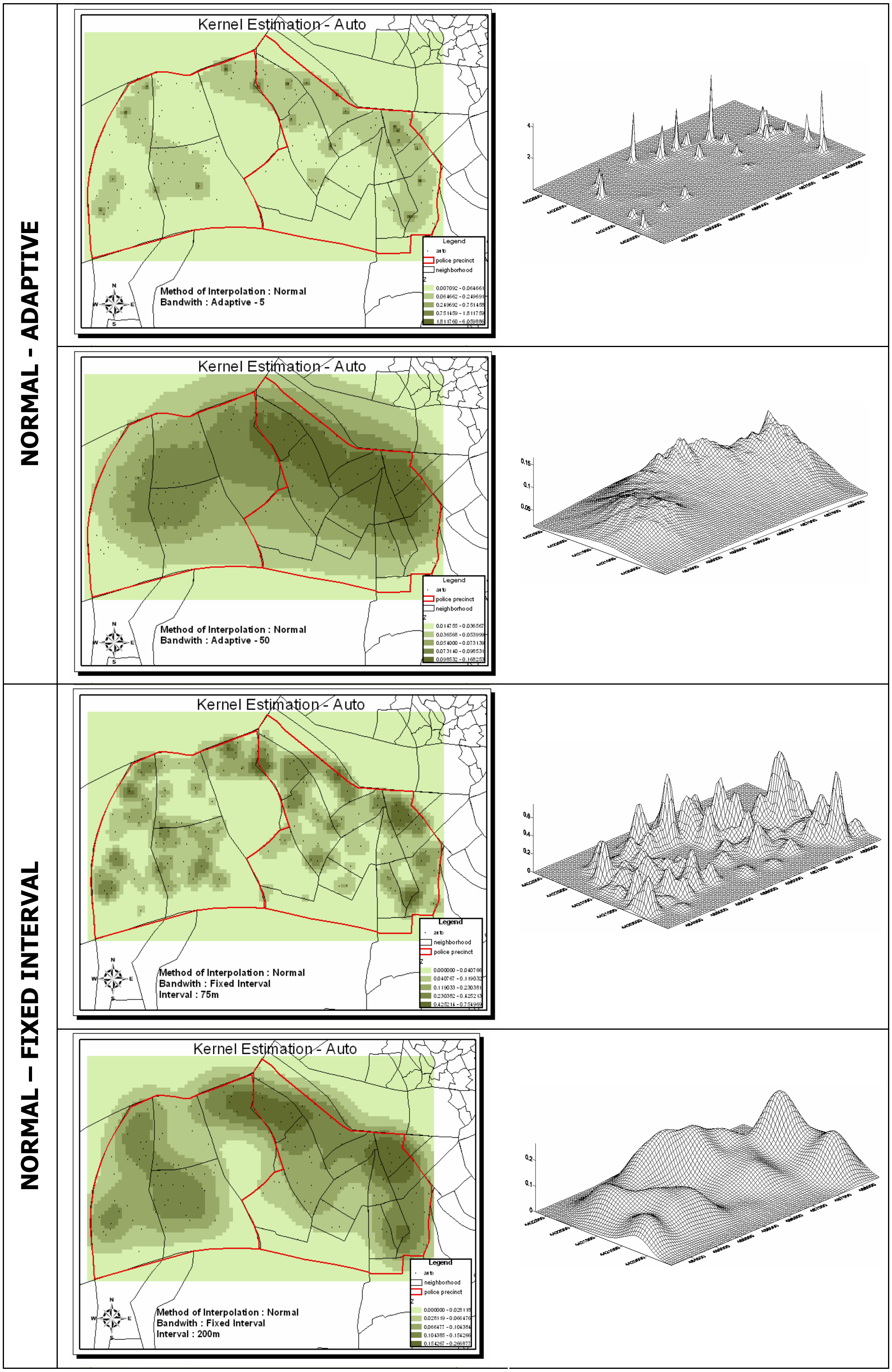
The other method for the summary of the pattern of the incidents is Kernel Estimation which has a number of different kernel functions such as normal, uniform, quartic, triangular and negative exponential. But Normal and Quartic functions are chosen for the study area. Adaptive and fixed interval kernels are tested for each function with different bandwidths. The results, shown in below maps, are obtained from CrimeStat and in order to visualize the estimated surface ArcGIS and Golden Software Surfer 8 are used.



Map 4.6: Kernel Estimation for Burglary Incidents

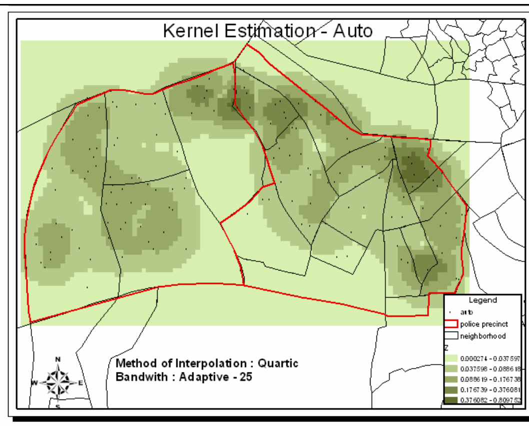


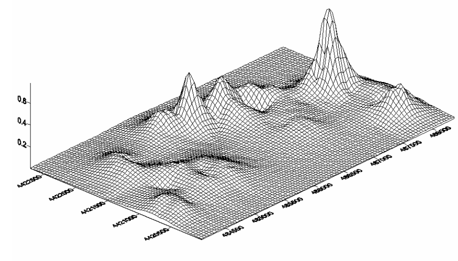
Map 4.6: Kernel Estimation for Burglary Incidents (cont'd)

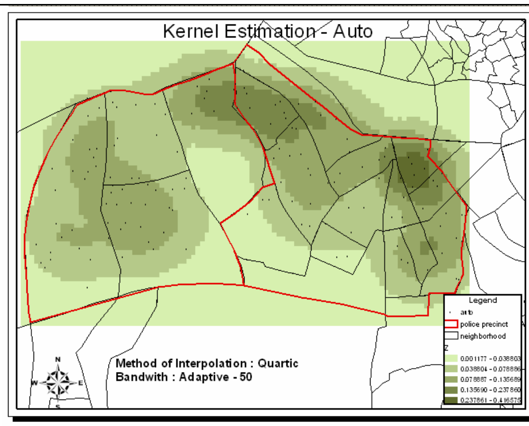


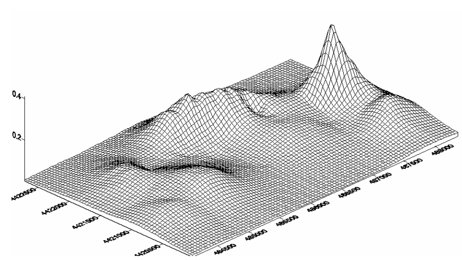
Map 4.7: Kernel Estimation for Auto Incidents

QUARTIC – ADAPTIVE

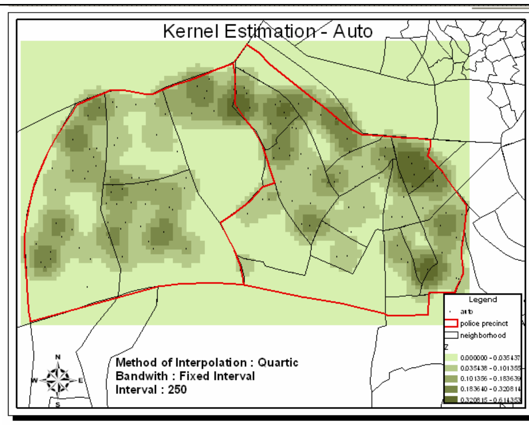


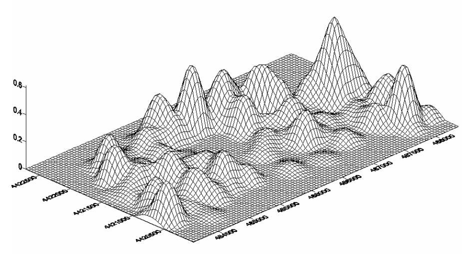


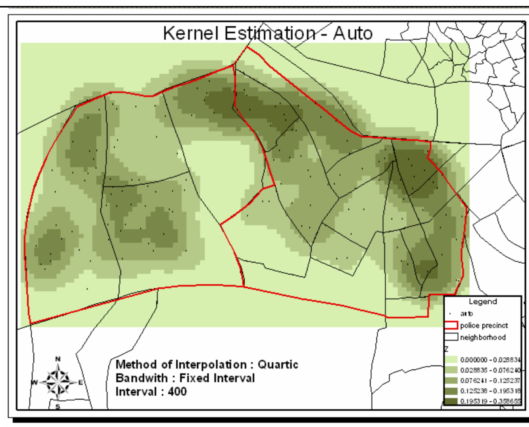


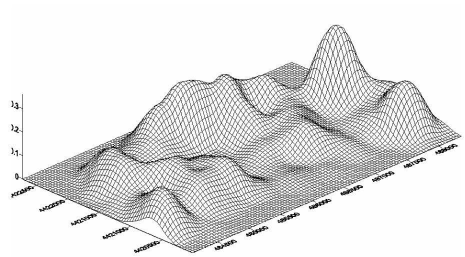


QUARTIC – FIXED INTERVAL

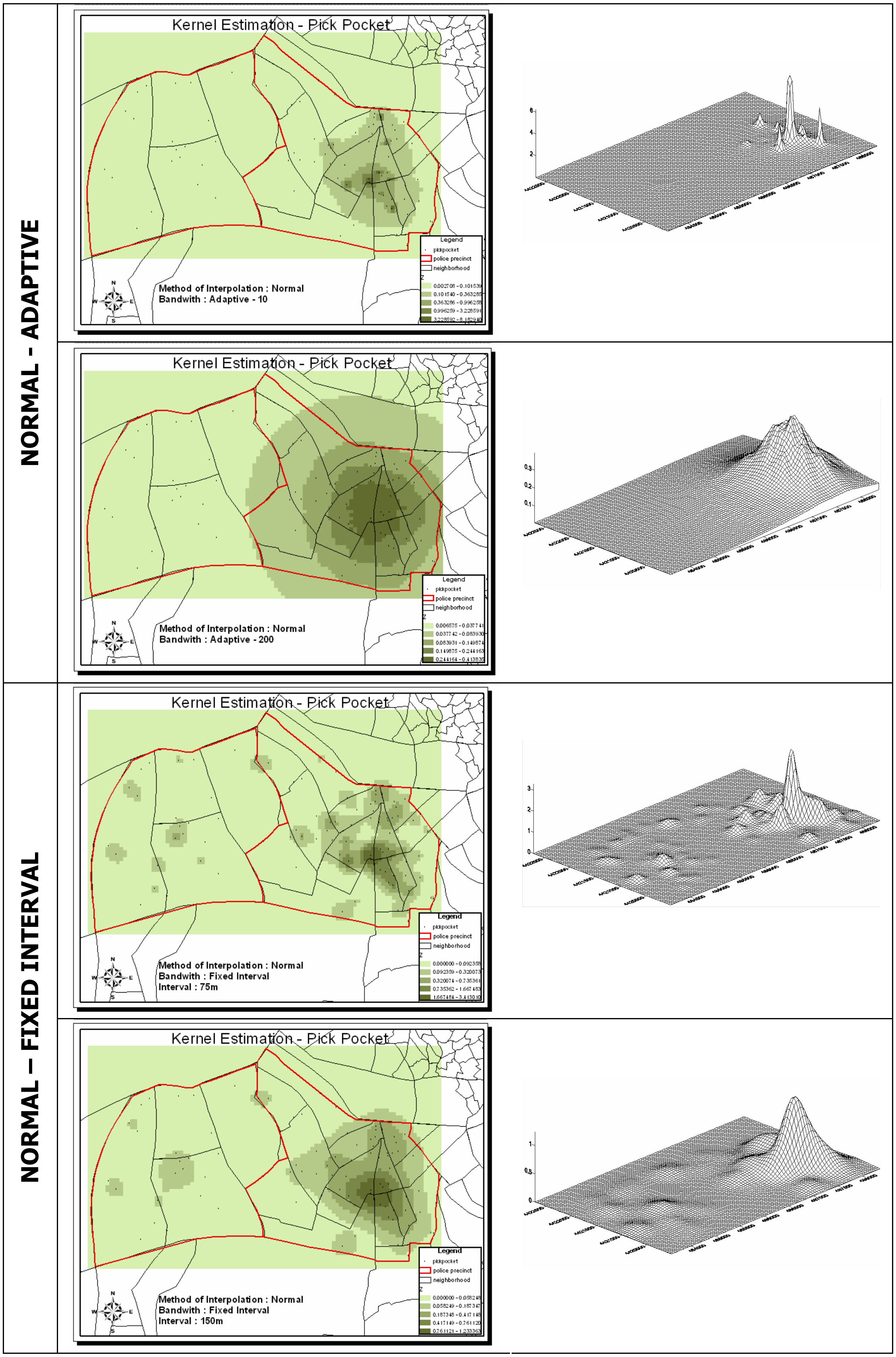






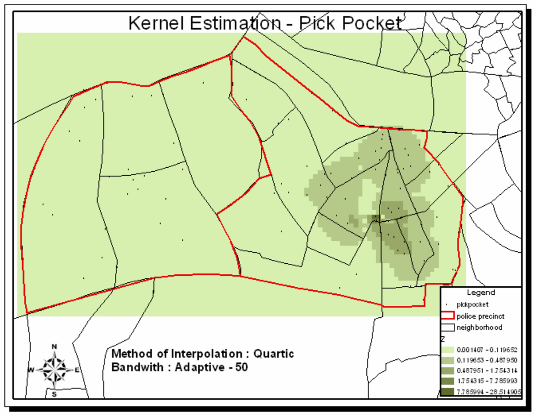


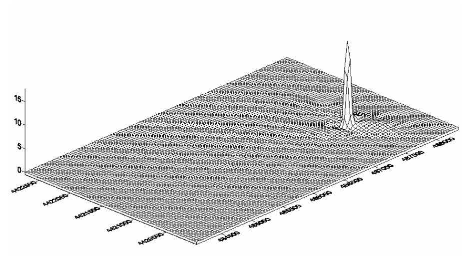
Map 4.7: Kernel Estimation for Auto Incidents (cont'd)

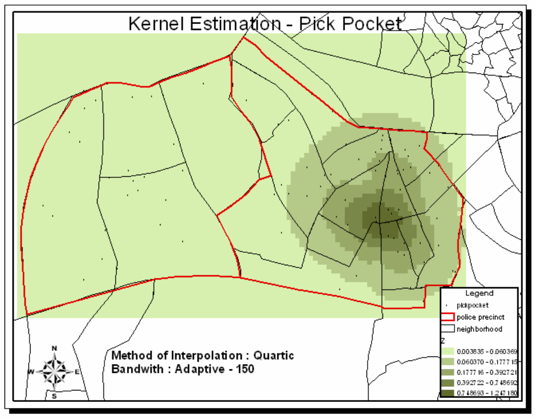


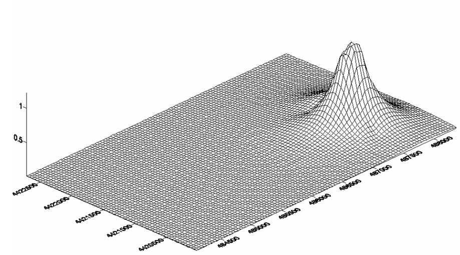
Map 4.8: Kernel Estimation for Pick Pocket Incidents

QUARTIC – ADAPTIVE

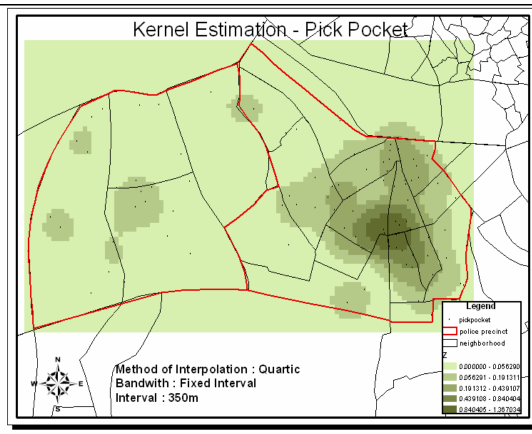


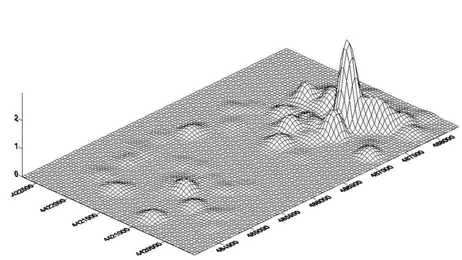


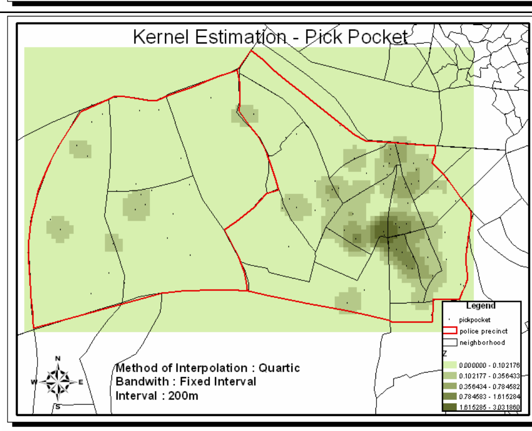


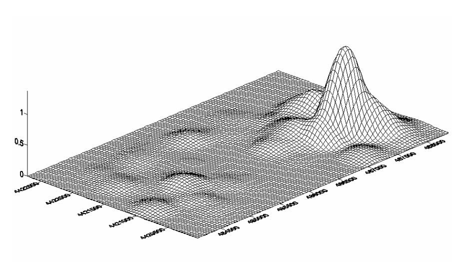


QUARTIC – FIXED INTERVAL

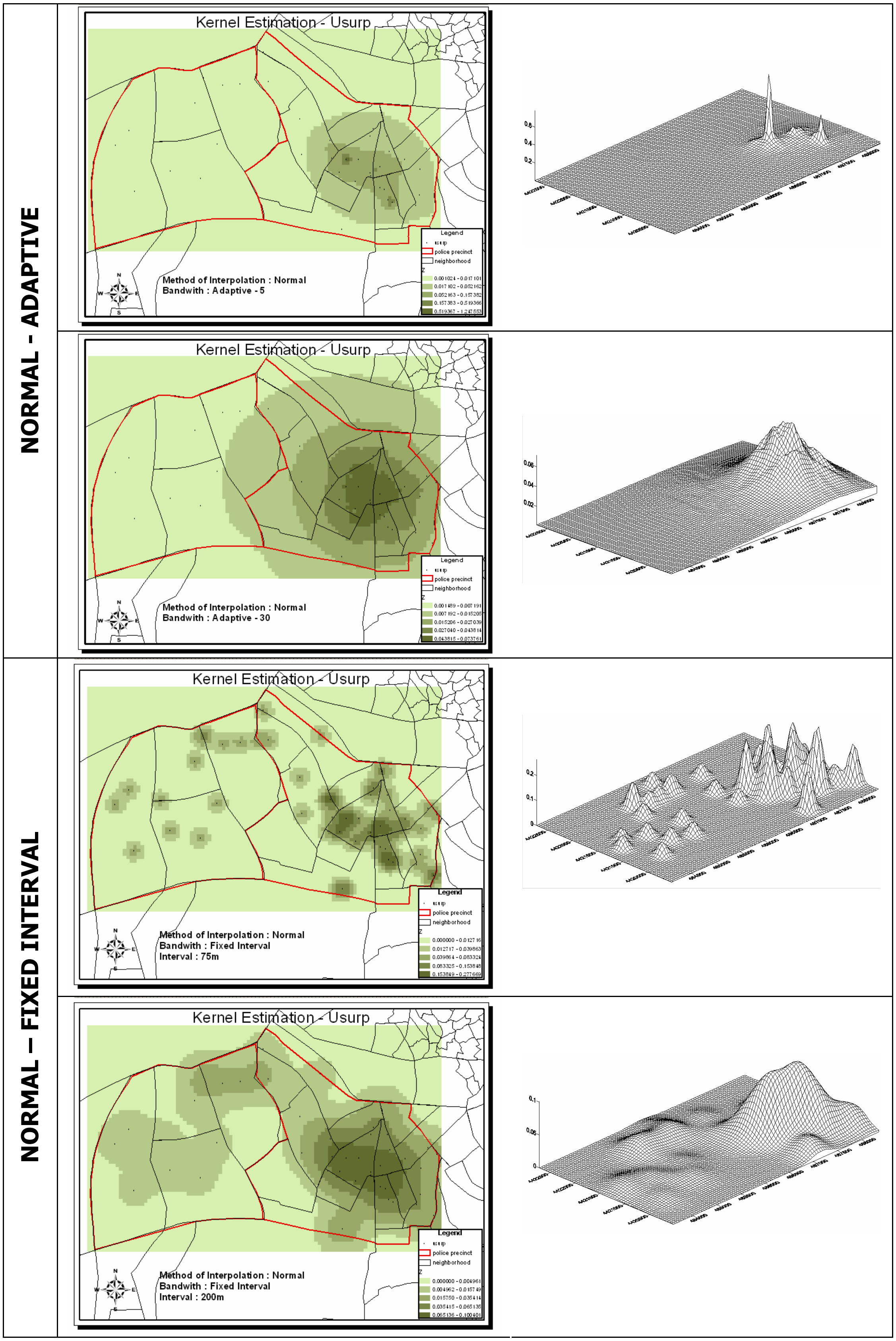






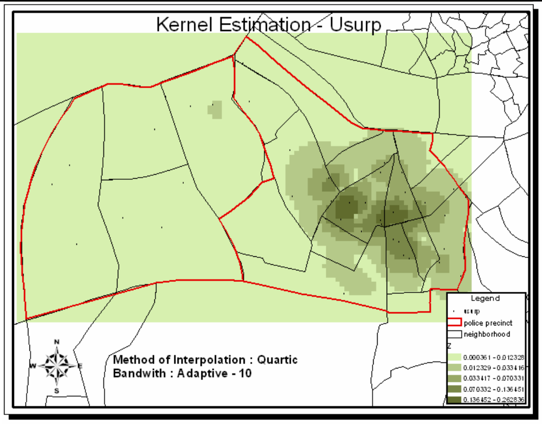


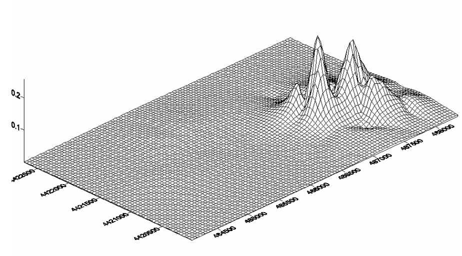
Map 4.8: Kernel Estimation for Pick Pocket Incidents (cont'd)

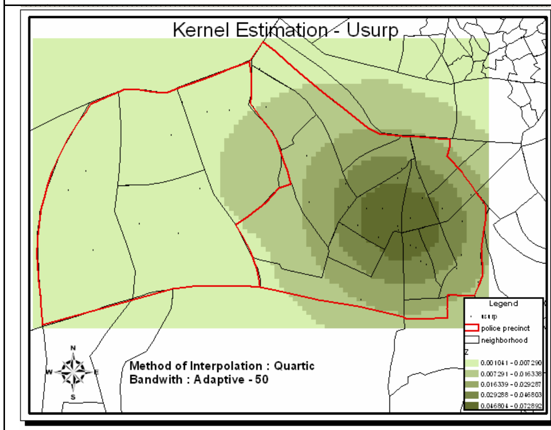


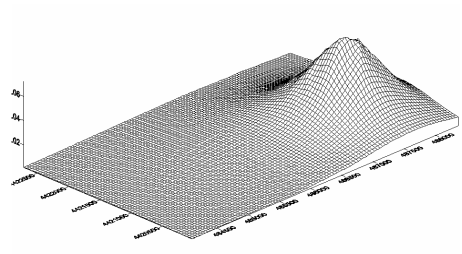
Map 4.9: Kernel Estimation for Usurp Incidents

QUARTIC – ADAPTIVE

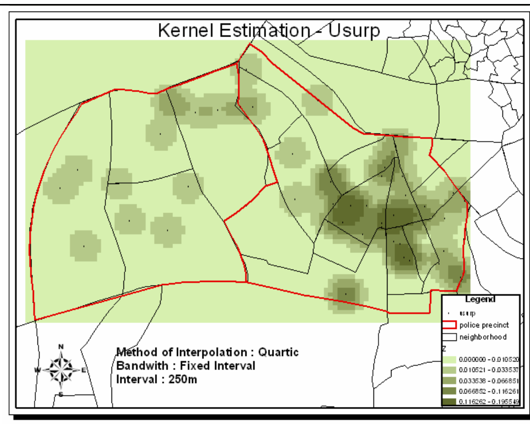


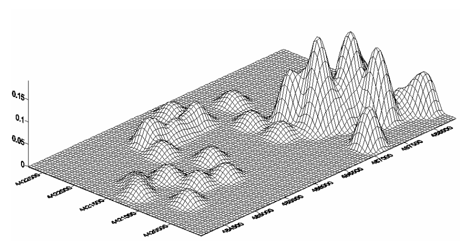


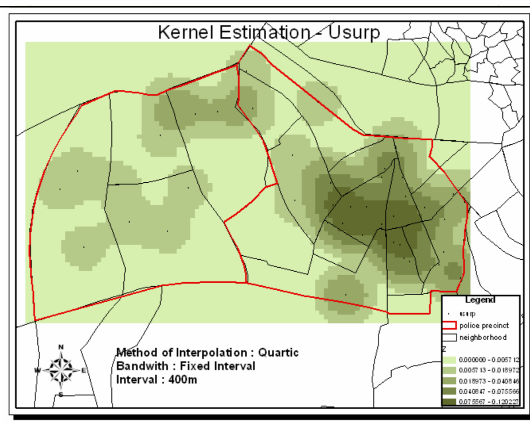


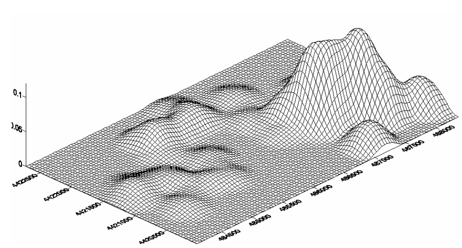


QUARTIC – FIXED INTERVAL

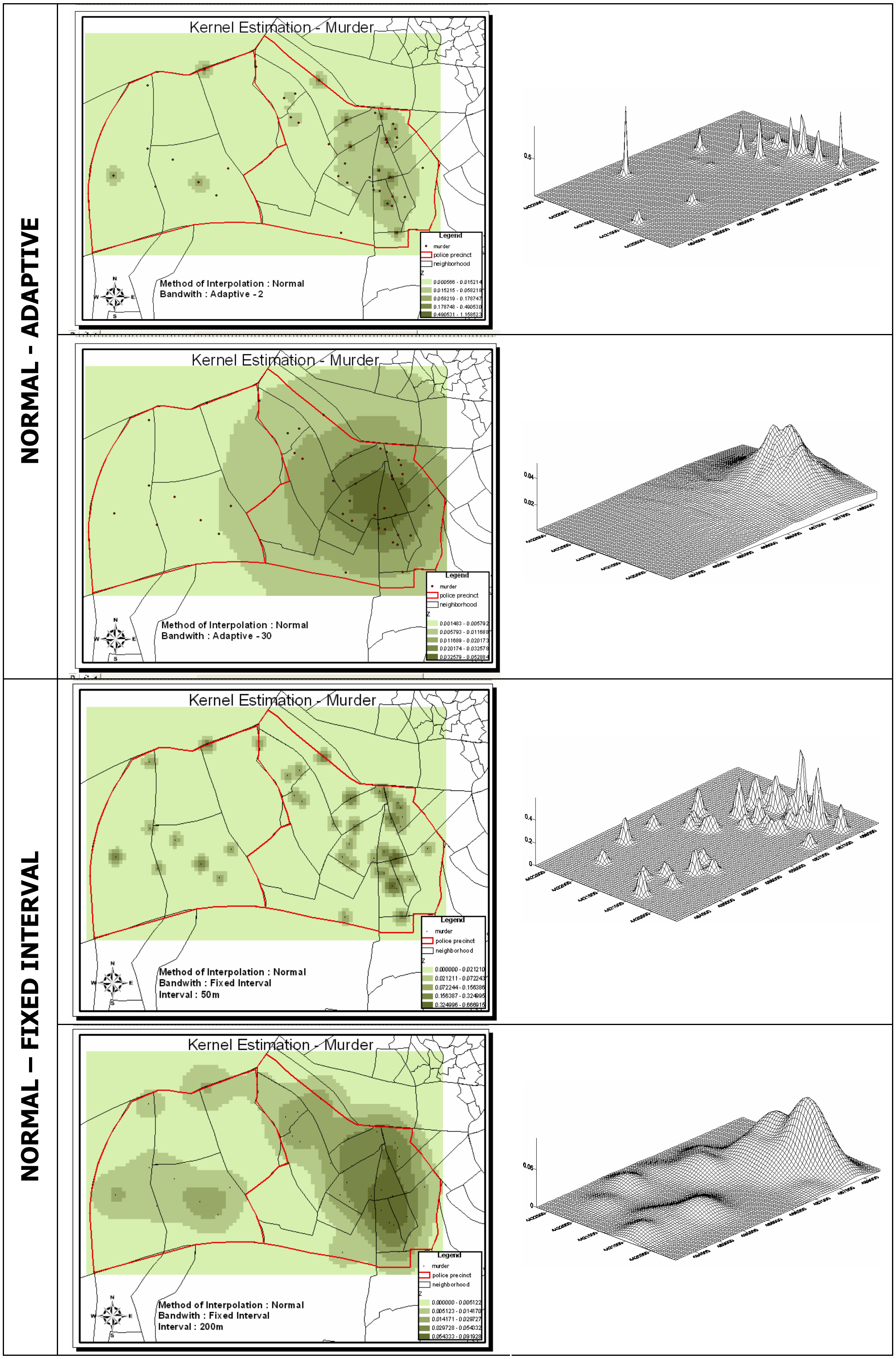




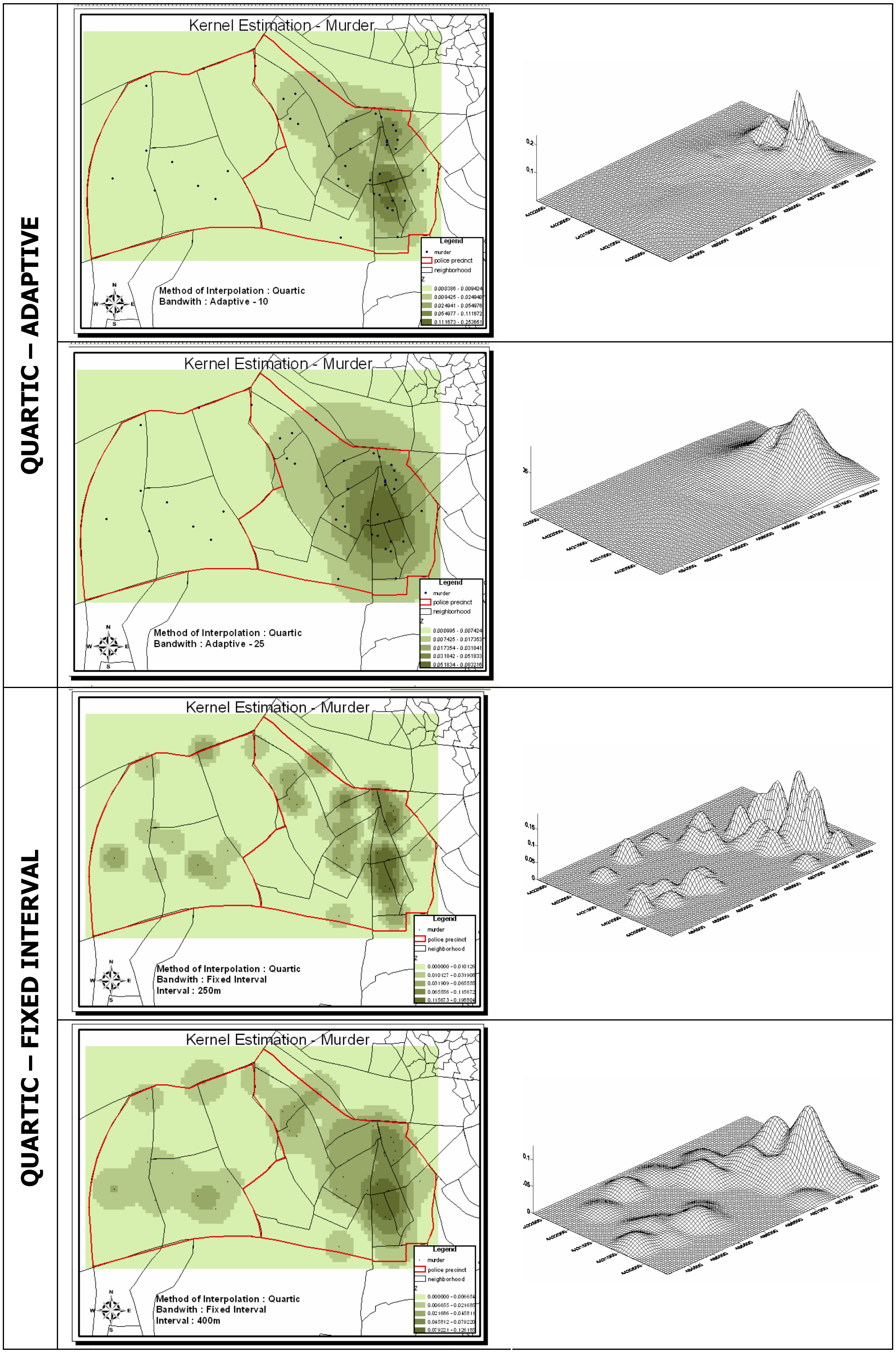




Map 4.9: Kernel Estimation for Usurp Incidents (cont'd)



Map 4.10: Kernel Estimation for Murder Incidents



Map 4.10: Kernel Estimation for Murder Incidents (cont'd)

In the above figures performed for each incident type, areas with higher densities are shown in darker tones and those with lower densities are shown in lighter tones. Figures show that while very small size give a very spiky appearance of the intensity, increasing the window size may cause the loss of the variation at that window size. For large bandwidths, number of distinct peaks decrease and single peaks increase.

The results obtained from normal distribution function provide similar visualizations with quartic function when adaptive bandwidths are placed. They generally give a single hill for the largest bandwidth based on the number of incidents, but the hill in normal function is not as smooth as the one in the quartic function. On the other hand, for fixed interval bandwidth, normal function yields similar outputs with quartic function. There are much more hills in fixed intervals than adaptive bandwidth for each function. For both adaptive and fixed interval bandwidths, the results of normal function are more generalized.

Both quartic and kernel estimation methods work for first order point pattern analysis but the outputs of kernel estimation indicate spatially more smooth, continuous and meaningful surfaces than quartic analysis. However, same as quartic analysis grid sizes have important effects on the results of kernel estimation. So, in order to choose the most appropriate bandwidth size, the same trials used in quartic method, are performed.

4.2.3. Application of Nearest Neighbor Distances

One of the distance statistics is the nearest neighbour index which uses distances between each point as its basis unlike quadrat and kernel methods and deals with second order properties. In addition it is not concerned with grid size problem. This analysis is performed in MS Excel program by the following below five stages;

- A distance matrix is calculated for each event
- The minimum values for each column are selected
- The minimum values (nearest values) which are defined as w, are counted
- G(w) values are calculated
- G(w) values are plotted against the nearest distances (w)

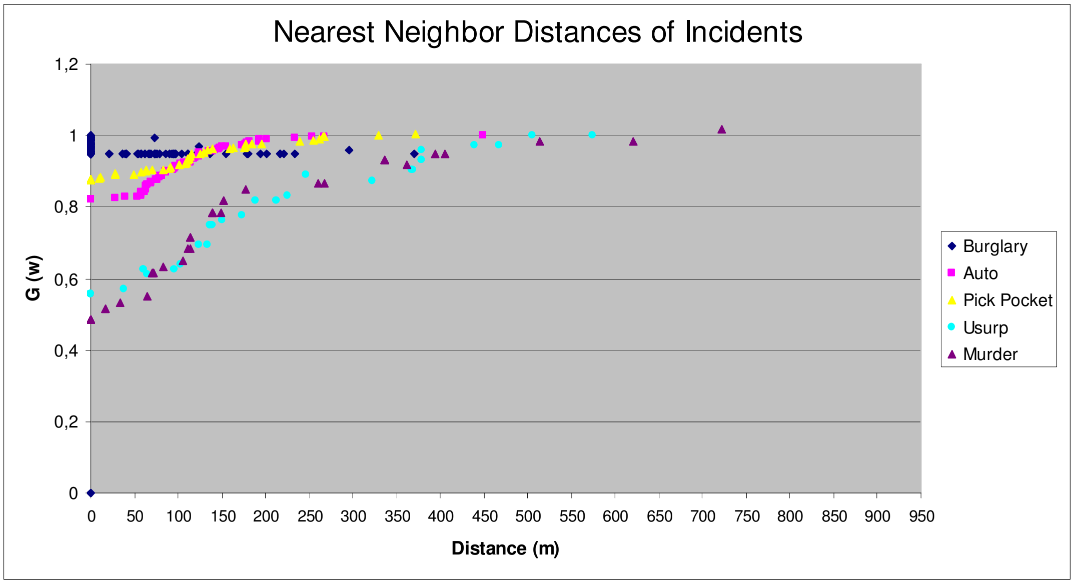


Figure 4.2: The Nearest Neighbor Distances for the Incidents

If the distribution function climbs very steeply in the early part of its range before flattening out, then the indication would be an observed high probability of short as opposed to long nearest neighbour distances, which would suggest clustering due to inter-event attraction. Alternatively if it climbs very steeply in the later part of its range, then the suggestion might be one of inter-event repulsion or regularity.

In both of these plots (Figure 4.2) a flattening is obtained after a certain distance. Since the plots indicate some climbs although not so steep, it is implied that there are cumulatively, rather some of short event-event distances; in other words, local clustering may be observed from these plots.

The curve for burglary incidents shows clustering for the first 100 meters while it shows nearly 150 meters for both auto and pickpocket incidents. For usurp incidents the curve indicates 250 meters and for murder it shows nearly 200 meters. This method demonstrates the possibility of the occurrence of another burglary event within 100 meters to the selected event. This may be the reason of any gang activity.

In order to be sure about the pattern of the incidents nearest neighbour index for each incident type are calculated for each incident for which the minimum distances have already computed in MS Excel program. The results of the computation, the NNI values are obtained as follows:

NNI (Burglary) = 0.126 which indicates cluster pattern
NNI (Auto) = 0.281 which indicates cluster pattern
NNI (Pick Pocket) = 0.228 which indicates cluster pattern
NNI (Usurp) = 0.615 which indicates cluster pattern
NNI (Murder) = 0.577 which indicates cluster pattern

If the results are compared for each incident, all nearest neighbour index values indicates a pattern between clustered and random, but while burglary, auto and pickpocket are more close to cluster distribution, the others are more close to random distribution.

4.2.4. Application of K Function

The other method called K function which is similar to but more comprehensive than nearest neighbour distance is applied to the study area. In order to analyze K function CrimeStat program is used which calculates the K statistic in terms of its linear counterpart L statistic.

The program takes each point in a set and draws a circle around it based on a specified search radius. CrimeStat takes the distance equal to about half of the region and divides it into 100 equal interval lengths, then creates a search radius of each length, from the smallest to the largest. It then counts the number of other points found inside the circle; it moves the search circle to the next point and continues the count until it visits all the points in the distribution.

After this process is completed, the radius of the circle is increased, and the entire process is repeated. Once it's done this around every point, it counts up the mean number of points within a given distance of each point. It then compares this number with the expected mean score for a random distribution.

The results obtained from Crime Stat are explained as;

- L_T: is the observed clustering for a specific distance (d)
- L_CSR: is the expected clustering for a specific distance (d) if the events were distributed randomly
- L_T_MIN: is the lower confidence interval
- L_T_MAX: is the upper confidence interval

An observed L_T value less than L_T_MIN would indicate that the events are much more dispersed than expected. An observed L_T value bigger than L_T_MAX would indicate that events are much more clustered than expected for that distance (interview with Scott, 2004). To summarize values of L that are greater than the upper limit of simulations indicate concentration while values of L less than the lower limit of simulations indicate dispersion.

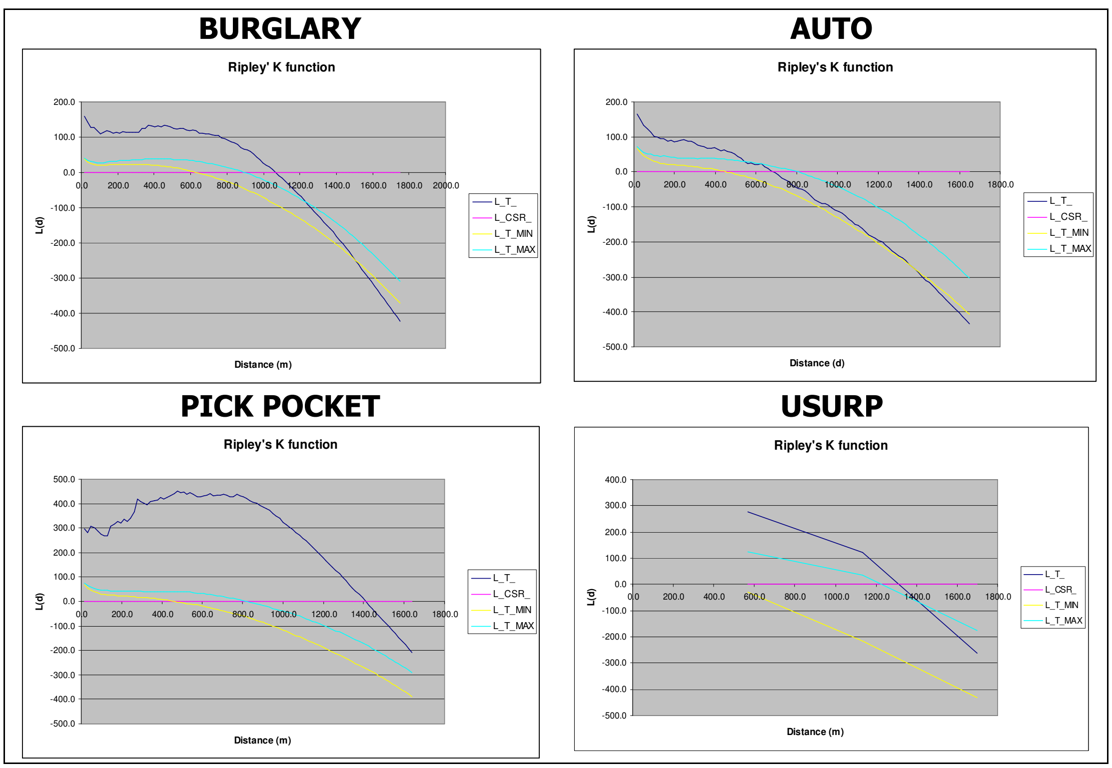


Figure 4.3: K-fuction for the Incidents

MURDER

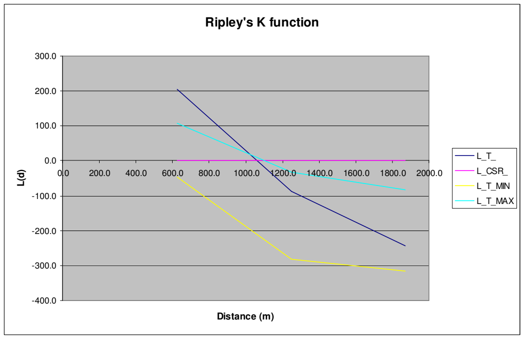


Figure 4.3: K-fuction for the Incidents (cont'd)

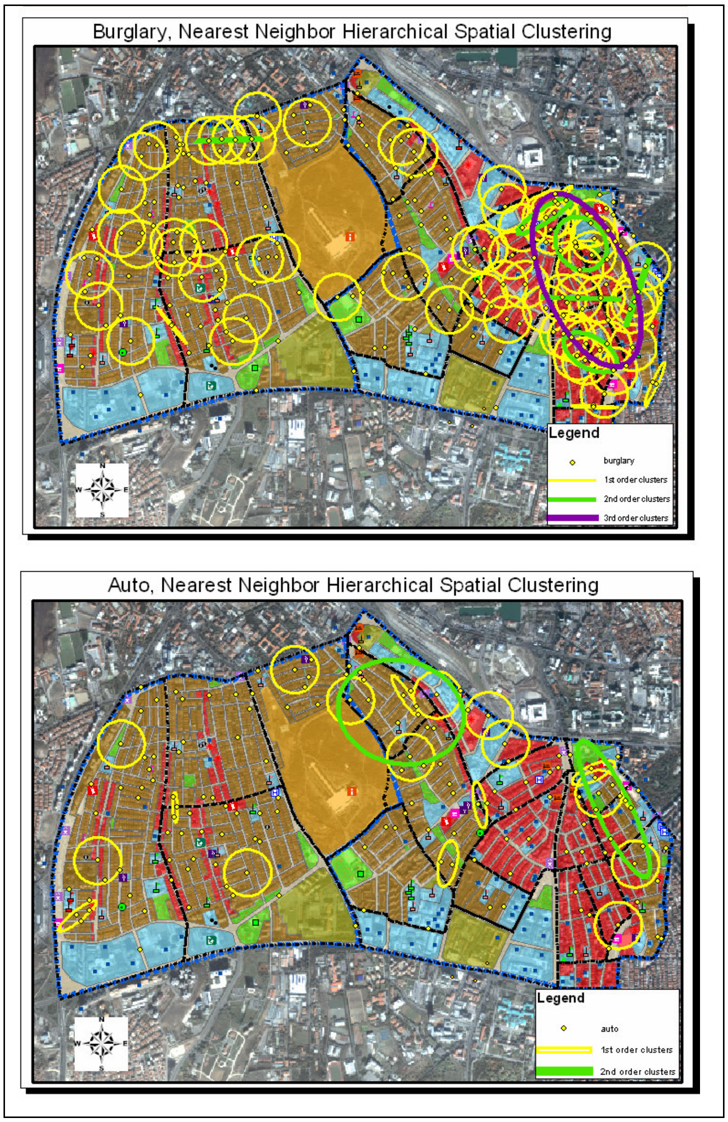
All the graphs (Figure 4.3) show clustering at the shorter distances for those distances where the dark blue line is higher than the light blue cyan colored line. K function of burglary incidents indicate clustering up to around 1200 meters, while auto incidents cluster within nearly 600 meters. On the other hand pickpocket incidents show clustering up to around 1400 meters while murder incidents cluster within 1000 meters. The methods point out that auto types are the most clustered incidents compared to the others. They become dispersed beyond 600 meters.

All three curves fall off because of edge effects since there are no data points outside the study area, in other words the search circle starts to go out of the study area where there are no events. Normally it's not interpreted for the entire scale instead it is appreciated for half or even a third of the scale (interview with Levine, 2004).

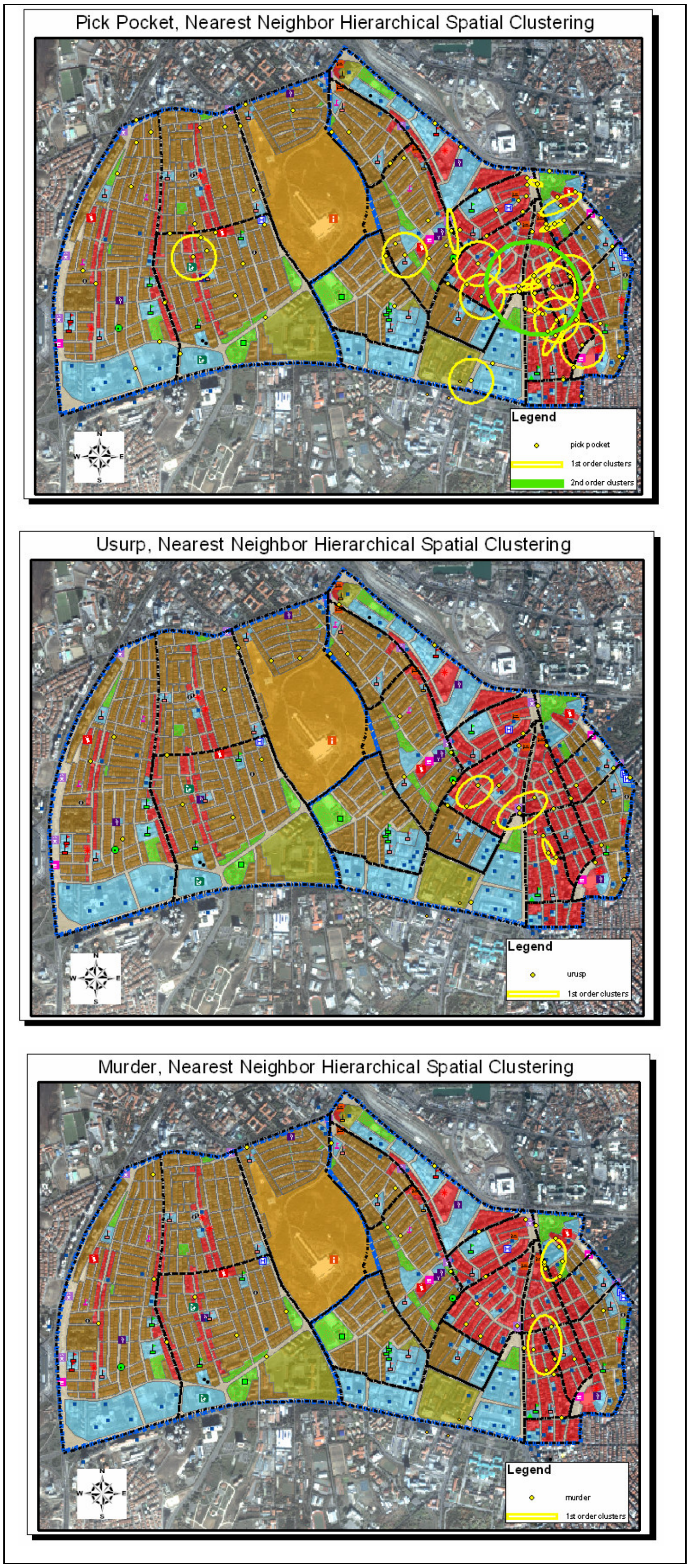
The two statistics – Nearest Neighbor Distance and K-function– measure related, but slightly different things. In other words, both use distances as a measure, but they use them differently. The Nearest Neighbour Distance measures the distance to the each nearest neighbor relative to a random distance while K-function measures clustering within a specified search radius. The first statistic compares neighbors, while the second compares distances. Although K function is more robust tool for second order effects most of the analysts prefer nearest neighborhood method (interview with Helms, 2004).

4.2.5. Application of Nearest Neighbor Hierarchical Clustering (Nnh)

In addition to the basic spatial point pattern methods, there are more methods to be applied for incident data which are obtained by CrimeStat program. One of them is Nearest Neighbor Hierarchical Clustering method. This technique identifies small geographical environments where there are concentrated incidents. Using minimum points per cluster as 5 for each incident types primary and secondary order clusters are determined by the method of Nearest Neighbor Hierarchical Clustering.



Map 4.11: Nearest Neighbor Hierarchical Clustering for the Incidents



Map 4.11: Nearest Neighbor Hierarchical Clustering for the Incidents (cont'd)

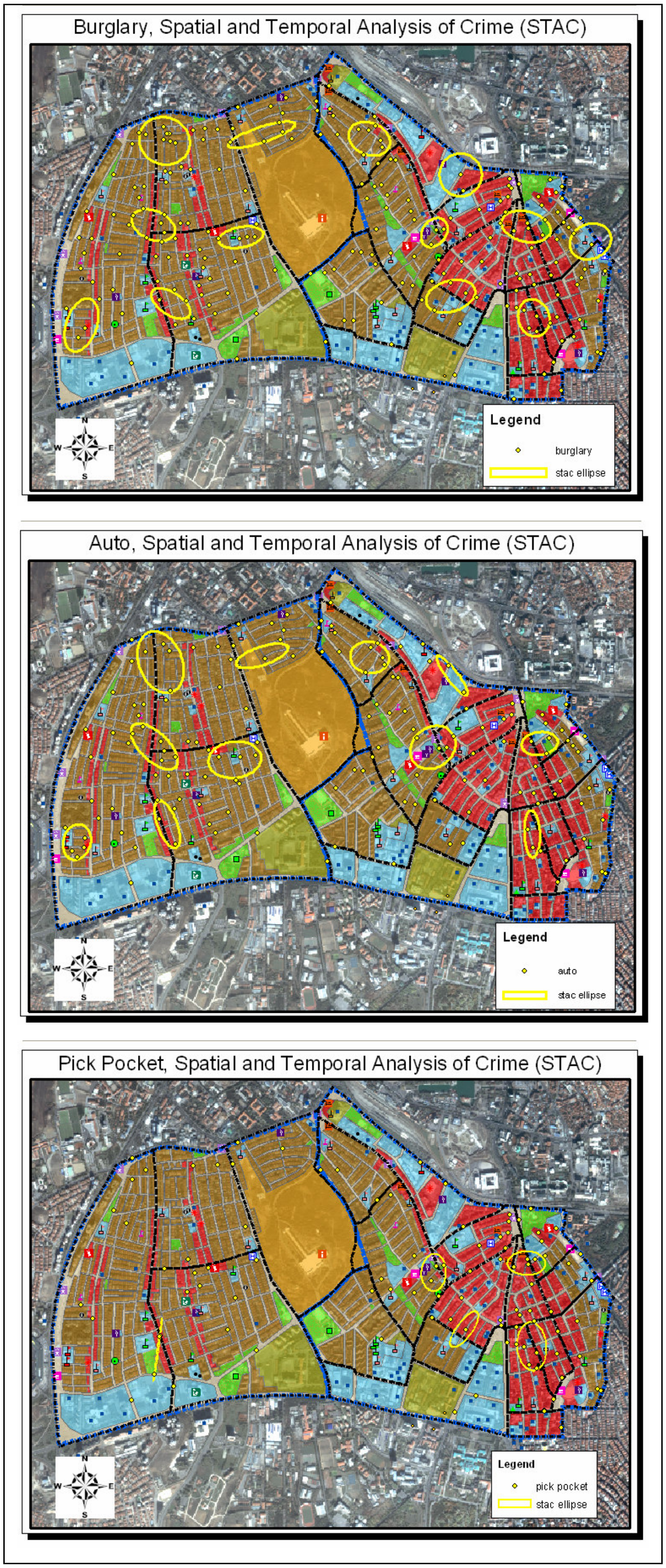
The above maps provide different levels of clustering for each incident type. Each of the level implies different policing strategies. For the smallest level, officers can intervene effectively in small neighborhoods. For example usurp and murder incidents in the study area form only first order clusters shown in yellow lines. Determining such areas can be useful for specific targeting. Second order clusters, seen in burglary, auto and pickpocket incidents as green lines, on the other hand, are more appropriate as patrol areas; these are larger than first order clusters, but include several first order clusters within them.

If third or higher order clusters are identified such as seen in burglary incidents in the study area and shown in purple lines, these are generally areas with very high concentrations of crime incidents over a fairly large section of the jurisdiction. The areas sometimes start to approximate precinct sizes and need to be thought of in terms of an integrated management strategy – police deployment, crime prevention, community involvement and long-range planning.

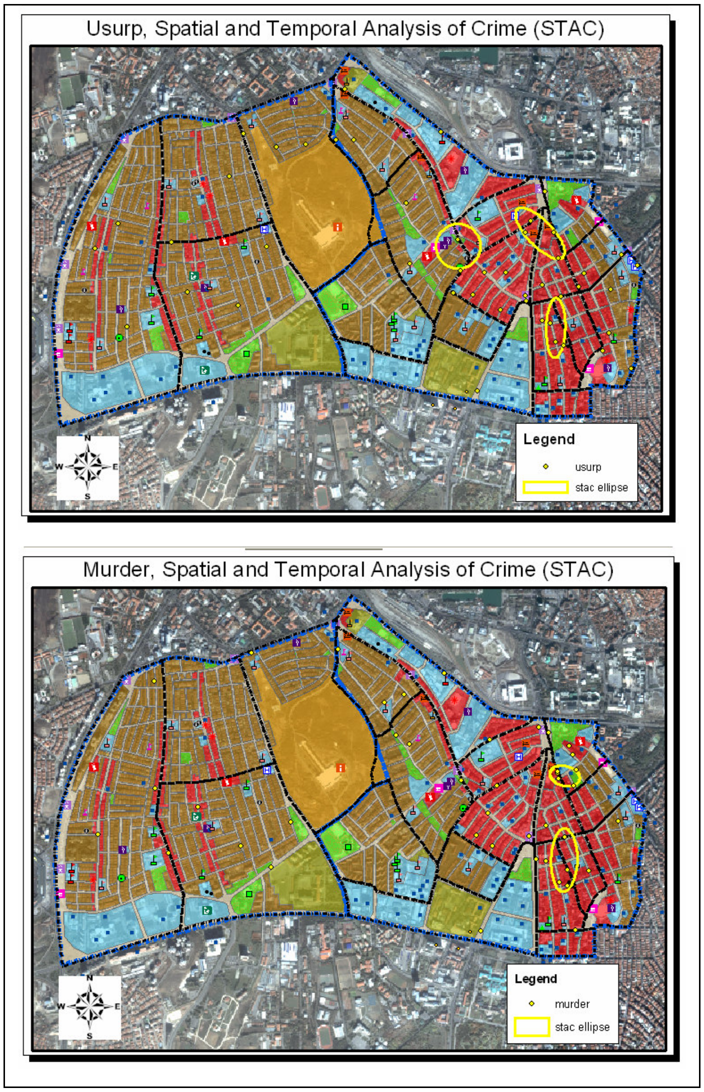
The outputs of Nearest Neighbor Hierarchical Clustering method are similar to kernel estimation method, especially to quartic kernel function with adaptive bandwidth, in which the highest crime concentration areas are mostly seen in the eastern part of the study area for each incident type.

4.2.6. Application of Spatial and Temporal Analysis of Crime (STAC)

Another method, applied to the study area is again obtained from CrimeStat which is a form of quadrat method. STAC as the first crime hotspot program (Levine, 2002) in which search radius is taken as 250 meters and minimum points per cluster are taken as 5 for the study area is a combination of a scan statistic (a circle is repeatedly laid over a grid and the number of points within the circle are counted) and a hierarchical clustering technique. The results are visualized as a standard deviational ellipse computed for the points identified to be a “hot spot”.



Map 4.12: Spatial and Temporal Analysis of Crime (STAC) for the Incidents

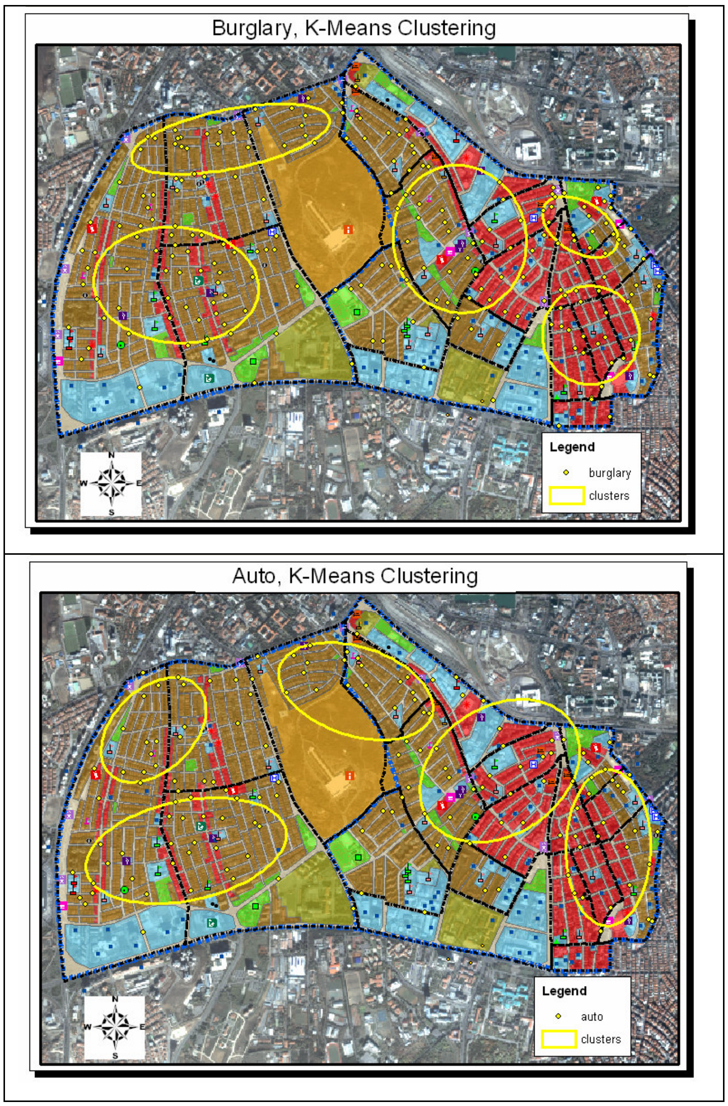


Map 4.12: Spatial and Temporal Analysis of Crime (STAC) for Incidents (cont'd)

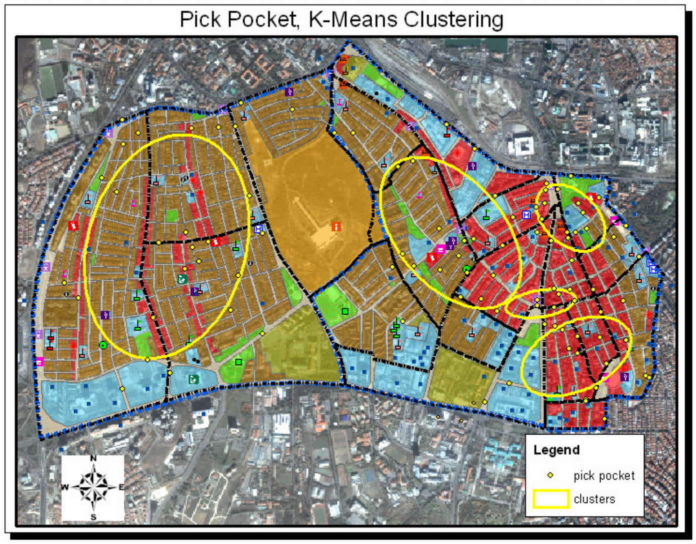
As noticed in the maps a number of densest Hot Spot Areas are determined for each incident. The results of STAC provide fewer and smaller clusters comparing to Nearest Neighbor Hierarchical Clustering. STAC provides 13 densest Hotspot Areas for 1041 burglary incidents, 11 areas for 377 auto incidents, 5 areas for 360 pick pocket incidents, 3 areas for 72 usurp incidents and 2 areas for 60 murder incidents. These results can be useful for specific targeting and prevention methods such as cameras, lightning, surveillance and control.

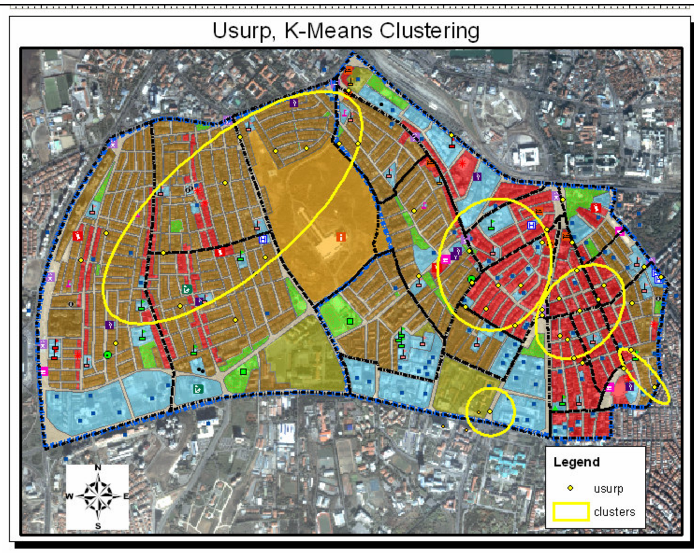
4.2.7. Application of K-Means Clustering

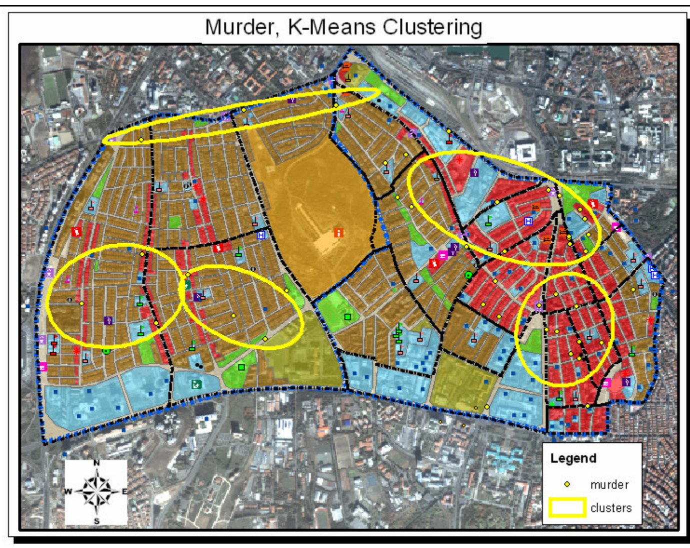
As a final method K-Means Clustering is performed which represents an attempt to define an optimal number of K locations where the sum of the distance from every point to each of the K centers is minimized. The data are grouped into K groups which are defined by the user and it is taken as 5 for the study area.



Map 4.13: K-Means Clustering for the Incidents







Map 4.13: K-Means Clustering for the Incidents (cont'd)

K-Means Clustering method provides clusters in a defined number for each incident as seen in the maps. In contrast to Nearest Neighbor Hierarchical Clustering this method only gives first order clusters and larger ellipses. These compact clusters are useful for the location of some facilities such as police station and hospitals.

4.2.8. Evaluation of the Outcomes

In this chapter, different methods are performed for each incident type in order to obtain information about the spatial distribution or the pattern of these incidents. Both the presentation of the points obtained form quadrat analysis and kernel estimation are spread out over the study area and both of the methods give an intensity value to each pixel and the results of these methods cover the entire study area.

If these two methods are compared, kernel estimation gives smoother results. So it can be said that kernel estimation is more preferable for finding the distribution of the incidents based on their density in the extension of the study area, in other words for finding their overall pattern.

As for second order properties nearest neighbour distances and K-function are performed for the study area. They give the information of the relationship between inter-event distances. However while nearest neighbour distances method uses distances only closest events and therefore only considers the smallest scales of pattern and ignore larger scales of the pattern, K-function provides more effective summary of spatial dependence over wider range of scales.

In order to find some specific areas in which the incidents are more clustered, there are three alternative methods; Nnh, STAC and K-Means Clustering. The results of each method provide rather different clusters.

If the amount of incidents is fewer, the results of Nnh and STAC become similar. But with the increase of the number of incidents, the ellipses – clusters are located in different areas.

On the other hand both of these three methods could be used for different purposes. For example, K-Means Clustering technique is useful when a user wants to control grouping. This method may be used for determining new police precincts or new locations for police stations.

As for determining specific targeting areas Nnh and STAC may be more helpful. Because, they define smaller areas than K-means clustering. The patrolling routes with the high priority could be determined based on these areas which have a certain boundaries.

In addition, these small clusters could be investigated according to the land use types which are located inside these clustered areas. Surely this investigation will light the way for further studies like done in this thesis.

For instance, for burglary incidents in the study area, it can be easily noticed that the 3^{rd} order cluster which is obtained from Nnh method, includes one subway stop, one primary school and one hospital. The commercial areas almost cover this zone which is located in the trade center of Çankaya and also Ankara. It covers the near environs of the known and busy roads as Atatürk Bulvarı, Ziya Gökalp Caddesi, Mithat Paşa Bulvarı, Sakarya Caddesi and Yüksel Caddesi. This zone includes various commercial usages like shops, bars, restaurants, cafes, banks, exchange offices, jewelers, private schools offering specialized courses and many offices with different purposes. In addition since the buildings are generally in apartment style the common usage of both residential and commercial could be found in this zone.

The two 2^{nd} order cluster zones, determined for auto incidents by Nnh completely contain one student dormitory, two primary schools, one subway stop and one open market. Generally residential areas and parks are covered by these zones which also include the neighborhoods of Anıttepe, Sağlık and Fidanlık.

As for pick pocket incidents Nnh method provide one 2^{nd} order clustered zone. Two subway stops, one primary school and one public square are completely within this zone. Kızılay square and Güvenpark which are always crowded in the all day times, especially at the weekends are covered by this zone. People who are living in Ankara often come to these areas to walk around, make shopping and go to a restaurant or a bar. Similar to the properties of the zone determined for burglary incidents this zone includes various commercial units too. Most importantly Yüksel Caddesi which is a pedestrian road and generally crowded is covered by this zone.

The three 1st order clusters, found for usurp incidents, completely contain two subway stops and one public square. This zone fully includes commercial areas. However these zones are smaller than the other zones, determined for burglary, auto and pick pocket incidents. Thus more specific locations could be detected. For example Konur Sokak and Kızılay Square are the most important or with the first priority places for usurp incidents.

Finally for murder incidents two 1st order cluster zones are found which are smaller like the zones determined for usurp incidents. These zones contain commercial areas and a park area.

On the other hand, the land use types are also investigated according to the clusters obtained from STAC method. Since these clusters are smaller than the clusters of Nnh, especially landmarks instead land use areas, inside these clusters, could be determined. For example, there are one high school, two primary schools, one university, two subway stops, one hospital and one retail center are found inside the 13 clusters found for burglary incidents. For auto incidents there are two universities, one subway stops, one retail center, one open market, one high school and one car park are determined in these 11 clusters. Also some park areas are noticed in these zones. For pick pocket incidents one subway stop is found in the 5 cluster zones. There are one subway stop and interestingly one police station are distinguished inside the 3 clusters of usurp incidents. As for murder incidents there is no any special landmark is detected in these 2 cluster zones. However the location of the clusters for both usurp and murder are almost same with the clusters obtained from Nnh method.

By referring to these results it can be said that schools, subway stops, hospitals, car parks and parks have important effects on the occurrence of the incidents. In addition residential and commercial areas have different effects on these incidents.

CHAPTER V

DETERMINATION OF THE RELATIONSHIP BETWEEN INCIDENTS LAND USE

In this chapter, correlation calculations are introduced in order to find the properties of the relationship between incidents and land use. The inputs of the correlation analysis are determined by referring to previous analyses which are explained in Chapter IV. These inputs are schools, subway stops, hospitals, car parks, parks, residential end commercial areas. Additionally, assuming that police stations and also minor roads have significant effects on the incidents, these variables are studied too.

In order to find the number of incidents inside or a certain distance to the related variable, ArcGIS 8.2 functions are utilized. Then, the correlation coefficient values are found and the correlations are presented again with the help of SPSS. At the end of the chapter potential risky areas are determined according to the results of the correlations and with the opportunities of ArcGIS 8.2 functions.

5.1. The Inputs of the Correlation Calculations

In order to determine the strength and the direction of the relationship between land use and crime incidents, land use and also landmark types are determined based on their potential risk on the occurrence of the incident. The land use and landmark types which are used for the relationship analysis are limited by the existing data. That means there may be more land use and also landmark types which have an effect on the amount of the crime incidents than used in this thesis but because of the limited data only existing information are used.

The land use types which are examined for the relationship between incidents and represented by polygon features in shapefile format are classified as residential, commercial, park and road areas. There are also military areas and public associations in the study area with a dense usage, but it is supposed that they are safer areas when compare to the others. However, despite there are some recorded incidents in public associations they are not taken into account since the rate is insufficient.

On the other hand schools, police stations, hospitals, subway stops and car parks are selected as landmark types and represented by point features in shapefile format which have significant effects on the occurrence of the incidents. Like land use types the same issue about the limited data is valid in determining the landmark types.

As the first stage of the correlation analysis, the size of the neighborhoods and also the size of the land use areas are computed. In order to generate this computation firstly the land use areas are selected by their location. The land use areas which are contained by the related neighborhood are selected in ArcGIS environment and then calculated in square meters. Thus the size and the percentage values of these land use areas are obtained based on their types (Table 5.1)

Similarly the incidents which are completely within the related neighborhood are selected in ArcGIS based on their types. Then the number and the proportion of each incident are computed based on its neighborhood which contains this incident. While calculating the incident number per square kilometer the size of military areas is subtracted from the total neighborhood area (Table 5.2). Because neighborhoods including larger military areas have naturally fewer incident proportions and this expected result may effect the further correlation calculations.

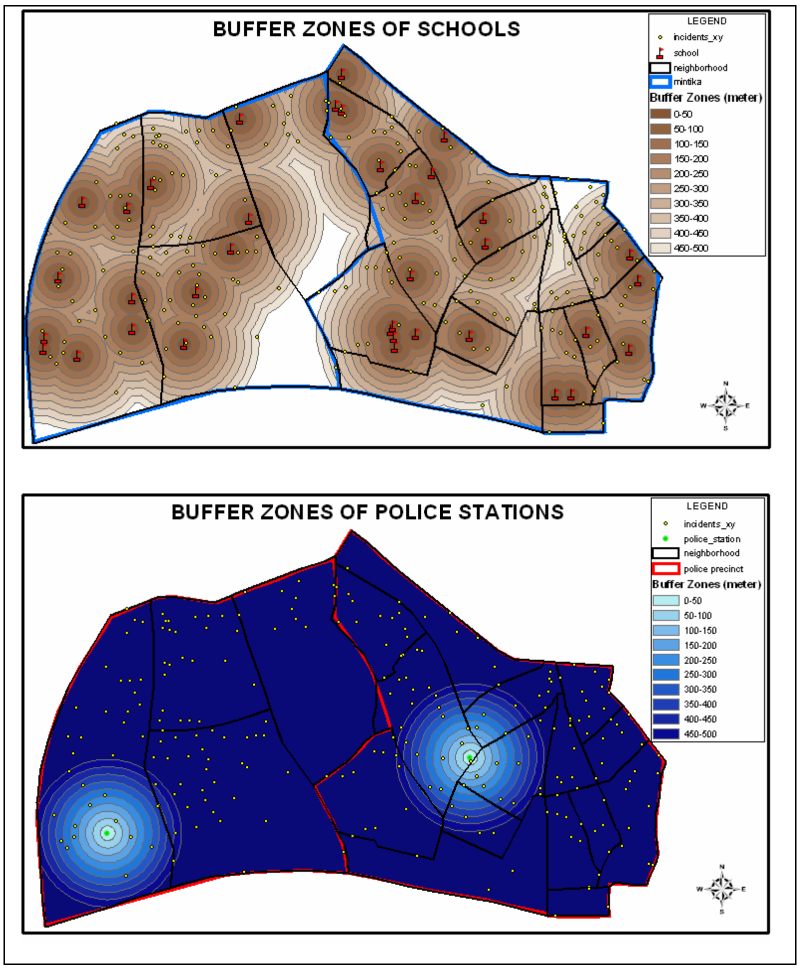
Table 5.1: The Size and Rate of the Land Use Types inside Each Neighborhood

Neighborhood	res (%)	res (m2)	com (%)	com (m2)	park (%)	park (m2)	road (%)	road (m2)	military (m2)	public (m2)	other (m2)	total (m2)
Anittepe	64.15	164628	2.38	6111	0.98	2520	24.70	63388	0	17691	2293	256632
Bahcelievler	60.22	351567	9.26	54087	1.79	10451	24.86	145139	0	21080	1503	583827
Cumhuriyet	0.00	0	61.24	81167	0.00	0	37.78	50062	0	1287	14	132529
Devlet	0.00	0	0.00	0	4.81	17719	28.92	106638	126570	239308	5106	495341
Emek	42.33	584855	6.76	93406	1.92	26578	27.71	382900	0	281969	11882	1381590
Eti	0.00	0	24.33	86088	6.19	21905	28.29	100087	19688	140262	5510	373539
Fidanlik	22.68	28038	29.00	35864	0.00	0	32.56	40261	0	19232	264	123660
Kavaklidere	0.00	0	61.48	47058	0.00	0	21.88	16744	0	12513	224	76539
Kizilay	0.00	0	65.13	175922	0.00	0	27.54	74404	0	19782	19	270127
Kocatepe	0.00	0	65.59	68878	0.00	0	27.91	29309	0	6831	0	105018
Korkutreis	0.00	0	51.12	101670	0.57	1124	28.45	56588	0	38941	558	198882
Kultur	30.31	87669	24.75	71588	2.48	7163	28.18	81519	0	37994	3350	289283
Maltepe	53.11	252800	6.88	32747	7.55	35932	20.50	97567	0	41517	15463	476025
Mebusevleri	68.67	182107	0.00	0	1.25	3320	30.00	79557	620303	225	0	885512
Mesrutiyet	0.00	0	65.15	112343	0.00	0	25.80	44489	0	15612	0	172444
Namikkemal	53.00	81032	0.00	0	0.00	0	13.04	19943	0	51925	3	152902
Saglik	10.10	15635	15.18	23500	26.30	40708	27.57	42670	0	28818	3446	154777
Yucetepe	47.89	206792	0.00	0	11.64	50283	21.44	92600	6973	82142	41	438832
Y.bahcelievler	50.31	461519	8.42	77192	7.028	64470	25.81	236760	186725	77313	20	1104000

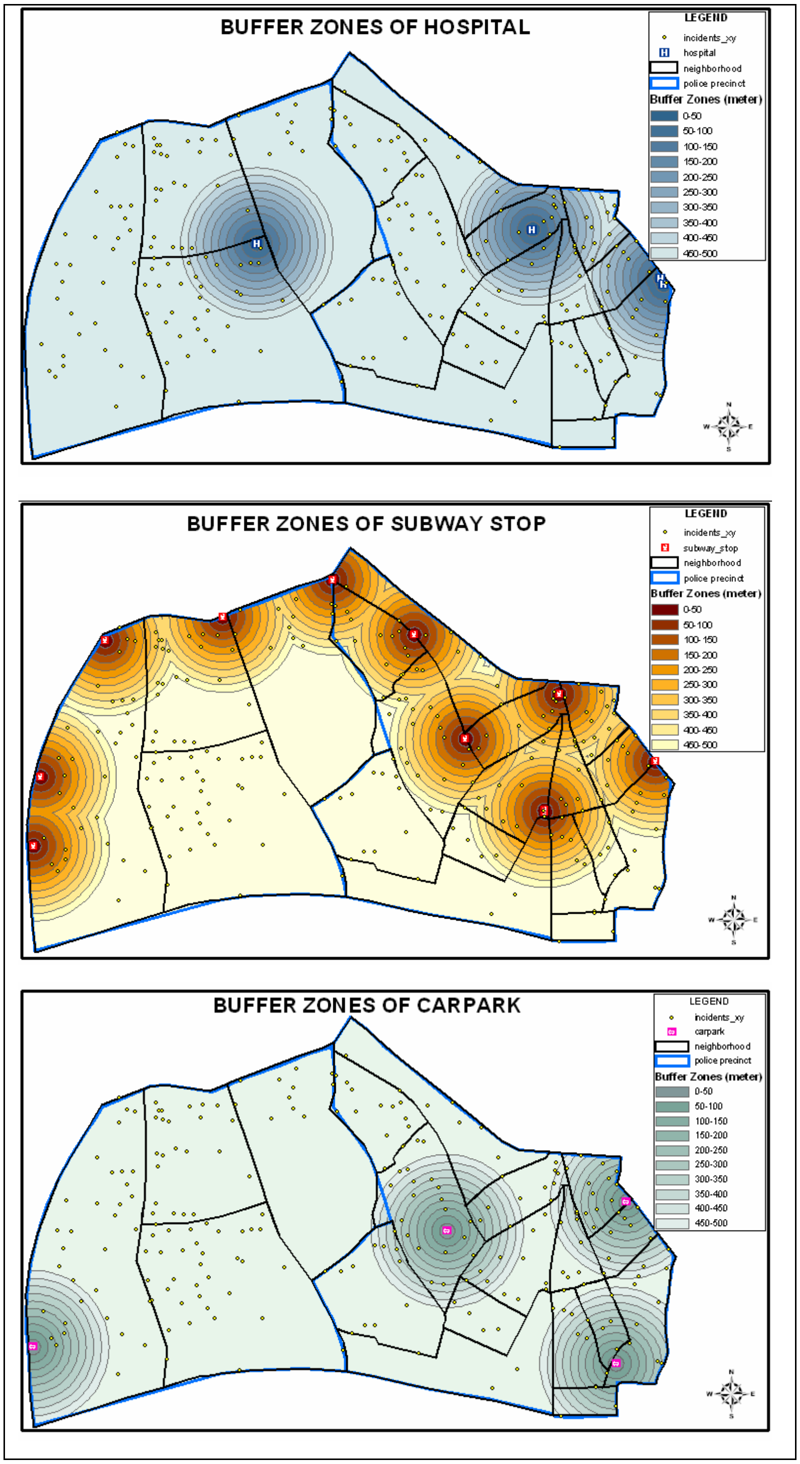
Table 5.2: The Amount and Proportion of the Incident Types Inside Each Neighborhood

Neighborhood	Cnt auto	auto/ km2	Cnt bur	bur/ km2	Cnt mur	mur/ km2	Cnt ppoc	ppoc/ km2	Cnt usurp	usurp/ km2	total-mil (m2)
Anittepe	35	136.38	36	140.28	3	11.69	7	27.28	2	7.80	**256632**
Bahcelievler	16	27.41	85	145.59	0	0.00	5	8.56	1	1.71	**583827**
Cumhuriyet	9	67.91	46	347.09	10	75.46	47	354.64	8	60.36	**132529**
Devlet	1	2.71	5	13.56	1	2.71	8	21.70	3	8.14	**368771**
Emek	62	44.88	141	102.06	5	3.62	20	14.48	3	2.17	**1381590**
Eti	29	81.96	16	45.22	2	5.65	3	8.48	2	5.65	**353851**
Fidanlik	12	97.04	31	250.69	0	0.00	11	88.95	4	32.35	**123660**
Kavaklidere	0	0.00	10	130.65	2	26.13	2	26.13	0	0.00	**76539**
Kizilay	6	22.21	67	248.03	6	22.21	100	370.20	15	55.53	**270127**
Kocatepe	14	133.31	19	180.92	1	9.52	12	114.27	0	0.00	**105018**
Korkutreis	13	65.37	85	427.39	4	20.11	15	75.42	2	10.06	**198882**
Kultur	23	79.51	54	186.67	1	3.46	8	27.66	7	24.20	**289283**
Maltepe	33	69.32	96	201.67	3	6.30	22	46.22	7	14.71	**476025**
Mebusevleri	27	101.81	52	196.07	2	7.54	3	11.31	4	15.08	**265209**
Mesrutiyet	7	40.59	77	446.52	8	46.39	49	284.15	7	40.59	**172444**
Namikkemal	6	39.24	12	78.48	0	0.00	4	26.16	1	6.54	**152902**
Saglik	28	180.91	60	387.65	7	45.23	24	155.06	2	12.92	**154777**
Yucetepe	7	16.21	29	67.15	0	0.00	2	4.63	0	0.00	**431859**
Y.bahcelievler	49	53.42	120	130.82	5	5.45	18	19.62	4	4.36	**917275**

As for landmark types firstly, multiple buffer rings are formed for each landmark points. The rings which have 50 meter intervals are than clipped according to the exterior boundary which is shaped by police precincts (Map 5.1). After generating the basic shape of the buffer areas the buffer rings are joined from incident data according to their spatial location. Thus points (incidents) are joined to polygons (buffers) and each polygon is given a summary of the numeric attributes of the points that fall inside it, and a count field showing how many points fall inside it.



Map 5.1: Buffer Zones for Landmark Points



Map 5.1: Buffer Zones for Landmark Points (cont'd)

As a consequence of these processes, the size of each buffer ring and the number of points inside these rings are computed and the incident number per square kilometer of each ring is calculated, shown in below tables.

Table 5.3: Incidents Inside Buffer Rings of Schools

ToBufDist (metre)	area (m2)	Cnt auto	auto/ km2	Cnt burg.	burg./ km2	Cnt murder	murder/ km2	Cnt ppocket	Ppocket/ km2	Cnt usurp	usurp/ km2
0-50	268770.26	14	52.09	24	89.30	2	7.44	7	26.05	1	3.72
50-100	726652.86	38	52.30	120	165.14	10	13.76	26	35.78	7	9.63
100-150	1021926.36	34	33.27	132	129.17	6	5.87	18	17.61	3	2.94
150-200	1148928.43	89	77.46	267	232.39	11	9.57	64	55.70	16	13.93
200-250	1105528.29	109	98.60	153	138.40	12	10.86	77	69.65	18	16.28
250-300	922024.35	12	13.02	77	83.51	0	0.00	14	15.18	6	6.51
300-350	765298.52	34	44.43	107	139.82	5	6.53	21	27.44	3	3.92
350-400	539283.84	12	22.25	61	113.11	3	5.56	84	155.76	11	20.40
400-450	411754.71	17	41.29	64	155.43	3	7.20	25	60.72	1	2.43
450-500	312889.62	17	54.33	25	79.90	6	19.18	13	41.55	4	12.78
500<	448401.76	1	2.23	11	24.53	2	4.46	11	24.53	2	4.46

Table 5.4: Incidents Inside Buffer Rings of Police Stations

ToBufDist (metre)	area (m2)	Cnt auto	auto/ km2	Cnt burg.	burg./ km2	Cnt murder	murder/ km2	Cnt ppocket	ppocket/ km2	Cnt usurp	usurp/ km2
0-50	15697.71	0	0.000	2	127.41	0	0.00	2	127.41	1	63.70
50-100	47093.12	1	21.24	9	191.11	0	0.00	0	0.00	1	21.24
100-150	78488.54	0	0.00	12	152.89	0	0.00	1	12.74	0	0.00
150-200	109883.96	12	109.21	56	509.63	2	18.20	22	200.21	10	91.01
200-250	141279.37	6	42.47	13	92.02	1	7.08	3	21.24	3	21.24
250-300	172674.79	7	40.54	29	167.95	2	11.58	5	28.96	0	0.00
300-350	204070.21	14	68.60	19	93.11	0	0.00	4	19.60	1	4.90
350-400	235465.62	19	80.69	60	254.81	4	16.99	35	148.64	1	4.25
400-450	266861.04	0	0.00	8	29.98	0	0.00	1	3.75	2	7.50
450-500	290631.52	6	20.65	26	89.46	0	0.00	4	13.76	0	0.00
500<	6109313.11	312	51.07	807	132.09	51	8.35	283	46.32	53	8.68

Table 5.5: Incidents Inside Buffer Rings of Hospital

ToBufDist (metre)	area (m2)	Cnt auto	auto/ km2	Cnt burg.	burg./ km2	Cnt murder	murder/ km2	Cnt ppocket	ppocket/ km2	Cnt usurp	usurp/ km2
0-50	24789.59	1	40.34	3	121.02	0	0.00	1	40.34	1	40.34
50-100	62089.34	2	32.21	22	354.33	0	0.00	4	64.42	1	16.11
100-150	99519.38	6	60.29	21	211.01	0	0.00	3	30.15	4	40.19
150-200	136851.78	6	43.84	21	153.45	1	7.31	3	21.92	0	0.00
200-250	174144.54	11	63.17	61	350.28	3	17.23	10	57.42	1	5.74
250-300	210080.60	14	66.64	35	166.60	5	23.80	22	104.72	2	9.52
300-350	233917.60	20	85.50	37	158.18	3	12.83	12	51.30	1	4.28
350-400	262308.83	19	72.43	40	152.49	2	7.63	8	30.50	1	3.81
400-450	291942.37	18	61.66	51	174.69	3	10.28	13	44.53	4	13.70
450-500	293314.69	11	37.50	66	225.01	2	6.82	22	75.01	10	34.09
500<	5882500.28	269	45.73	684	116.28	41	6.97	262	44.54	47	7.99

Table 5.6: Incidents Inside Buffer Rings of Subway Stops

ToBufDist (metre)	area (m2)	Cnt auto	auto/ km2	Cnt burg.	burg./ km2	Cnt murder	murder/ km2	Cnt ppocket	ppocket/ km2	Cnt usurp	usurp/ km2
0-50	67645.55	4	59.13	20	295.66	4	59.13	76	1123.50	5	73.92
50-100	177505.73	16	90.14	56	315.48	1	5.63	11	61.97	5	28.17
100-150	278425.48	30	107.75	90	323.25	5	17.96	30	107.75	8	28.73
150-200	375980.12	14	37.24	44	117.03	2	5.32	9	23.94	0	0.00
200-250	459004.30	41	89.32	106	230.94	12	26.14	32	69.72	10	21.79
250-300	525702.37	20	38.04	68	129.35	3	5.71	39	74.19	2	3.80
300-350	592397.20	34	57.39	128	216.07	10	16.88	34	57.39	8	13.50
350-400	571932.65	42	73.44	116	202.82	9	15.74	41	71.69	14	24.48
400-450	428290.64	44	102.73	85	198.46	3	7.01	27	63.04	4	9.34
450-500	381945.47	25	65.45	44	115.20	1	2.62	7	18.33	4	10.47
500<	3812629.49	107	28.07	284	74.49	10	2.62	54	14.16	12	3.15

Table 5.7: Incidents Inside Buffer Rings of Carparks

ToBufDist (metre)	area (m2)	Cnt auto	auto/ km2	Cnt burg.	burg./ km2	Cnt murder	murder/ km2	Cnt ppocket	ppocket / km2	Cnt usurp	usurp/ km2
0-50	30148.71	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
50-100	78180.46	25	319.77	17	217.45	0	0.00	11	140.70	0	0.00
100-150	123601.41	9	72.82	11	80.00	0	0.00	5	40.45	0	0.00
150-200	166753.54	11	65.97	57	341.82	1	6.00	13	77.96	5	29.98
200-250	211601.40	24	113.42	84	396.97	5	23.63	31	146.50	3	14.18
250-300	255011.44	15	58.82	45	176.46	3	11.76	11	43.14	2	7.84
300-350	287235.55	13	45.26	75	261.11	8	27.85	22	76.59	8	27.85
350-400	311628.34	15	48.13	39	125.15	2	6.42	14	44.93	9	28.88
400-450	340239.32	11	32.33	48	141.08	6	17.64	32	94.05	7	20.57
450-500	370257.17	17	45.91	66	178.25	2	5.40	19	51.32	6	16.21
500<	5496801.66	237	43.12	599	108.97	33	6.00	202	36.75	32	5.82

After the amount of incidents and the size of land use areas for each neighborhood and also, the buffer ring areas of the landmark points and the amount of incidents inside these rings are obtained, the correlation between incidents and this land data is computed. In order to perform correlation calculation and generate the scatter graphs, the SPSS software is used in which Pearson's product moment correlation coefficient is chosen as a correlation technique because of the properties of the data. The value that is acquired from the calculation and lies between -1 and +1 indicates the strength and the direction of the related relation.

5.2. Correlation between Land Use and incidents

The correlation coefficient values are calculated for each incident type separately. In order to animate these correlations between the incidents and land data the correlations are represented by scatter graphs which give an initial idea about the direction and the strength of the relation. Following parts include the correlation calculations, generalized for five incident types: burglary, auto, pick pocket, usurp and murder.

5.2.1. Determining the Relationship between Burglary incidents and Land Use

By looking at the scatter graphs constructed for denoting the correlation between the variables the first idea about the relationship is obtained. The graphs performed for burglary events show positive correlation while sometimes they show negative or no correlation between the land data (Figure 5.1). For instance for commercial, park areas and road and park areas and roads the correlation is in positive direction which could be explained as; if commercial areas or parks or roads increase the rate of the incidents will increase.

On the contrary, for residential areas the correlation is in negative direction which means if residential areas increase, the rate of the burglary incidents will decrease. As for schools, police stations, car parks, hospitals and subway stops the rate of the incidents decreases where places far from these landmarks. This signifies that the risk for the occurrence of the burglary incidents near to these landmarks is higher.

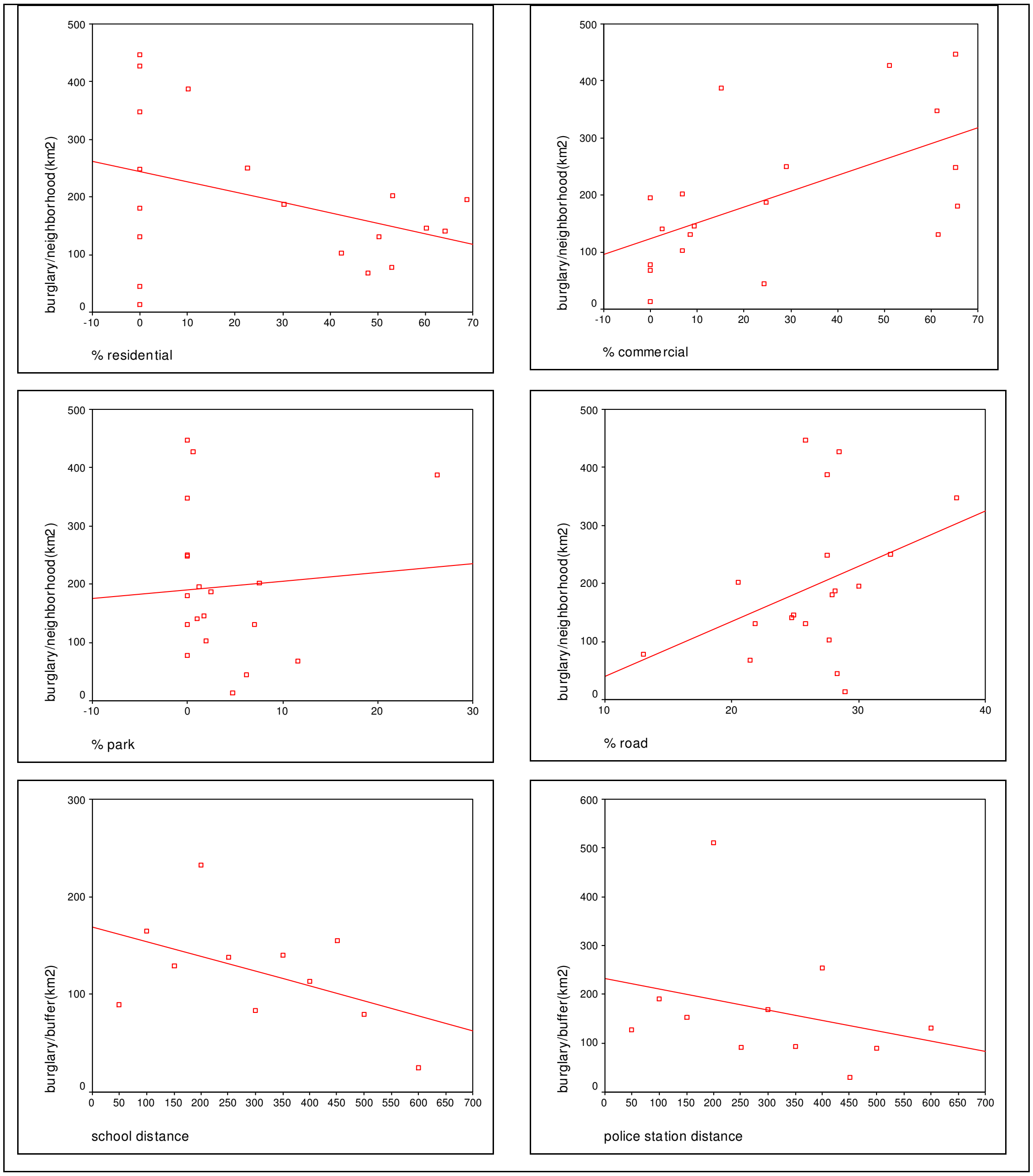


Figure 5.1: Scatter Graphs for Burglary between Land Use and Landmarks

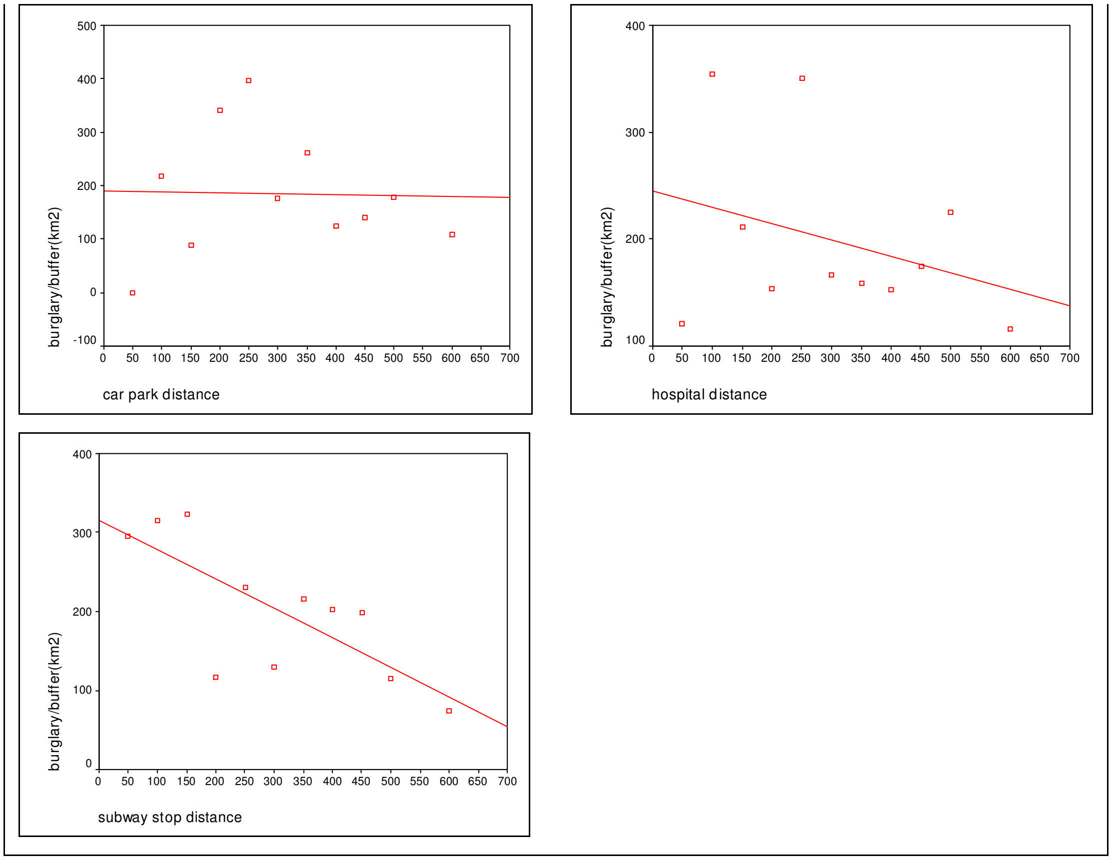


Figure 5.1: Scatter Graphs for Burglary between Land Use and Landmarks (cont'd)

In order to be sure about the strength of the relationship, the values of the correlation coefficient for burglary between each land use and landmark types are given in below table (Table 5.8).

Table 5.8: Correlation Coefficient Results for Burglary Incidents

Car Park	0.007
Park	0.074
Police Station	- 0.296
Hospital	- 0.308
Residential	- 0.377
Road	0.380
School	- 0.455
Commercial	0.573
Subway	- 0.746

A correlation coefficient of -0.746 indicates negative and strong relationship between burglary incidents and subway stops. This means that near to the subway stops the rate of the incidents increases. This may be the reason of the preference of the burglars on the places which are near to the subway stops which may help them to escape easily.

On the other hand commercial areas with that of 0.573 coefficient have positive and strong relationship between burglary incidents. This may be the reason of the choice of the burglars on the commercial areas, since the security of these buildings which includes dense commercial units but also residential flats are partly insufficient.

In addition in the study area, commercial areas are generally formed as small units, instead of large trade centers in which the security functions such as cameras, alarms and security guards may not be efficient enough. That means the apartments in the study area are generally shared with different commercial units and this may cause a low security.

However for regions in which the usage of residential areas is dense, the security precautions such as locked building doors or doorkeepers may also higher than the regions that have commercial units. In addition because of the neighborhood relation between people who are living in the same neighborhood know each other and they could have easily suspicion if there is stranger person. These factors may be the causes of - 0.377 coefficient value which denotes that the risk for the occurrence of the burglary incident is lower in residential areas.

A correlation coefficient of -0.455 between schools and burglary incidents represents a relationship in negative direction which implies that in the places near to the schools the rate of the burglary incidents increases. In the evenings, for instance, since there is no any activity in the surroundings of the schools, in other words they are more desolated in the evenings comparing to the day times, these places may become more suitable for burglars.

For roads, the correlation coefficient value is obtained as 0.380 which indicates a relation in positive direction but not so strong. This is an anticipated reason since there is not any potential property to be robbed on the roads so the result provide weak relation. But when the correlation coefficient value is calculated for both major roads and minor roads in order to find their effects on the surrounding places, the value is obtained as 0.438 for minor roads, it is -0.177 for major roads. One reason is when it is the main corridor the escaping ways are limited compared to the interior street networks where there are lots of twists and turns which make the escape an easier one. Also patrolling may be more in case of main roads.

When burglary incidents are taken into consideration as a crime type against to properties the weak relationship between burglary and both hospitals, police stations and parks are not surprising since there is nothing valuable to be robbed in these places. In addition it is normal that there is almost no correlation between burglary and car parks. Only that of -0.296 coefficient is interesting between burglary and police station which is the indicator of the higher rate of burglary in the places near to the police stations. But when the location of police stations is considered which generally prefers the places in city centers with the high income level this kind of correlation is partly normal.

5.2.2. Determining the Relationship between Auto incidents and Land Use

Similarly the scatter graphs are performed for auto incidents with the same land use and landmark types (Figure 5.2). As seen in the figures some of the correlations are in positive direction and some of them are in negative direction. For land use types while residential, park areas and roads show positive correlation, commercial areas shows negative. That means that the rate of the auto incidents increase in residential, park areas and roads in contrast to commercial areas.

As for landmark types police stations and hospitals have positive direction when schools, car parks and subway stops have negative direction. This is the indicator of higher auto incidents in the places close to the schools, car parks and subway stops.

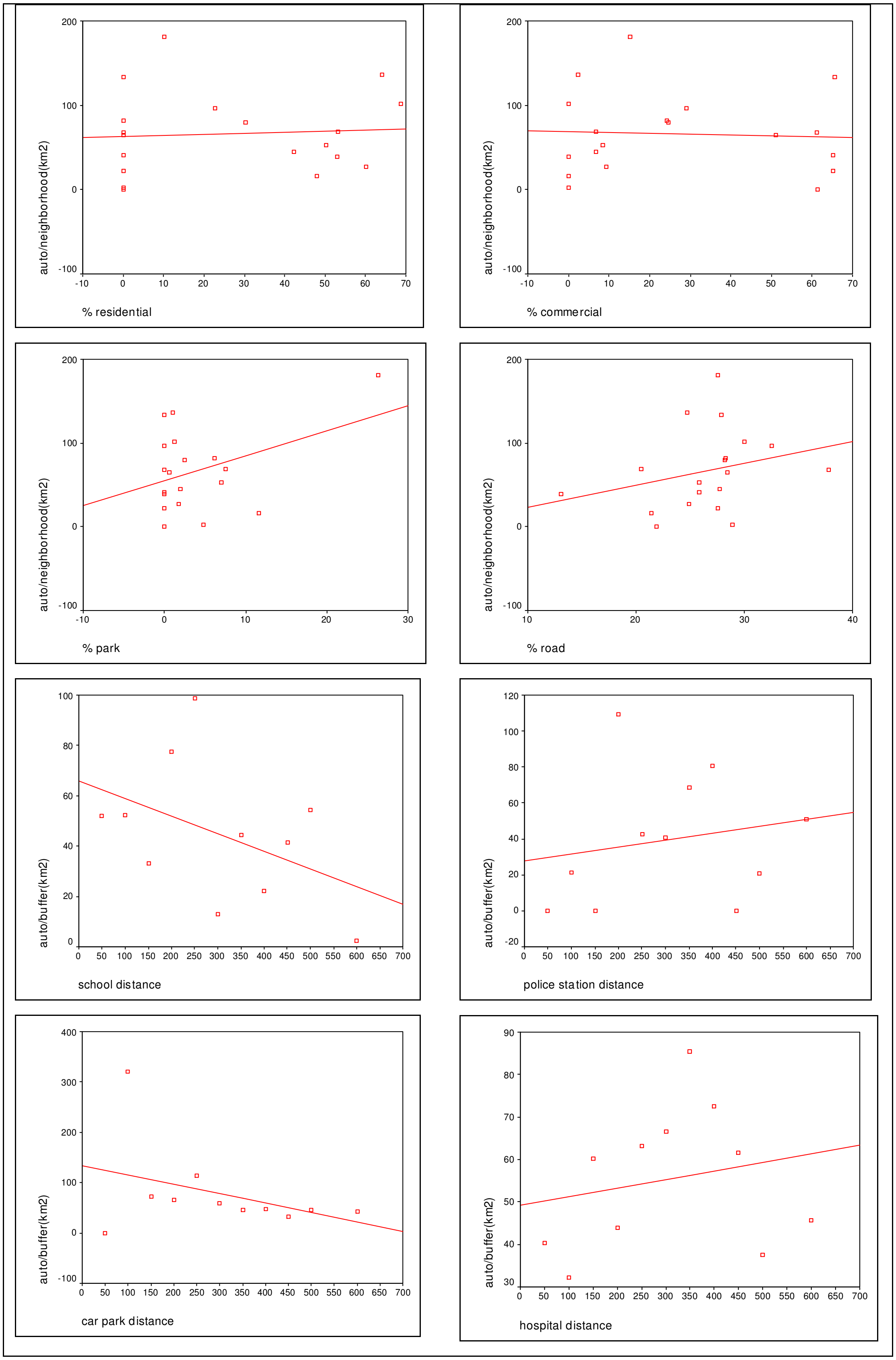


Figure 5.2: Scatter Graphs for Auto between Land Use and Landmarks

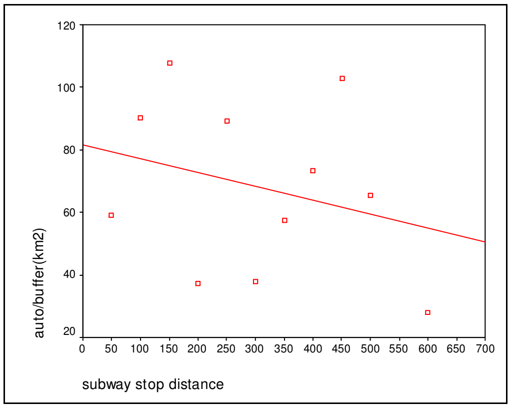


Figure 5.2: Scatter Graphs for Auto between Land Use and Landmarks (cont'd)

For the strength of the correlation the values of coefficient are examined for each variable between auto incidents (Table 5.9).

Table 5.9: Correlation Coefficient Results for Auto Incidents

Commercial	- 0.056
Residential	0.069
Police Station	0.186
Hospital	0.239
Subway	- 0.253
Road	0.281
Car park	- 0.391
Park	0.396
School	- 0.416

The values of correlation coefficient indicate that auto criminals prefer mostly open spaces as excepted. The strongest correlation in schools denotes that the rate of the auto incidents is higher in places near to the schools. The usage of the gardens of the schools as parking lots, especially in the evenings, may be the result of strong correlation. This result is also verified by the previous analysis which shows that auto incidents are mostly occurred in the evenings and nights from the evening 18.00 to the morning 09.00.

Parks and car parks are also risky for auto incidents since these areas are mostly dark and may not be protected effectively. Pavements close to the parks are generally used for parking and surroundings of the parks may not be light and safe. Similarly for car parks although there is at least one official person, this result shows that they could not protect the car parks sufficiently or these car parks may not be planned correctly.

Roads are also potential areas for auto incidents. Major roads and minor roads are examined separately and the correlation coefficients are obtained as 0.285 for minor roads, 0.001 for major roads. There are more parking cars on the interior streets than main corridors since it is generally forbidden to park the cars on the major roads. So it is normal that the correlation is stronger in minor roads than major roads.

The other variables such as subway stops, hospitals, police stations, residential and commercial areas do not have so strong correlation with auto incidents.

5.2.3. Determining the Relationship between Pick Pocket incidents and Land Use

The scatter graphs, given in below figures and performed for pick pocket incidents between land use and landmark types, show positive correlation in commercial areas, roads, schools and hospitals. On the other hand they show negative correlation in residential areas, parks, police station, car parks and subway stops. This denotes that the risk of the occurrence of the pick pocket incidents is higher in commercial areas, roads and places near to the police stations, car parks and subway stops.

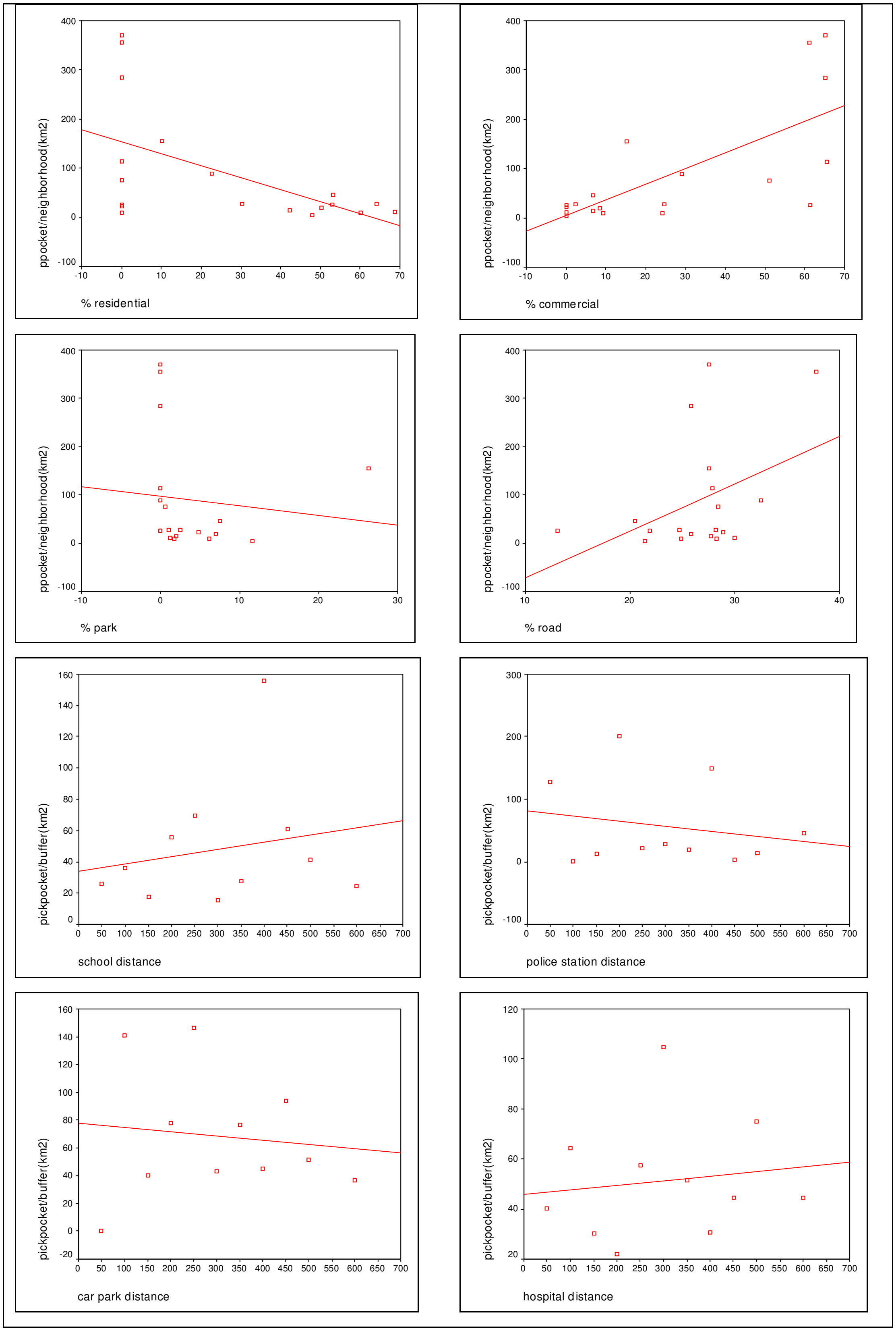


Figure 5.3: Scatter Graphs for Pick Pocket between Land Use and Landmarks

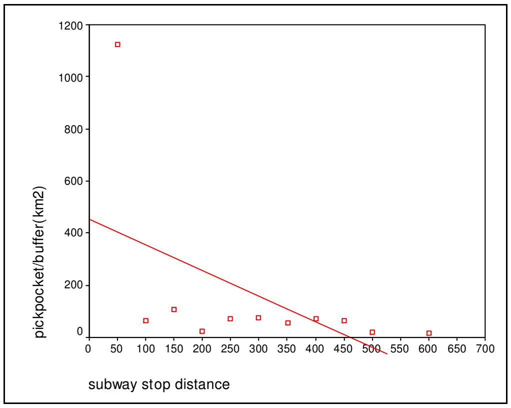


Figure 5.3: Scatter Graphs for Pick Pocket between Land Use and Landmarks (cont'd)

In order to obtain the information of the strength of the correlation between pickpocket incidents and land use and landmark types the correlation coefficients are computed (Table 5.10):

Table 5.10: Correlation Coefficient Results for Pick Pocket Incidents

Carpark	- 0.100
Park	- 0.107
Hospital	0.150
Police Station	- 0.211
School	0.229
Road	0.421
Subway	- 0.539
Residential	- 0.552
Commercial	0.711

The strongest correlation is obtained in commercial areas. This is an expected result because there are too much people in commercial areas to walk around for shopping. Similarly cafes, restaurant and shops are too attractive for pocket pickers especially if they are full of people. In addition since they come to commercial areas to spend money they carry too much money in their pockets or bags. In previous analysis it is obtained that most of the pick pocket incidents occur in the day time and on Saturday which is the time for shopping or walking around for people who are working at weeks.

As a normal result the strong and positive correlation with commercial areas, the coefficient value in residential areas give gives strong and negative correlation. This is why commercial areas too risky for pick pocket incidents. Pocket pickers generally prefer crowded areas so residential areas are not suitable for them.

Likewise subway stops are generally crowded in the day time and they may help pocket pickers to steal and escape quickly and nobody could notice them. For the surroundings of the subway stops are also active and crowded.

The coefficient value between roads and pickpocket incidents shows positive and strong association too. In the same way major roads and minor roads are examined separately. For minor roads the coefficient is 0.245 while it is 0.003 for major roads. The reason of being more risky of minor roads is similar with burglary which explains that escaping is easier in minor roads. Since the victims who are walking in the streets are more defenceless it may be easy for pocket pickers who may be pedestrian or use a car, to steal the victim's bag.

The other variables as schools, police stations, hospitals, parks and car parks have weak correlation. It can be explained that in parks or schools, for instance, people mostly do not carry valuable things in their pockets. So these areas are not attractive for pocket pickers.

5.2.4. Determining the Relationship between Usurp incidents and Land Use

Below figures, including scatter graphs between usurp incidents and land use and landmark types, show positive correlation in commercial areas, roads and places close to schools and car parks (Figure 5.4). On the other hand the correlations between usurp incidents and residential areas, parks and places near to the police stations, hospitals and subway stops have negative direction. These correlations are explained as if the size of the commercial areas and the distance to subways, police stations and hospitals increase in a neighborhood the rate of the usurp incident will also increase in this neighborhood.

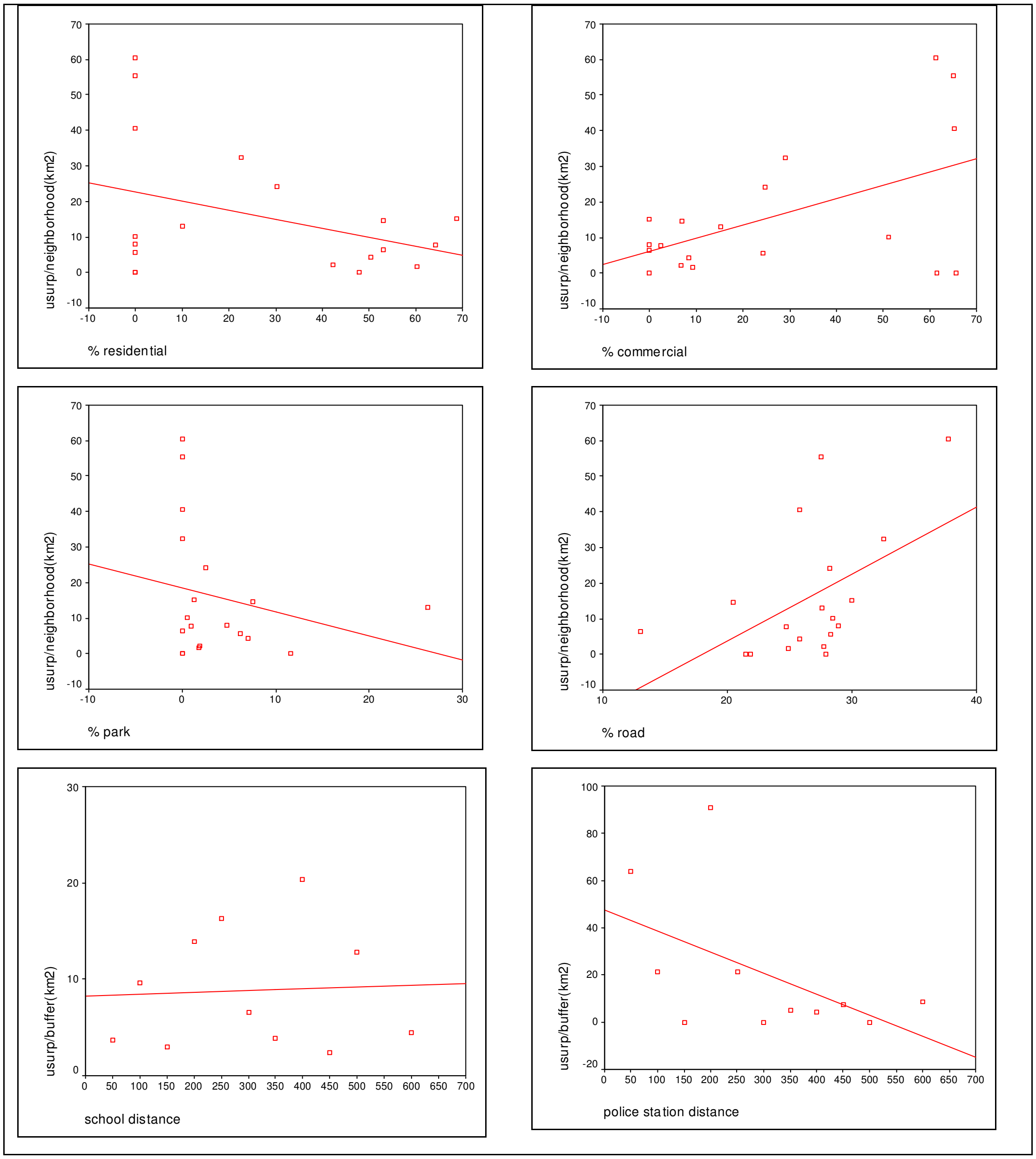


Figure 5.4: Scatter Graphs for Usurp between Land Use and Landmarks

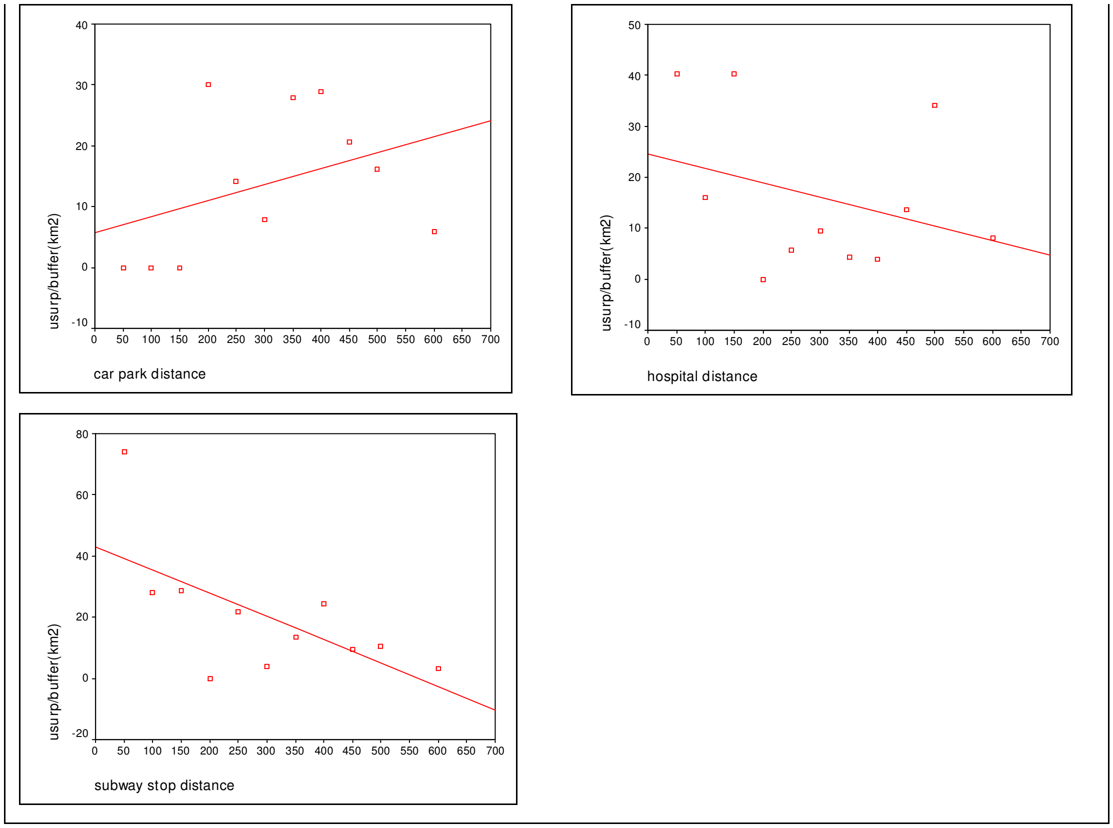


Figure 5.4: Scatter Graphs for Usurp between Land Use and Landmarks (cont'd)

Like previous correlation calculations the coefficient values are computed for usurp incidents between land use and landmark data in order to understand how these variables affect the rate of the incidents (Table 5.11).

Table 5.11: Correlation Coefficient Results for Usurp Incidents

School	0.075
Park	- 0.235
Hospital	- 0.327
Residential	- 0.373
Car park	0.424
Road	0.518
Commercial	0.529
Police Station	- 0.533
Subway	- 0.645

Since usurp is a crime type against human and also property, the offenders normally choose places such as subway stops which are more desolate in the evenings than day time. The coefficient of -0.645 for subway stops and the rate of usurp incidents in the evenings are the indicator of this strong correlation which indicates that places near to subway stops are more risky for usurp incidents.

However the coefficient value of -0.533 is surprising that the rate of the usurp incident increase where the places are near to the police stations. This may be reason of the types of the usurp incidents which includes the events of resist the police and bribe are occurred in high rate and these incidents are mostly seen in police stations. In addition since police stations are not the places to protect their surroundings it may not be the guaranty that places near to the stations are safety. The location of the police stations is only important to reach the incident place at the right time when receive any call for service.

The correlations between commercial areas with 0.529 and the roads with 0.518 coefficients have nearly same strength with police stations. Naturally commercial areas and roads are suitable places since they are more desolated in the evenings. The correlation calculation for major roads and minor roads gives 0.311 coefficient value for minor roads while it gives 0.040 value for major roads which means nearly no correlation. This may be reason of the less patrolling in case of minor roads.

Since commercial areas are more appropriated than residential areas it is normal that the coefficient value is acquired as -0.373 which means if the usage of the residential areas increase the number of usurp incidents decreases.

The coefficient of 0.424 indicates that the rate of the usurp incidents decreases where the places are close to the car parks. However hospitals which provide -0.327 coefficient are also potential places for usurp incidents in which people may more weakness and defenceless in hospitals.

Unlike these strong correlations, explained above, hospitals, parks and schools have weaker associations between the usurp incidents. Especially, the correlation between schools and usurp incidents is almost nonexistent.

5.2.5. Determining the relationship between Murder incidents and Land use

Finally the scatter graphs are performed for murder incidents to the same land use and landmark types (Figure 5.5). These graphs show positive correlation in commercial areas, roads, parks and police stations, car parks and hospitals. For residential areas, schools and subway stops the correlations have negative direction. In other words while murder incidents increase in commercial areas, roads, parks and places close to the subway stops and schools; they decrease in residential areas and in places near to the police stations, car parks and hospitals.

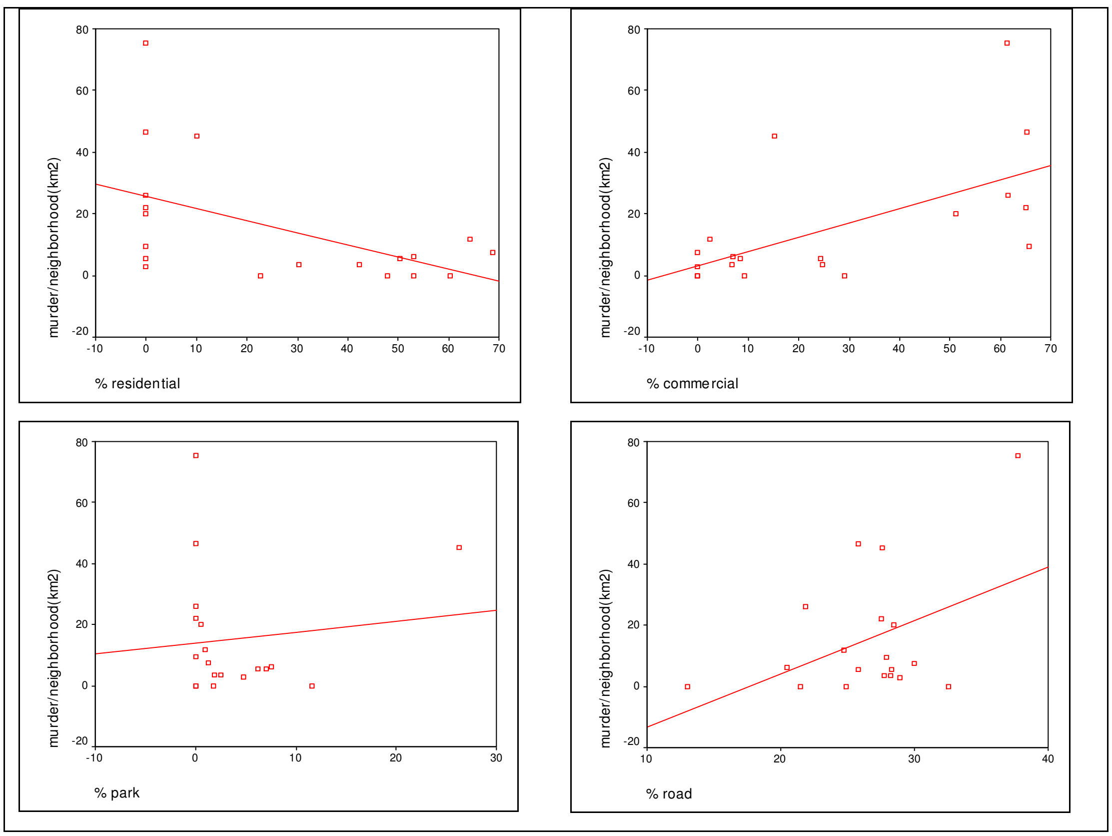


Figure 5.5: Scatter Graphs for Murder between Land Use and Landmarks

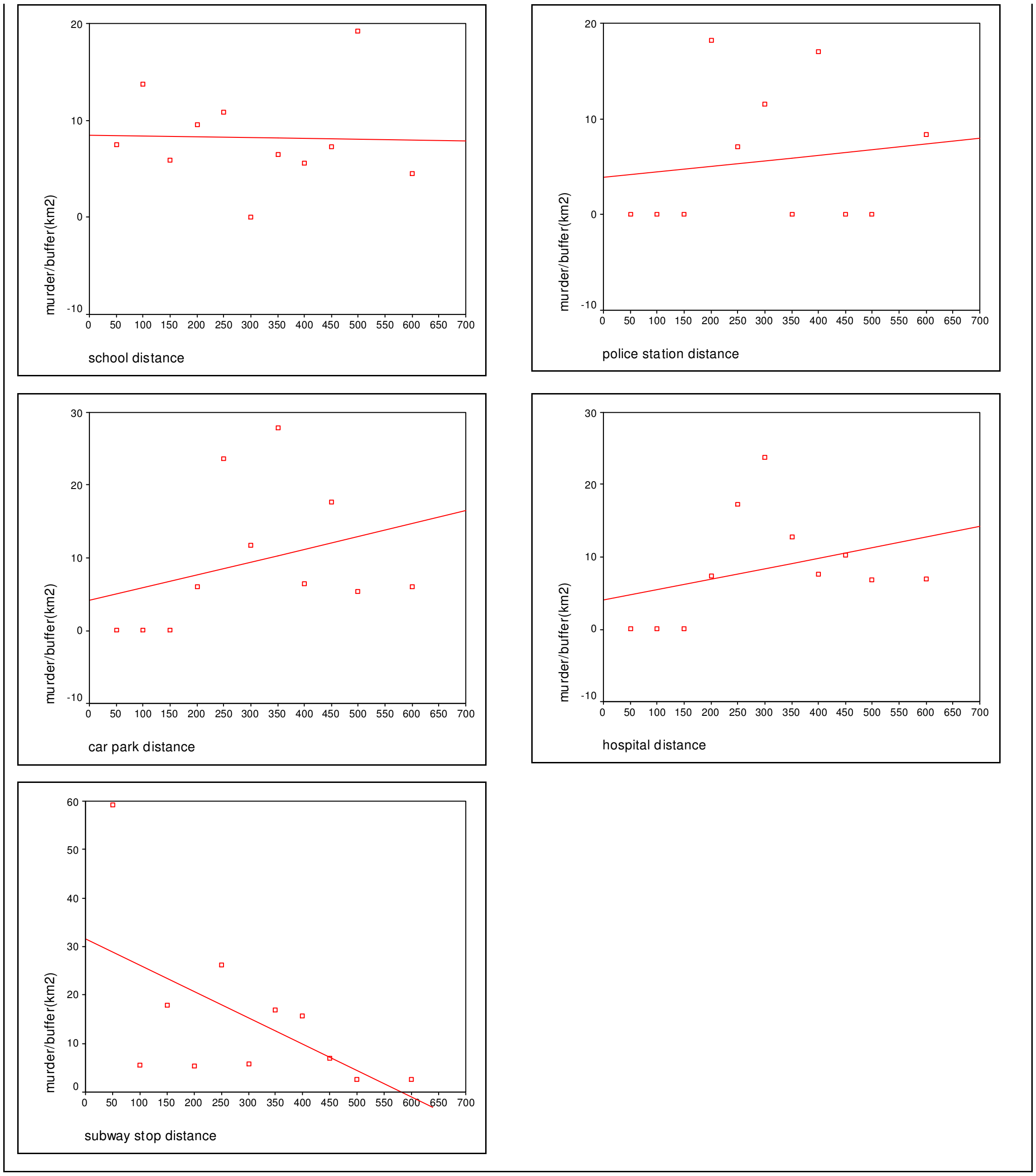


Figure 5.5: Scatter Graphs for Murder between Land Use and Landmarks (cont'd)

Similarly, as seen in the below table (Table 5.12) the correlation coefficient values are calculated for finding the strength of the relationship between the variables.

Table 5.12: Correlation Coefficient Results for Murder Incidents

School	- 0.008
Park	0.111
Police Station	0.135
Carpark	0.343
Hospital	0.361
Road	0.440
Residential	- 0.523
Subway	- 0.577
Commercial	0.601

The correlation coefficient of 0.601 in commercial areas denotes that the rate of the murder incidents is higher than residential areas of which the coefficient value is -0.523. The bars, restaurants with alcohol, located in commercial areas may be potential for the occurrence of murder incidents especially in the nights. As a result of the arguments or brawls which are generally started after taking too much alcohol in these places people who have knife or guns may injure the others, and this events sometimes finish with murder. In addition most of the murder incidents happened because of the disagreements on materialistic things, seen in commercial areas.

Similar to usurp, subway stops are risky areas for murder incidents since these places are more desolated. This result is confirmed by the coefficient of -0.577 which indicates that the rate of the murder incident increase in places which locate near to the subway stops.

For roads which give 0.440 coefficient value it can be understood that these areas are also potential areas for murder incidents. However the calculations performed for major roads and minor roads specify that minor roads with 0.111 are more in correlation than major roads with 0.047 coefficient. As for the places far away from hospitals and car parks of which coefficient is obtained as 0.361 for the initial, and 0.343 is obtained for the other, the rate of the murder incidents increase.

For parks and police stations there is no any noteworthy correlation with murder incidents. In addition it can be easily said that the correlation between schools and murder incidents do not exist.

5.3. Determining the Risky Areas according to the Correlation Results

As the results of the correlation calculations, the strength and the direction of the relationship between incidents and land use and also landmarks are obtained. By using these correlation coefficient values it can be possible to establish the potential areas for the each incident type to occur. Instead of taking into account the amount and the location of the incidents, the risky areas are determined by the existence of the related land use type and the distance to the related landmarks. The information of the effects of land use and landmark types on the occurrence of the incidents are obtained in previous analysis including the calculation of correlation coefficient.

For this purpose convert features to raster, straight line, reclassify and raster calculation which are the tools of ArcGIS 8.2 are used. Straight line distance contains the measured distance from every cell to the nearest source. It shows how to calculate how far each cell is from the source. Reclassifying means replacing input cell values with new output cell values based on new information. Raster calculation is used to weight the rasters by mathematical calculations.

For residential, commercial, park areas and roads which are represented by polygon features firstly the polygons are converted into raster format. Than each pixel value is reclassified according to their correlation results between the related incidents.

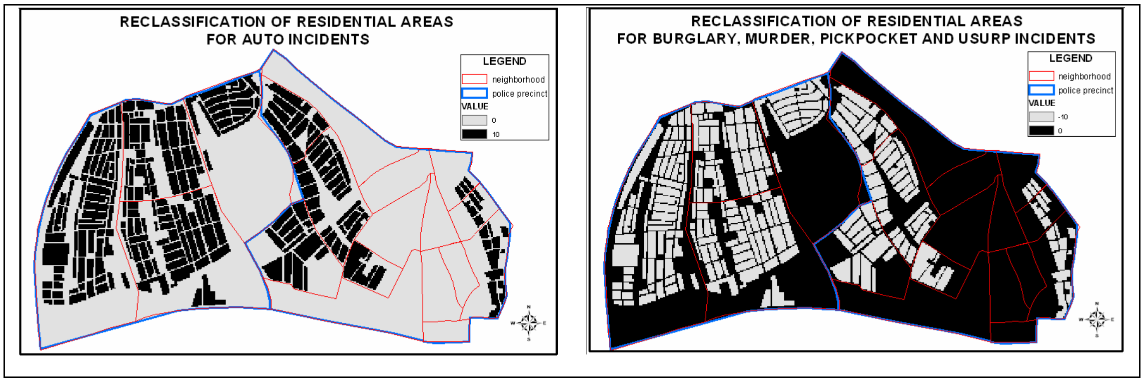
Although some of the correlations between land features and incidents do not exist so strongly, almost all features are reclassified. The reclassification values are given by taking the direction of the correlation into consideration (Table 5.13). However the weights which are given in raster calculation as a final analysis reflect the strength of the correlation.

Table 5.13: The Directions of the Correlations

		Burglary	Auto	Pick Pocket	Usurp	Murder
Land use	Residential	-	+	-	-	-
	Commercial	+	-	+	+	+
	Park	+	+	-	-	+
	Minor Roads	+	+	+	+	+
Landmark	Schools	-	-	+	+	NA
	Police Stations	-	+	-	-	+
	Car Parks	NA	-	-	+	+
	Hospitals	-	+	+	-	+
	Subways	-	-	-	-	-

In the table, the sign (+) for land use areas indicates that these areas are more prone to the incidents while the sign (-) for land use areas denotes that the rate of the incidents are lower in these areas. As for landmarks the sign (-) shows that the rate of the incidents declines in the areas far from that landmarks and denotes that closer areas are more risky. Moreover (+) sign for landmarks shows that the areas far from that landmarks are more risky than the areas closer to that landmark. On the other hand NA (Not Applicable) specifies that there is no any correlation between that incident and that land feature.

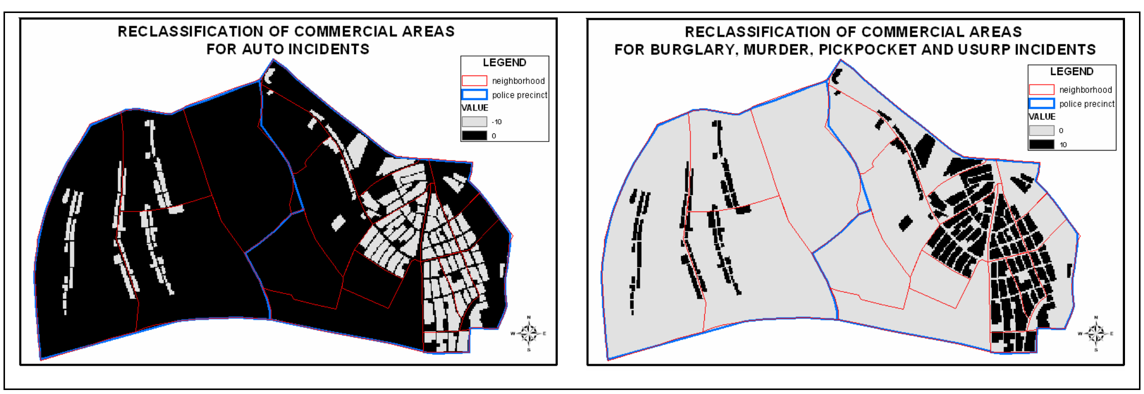
For residential areas it is obtained that these areas are attractive for the criminals who create auto incidents, but in contrast these areas are not preferred by the criminals of burglary, murder, pick pocket and usurp. So, in order to realize the reclassification the value is given 10 to the residential areas and 0 to the others, if there is a correlation in positive direction which is seen in auto incidents. On the other hand the value is given -10 to the residential areas and 0 to the others, if there is a correlation in negative direction which is seen in burglary, murder, pick pocket and usurp incidents (Map 5.2).



Map 5.2: Reclassification of Residential Areas

In above map while left part summarizes that residential areas which are shown in dark color are attractive for auto incidents, right part indicates that these areas which are shown in light color are not suitable for burglary, murder, pick pocket and usurp incidents.

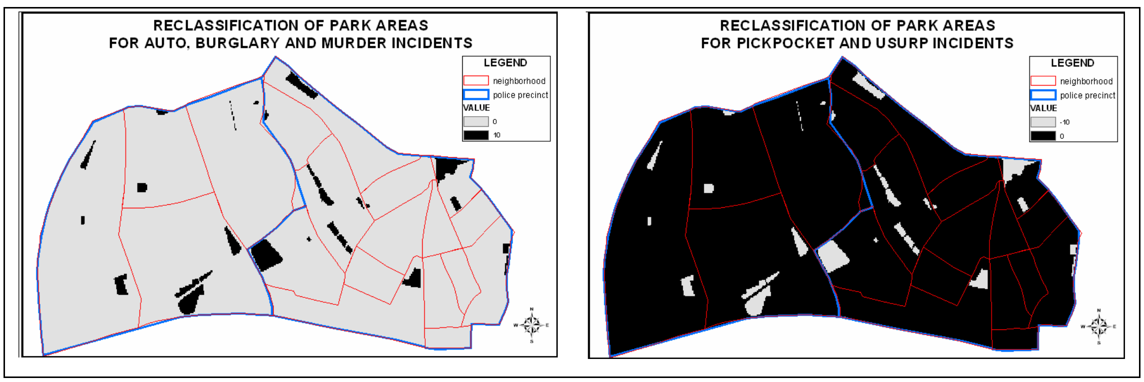
Naturally the representation of the reclassification for commercial areas is in the reverse way in respect of residential areas. Since the coefficient values denote that commercial areas are inappropriate for auto incidents but they are suitable for burglary, pick pocket, murder and usurp incidents, these areas are reclassified as -10 for auto incidents and 10 for the other incidents (Map 5.3).



Map 5.3: Reclassification of Commercial Areas

Left image indicates that commercial areas which are shown in light color have negative effects on auto incidents while they have positive effects on burglary, pick pocket, murder and usurp incidents as seen in right part with dark color representation.

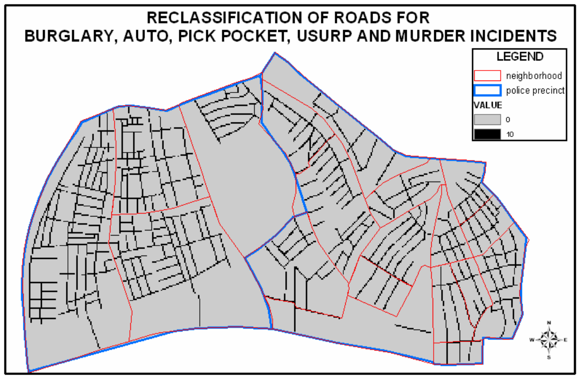
The results of the correlation coefficient calculations denote that the rate of the auto, burglary and murder incidents increase in park areas while pick pocket and usurp incidents decrease. In this case 10 is given to the parks for auto, burglary and murder incidents for the reclassification process, while -10 is given to the parks for pick pocket and usurp incidents (Map 5.4).



Map 5.4: Reclassification of Park Areas

As seen in the left image, dark areas, representing the park areas, are potential for auto, burglary and murder incidents. But on the contrary they are not attractive for pick pocket and usurp incidents as shown in the right image in which parks are represented in light color.

It is obtained in coefficient calculations that there is a correlation in positive direction between roads and the incidents. In addition both five incident types are higher in minor roads than major roads. That means minor roads are more preferable and more appropriate for the incidents to occur. Thus in order to reclassify the roads feature 10 is given to the minor roads and 0 is given to the other areas.

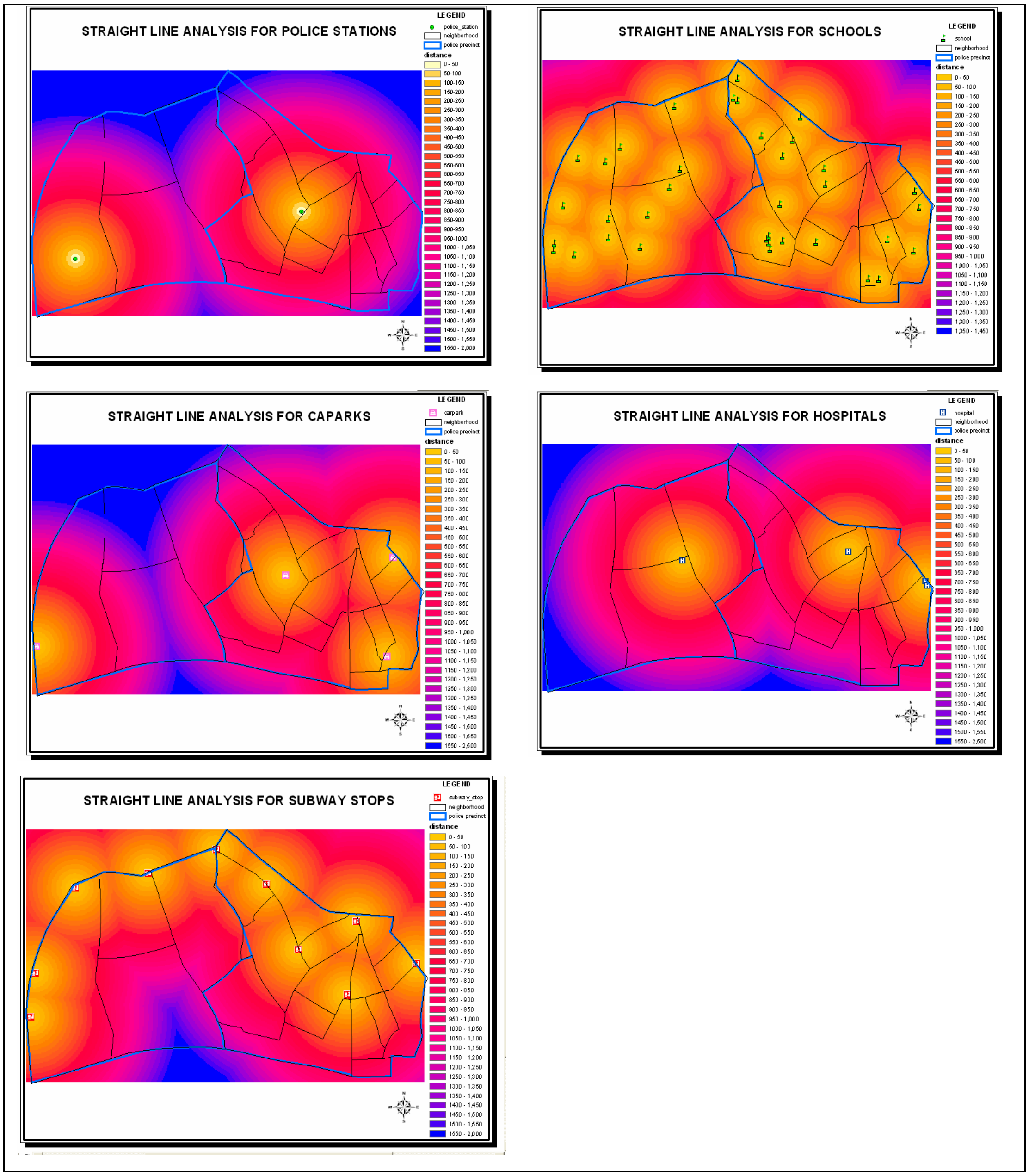


Map 5.5: Reclassification of Roads

As seen in the figure, dark areas with high values which represent the minor roads indicate that these areas are potential for the occurrence of both five incident types.

For schools, police stations, car parks, hospitals and subways which are represented by point features, straight line is applied based on the distance from the related point feature.

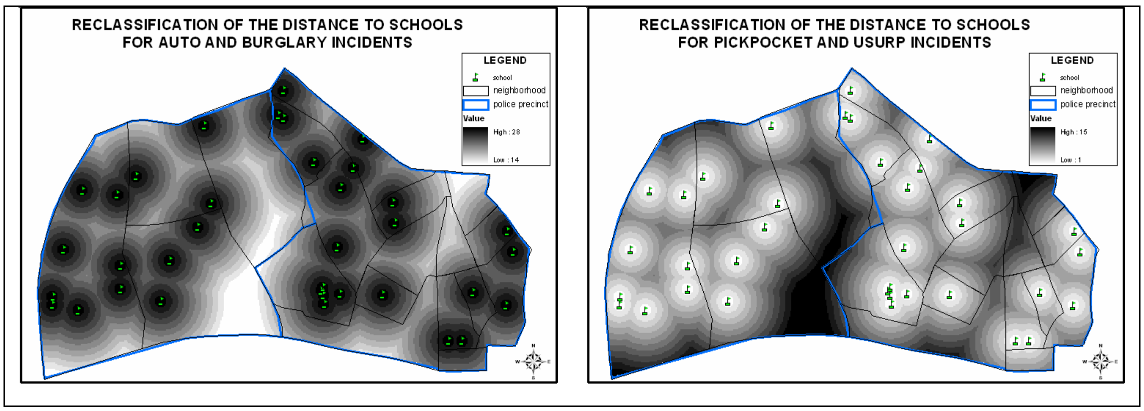
For both landmark features the range of the straight line of which the representation is similar with buffer analysis but it is different in format (straight line provides the results in raster format) is adjusted according to the distance from the related point. Thus a distance value is given to every pixel (Map 5.6).



Map 5.6: Straight Line for Point Features

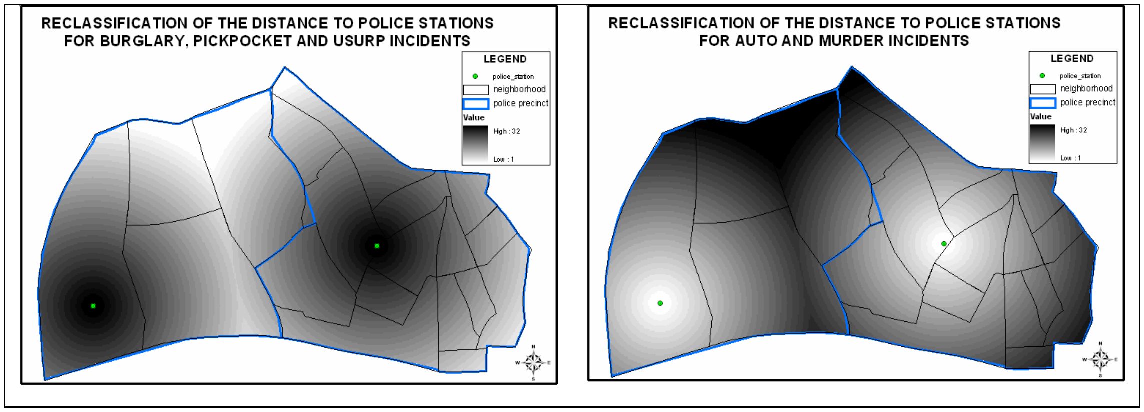
After forming straight line these areas are reclassified according to the effects of their distances on the occurrence of the incidents. As a result of the coefficient calculations the correlation information is obtained for each landmark with the incidents. For reclassification high values are given to risky distances which are shown in darker colors, and low values are given to safe areas which are shown in lighter colors.

Since it is achieved that the rate of the auto and burglary incidents increase in the places near to the schools, the distances closer to the schools are given higher values. In the mean time the places far from schools in which auto and burglary incidents decrease are given lower values. On the contrary the places near to the schools are given lower values for pick pocket and usurp incidents because these incidents are higher in the places far from schools (Map 5.7). However distances to the schools are not reclassified for murder incidents since it is noticed that there is no any correlation between schools and murder.



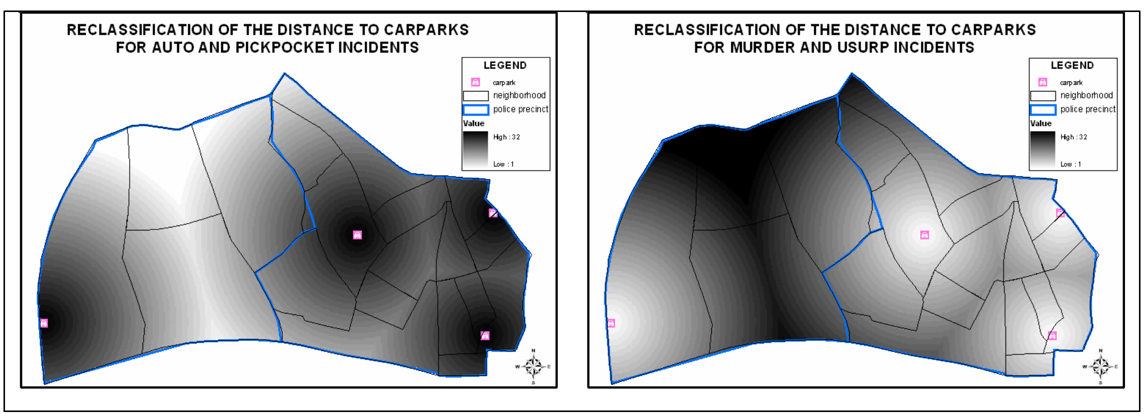
Map 5.7: Reclassification of the Distances to Schools

For police stations it is determined that burglary, pick pocket and usurp incidents occur in the places near to the police stations. So the areas far from police stations are given lower values while the areas closer to the police stations are given higher values for burglary, pick pocket and usurp incidents. As for auto and murder incidents, since places near to the police stations includes fewer auto and murder incidents the areas closer to the police station are given lower values (Map 5.8).



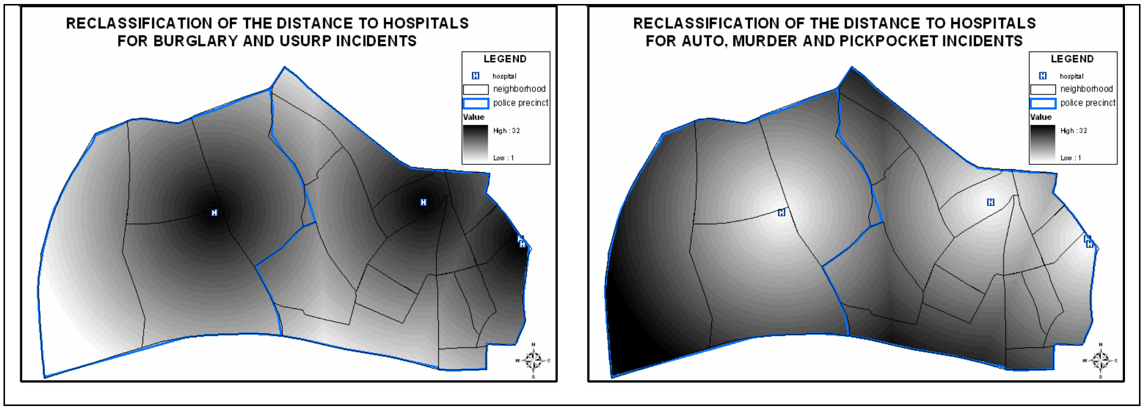
Map 5.8: Reclassification of the Distances to Police Stations

The areas near to the car parks, given higher values, which are represented by darker colors, indicate that these areas are more risky than the areas far from car parks for auto and pick pocket incidents. However for murder and usurp incidents areas closer to the car parks are given lower values since these areas are inappropriate for the occurrence of these incidents (Map 5.9). As for burglary incidents the distances to the car parks are not reclassified since there is no any correlation detected.



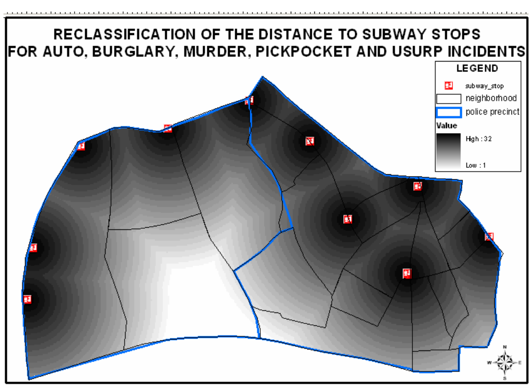
Map 5.9: Reclassification of the Distances to Car Parks

For the reclassification of the distances of hospitals, the areas that are closer to these landmarks are given higher values and shown in dark colors for burglary and usurp incidents. On the contrary the areas far from hospitals are given higher values for auto, murder and pick pocket incidents which indicate that these areas are more prone to these incidents (Map 5.10).



Map 5.10: Reclassification of the distances to Hospitals

Lastly the distances to the subway stops are reclassified according to the direction of the correlation between the incidents and the subway stops. Both five incidents show increasing in the areas near to the subway stops. Thus while the areas closer to the subway stops are given higher values and shown in dark colors, the areas far from subway stops are given lower values and shown in lighter colors (Map 5.11).



Map 5.11: Reclassification of the Distances to Subway Stops

At this point it is possible to analyze the risky areas for the incidents. This is performed by raster calculation of which the weights are given according the coefficient values which denote the strength of the correlation. Below table shows the given weights which are determined based on the importance of the variables (land use, landmarks) - on the occurrence of the incidents. Greater weights indicate that these variables have strong effects on the incidents (Table 5.14). Weights are obtained by calculating the percentage of the coefficient values.

Table 5.14: Correlation Coefficient Values and Weights

Land Use	Coefficient	Weight
BURGLARY		
Park	0.074	**0.023**
Police Station	(-) 0.296	**0.091**
Hospital	(-) 0.308	**0.094**
Residential	(-) 0.377	**0.115**
Minor Road	0.438	**0.134**
School	(-) 0.455	**0.139**
Commercial	0.573	**0.175**
Subway	(-) 0.746	**0.228**
		1.000
PICKPOCKET		
Carpark	(-) 0.100	**0.035**
Park	(-) 0.107	**0.038**
Hospital	0.150	**0.074**
Police Station	(-) 0.211	**0.081**
School	0.229	**0.086**
Minor Road	0.245	**0.139**
Subway	(-) 0.539	**0.190**
Residential	(-) 0.552	**0.194**
Commercial	0.711	**0.250**
		1.000
MURDER		
Park	0.111	**0.040**
Minor Road	0.111	**0.040**
Police Station	0.135	**0.049**
Carpark	0.343	**0.124**
Hospital	0.361	**0.131**
Residential	(-) 0.523	**0.189**
Subway	(-) 0.577	**0.209**
Commercial	0.601	**0.218**
		1.000

Land Use	Coefficient	Weight
AUTO		
Commercial	(-) 0.056	**0.024**
Residential	0.069	**0.030**
Police Station	0.186	**0.081**
Hospital	0.239	**0.104**
Subway	(-) 0.253	**0.110**
Minor Road	0.285	**0.124**
Carpark	(-) 0.391	**0.171**
Park	0.396	**0.173**
School	(-) 0.416	**0.182**
		1.000
USURP		
School	0.075	**0.022**
Park	(-) 0.235	**0.068**
Minor Road	0.311	**0.090**
Hospital	(-) 0.327	**0.095**
Residential	(-) 0.373	**0.108**
Carpark	0.424	**0.123**
Commercial	0.529	**0.153**
Police Station	(-) 0.533	**0.154**
Subway	(-) 0.645	**0.187**
		1.000

As a consequence of the computations of the weights the effects of the variables on the occurrence of the incidents or the relations between these variables and the incidents may be recognized easily.

In order to test the straightness of the determined relationships between these variables and the incidents, the risky areas are determined by raster calculation based on the location and the size of the land use and landmarks without considering the location of the incidents. For this purpose each pixel which has a new value, given in the reclassification process, is multiplied with a weight which is determined in Table. Below figure (Figure 5.6) gives an example for raster calculation which is done for burglary incidents. Then these areas are compared with the hotspot areas which are obtained from kernel estimation, generated in Chapter IV (Map 5.12).

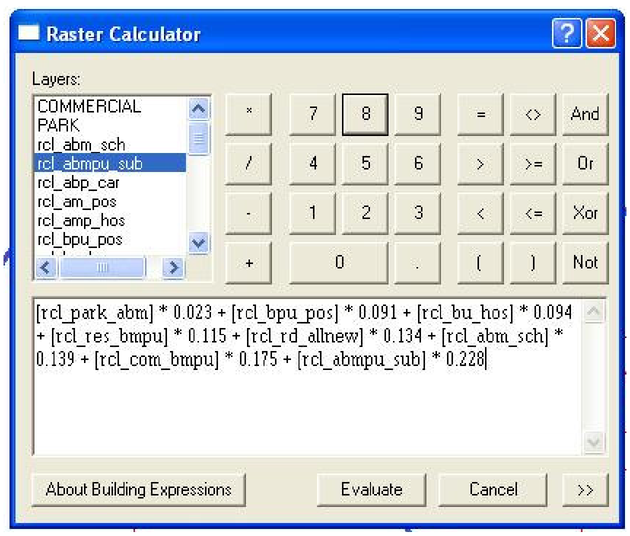
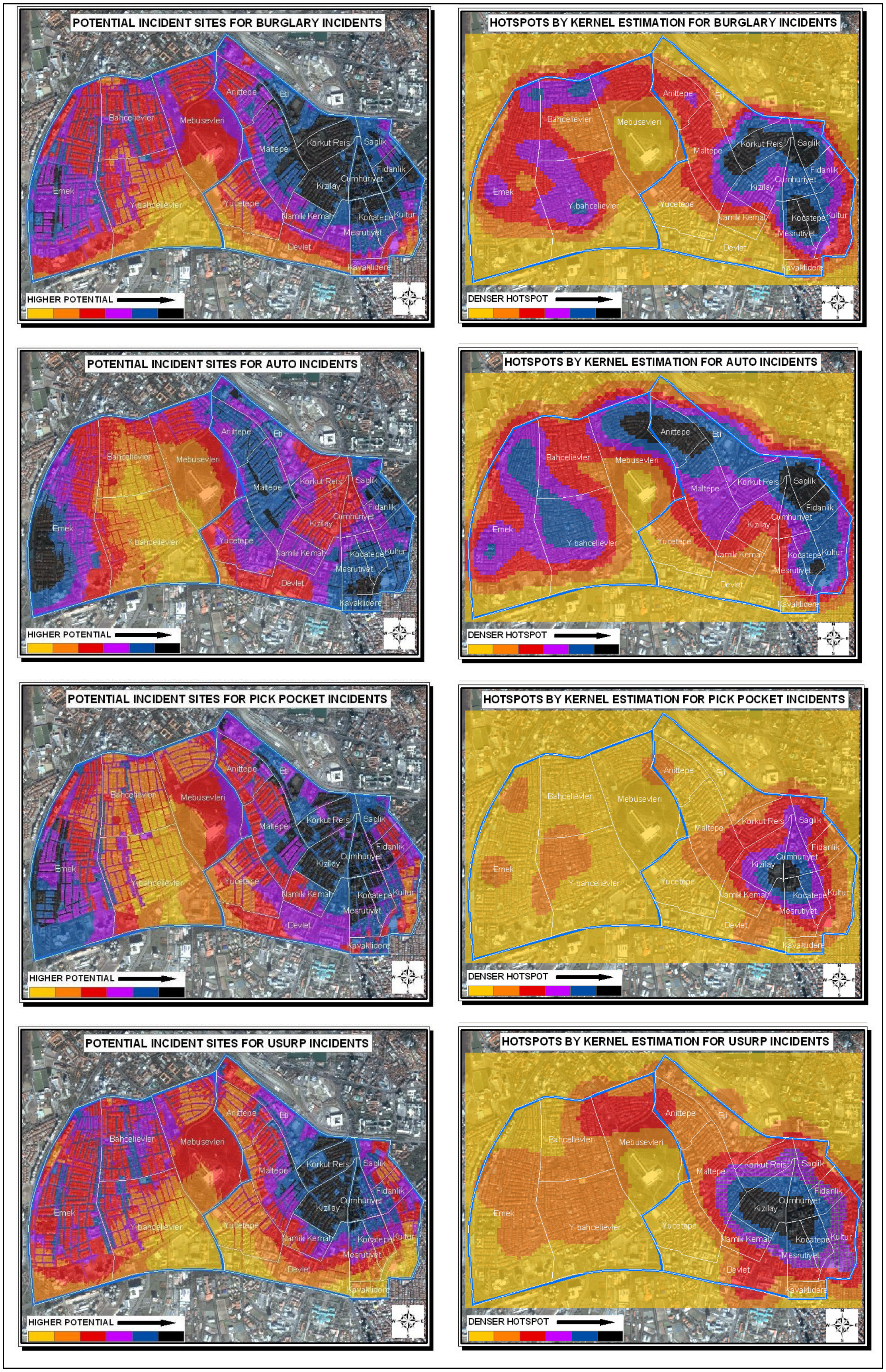
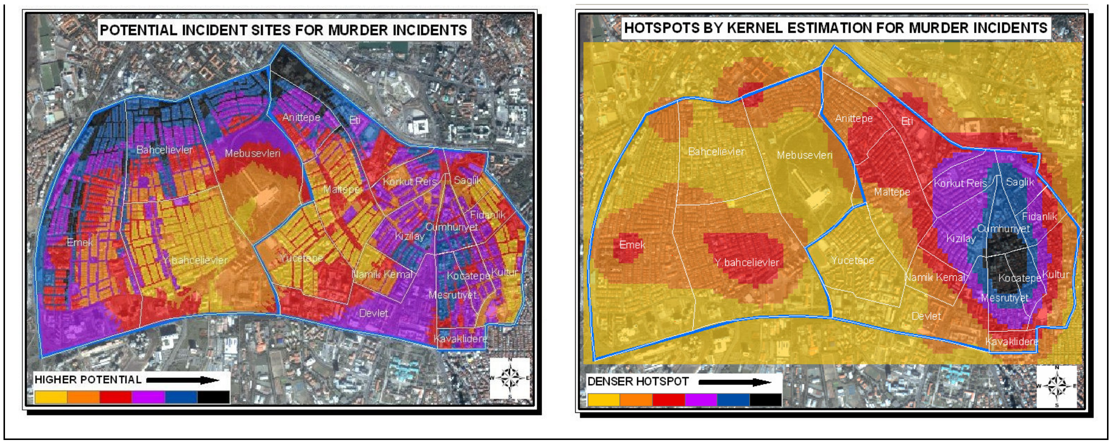


Figure 5.6: Raster Calculation for Burglary Incidents



Map 5.12: Comparison of Potential Crime Areas with Hotspot Areas



Map 5.12: Comparison of Potential Crime Areas with Hotspot Areas (cont'd)

Left column of the figure shows the areas which are more potential for the occurrence of the incidents. In the result of the raster calculation dark areas in the left column denote the best places for the offenders to commit a crime. On the other hand right column of the figure indicates the hotspot areas in which most of the incidents happened in the year of 2003.

In the case of comparing these two columns - the potential risky areas, acquired by raster calculation, with the hot spot areas, acquired by kernel estimation - similar results can be detected. Surely, completely same results could not be expected since one analysis provides the probability of the occurrence; the other one presents the existing structure. The results of the hotspot analysis naturally point specific locations which include happened incidents, while the results of the raster calculation evaluate every point – pixel of the area according to the possibility of an occurrence of an incident. The darker areas in the left column show the areas in which there is a possibility of an incident to occur. Hence the darker areas in the left column cover larger areas than the darker areas in the right column.

Especially for burglary, auto and usurp incidents, the results of both raster calculation and hotspot areas are closer to each other. Thus it can be said that the responsibility of the selected land use and landmark variables on these incidents for which the weights are determined by correlation calculations are partly more accurate to estimate the risky areas. Furthermore it can be possible to say that if there is a place with similar land use and also similar socio-economic structure with the study area, by applying such processes , it is probable to determine the potential crime areas without incident data.

5.4. Evaluation of the Outputs

In this chapter, the relationship between land features, which are detected in Chapter IV, and five incident types are tried to be determined. The land use variables are grouped into four classes; residential, commercial, park areas and roads (as major and minor roads). On the other hand land mark types are grouped into five classes; schools, subway stops, hospitals, carparks and police stations.

The strongest relation is distinguished in commercial areas with the exception of auto incidents, as predictable. Since offenders travel shorter distances and prefer the places where quick escape is possible, it is found that the subway stops have strong relations with all incidents especially burglary and usurp. Thus it can be said that commercial areas and the places closer to the subway stops are the most attractive places for the offenders of burglary, pick pocket, usurp and murder.

Regarding to the results, if commercial areas are more preferred it is unsurprising that residential areas have pushing effects on the occurrence of the incidents. They have a relation only with the auto incidents which is not so strong.

As an expected result minor roads have a remarkable relation with the all five incidents. Similar to subway stops they may provide an easy way to escape rapidly. Also victims are partly lonelier in the streets and streets are more desolated and uncontrolled comparing to the major roads. Thus most of the offenders prefer places near to the minor roads.

Since the incident data, used in this thesis, do not include the crime types such as artificial foodstuffs or drugs which are the most offenses seen around schools, the strong relation between incidents and schools could not be obtained. However the results provide information that schools have a relation with burglary and auto incidents. Although this is not an expected result, it may be explained with the properties of the surrounding neighborhoods of these schools such as income or social level. In addition with a detail research it may be found that the students in these schools have a tendency to realize a crime, may be burglary or auto incidents. In other words, there may be some peculiar results which are only seen in this area but are not the scope of this thesis.

On the other hand car parks have a relation with the incidents of auto and pick pocket though they are not so strong. This relation is not surprising but the direction of the relation demonstrates that car parks in this study area may be uncontrolled and distrustful.

In spite of the weak correlation between parks and the incidents, a partly relation is seen with auto, murder and burglary incidents. Similarly hospitals have not a strong relation with any incidents. Only burglary and usurp have a relation with hospitals but not remarkable.

However the most surprising result is obtained between the police stations and the incidents. There is not any strong correlation between them but the results indicate that usurp, burglary and pick pocket incidents occur in the places closer to the police stations. Whereas it is expected that police stations are dissuasive function on the occurrence of the incidents. But in this study area it is noticed that police stations have not important effects on the incidents.

By referring to these relationships, the potential areas for the five incident types are tried to be found by raster calculation for which the required weights are determined based on the correlation results.

As a result of the raster calculation some parts of the study area are detected as the risky areas for the incidents. That means, these areas may require more security control, precautions and maybe planning in the first priority.

After determining these risky areas, in order to understand if such a way which is realized in this study could be applied in another study area, a comparison with the hotspot areas are achieved. Some similarities between the distribution of the incidents and risky areas are detected in this comparison.

Of course, this way could not be the unique or the best method to determine the risky areas, since there must be much more criterions that have an effect on the incidents. But it can be said that using such this way with adding more variables provide better and more realistic information about the risk of the land use in an area.

CHAPTER V

CONCLUSION AND RECOMMENDATIONS

This study aimed to apply GIS to crime incident analysis and to present and evaluate the contributions of GIS in order to understand the status of incidents' pattern and also to predict the possible occurrence of the future incidents based on the structure of the land use.

The study consists of four stages to realize this aim. Firstly the status of the occurred incidents is identified based on their rate, occurrence time and place, by descriptive analysis. Additionally, the structure of the study area is examined. Secondly, several incident pattern methods are carried out within the study area by making use of GIS in order to provide a comprehensive profile about the pattern of the incidents in the area. The best method can be found based on the requirements of the user. In the study, for the representation of the incident pattern over the study area kernel estimation method is chosen. As for the determination of land use and landmarks inside the hotspot areas Nearest Neighbor Hierarchical Clustering method is preferred.

In the third stage, the relationship between incidents the land uses in the study area is determined and it is found that incidents generally prefer commercial areas and places near to the subway stops and streets. In the last stage, correlation based parameters are integrated into GIS to generate the potential areas for the incidents. Than the results are compared with the determined hotspot areas in order to discover the usage of this way in some other areas which have a similar land use structure.

Main findings of the study can be classified into two themes. The first one is on the contributions of GIS-based methods in incident pattern analysis which provides some clustered areas include more incidents. Several different visualization techniques for visualizing spatial incident data on a GIS are presented in this first theme. The second one is on the advantages of using GIS integrated correlation results for the relationship between incident locations and land use.

In the thesis, it is observed that the abilities of GIS such as managing data, organizing in different spatial units, providing quick and objective results and linking the spatial and other kind of information satisfy the better results for incident analysis. The findings of this study showed that GIS is compatible with many statistical tools which are generated for incident analysis. Crime analysts can make greater use of GIS to analyze and display geographic concentrations of the crime events.

By integration capability of GIS with statistical tools realistic results are obtained which allows police personnel to plan effectively for emergency response, determine mitigation priorities, analyze historical events, and predict future events. According to the results of incident area analysis urban policy priority areas can be determined by the police. Thus, finding out the areas which are required emergency intervene saves the individuals or values before being damaged.

Study of incident trends in some specific areas like commercial areas or places near to the subway stops would be guiding pole for crime prevention methods. By analyzing those factors it can be possible to provide accurate and pertinent information relative to crime patterns and correlations to aid in the investigative process of crime analysis.

Certainly the significance on the occurrence of the incidents of the social, economic and demographic structure such as population (day and night) of the study area, age, income level, household, sex, race, employment and educated level of both victims and offenders could not be discussed. But the thesis deals only with the structure of the land use, since it is more concrete and it is more possible to utilize it than the social facts. Most importantly, to acquire the land use data is relatively easier than to obtain social, economic and demographic data.

Although GIS is a useful tool in the investigation of spatial patterns of crime, it is inevitably limited by the amount of spatial data that are available. In the process of collecting the necessary data the most important problem that met in the thesis is to get the incident data. It is a very difficult process to receive the required data from the related institutions in Turkey. Especially police would not like to provide the incident data to any outsource since they keep them in a secret and think that if the information of where the incidents take place is published, personal privacy may be damaged or they do not want to worry inhabitants about the safety of their neighborhood or they do not want people to doubt about their performance or the efficiency of their work.

Furthermore, the information of the incidents is only related to the happened but only reported incidents. That means the amount of the incidents may not reflect the existing status since there may be more incidents happened but not reported because of the various results which are tried to be explained in Chapter III. In addition, the definite rate of the victimization is not known. For this problem in order to find the amount of wronged people, a victimization survey in which questionnaires are prepared and ask people if they are exposed to any incidence should be realized.

In addition for land use data it is also not possible to obtain detail information about the usage of the functions except site surveys which is done for this study. For instance if the information of the number and locations of the restaurants or cafes with alcohol or the ratio of the commerce and residence usage of a building or the type of the commercial usage such as shop, market, stadium, factory, restaurant, bar or café are received the results will be more meaningful, accurate and realistic.

Thus in the process of determining the relationship between incidents and land use, the selection of a criteria which influence the extent and the nature of the incidents is dependent on the limited data. More comprehensive data collection on the site related to the land use will certainly bring new parameters.

It is observed that giving similar weights to the land use and landmark functions in another study area gives information about potential crime areas. However the study area should include similar usages, similar structure and also similar social characteristics. In order to test the accuracy of these weights it will be more realistic if these weights are applied in another study area. Than it will be more than appropriated if the results are compared with the incident data of that study area. But since it is hard to obtain the incident data for another area, it was not possible to realize this second test in this thesis.

The information of patrol routing or the locations of cameras are also necessary and useful information. Thus it is possible to integrate these preventive functions with the existing status and find out the uncontrolled areas and test the impression of these functions on the occurrence of the incidents. If it is detected that the hotspot areas are located where the cameras are or what else, it will be understood that these precautions are not the best ways to reduce the rate of the incidents.

In addition to the police, many organizations can influence the extent of the incidents. Professional Organizations such as Chamber of City Planners, Chamber of Architectures and Chamber of Engineers come into picture when one is talking about physical environment within which crime takes place. Both short and long-term measures can be taken by these organizations which will reduce the likelihood of crime and the fear of crime. For instance highway engineers may improve street lightning in a particularly vulnerable area, or the planners might design a residential layout so as to avoid the creation of crime-prone areas.

City planners may plan their cities with not only considering the direction of the cities' growing but also the function of crime. When they determine the location and the usage of a facility they may take into account its security and also its effect on the safety of its near surroundings. For instance, for the neighborhood of Emek in the study area which is covered with mostly residential areas the city terminal station (AŞTİ) is located very close. The most important result for the occurrence of burglary and auto incidents in this neighborhood is the terminal. Offenders may come to Ankara for a day and in order to realize their offense they prefer places closer to the facility which helps them to escape easily and rapidly. Thus it can be said that terminals which have irregular population in a day should be located far from inhabited areas. Similarly places near to the subway stops are also risky but it is not possible to locate them far from inhabited areas since it may loose its function. But it is possible to take more precautions in the places near to the subway stops.

While building police beats the communities or neighborhoods are generally divided according to boundaries of neighborhoods or roadways. But the level of service needed, number of incidents and also the land use are more important factors that should be looked at while building the police beats. In addition dividing the beats based on only demographics the resources may be allocated improperly. Because there may be several areas where there are thousand inhabitants but they generate few calls for service, while there are less people that require much greater levels of service. For example commercial areas do not have any inhabitants especially at the nights but they are more prone to incidents. As a utopian scheme the dense commercial areas, based on its safety and size, may be assigned as one beat.

Further studies may focus directly on physical design components of crime prevention. The physical characteristics of an area can influence the behavior of both residents and potential offenders. Increase lighting, use of surveillance equipment, access control, alarms and other physical changes are intended to bring about greater social cohesion, citizen concern and involvement, and ultimately, reduced crime and fear of crime.

In addition, in the further studies the rate of the incidents may be examined in certain periods which are determined by the time of taking a precaution. For example; in the potential target areas such as, commercial areas and places near to subway stops which are determined in this thesis, patrolling routes may be designed again and the rate of the incidents are examined after three months. Thus it will be possible to test the impression of this precaution.

Considering the outcomes of this study it can be stated that the significance of adopting an interagency approach to crime prevention and fostering the development of co-ordination and co-operation among organizations are important for reducing crime. All organizations will not necessarily be relevant to every crime prevention initiative, but there are times when man will have a part to play. Looking first at the statutory bodies, most lists of organizations with an interest in crime prevention would include police, probation and social services, and education, planning, highways, housing and architects' organizations. For this co-operation, utilizing the GIS-based methods for the data suppliers and decision makers is unavoidable.

REFERENCES

Ahmadi, M. (2003) **Crime Mapping and Spatial Analysis**, Thesis Submitted to the International Institute for Geo-information Science and Earth Observation, Enschade, The Netherlands.

Ankara Emniyet Müdürlüğü (2002) **Ankara Geneli Cinayet, Gasp, Hırsızlık ve Yankesicilik-Dolandırıcılık Bürolarının 2002 Aralık Ayı Verileri**.

Bailey, T.C. and Gatrell, A.C. (1995) **Interactive Spatial Data Analysis**, Addison Wesley Longman Limited, England.

Block, C.R. (1998) "The Geoarchive: An Information Foundation for Community Policing" in D. Weisburd and T. McEwen (eds.), **Crime Mapping and Crime Prevention,** , Criminal Justice Press, Monsey, New York, USA, 27-81.

Block, R.L. and Block, C.R. (1995) "Space, Place and Crime: Hot Spot Areas and Hot Places of Liquor – Related Crime", in J.E. Eck and D. Weisburd (eds.), **Crime and Place**, Monsey, NY: Criminal Justice Press.

Boba, R. (2001) **Introductory Guide to Crime Analysis and Mapping**, Unpublished Report to the Office of Community Oriented Policing Services, US Department of Justice.

Boggs, S.L. (1966) "Urban Crime Patterns", **American Sociological Review**, 30:899-908.

Bowers, K. and Hirschfield, A. (1999) "Exploring Links Between Crime and Disadvantage in North-west England: An Analysis Using Geographical Information Systems", **International Journal of Geographical Information Science**, Vol. 13, No.2, 159-184.

Brantingham, P.J. and Faust, F.L. (1976) "A conceptual model of crime prevention", **Crime and Delinquency**, 22:284-296.

Brantingham, P.J. and Brantingham P.L. (1991) "Introduction: The Dimensions of Crime" in P.J. Brantingham and P.L. Brantingham (eds.), **Environmental Criminology**, Waveland Press, Inc., Illinois.

Canter, P.R. (1998) "Geographic Information Systems and Crime Analysis in Baltimore Country, Maryland" in D. Weisburd and T. McEwen (eds.), **Crime Mapping and Crime Prevention,** Criminal Justice Press, Monsey, New York, USA, 157-190.

Canter, P. (2000) "Using Geographic Information System for Tactical Crime Analysis", in V. Goldsmith, P.G. McGuire, J.H. Mollenkopf and T.A. Ross (eds.), **Analyzing Crime Patterns**, Sage Publications, Inc., United States of America.

Chakravorty, S. and Pelfrey, Jr. W.V. (2000) "Exploratory Data Analysis of Crime Patterns", in V. Goldsmith, P.G. McGuire, J.H. Mollenkopf and T.A. Ross (eds.), **Analyzing Crime Patterns**, Sage Publications, Inc., United States of America.

Çankaya Belediyesi (2001) **Çankaya Kent Sağlık Profili ve Sağlıklı Kent Göstergeleri**.

Çankaya Belediyesi (2002) **Kent Sağlık ve Gelişim Planı (2002-2007)**.

Devlet İstatistik Enstitüsü Başkanlığı (2000) **2000 Genel Nüfus Sayımı – Nüfusun Sosyal ve Ekonomik Nitelikleri.**

Harries, K. (1999) **Mapping Crime: Principles and Practice,** Crime Mapping Research Center, National Institute of Justice, Office of Justice Programs, U.S. Department of Justice, Washington DC.

Lab, P.S. (2000) **Crime Prevention: Approaches, Practices and Evaluations**, Anderson Publishing Co., Cincinnati, OH.

Levine, N. (2002) **A Spatial Statistics Program for the Analysis of Crime Incident Locations**, the National Institute of Justice, Washington, DC.

McLafferty, S., Williamson, D. and McGuire P.G. (2000) "Identifying Crime Hot Spots Using Kernel Smoothing", in V. Goldsmith, P.G. McGuire, J.H. Mollenkopf and T.A. Ross (eds.), **Analyzing Crime Patterns**, Sage Publications, Inc., United States of America.

Nelson, A.L., Bromley, R.D.F., Thomas, C.J. (2001) "Identifying Micro-Spatial and Temporal Patterns of Violent Crime and Disorder in the British City Centre" **Applied Geography**, 21, 249–274.

Phillips, P.D. (1972) "A Prologue to the Geography of Crime" **Proceedings of the Association of American Geographers**, 4:86-90.
Pyle, G.F. (1974) **The Spatial Dynamic of Crime**, Research Paper in Department of Geography, University of Chicago, Chicago, No.159.

Ratcliffe, J.H. (2001) "Damned If You Don't, Damned If You Do: Crime Mapping And Its Implications In The Real World", **Policing and Society**, Vol. 12, No. 3, pp. 211-225.

Ratctiffe, J.H. and McCullagh, M.J. (1999) "Hotbeds of crime and the search for spatial accuracy", **Geographical Systems**, 1, 385-398.

Rengert, G.F. (1997) "Auto Theft in Central Philadelphia", in R. Nomel (ed.), **Policing for Prevention: Reducing Crime, Public Intoxication and Injury**, Monsey, NY: Criminal Justice Press.

Vellani, K.H and Nahoun, J.D. (2001) **Applied Crime Analysis**, Butterworth-Heinemann, Woburn, MA.

Walford, N. (1994) **Geographical Data Analysis**, John Wiley and Sons Ltd, England.

Weisburd, D. and McEwen, T. (1998) **Crime Mapping and Crime Prevention**, Crime Prevention Studies, Vol. 8, Criminal Justice Press, Monsey, New York, USA.

Wieczorek, W. and Hanson C.E. (1997) "New Modeling Methods, Geographic Information Systems and Spatial Analysis", **Journal of Alcohol Health and Research World**, Vol. 21 (4), pp. 331-339.

ONLINE REFERENCES

web1:http://www.icjia.state.il.us/public/index.cfm?metaSection=Publications&metapage=AR98_p09, Illinois Criminal Justice Information Authority, “1998 Annual Report” (visited on 02.05.2003)

web2: http://www.icpsr.umich.edu/NACJD/crimestat.html, National Archive of Criminal Justice Data, “Crime Stat Spatial Statistics Program: Version 2.0” (visited on 22.09.2003)

web3:http://www.goldensoftware.com/, Golden Software (visited on 15.05.2003)

web4:http://sal.agecon.uiuc.edu/csiss/pdf/points.pdf, Anselin, L. (2003) “An Introduction to Pattern Analysis Using Crime Stat” (visited on 11.11.2003)

web5:http://die.gov.tr/esg_Il_Bolge.htm (visited on 10.09.2004)

INTERVIEWS

Interview (e-mail) with Dan Helms (2004) GIS and Crime Analysis Specialist, National Law Enforcement and Corrections Technology Center, Rocky Mountain Region.

Interview (e-mail) with Lauren Scott (2004) Pd. D., ESRI Geoprocessing, ESR Inc., California.

Interview with Mesut Demirbilek (2004) Director of Place Investigation of Incident and Determination of Identity Department, General Directorate of Security, Ankara.

Interview (e-mail) with Ned Levine (2004) Urban Information and Policy Evaluation Specialists, The National Institute of Justice, Washington.

Interview with Sevil Atasoy (2004) Prof. Dr., Institute of Forensic Sciences, Istanbul University, Cerrahpasa Campus, Istanbul.

Interview (e-mail) with Spencer Chainey (2004) Director of Geographic Information Science, The Jill Dando Institute of Crime Science University College London, London.