Louisiana State University
LSU Digital Commons

LSU Master's Theses Graduate School

2007

Accuracy assessment of individually calibrated Journey-To-Crime Geographic Profiling models

Tania Pal
Louisiana State University and Agricultural and Mechanical College, tpal1@lsu.edu

Follow this and additional works at: http://digitalcommons.lsu.edu/gradschool_theses

Recommended Citation
Pal, Tania, "Accuracy assessment of individually calibrated Journey-To-Crime Geographic Profiling models" (2007). *LSU Master's Theses*. 309.
http://digitalcommons.lsu.edu/gradschool_theses/309

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gcoste1@lsu.edu.

ACCURACY ASSESSMENT OF INDIVIDUALLY CALIBRATED JOURNEY-TO-CRIME GEOGRAPHIC PROFILING MODELS

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science

in

The Department of Geography and Anthropology

by
Tania Pal
B.S., University of Calcutta, 2001
December 2007

DEDICATION

This thesis is dedicated to my parents, sisters, nephew and my uncle for their unconditional and endless love and support that they shower on me. My family is my strength.

ACKNOWLEDGEMENTS

I would like to take this opportunity to acknowledge the people who immensely contributed to this thesis and without each one of them, this research would not have been possible.

I sincerely thank my major professor, Dr. Michael Leitner for his unending patience, assistance, support, knowledge and encouragement that he provided me at every stage of my thesis. This research would not have been possible without his valuable guidance. I also extend my sincere gratitude and appreciation to Dr. Anthony Lewis and Dr. Lei Wang for accepting my request to be in my committee and always offering their help, support and motivation to complete this thesis.

I am grateful to the LSU START Program and the department of Geography and Anthropology for funding my education and research here at LSU.

My friends in the department of Geography and Anthropology, in other departments, and in ESRI, Redlands, have played a major supportive role in the completion of my thesis. I thank them all.

Last but not the least, special thanks to my wonderful family from the bottom of my heart for their encouragement, support, help and confidence they gave me at each step in accomplishing my goal of earning this masters degree.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES vi

LIST OF FIGURES viii

ABSTRACT ix

CHAPTER 1: INTRODUCTION 1
- 1.1 What is Geographic Profiling? 1
- 1.2 Accuracy Measures for Geographic Profiling 2
- 1.3 Hypothesis and Research Questions 3
- 1.4 Significance of Research Work 3

CHAPTER 2: LITERATURE REVIEW 5
- 2.1 Spatial Analysis of Crime 5
- 2.2 Journey-to-Crime 7
- 2.3 Distance Decay Functions 8
- 2.4 Environmental Criminology Theories Underlying Geographic Profiling 9
  - 2.4.1 Awareness and Activity Spaces 10
  - 2.4.2 Routine Activity Theory 11
  - 2.4.3 Least Effort Principle 11
  - 2.4.4 Rational Choice Theory 12
  - 2.4.5 Buffer Zone 12
- 2.5 History of Geographic Profiling 13

CHAPTER 3: DATA & METHODOLOGY 15
- 3.1 Data & Study Area 15
- 3.2 Methodology 20
- 3.3 Distance Decay Functions 22
  - 3.3.1 Linear 22
  - 3.3.2 Negative Exponential 22
  - 3.3.3 Truncated Negative Exponential 23
  - 3.3.4 Normal 24
  - 3.3.5 Lognormal 25
- 3.4 Journey to Crime Routine 26
- 3.5 Accuracy Assessment Measures 29
  - 3.5.1 Euclidean Distance Error 29
  - 3.5.2 Top Profile Area 30
  - 3.5.3 Hit Score Percentage 31

CHAPTER 4: RESULTS & DISCUSSIONS 32
- 4.1 Results for Auto Theft Serial Offenders 34
- 4.2 Results for Larceny Serial Offenders 36

4.3 Results for Residential Burglary Serial Offenders 39
CHAPTER 5: CONCLUSIONS 43
REFERENCES 48
VITA 52

LIST OF TABLES

Table 1. Crime Series by Crime Type for Baltimore County, 1994 – 1997 18

Table 2. Sample Attribute Table for Property Crime Incidents 19

Table 3. Sample Data Set with Study Area Information of Entire Data Set 19

Table 4. Intercept and Slope Values for Linear Distance Decay Function 22

Table 5. Intercept and Slope Values for Negative Exponential Distance Decay Function........... 23

Table 6. Peak Distance, Peak Likelihood and Exponent Values for Truncated Negative Exponential Distance Decay Function 24

Table 7. Mean, Standard Deviation and Intercept Values for Normal Distance Decay Function 25

Table 8. Mean, Standard Deviation and Intercept Values for Lognormal Distance Decay Function.. 26

Table 9. Auto Theft – Summary of Regression Results.. 35

Table 10. Auto Theft – Maximum, Minimum and Median Values of Accuracy Measures 35

Table 11. Auto Theft – Paired Sample T-Test Results of Euclidean Distance Error................... 36

Table 12. Auto Theft – Paired Sample T-Test Results of Hit Score Percentage 36

Table 13. Auto Theft – Paired Sample T-Test Results of Top Profile Area 36

Table 14. Larceny – Regression Result Summary 37

Table 15. Larceny – Maximum, Minimum and Median Values of Accuracy Measures............. 38

Table 16. Larceny– Paired Sample T-Test Results of Euclidean Distance Error 39

Table 17. Larceny – Paired Sample T-Test Results of Hit Score Percentage............................... 39

Table 18. Larceny – Paired Sample T-Test Results of Search Area 39

Table 19. Residential Burglary – Regression Result Summary 40

Table 20. Residential Burglary – Maximum, Minimum and Median Values of Accuracy Measures.. 41

Table 21. Residential Burglary - Paired Sample T-Test Results of Euclidean Distance Error 42

Table 22. Residential Burglary – Paired Sample T-Test Results of Hit Score Percentage 42

Table 23. Residential Burglary –Paired Sample T-Test Results of Search Area 42

LIST OF FIGURES

Figure 1. Example of a Pin Map of an Area in Baltimore County, Maryland 6

Figure 2. Awareness and Activity Spaces 10

Figure 3. Routine Activity Theory 11

Figure 4. Study Area of Baltimore County, Maryland.. 15

Figure 5. Auto Theft Crime Scenes & Haven Locations Distribution, Baltimore County, 1994 – 1997. 16

Figure 6. Larceny Crime Scenes & Haven Locations Distribution, Baltimore County, 1994 – 1997 17

Figure 7. Residential Burglary Crime Scenes & Haven Locations Distribution, Baltimore County, 1994 – 1997. 18

Figure 8. Journey-to-Crime Distance Decay Functions 21

Figure 9. Journey-to-Crime Routine in the Spatial Modeling Window in CrimeStat® 3.1 28

Figure 10. Measuring the Euclidean Distance Error in a sample JTC GP map 29

Figure 11. Top Profile Area in a Sample Journey-to-Crime Geographic Profiling Map............. 30

Figure 12. Top Profile Area in a Sample Journey-to-Crime Geographic Profiling Map............. 31

Figure 13. Auto Theft - Crime Scene Distribution using a 0. 25 Mile Bin Distance Interval 34

Figure 14. Larceny - Crime Scene Distribution using a 0. 25 Mile Bin Distance Interval........... 37

Figure 15. Residential Burglary - Crime Scene Distribution using a 0. 25 Mile Bin Distance Interval 40

ABSTRACT

Technological advances are fundamental to the development of spatial analysis tools and methodologies available and used within the criminal investigative process. This research focuses on one such methodology for serial crime analysis: Journey-to-Crime (JTC) Geographic Profiling (GP).

JTC or the study of the travel behavior between an offender's residence to and from the crime scene has been a subject of study within criminology for many years. GP, based on such travel behavior, is a spatial analysis and decision support tool that is used by law enforcement agencies to determine or predict the likely location of a serial offender's residence or 'haven'. The tool uses locations of a connected series of crimes and applies various functional distance measures to them which have been avoided by traditional analytical methodologies. GP models are probability density distributions of crime trips, which help to narrow down the geographical search area or the offense domain for an offender.

This research uses 135 serial property crime incidents from Baltimore County, Maryland between 1994 and 1997 for three different crime types - **auto theft**, **larceny** and **burglary.** The objective is to analyze the accuracy of individually (i.e., by crime type and distance decay functions) calibrated JTC GP models by comparing them with the default-valued (available in CrimeStat® 3.1) JTC GP models. The JTC GP accuracy assessment is conducted on the following three measurements:

- Euclidean distance error – the straight-line distance between the actual home location and the predicted home location.
- Top profile area – the area of all cells with a probability score equal to or higher than the probability score assigned to the actual haven.

- Hit score percentage – the ratio of the area searched before the offender's residence is found, to the total study area.

The smaller the value of the above measures, the better the model predicts. Results indicate that for most cases there are no statistically significant differences between the individually calibrated and default valued JTC GP models. Thus it could be concluded that police department and other investigative agencies using CrimeStat® 3.1 will save resources (personnel, time and financial) if they use the default values for the JTC distance decay functions parameters instead of individually calibrating the data while creating GP models for serial offenders.

CHAPTER 1: INTRODUCTION

1.1 What is Geographic Profiling?

According to Rossmo (2000), geographic profiling is a spatial analysis and decision support tool, consisting of various investigative and analytical methodologies, that is used by criminologists or law enforcement agencies to predict the most probable area of offender residence by analyzing the locations of a connected series of crimes. It is typically used in cases of serial murder or rape (but also arson, larceny, robbery, and other crimes). Geographic profiling could be part of the forensic analysis of a crime case, which also includes the development of a criminal *modus-operandi,* (MO), psychological and behavioral profile, ballistics, fiber analysis and DNA analysis, to name a few. Forensic analysis is a multi-disciplined collection of scientific techniques in which investigators attempt to coherently relate various elements of a crime in order to successfully prosecute an offender (Kent, 2003). As such geographic profiling alone cannot solve a crime, but it helps to narrow down the search area of an unknown offender thus saving a lot of resources (personal and financial) and is therefore referred to as geographic prioritization (Rossmo, 2000).

Geographic profiling is based on the rich conceptual framework developed by Brantingham and Brantingham (1981). The framework describes the journey-to-crime of potential offenders to search for targets in their environment or activity space to help predict where the offender will commit crimes and how the spatial distribution of potential targets and the activity areas they traverse during their routine activities influence the offenders' choices. According to their research, in general offenders commit crimes where there is an overlap between suitable targets and their personal awareness space. The theoretical work of the Brantingham and Brantingham (1984) and many others (e.g., Brown & Altman, 1981; Clarke & Cornish, 1995; Bernasco, 2006; Rengert, 1980, 1981) is complemented by even larger numbers of empirical research on spatial crime pattern, offender mobility and criminal target choice. Such studies have shown that most offenders commit crimes close to their homes and as the distance from their homes increases, the number of crimes committed decreases (Baldwin & Bottoms, 1976; Capone & Nichols, 1975; Gabor & Gontheil, 1984; Hesseling 1992; LeBeau, 1987; Philips, 1980; Rengert *et al.,* 1999; Snook, 2004, Turner, 1969; Van Koppen & James, 1998; Wiles & Costello, 2000) and thus offender search patterns for targets usually follow a distance decay function in which there is an inverse relationship between the number of crimes committed and the distance from an offender's haven (Rossmo, 2000).   Geographic profiling essentially inverts these ideas to locate where an offender lives by using information about where the offender has chosen to commit crimes (Paulsen, 2006).

1.2 Accuracy Measures for Geographic Profiling

The accuracy of the geographic profiling models to be discussed and analyzed in this thesis is measured by the following:

- Euclidean distance error – the straight-line distance between the actual home location and the predicted home location. The shorter the distance, the better the model is.
- Top profile area – also called the priority search area is a part of the offense domain, where investigators should focus in looking for the home base of an offender. It is the area of all cells with a probability score equal to or higher than the probability score assigned to the actual haven. The smaller the area, the lesser resources are required to search for the offender, the better the model predicts.
- Hit score percentage – It is the ratio of the area searched (following the geographic profiling prioritization) before the offender's residence or haven is found, to the total hunting area. The smaller this ratio, the better the geoprofile's focus and the better the model predicts. There are no intrinsic disadvantages to this measure.

1.3 Hypothesis and Research Questions

This research is an empirical study and analysis of the accuracy of individually calibrated journey to crime functions used to define geographic profiling models for serial offenders. The data analysis involves comparing the JTC GP models created from the default values in the journey-to-crime module in the Crime Stat® 3.1 (Levine, 2007) with the models created from the individually calibrated values for the same data set. The hypotheses tested are as follows:

Null Hypothesis (Ho): There is no statistical difference between the results derived from either the default or the individually calibrated JTC GP models.

Alternate Hypothesis (H_a): There is a significant statistical difference between the results derived from the default and the individually calibrated JTC GP models. More specifically, it is expected that the individually calibrated models yield better and more accurate results as compared to the results obtained from JTC GP models that are based on default parameters.

1.4 Significance of Research Work

Theoretically, it was as early as 1986 when Le Beau (1987) through his research in crime pattern theories recognized the investigative potential of geostatistical analysis for reducing offender search areas. Technological advances in desktop computer mapping provided a major breakthrough in the way investigators were able to visualize the occurrence of crimes and also analyze criminal activity in a variety of contexts in which they occurred. The use of geographic information systems (GIS) to store and analyze discrete data points relative to other intelligence assets facilitated the criminal investigation process. The flexibility of GIS technologies also enabled combining spatial analysis, statistics, and report generation to help investigators with the ability to identify change, reveal patterns and trends, and model possible methods of mitigation.

To this day, law enforcement agencies have come to rely on geographic analysis to quickly analyze and disseminate information in order to provide meaningful and coherent investigation and apprehension strategies (Kent, 2003). The first true geographic profiling was developed in 1990 when crime pattern theory was utilized as a heuristic for the construction of an algorithm model for locating offender residence (Rossmo, 2000). Since then, there has been an increased interest in and use of geographic profiling by law enforcement agencies (Paulsen, 2006). Law enforcement agencies ranging from the RCMP (Royal Canadian Mounted Police) in Canada, the National Crime Faculty in England, the BKA (Bundeskriminalamt) or Federal Criminal Investigation Office in Germany, the BATF (Bureau of Alcohol, Tobacco and Firearms) in the US and numerous local jurisdictions all use geographic profiling to assist in serial crime investigations (Rossmo, 2003). Despite all the publicity and support that GP has received in the last few years (Paulsen, 2004; Ramsland, 2005), almost no empirical research exists as to the accuracy of GP software programs, including Journey-to-Crime (Crime Stat® 3.1), Rigel, or DRAGNET, in predicting the location of the serial offender's residence. This research attempts to assess the accuracy of the JTC GP methods in Crime Stat® 3.1 by comparing the results derived from distance decay functions that use the default parameters with individually calibrated distance decay functions. If results prove that the null hypothesis cannot be rejected, then,

1. Default parameter values should be used when creating JTC GP.
2. Distance decay functions do not need to be individually calibrated.
3. Time and resources (personal, money) would be saved.

It should also be noted that this comparative analysis has never been done before.

CHAPTER 2: LITERATURE REVIEW

2.1 Spatial Analysis of Crime

Criminology, the study of crime, has long been a part of other disciplines such as sociology and psychology (Georges, 1978). Since the late 1970s there has been a realization that there is a spatial aspect associated with crime, as crime has an inherent geographical quality (Chainey & Ratcliffe, 2005). Geographers became interested and began to study how crime occurrence can be modeled in a geographical context, to better understand the patterns exhibited by the distribution of crime in any particular place or location (Taras, 1996). As stated by Georges (Georges, 1978, pp. 4):

> "The objectives of the geography of crime are to describe and map the spatial distribution of crime in greater detail and meaning than has been done before. This field of research attempts to relate the spatial patterns of crime to the environmental, social, historical, psychological (cognitive), and economic variables that may explain crime manifestation in regard to locale. Last but not the least, it is hoped that its contribution to the analysis of the dynamics of crime manifestation will help those charged with responsibility of crime control to assess better the effectiveness of programs they currently use."

The interest in the spatial analysis of crime spans from the perspective of understanding the etiology of crime and to develop criminal justice methods and practices to reduce crime (Anselin, et al. 2000). It is not limited to criminologists, but urban geographers, police officers, crime analysts and other researchers in the public and private sector have long been interested in the spatial dimension of crime (Gaile and Wilmott, 2003). Geographic Information Sciences (GISc), cartography, remote sensing and quantitative methods and mathematical models in geography have facilitated the study of the spatial dimension of crime. The identification of crime hot spots

(Sherman & Buerger, 1989), theoretical concepts in routine activities theory, rationale choice theory and research into mental maps, awareness space and journey-to-crime all refocus attention on spatial/locational features of crime. Technological advances, primarily in computer capabilities, are fundamental to recent analytical advances in the methods available for analyzing place-based crime data. The advent of computer mapping applications and accompanying geographic information systems (GIS) are crucial to being able to measure and represent the spatial relationships in data (Anselin, et al. 2000).

Crime mapping has been a very useful tool in the process of crime analysis. For example, the New York Police department has traced back the use of maps to at least 1900 (Gaile and Wilmott, 2003). The traditional form was using pin maps (Figure 1 - Harries, 1999) to show where crime occurred, but it had its own limitations. The difficulty to read the crime pattern for several different types of crime, loss of data, large space requirement, no capability to query the data as they were static maps, are few of those limitations (Harries, 1999). Thus, with the advent of desktop mapping, crime research has been revolutionized and which influenced the technology of policing.

Figure 1. Example of a Pin Map of an Area in Baltimore County, Maryland (Source: Harries, 1999)

The criminal investigative process involves a variety of analytical techniques that support the apprehension and successful prosecution of an offender (Kent, 2003). Geographic profiling is an advanced investigative technique that forms part of a crime scene forensic analysis which includes development of a criminal *modus-operandi* (MO), psychological and behavioral profile, ballistics, fiber analysis, DNA analysis, just to name a few (Kent, 2003). GP is based on the principles of journey-to-crime that have been a subject of study within criminology for many years.

2.2 Journey-to-Crime

JTC, a term first coined by Philips (1980), is the study of the travel behavior between an offender's residence to and from the crime scene. The journey-to-crime approach is a precursor to geographic profiling techniques and has been used for years to locate the likely origin of a serial offender based on the properties associated with the distribution of crime incidents (Levine, 2007). The JTC model algorithm is based on a combination of location theories, which attempts to find an optimal location for any particular distribution of activities or population over a region, and travel demand models developed for transportation planners (Levine, 2002). In JTC models criminal travel behavior is observed by measuring the distance between the known crime site and the offender's known residence. The behavior is quantified by plotting the statistically aggregated distances against the number of crime committed to illustrate the percentage of crime for a given distance unit (Kent, 2006). Different variations of the journey-to-crime models are utilized by contemporary criminologists to study case-specific criminal spatial characteristics. In terms of its descriptive statistical capabilities, journey-to-crime models are dependent upon numerous conditions, including the scale of observation.

Traditional journey-to-crime techniques were founded from sociological research developed from the Chicago School of the 1920s (Anselin et al., 2000). Significant results were obtained through the applications of journey-to-crime distance analysis from research done by Capone and

Nichols (1975) who noted that property crime offenders generally traveled farther distances than offenders committing crimes against people; by Lottier (1938), who analyzed the ratio of chain store burglaries to the number of chain stores by zone in Detroit; and by Turner (1969), who analyzed delinquency behavior by a distance decay travel function showing how more crime trips tend to be close to the offender's home with the frequency dropping off with distance.

Some of the commonly utilized algorithms in journey-to-crime studies include: mean and median crime trip distances, medial circles, mobility triangles, and distance decay functions, just to name a few (Rossmo, 2000). As each approach has its own unique qualities, selecting the most appropriate modeling application will depend entirely on the characteristics of the environment in which a crime occurs; usually requiring a trial-and-error approach (Levine, 2007).

2.3 Distance Decay Functions

The distance decay approach is one of the most useful presentations of journey-to-crime data. As the name suggests, distance decay, in the context of crime mapping, refers to the decrease in the frequency of crimes committed by an offender as the distance from the haven increases. Thus, in general there is an inverse relationship between the number of crimes committed and the distance the offender travels to commit the crimes. It also has been suggested that as an offender's criminal career matures, such crime trip distances lengthens and the size of the hunting area increases (Brantingham & Brantingham, 1981; Canter & Larkin, 1993).

In order to understand the underlying significance of the distance decay approach for criminal activity, various theoretical models have been suggested based on existing migration algorithms and intervening opportunity theory, which are founded on the ecological context of Sir Isaac Newton's gravity function (Levine, 2002). Rengert (1981) developed a mathematical equation that defined journey-to-crime based on a modified general opportunities model (Gore & Tofiluk, 2002; Levine 2007):

$$P_{ij} = K.E_i.V_j.f(d_{ij}) \tag{2.1}$$

where the probability, P, that an offender from zone i, committed a crime in location j is related to the product of the enumerated number of trips produced (emissiveness) from the origin, E_i, and the number of potential targets (attractiveness) at the destination, V_j, for travel cost, $f(d_{ij})$ (Levine, 2007). Basically, this model theorized that the probability an offender would commit a crime at a given location is entirely dependent on both the production cost, what Rengert (1981) called emissiveness, and the attractiveness for that destination. Rengert's (1981) cost value is an undefined functional distance metric that is, presumably, a straight-line Euclidean measure of the distance between origin and destination. While not empirically defined, the hypothetical results of his model were compared against observed burglaries in Philadelphia, PA in order to measure its effectiveness. As noted by Gore & Pattavina (2001), the theoretical value of Rengert's model (1981) is that it can be used to predict crime patterns for locations that have empirically quantified the observable travel production and zone attractiveness. These are essential components used within travel demand models.

The research presented in this thesis utilizes five probability density distribution functions available in Crime Stat® 3.1 (Levine, 2007) for journey-to-crime modeling. This renders more flexibility in describing an accurate simulation of offender travel behavior under different conditions such as crime type, time of the day, method of operation and other variables. The five functions are linear, negative exponential, truncated negative exponential, normal and lognormal. Each of these functions is explained in detail in Chapter 3.

2.4 Environmental Criminology Theories Underlying Geographic Profiling

The spatial distribution of crime is influenced by three general factors as suggested by Rengert (1981):

1. the location of crime prone populations;
2. the location of opportunities for crime; and
3. the relative accessibility of potential offenders to opportunities.

The environmental criminology theories underlying geographic profiling are based on these three factors.

2.4.1 Awareness and Activity Spaces

An awareness space is defined as, "all the locations about which a person has knowledge above a minimum level even without visiting some of them….Awareness space includes activity space and its area enlarges as new locations are discovered and/or new information is gathered" (Clark, 1990, pp. 24-25). In general offenses should occur within a criminal's awareness space.

An activity space is defined as, "the area within which most of a person's activities are carried out, within which the individual comes most frequently into contact with others and with the features of the environment and its area enlarges as new locations are discovered and/or new information is gathered" (Clark, 1990, pp. 24 - 25). An activity space thus includes those areas that are well known to the offender and/or target through routine (daily or weekly) activities such as traveling to school, shopping and/or seeking out entertainment, etc. and is contained within the awareness space. These locations are referred to as activity nodes (Figure 2).

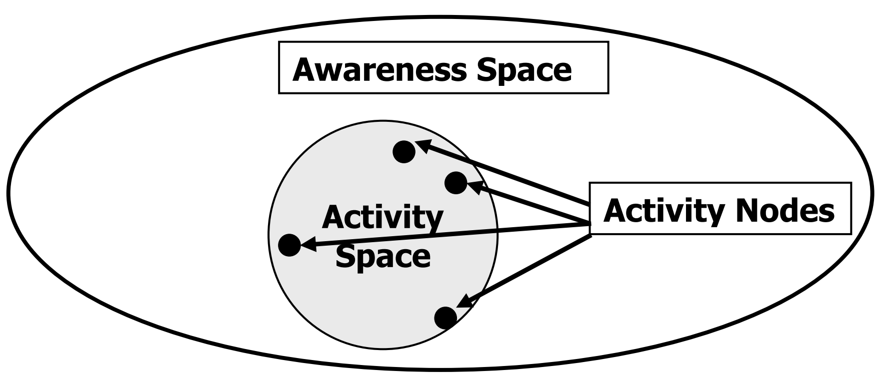


Figure 2. Awareness and Activity Spaces (Source: Brantingham & Brantingham, 1981)

2.4.2 Routine Activity Theory

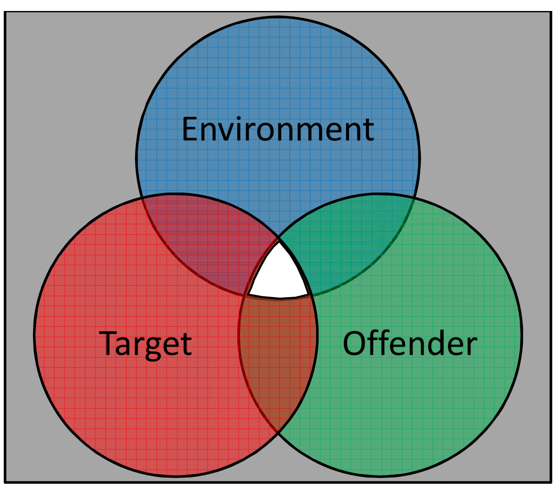
There are three elements in Routine Activity Theory:

a) motivated offenders, b) suitable targets, and c) an environment with an absence of capable guardians against a violation (Felson & Clarke, 1998). According to this theory, for a direct-contact predatory crime to occur the offender's activity space, the target's activity space must intersect in time and space, within an environment considered appropriate for criminal activity. (Rossmo, 2000).

According to Rossmo (2000), the opportunity structure of crime can be summarized as following:

$$\text{crime} = (\text{offender} + \text{target} - \text{guardian})\ (\text{place} + \text{time}) \tag{2.2}$$

The level of convergence in space and time of the three elements of the Routine Activity theory could influence the crime rates (Cohen & Felson, 1979). Figure 3 is a graphic representation of the theory.



Offender: Offender's Awareness Space
Target: Target's Awareness Space
Environment: Shared landscape, situation, neighborhood
Offender/Target: Shared Activity Space
Offender/Environment: Offender's Activity Space
Target/Environment: Target Activity Space
Offender/Victim/Environment: Intersection of Offender and Victim Activity Space

Figure 3. Routine Activity Theory (Source: Cohen & Felson, 1979)

2.4.3 Least Effort Principle

The least effort principle (Zipf, 1949) or the nearness principle is the underlying law governing human activity and perhaps the most basic heuristic in geography (Rossmo, 2000).

According to this theory, a person who is “given various possibilities for action….will select the one requiring the least expenditure of effort” (Reber, 1985, pp. 400). Thus, in terms of criminology, the theory suggests that all things being equal, an offender will choose to commit crimes closer to their homes than further away.

2.4.4 Rational Choice Theory

The Rational Choice theory is based on a decision making approach. It is a “voluntaristic, utilitarian action theory in which crime and criminal behavior are viewed as the outcome of choices. These, in turn, are influenced by a rational consideration of the efforts, rewards, and costs involved in alternative courses of action” (Cornish, 1993, pp. 362).

The rational choice perspective as presented by Cornish and Clarke (1985) is based on three concepts: (1) criminal offenders are rational and make choices and decisions that benefit themselves; (2) a crime-specific focus is required; and (3) there is a distinction between choices related to criminal involvement and decisions related to criminal events (Rossmo, 2000).

2.4.5 Buffer Zone

Buffer zone is referred to an area surrounding a particular activity node, most notably the residence of the offender, from which little to no criminal activity will be observed (Brantingham & Brantingham, 1980). It is assumed that such an area would represent an elevated level of risk associated with operating too close to the home. This characteristic is also called the coal-sack effect by Newton & Swoope (1987) whereby the offender, either intentionally or otherwise, avoids committing an offence in particular areas surrounding his or her residence. Notably, the buffer zone is seldom observed for spontaneous and/or passion crimes (LeBeau, 1987), and most likely occur for predatory offences which can be characterized as pre-meditated (Canter & Larkin, 1993). A specific consideration for the existence of the buffer zone is that it may not always be applied around the offender’s residence, but may also refer to any particular node that represents any single or shared (i.e., home and work) location in the criminal's routine activity space and is termed as the criminal's 'haven' (Newton & Swoope, 1987). The idea of a safety zone, however, can be misleading in that some criminal activity may exist if the offender perceives conditions and circumstances to be favorable for the commission of a crime (Rossmo, 2000) – a rationale that was supported by Godwin & Canter (1997) for United States (US) serial offender body dump sites. Other examples can include peeping, stalking, and other illegal surveillance activities. Support for the existence of buffer zone-like features can be observed quantitatively. According to Rossmo (2000), combining the linear increase in an offender's opportunity to commit crimes with the decrease in travel desire, a criminologist should be able to observe a buffered distance decay function (Rossmo, 2000). This application was substantiated by Canter & Larkin (1993). Using regression equations, the researchers were able to approximate a one-kilometer buffer zone around the havens of United Kingdom (UK) serial rapists. Rossmo (2000) notes that such zones also existed for similar studies of US and UK serial killers (Godwin & Canter, 1997) and Levine (2002) cites similar characteristics for various offences in the US. This linear increase of an offender's opportunity to commit crime however, assumes an equally available distribution of opportunities and targets. An offender's hunting ground, target selection, spatial travel preferences, and buffer zone can be estimated using available geographic modeling applications. As proposed by this thesis, these elements can be modeled by calculating the measurable travel characteristics expressed by the distribution of a serial offender's known linked crime scenes. To accomplish this task, criminologists utilize one of the most relevant modeling applications available: journey-to-crime.

2.5 History of Geographic Profiling

Geographic profiling (GP) is one of the more recent analytical advances in the spatial study of crime. It is an investigative methodology that uses the locations of a connected series of crimes to determine the most probable area of an offender residence or 'haven'. It is based on a probability density map with the cell having the highest probability indicating the likely (or predicted) residence or "haven" of the offender. The map is usually a colored isometric map. GP is generally applied in cases of serial murder, rape, arson, and robbery, though it can be used in single crimes (auto theft, burglary, bombing, etc.) that involve multiple scenes or other significant geographic characteristics (Crime Mapping Research Center, 1999).

The history of geographic profiling dates back to as early as 1979. Holt was the first to develop a geographic profile with the application of spatial analysis and mapping (Rossmo, 2000). He was followed by LeBeau (1987) in 1986, who recognized the investigative potential of geostatistical analysis and crime pattern research for reducing offender search areas (Paulsen, 2006). In 1990 a comprehensive geographic profiling model was developed by Rossmo (2003). Until today the biggest influence on geographic profiling could be attributed to crime pattern theory and the research done by Paul and Patricia Brantingham (1981). Their research indicated that, in general, offenders commit crimes where there is an overlap between suitable targets and their personal awareness space (Brantingham and Brantingham, 1981). Thus, offender search patterns usually follow a distance decay function in which there is an inverse relationship between the number of crimes committed and the distance from an offender's haven (Rossmo, 2000). Journey to crime research supports these ideas, indicating that most criminals travel relatively short distances from home to commit a majority of crimes (Phillips, 1980; Ratcliffe, 2003). Geographic profiling essentially takes these ideas and inverts them (Paulsen,2006). Using information about where an offender has chosen to commit crimes, geographic profiling attempts to determine where the offender is most likely to reside (Rossmo, 2000). Using crime site location information and distance decay analysis, geographic profiling then seeks to help narrow the search area through the creation of a geographic profile region.

CHAPTER 3: DATA & METHODOLOGY

3.1 Data & Study Area

The data for this research consists of a complete set of all offenders arrested for three or more of the same crimes in Baltimore County, Maryland (Figure 4), between 1994 and 1997. It includes 135 solved serial property crimes provided by the Baltimore County Police Department.

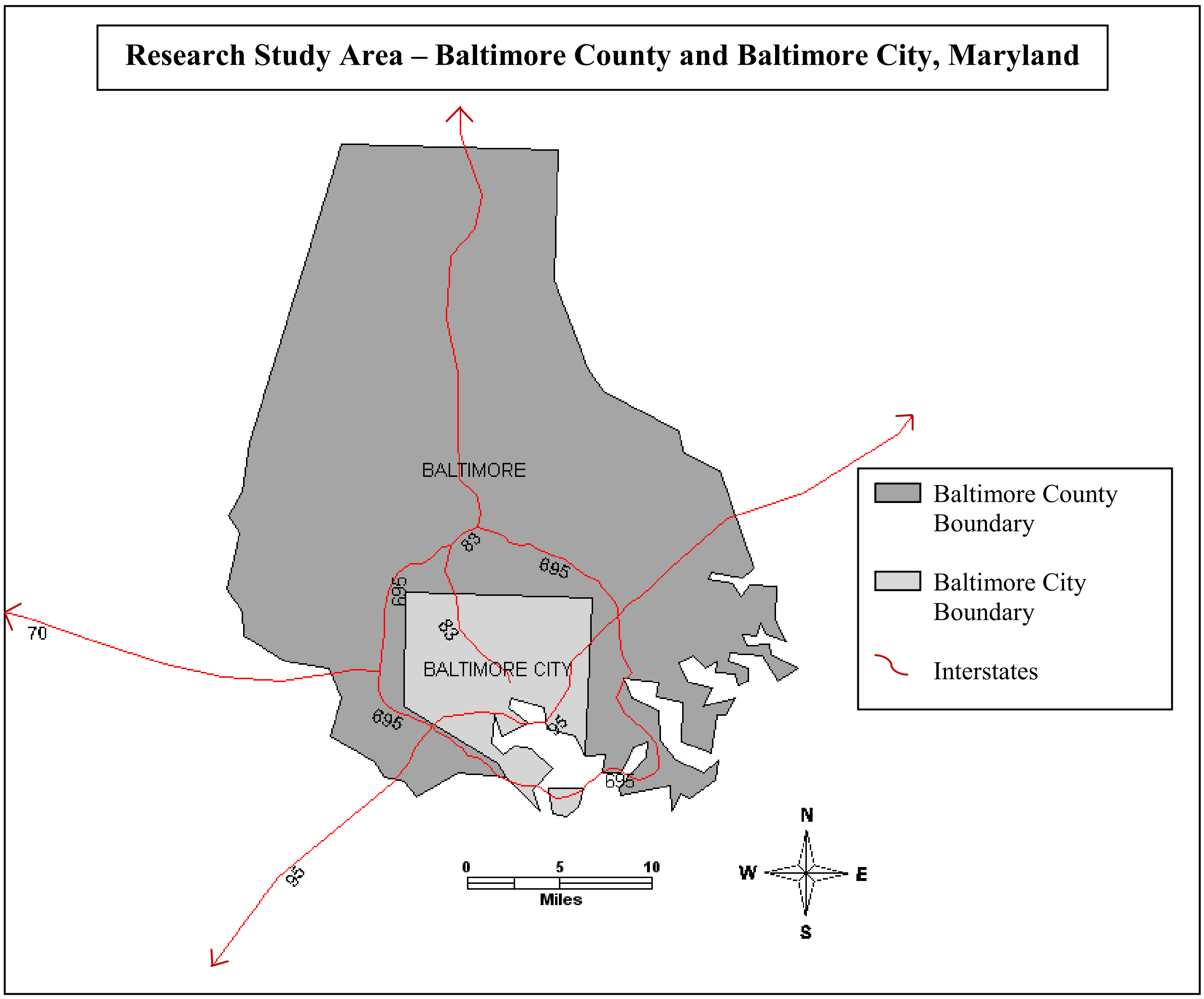


Figure 4. Study Area of Baltimore County, Maryland

The property crimes analyzed in this research are:

- Auto Theft - defined as the act of theft or attempted theft of a motor vehicle, including joy riding. A motor vehicle is self-propelled and runs on land surface and not on rails.

Motorboats, construction equipment, airplanes, and farming equipment are specifically excluded from this category (Federal Bureau of Investigation, 2006). Figure 5 shows the spatial distribution of the incidents and haven locations of corresponding offenders for auto theft serial cases that occurred in Baltimore County between 1994 and 1997.

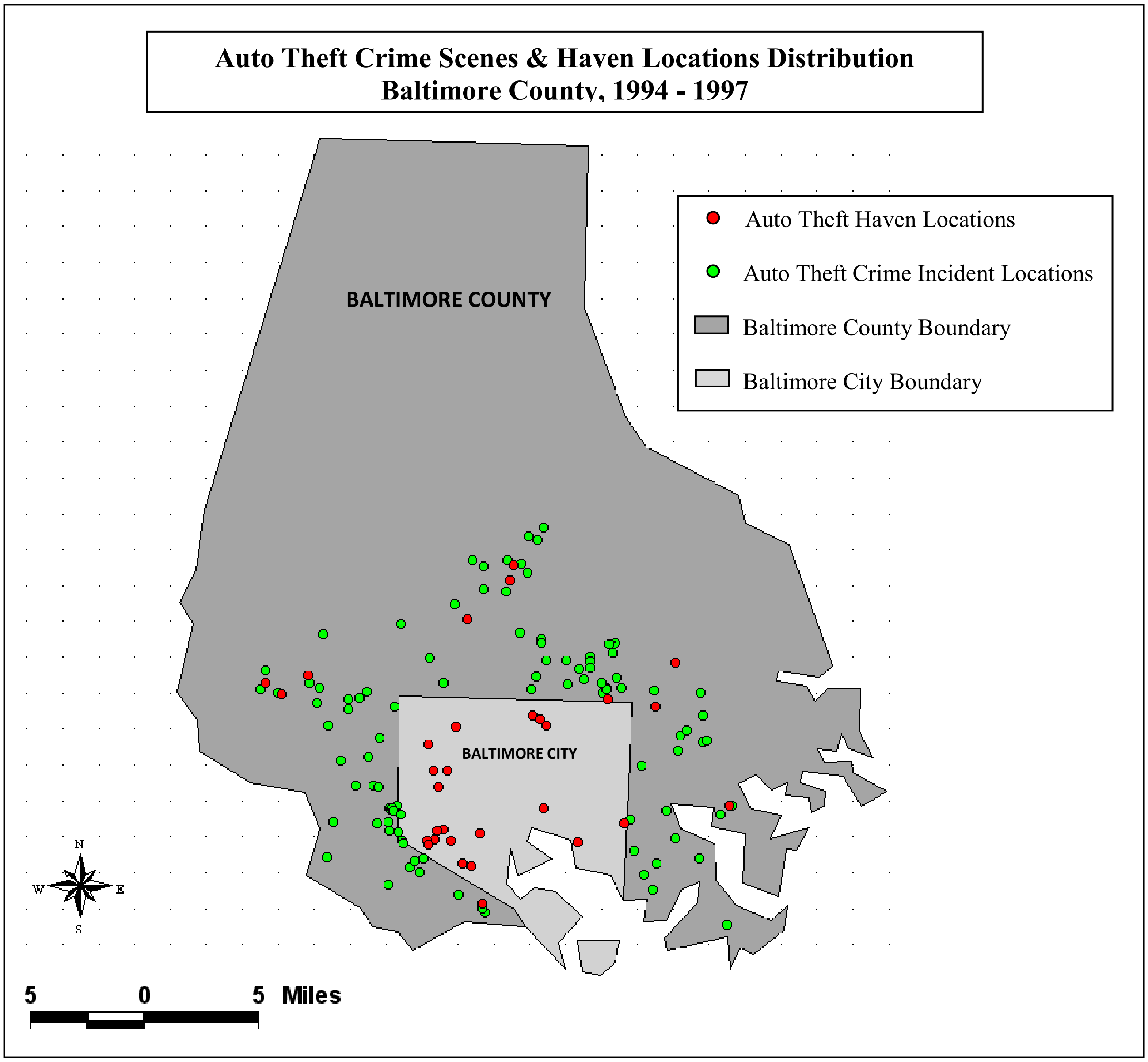


Figure 5. Auto Theft Crime Scenes & Haven Locations Distribution, Baltimore County, 1994 – 1997 (Source: Baltimore County Police Department, Maryland).

- Larceny - defined as the unlawful taking, carrying, leading, or riding away of property from the possession or constructive possession of another (Federal Bureau of Investigation, 2006).

Figure 6 shows the spatial distribution of the incidents and haven locations of corresponding offenders for larceny serial cases that occurred in Baltimore County between 1994 and 1997.

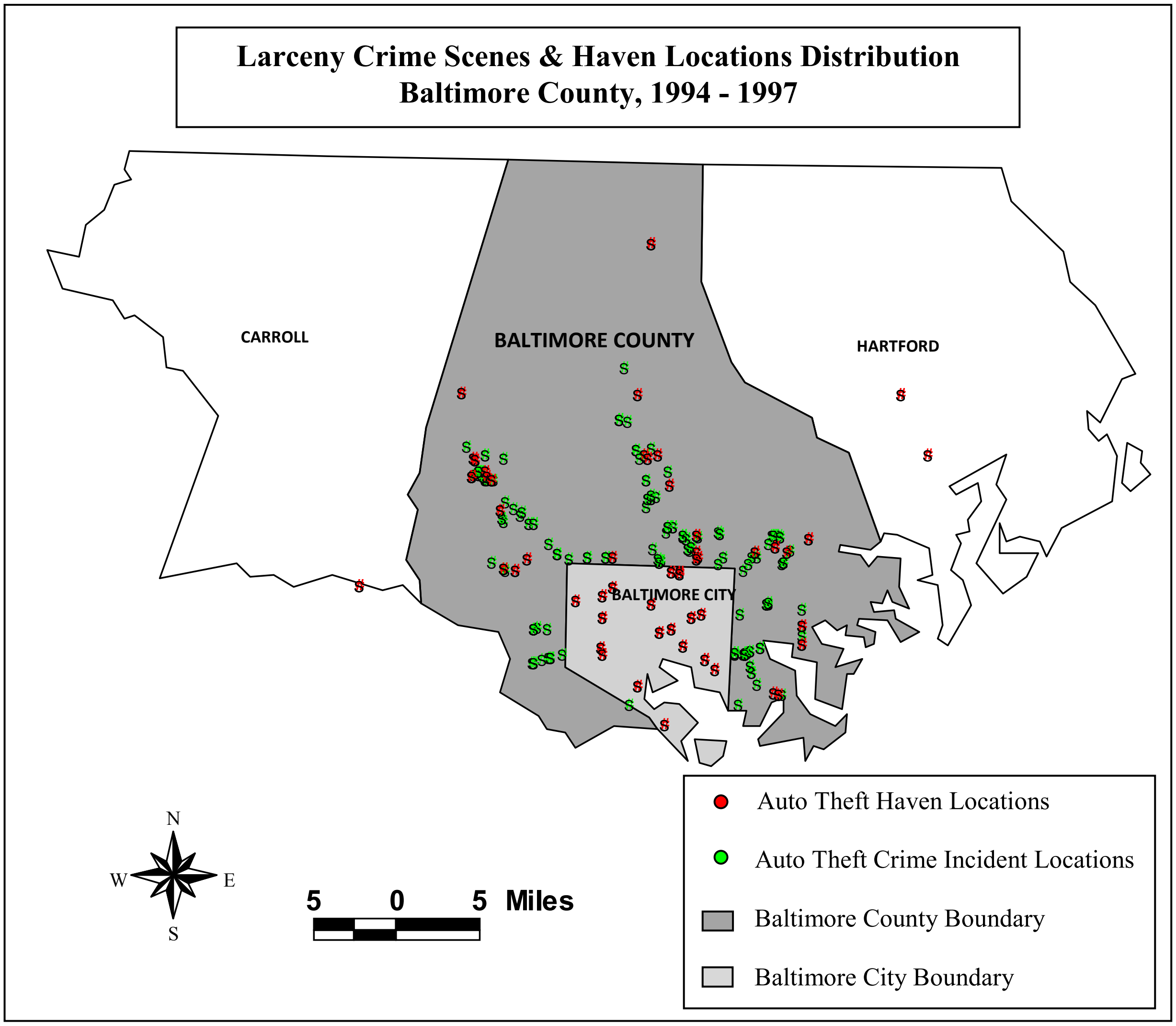


Figure 6. Larceny Crime Scenes & Haven Locations Distribution, Baltimore County, 1994 – 1997 (Source: Baltimore County Police Department, Maryland).

- Residential Burglary - Burglary is defined as the unlawful entry into a building or other structure with the intent to commit a felony or a theft (Federal Bureau of Investigation, 2006). Figure 7 shows the spatial distribution of the incidents and haven locations of corresponding offenders for larceny serial cases that occurred in Baltimore County between 1994 and 1997.

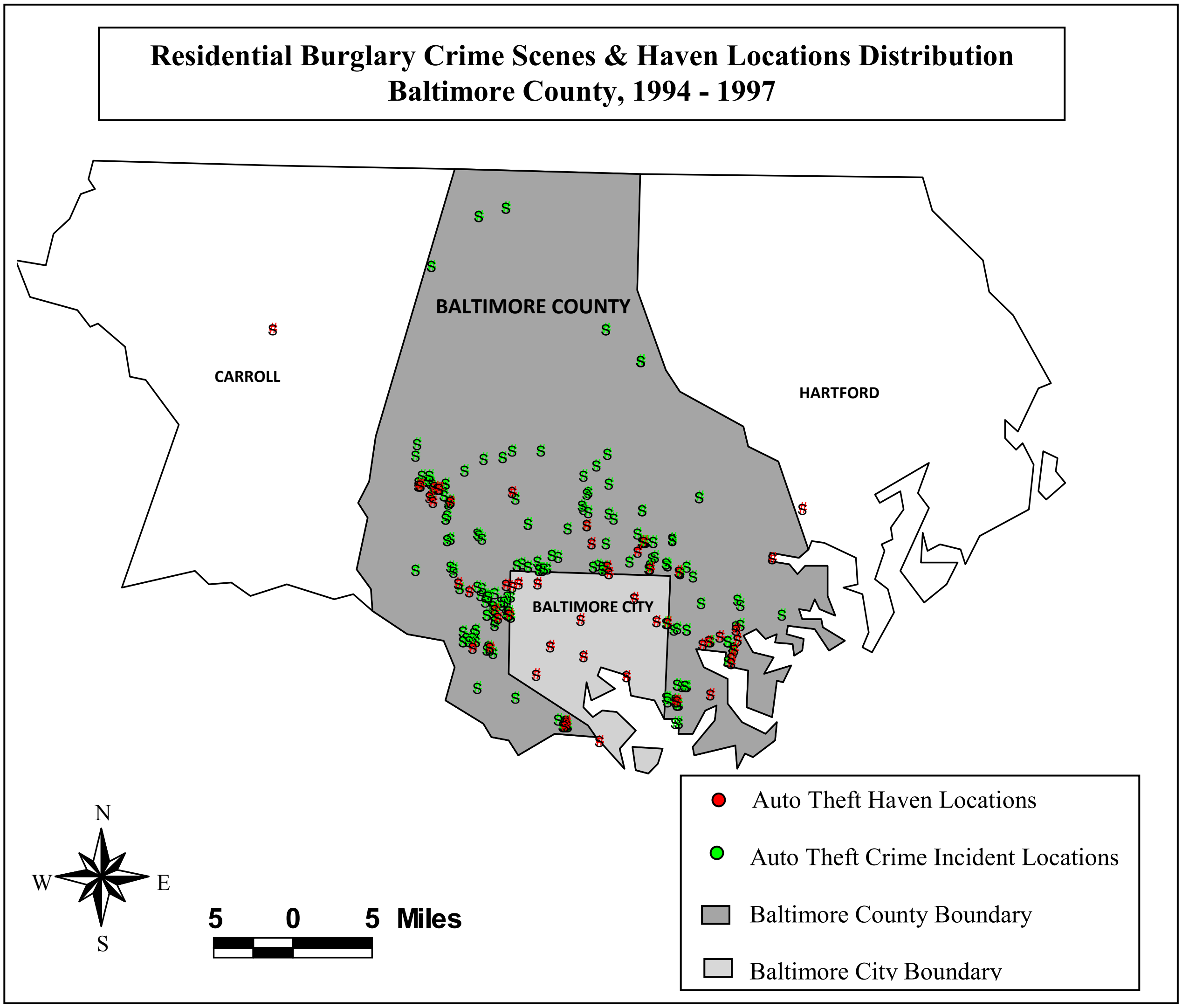


Figure 7. Residential Burglary Crime Scenes & Haven Locations Distribution, Baltimore County, 1994 – 1997 (Source: Baltimore County Police Department, Maryland).

Table 1 provides descriptive information about the different crime types used in this study.

Table 1. Crime Series by Crime Type for Baltimore County, 1994 – 1997 (Source: Paulsen, 2006)

Crime Type	Number of Series	Total Crime Incidents	Average Crimes Per series
Auto Theft	31	143	4.61
Larceny	59	240	4.06
Residential Burglary	56	241	4.30
All Crime Series	135	624	4.62

All series included in this research were verified by the arresting agency in Baltimore County, MD (Paulsen, 2006). Table 2 lists the data set's attributes which included a crime series ID, x- and y-coordinates for each crime incident location and the haven location of arrested offenders, and the start date and time for each property crime case. All coordinates in the data set have been projected to UTM NAD 1983 Zone 18 with measurement units in meters (approximately 3.3 feet).

Table 2. Sample Attribute Table for Property Crime Incidents (Source: Paulsen, 2006)

CRIMECOD	DATE_	TIME	INCIDX	INCIDY	HOMEX	HOMEY
bat004	5/26/1995	0	-76.5409	39.4049	-76.4918	39.3932
bat004	5/29/1995	2344	-76.5435	39.4035	-76.4918	39.3932
bat004	6/2/1995	0	-76.6015	39.4042	-76.4918	39.3932
bat004	6/26/1995	2210	-76.4695	39.2691	-76.4918	39.3932
bat004	6/28/1995	0	-76.4675	39.3431	-76.4918	39.3932
bat004	2/25/1996	240	-76.4684	39.3597	-76.4918	39.3932

For the purposes of calculating the JTC GP, the study area for each serial crime was defined as a rectangular grid by measuring the bottom left hand coordinates of the location closest to the most western and most southern crime incident and the top right hand coordinates of the location closest to the most eastern and most northern crime incident, so that all crime incidents lie within the rectangular study grid. This study area information was provided by Paulsen, who used the same information in his research (Paulsen 2006), as an SPSS® 15.0 file (Table 3).

Table 3. Sample Data Set with Study Area Information of Entire Data Set (Source: Paulsen, 2006)

alldata.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

12 : caseid br009

	caseid	crime	offenses	jtcavg	jtcmin	jtcmax	dispersi	searchar	leftx	lefty	rightx	righty
12	br009	commrob	8	.2814	.1494	.4910	.5300	.2300	76.60604	39.39886	76.59568	39.40537
13	br010	mixed	4	1.7765	.1593	5.4046	5.6000	7.6500	76.66746	39.37008	76.52854	39.38748
14	br011	streetro	4	1.8893	.9082	3.3098	4.2100	16.7400	76.73544	39.36667	76.65725	39.42761
15	br012	mixed	5	.6137	.1029	1.3839	1.5300	1.9300	76.82885	39.36961	76.79149	39.38536
16	br013	commrob	11	2.5520	.2185	4.5262	5.0500	10.3800	76.82876	39.35664	76.72515	39.38655
17	br014	commrob	3	2.0021	1.7941	2.4182	1.8800	4.1400	76.49097	39.38913	76.45472	39.42093
18	br015	commrob	5	2.6063	1.3633	4.9319	6.0000	36.8200	76.55460	39.33041	76.43848	39.41977
19	br016	streetro	6	6.4891	4.4277	9.4108	6.7600	31.3100	76.78110	39.27244	76.70787	39.38728

Data View / Variable View

SPSS Processor is ready

The bottom left coordinates of the study area for each serial crime are labeled 'leftx' and 'lefty'; the top right coordinates, 'rightx' and righty' in Table 3. The coordinates of all crime sites and of the offender's 'haven' are expressed in decimal degrees. From this data set, calibration data sets are created and stored as Excel® 2007 spreadsheets.

3.2 Methodology

The process for calibrating a journey-to-crime distance decay function uses the traveled distances measured between each origin and destination stored within the calibration-sample data set. The origin represents the offender's residence while the destination represents the offender's crime incident location. The calibration routine is executed in six steps (Levine, 2007),

1. The data set is checked to ensure that there are X and Y coordinates for both the arrested individual's residence location (origin) and the crime incident location (destination) for which the individual is being charged.
2. The origin-to-destination (O-D) locations for each crime series are imported from ArcMap® 9.2 to Excel® 2007 spreadsheet. Thus, the data are sorted into sub-groups based on different types of crimes –auto theft, larceny and residential burglary. Each sub-group is saved as a separate file.
3. For each crime type, the distances are grouped into intervals (referred to as bins) of 0.25 miles each. This was accomplished in two steps: first by sorting the data in ascending order and second a frequency distribution is applied for each O-D distances and grouped into 0.25 mile intervals or bins. The selection of the bin interval is dependent on the size of the data set.
4. For each crime type, a new file is created which includes only the frequency distribution of the distances broken down into quarter mile distance intervals, (d_i).

5. Frequency intervals measured in step-3 are converted into relative frequencies by dividing the frequency values for each interval by the total number of incidents, *n (since it is a sample)*, and multiplying by 100. Second, the distance intervals are adjusted to the mid-point of each bin in order to provide a better representation for the bin's contribution for the distribution (McGrew & Monroe, 1993).
6. Using SPSS® 15.0, a series of univariate regression equations are executed to model each frequency as a function of the distance. The percentage of incidences within each frequency interval (Pct_i) is used as the dependent variable. Five equations are mathematically calibrated to obtain the best fit for the given distributions. Figure 8 illustrates the five functions utilized by this research.

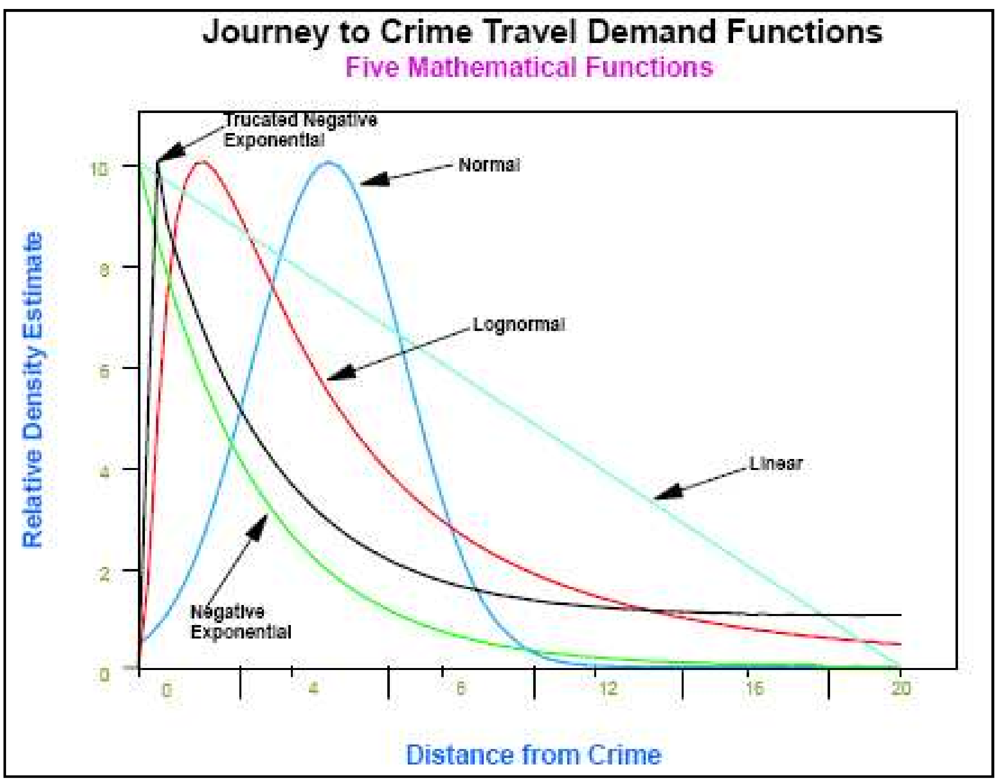


Figure 8. Journey-to-Crime Distance Decay Functions (Source: Levine 2007)

3.3 Distance Decay Functions

3.3.1 Linear

The linear function is the simplest type of distance decay model. According to this model, the likelihood of committing a crime at any particular location declines by a constant amount with distance from the offender's home. It is highest at the offender's home but drops off by a constant amount for each unit of distance until it falls to zero. The form of the linear equation is:

$$Pct_i = A + Bd_{ij} \tag{3.1}$$

where Pct_i is the likelihood that the offender will commit a crime at a particular location, *i*, d_{ij} is the distance between the offender's residence at location j and crime location *i*, *A* is the y-intercept, and *B* is the slope coefficient which defines the fall off in distance with an expected negative sign since the likelihood should decline with distance. The user must provide values for *A* and *B*. This function assumes no buffer zone around the offender's residence. When the function reaches 0 (the X axis), the routine automatically substitutes a 0 for the function.

Table 4. Intercept and Slope Values for Linear Distance Decay Function (Source: Levine, 2007)

Property Crime Type	A (Intercept) Individually Calibrated	B (Slope) Individually calibrated	A (Intercept) Default Value	B (Slope) Default Value
Auto Theft	2.006 – 2.03	-0.0074 – (-)0.0069	1.9	-0.06
Larceny	0.697 – 1.977	-0.0064 – 2.672	1.9	-0.06
Residential Burglary	2.29 – 2.47	-0.10 – (-)0.09	1.9	-0.06

3.3.2 Negative Exponential

A slightly more complex function, this model describes how the occurrence of crime is highest near the offender's 'haven' and drops off at a constant rate with distance. The mathematical expression is:

$$Pct_i = A * e^{-B*d_{ij}} \tag{3.2}$$

where Pct_i is the likelihood that an incident will occur at a particular location, *i*, d_{ij} is the distance between each reference or 'haven' location j and each crime location, *i*, *e* is the base of the natural logarithm, *A* is the coefficient, and *C* is the exponent. Like the linear function above, it assumes no buffer zone around the offender's residence. The function parameter values used are as follows:

Table 5. Intercept and Slope Values for Negative Exponential Distance Decay Function (Source: Levine, 2007)

Property Crime Type	A (Intercept) Individually Calibrated [Default]	B (Slope) Individually Calibrated [Default]
Auto Theft	0.069 – 0.648 [1.89]	-0.407 – 0.403 [-0.06]
Larceny	0.36 – 14.33 [1.89]	-0.83 – (-)0.45 [-0.06]
Residential Burglary	0.08 – 0.82 [1.89]	-0.53 – 0.45 [-0.06]

3.3.3 Truncated Negative Exponential

This is a complex function consisting of two distinct decay equations: linear and exponential. For locations in close proximity to the residences, a positive linear function is defined (Equation 3.3), starting at zero (the 'haven') and increasing to a peak distance, $Maxd_p$. Thereupon, the function follows a negatively signed exponential function, declining quickly as the distance increases (Equation 3.4).

Linear: $$Pct_i = A + Bd_{ij} \text{ for } d_{ij} >= 0,\ d_{ij} <= Maxd_p \quad (3.3)$$

Negative Exponential: $$Pct_i = A * e^{-B*d_{ij}} \text{ for } d_{ij} > Maxd_p \quad (3.4)$$

where d_{ij} is the distance from the 'haven' location j to the crime location i, *B* is the slope of the linear function, *A* is the coefficient for the negative exponential function, *C* is the exponent and d_p is the peak distance. This model can be used to approximate the often-observed buffer zone effect surrounding an offender's residence. The function parameter values used are as follows:

Table 6. Peak Distance, Peak Likelihood and Exponent Values for Truncated Negative Exponential Distance Decay Function (Source: Levine, 2007)

Property Crime Type	d_p (Peak Distance) Individually Calibrated [Default]	Peak Likelihood Individually Calibrated [Default]	C (Exponent) Individually Calibrated [Default]
Auto Theft	1.36 – 2.86 [0.4]	5.8394 – 7.6923 [13.8]	-0.38 – (-)0.30 [-0.2]
Larceny	0.62 – 12.62 [0.4]	0.82 – 9.98 [13.8]	-0.46 – (-)0.44 [0.2]
Residential Burglary	0.125 [0.4]	21.21 – 23.50 [13.8]	-0.56 – (-)0.42 [0.2]

3.3.4 Normal

According to the normal function model, the peak likelihood is at some optimal distance from the offender's home base. Thus, the function rises to that distance and then declines. The rate of increase prior to the optimal distance and the rate of decrease from that distance are symmetrical in both directions. The mathematical expression is:

$$Pct_i = A * 1/(S_d * sqrt(2\pi)) * e^{-0.5 * Z_{ij}^2} \quad (3.5)$$

The estimation of parameters can be solved in 3 steps. First, a standardized variable Z_{ij} is created for the distance d_i, and is calculated as:

$$Z_i = (d_i - MeanD\ ) / S_d \quad (3.6)$$

where, MeanD is the mean distance and S_d is the standard deviation of the distance. Second a normal transformation of Z_{ij} is constructed with

$$Normal\ (Z_i) = 1/(S_d * sqrt(2\pi)) * e^{-0.5 * Z_{ij}^2} \quad (3.7)$$

where $\pi = 3.14$, e = 2.72 and Z_i is the standardized variable. And finally, the normalized variable is regressed against the percentage of all crimes of that type falling into the interval, Pct_i with no constant

$$Pct_i = A * Normal\,(Z_i) \quad (3.8)$$

A, the y-intercept is estimated by the regression coefficient. By carefully scaling the parameters of the model, the normal distribution can be adapted to a distance decay function with an increasing likelihood for near distances and a decreasing likelihood for far distances. For example, by choosing a standard deviation greater than the mean (e.g., MeanD = 1, S_d = 2), the distribution will be skewed to the left because the left tail of the normal distribution is not evaluated. The function parameter values used are as follows:

Table 7. Mean, Standard Deviation and Intercept Values for Normal Distance Decay Function (Source: Levine, 2007)

Property Crime Type	MeanD Individually Calibrated [Default]	Standard Deviation (S_d) Individually Calibrated [Default]	A Individually Calibrated [Default]
Auto Theft	0.92 – 19.74 [4.2]	1.77 – 13.38 [4.6]	48.58 – 51.40 [29.5]
Larceny	2.13 – 2.85 [4.2]	0.78 – 11.65 [4.6]	0.56 – 48.09 [29.5]
Residential Burglary	1.67 – 2.48 [4.2]	0.89 – 10.14 [4.6]	46.58 – 59.72 [29.5]

3.3.5 Lognormal

The lognormal function is similar to the normal except it is more skewed, either to the left or to the right. It has the potential of showing a very rapid increase near the offender's home base with a more gradual decline from a location of peak likelihood. It is also similar to the Brantingham and Brantingham (1981) model. The mathematical form of the function is:

$$Pct_i = A * 1/(d^2_{ij} * S_d * sqrt(2\pi)) * e^{-[\ln(d^2_{ij}) - MeanD]^2 / 2 * S_d^2} \qquad (3.9)$$

Four intermediate variables, L, M, O and P are created to facilitate the breaking down of this complex transformation into simpler units, where

$$L = \ln(d_i^2) \qquad (3.10)$$

$$M = (L - MeanD)^2 \qquad (3.11)$$

$$O = M / (2 * S_d^2) \qquad (3.12)$$

$$P = e^{O} \tag{3.13}$$

The lognormal conversion is calculated according to the following formula:

$$Lnormal(d_i) = 1/(d^2{}_{ij} * S_d * sqrt(2\pi)) * P \tag{3.14}$$

where $\pi = 3.14$. Finally, the lognormal variable is regressed against the percentage of all crimes of a particular type falling into the interval, Pct_i, with no constant according to the expression:

$$Pct_i = A * Lnormal(d_i) \tag{3.15}$$

where A, the y-intercept is estimated by the regression coefficient. Once completed, each mathematically calibrated distance decay function can be utilized to run the journey to crime routine in CrimeStat® 3.1 as detailed in the following section. The function parameter values used are given in Table 8:

Table 8. Mean, Standard Deviation and Intercept Values for Lognormal Distance Decay Function (Source: Levine, 2007)

Property Crime Type	MeanD Individually Calibrated [Default]	Standard Deviation (S_d) Individually Calibrated [Default]	A Individually Calibrated [Default]
Auto Theft	19.75 [4.2]	11.74 [4.6]	-0.35 – 0.62 [8.6]
Larceny	1.96 – 2.78 [4.2]	0.47 – 11.65 [4.6]	20.58 – 26.30 [8.6]
Residential Burglary	0.89 – 1.96 [4.2]	0.45 – 10.14 [4.6]	85.30 – 96.76 [8.6]

3.4 Journey-to-Crime Routine

The Journey-to-Crime (JTC) routine in CrimeStat® 3.1 is used to make estimates about the likely location of the residence of a serial offender for a given crime type using the corresponding incident locations and a distance decay function. The likely location of the residence will be estimated twice, first with the distance decay function default values and second with individually calibrated distance decay functions. In both cases, the JTC routine assigns a value to each point of the regular grid that is superimposed over the study area. These values are referred to as probability-scores, and indicate the likelihood that any location within the study area is the offender's likely residence. The JTC procedure (Figure 9) is executed in five steps. In the following example a randomly chosen serial crime (larceny #8) is used to explain the JTC procedure.

1. The primary data file is selected which for any particular serial crime (e.g., larceny #8) is the incident database (.dbf) file for the corresponding ArcMap® 9.2 shape file. The coordinate system is also defined to be longitude –latitude (spherical) with decimal degrees as unit.
2. The study area for the particular set of incidents (serial crime, larceny #8) is defined in the reference file window. It is constructed from a rectangular matrix or minimum bounded rectangle consisting of 100 columns. 100 columns is the default value that can be changed. The coordinates for this rectangle have been provided along with the data set in an SPSS® 15.0 file. These are the same coordinates that Paulsen used in his research (Paulsen 2006).
3. The Journey-to-Crime routine is part of the 'Spatial Modeling' Window in Crime Stat® 3.1. Depending on the mathematical function and whether the default or the individually calibrated distance decay functions are applied, the function parameters are populated. It should be noted that the default values are those that the software uses automatically, whereas the individually calibrated values are those that have been calculated for the specific distance decay function for each serial crime.
4. The output after running the JTC routine can be saved in two forms – shape file (a density probability map) and text file.
5. These above procedures are iterated twice for each serial crime – first with the default parameter values for each of the five distance decay functions and second with the mathematically calibrated values for the same five distance decay functions. Thus, the JTC GP routine will be executed 10 times (5 distance decay functions times 2 different parameter settings (default and calibrated) for each serial crime. For example, the larceny data set consists

of 56 serial crimes. This will result in 560 (56 x 5 x 2) different JCT geographic profiles. The shape file generated by the JTC routine is a density probability map with each grid cell having a probability score or z-value, indicating the likelihood that a particular location is the offender's residence. This density surface estimating the likely offender's residence is termed a geoprofile (Rossmo, 2000). The highest scored grid cell represents the estimated residence (peak likelihood). These shape files or density maps are brought into ArcMap® 9.2 to perform accuracy assessment of each modeling function. Because a variety of modeling functions (i.e., crime types, distance decay functions, and default or calibrated distance decay parameters) will be examined, it is necessary to measure each technique's effectiveness based on its ability to prioritize a cost-effective search area from which to identify the individual's residence (Canter *et al.*, 2000). Three different accuracy assessments will be applied in this research. They are discussed in the following sections.

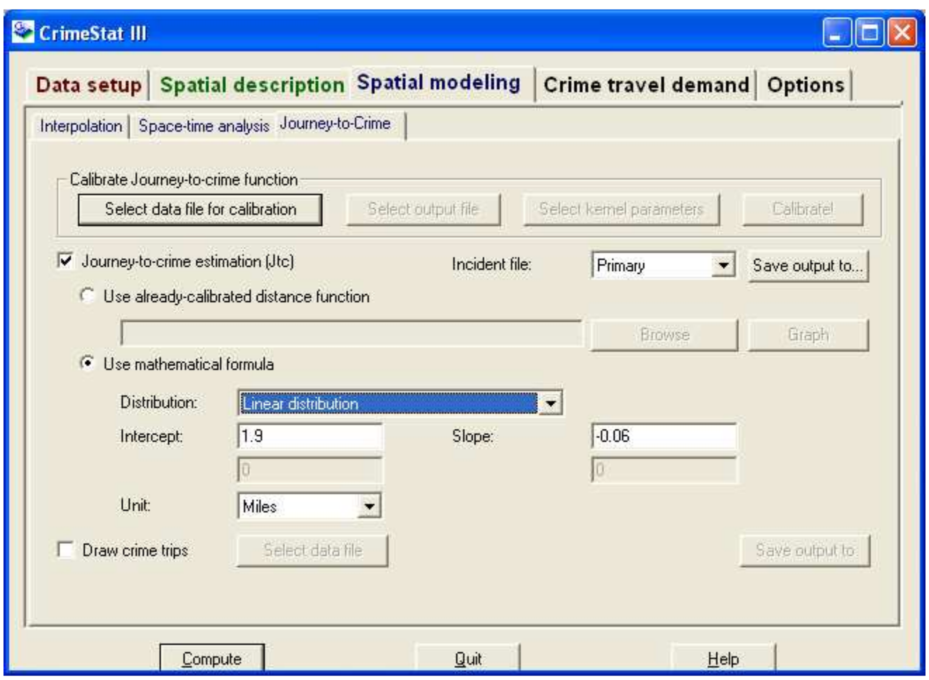


Figure 9. Journey-to-Crime Routine in the Spatial Modeling Window in CrimeStat® 3.1 (Source: Levine, 2007)

3.5 Accuracy Assessment Measures

3.5.1 Euclidean Distance Error

Contemporary journey-to-crime models assess error by measuring the distance between the predicted and the actual residence. Distance error provides a good measure for assessing a geographic profile's spatial precision (Kent, 2006). For this research, the straight-line distance between the grid cell representing the peak likelihood and the grid cell representing the serial offender's actual residence (hit-score) is measured (Figure 10). This is accomplished in the following way:

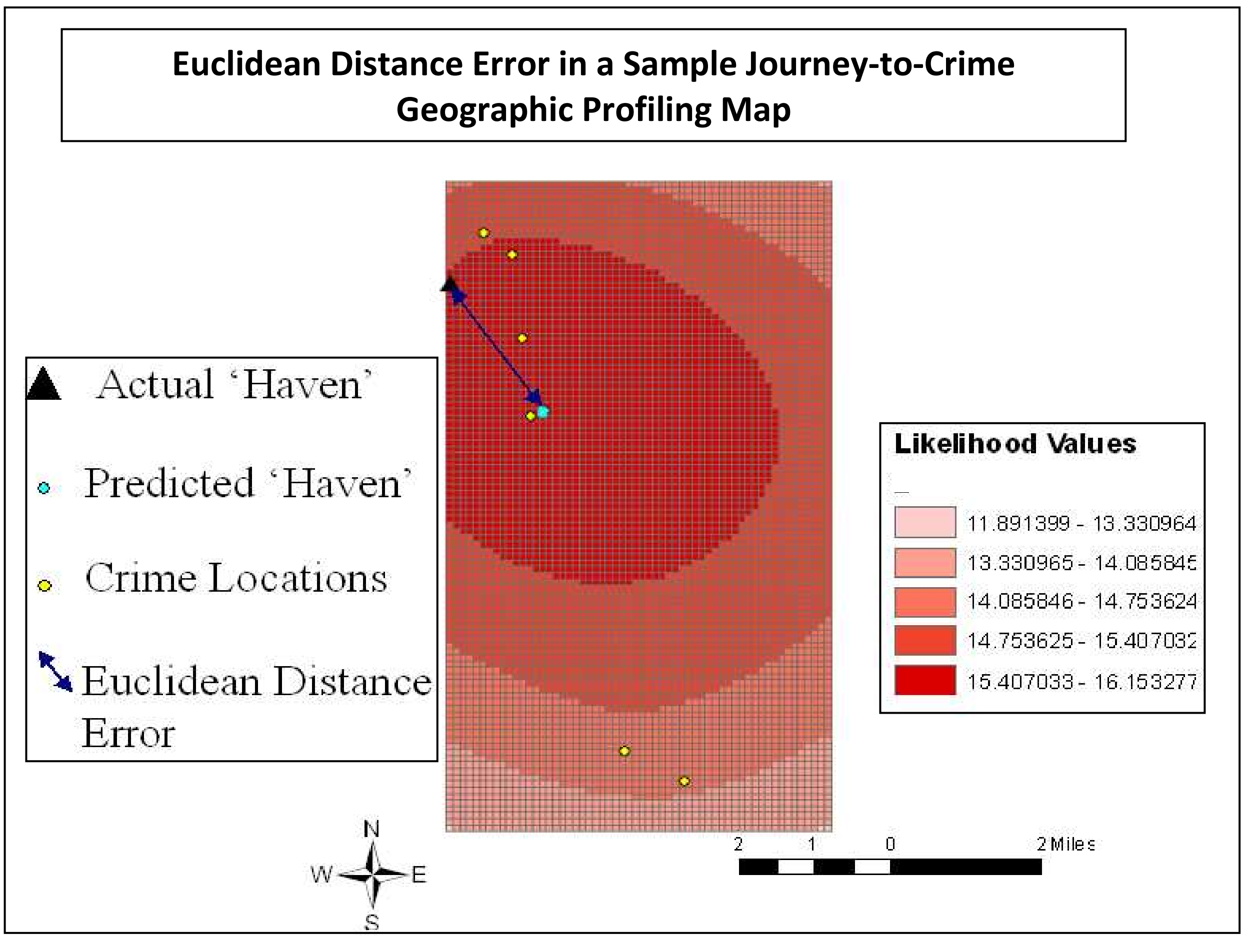


Figure 10. Measuring the Euclidean Distance Error in a sample JTC GP map

The JTC output (a density map) is input into ArcMap® 9.2. Using an existing script in ArcMap® 9.2, the centroid coordinates for each cell of the density map are calculated and the X- and Y-coordinates added to the file. The Euclidean distance is now measured between the centroid of the grid cell with the highest probability (predicted 'haven') and the actual 'haven' of the serial offender (Figure 10). This distance is mathematically calculated in CrimeStat® 3.1 using the Distance Analysis routine.

3.5.2 Top Profile Area

The top profile area – also called the priority search area is a part of the offense domain, where investigators should focus in looking for the home base of an offender. It is the area of all cells with a probability score equal to or higher than the probability score assigned to the actual 'haven' (Figure 11).

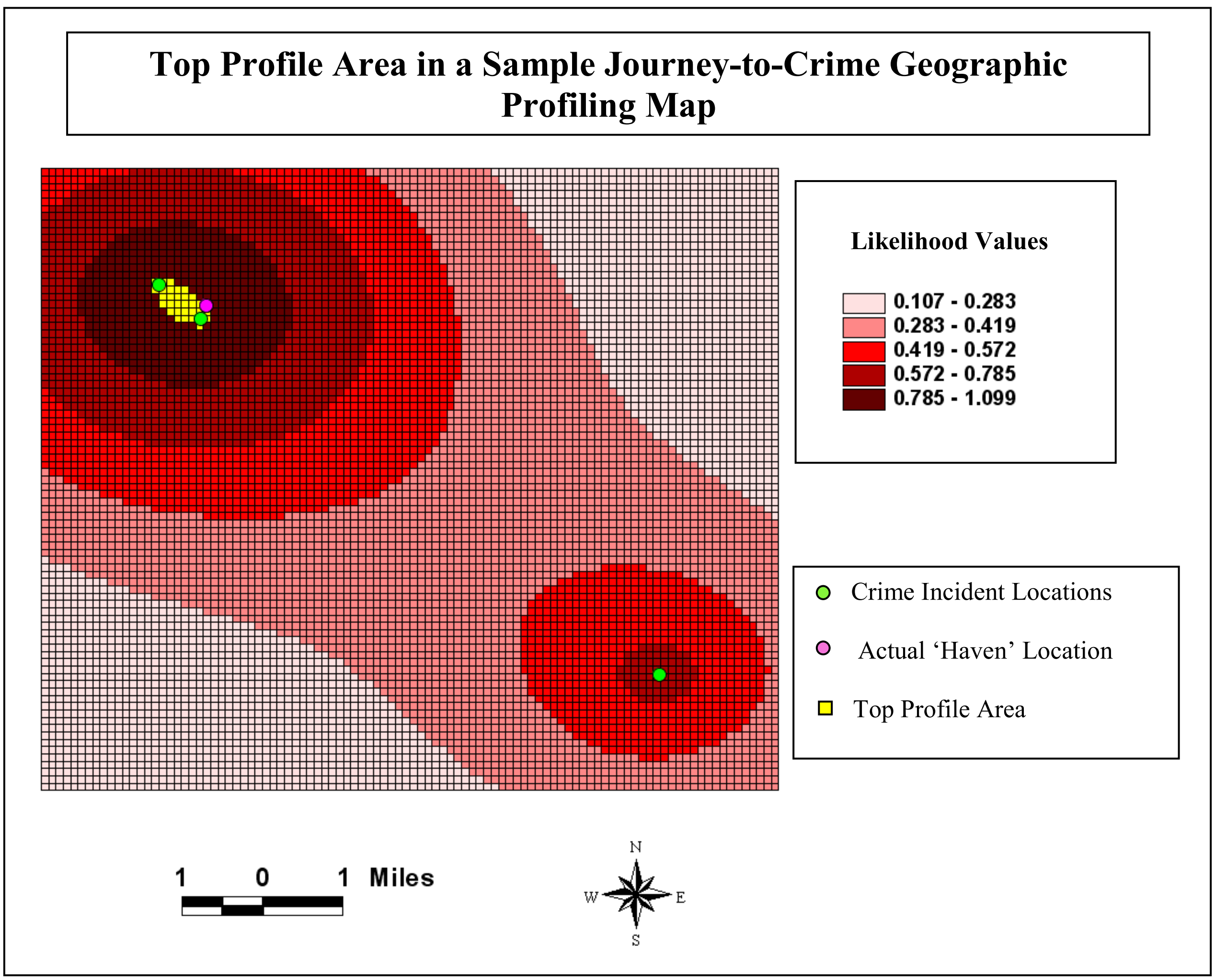


Figure 11. Top Profile Area in a Sample Journey-to-Crime Geographic Profiling Map

The smaller the area, the lesser resources are required to search for the offender, the better the model predicts. This method estimates accuracy by identifying the proportion of the area that must be searched in order to successfully identify the offender's residence.

3.5.3 Hit Score Percentage

Hit score percentage is the ratio of the area searched before the offender's residence is found to the total study area. The search area is estimated using the geographic profiling prioritization where the cells with the likelihood or probability score higher than or equal to the likelihood or probability score of the cell containing the actual 'haven' are only considered (Figure 12). The smaller the ratio, the better the geoprofile's focus and the better the model predicts. A low hit-score percentage indicates a more accurate prediction. The hit-score percentage is the best measure of a geographic profile's predictive utility as there are no intrinsic disadvantages to this measure.

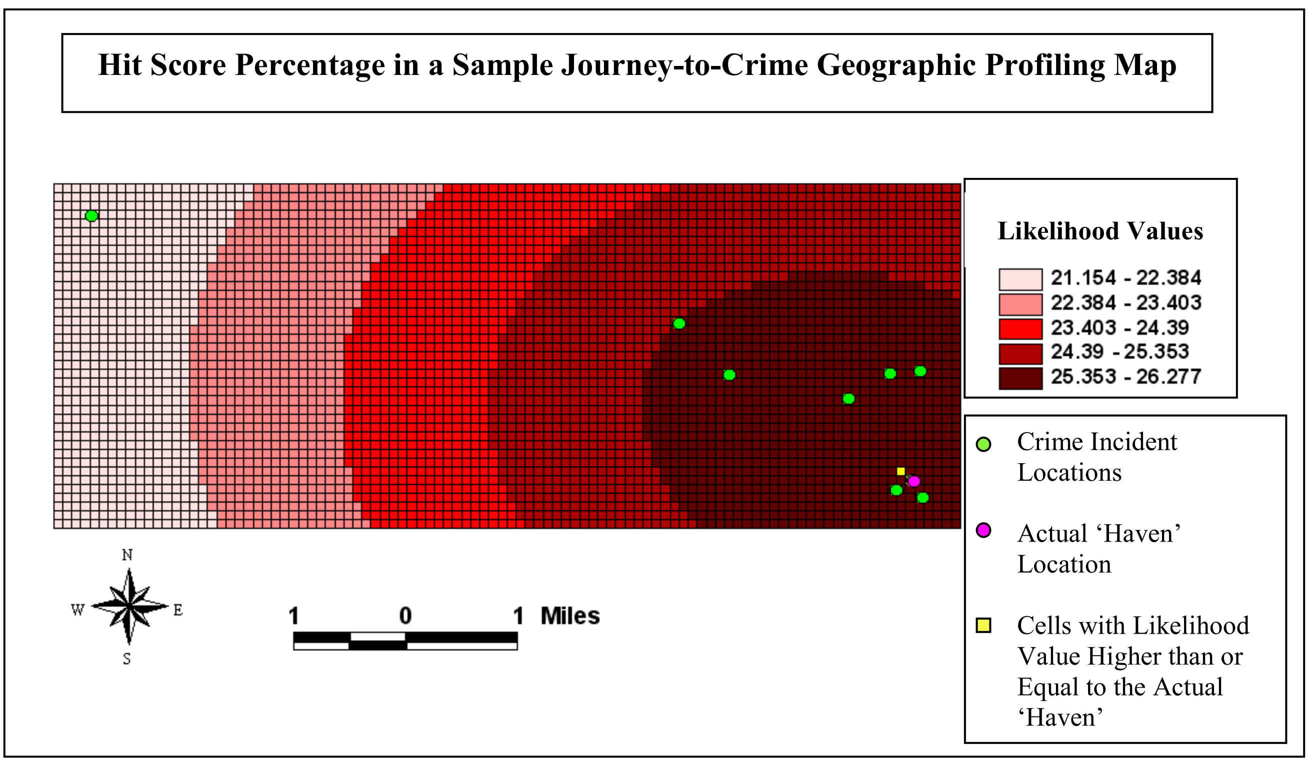


Figure 12. Top Profile Area in a Sample Journey-to-Crime Geographic Profiling Map

CHAPTER 4: RESULTS & DISCUSSIONS

This chapter discusses the results obtained from the analysis. In this analysis two different sets of parameter values for each of the five distance decay functions have been used: first, the default and second, the individually calibrated parameter values. Both sets of parameter values for the five different distance decay functions were applied to each serial crime for each crime type (i.e., auto theft, burglary and larceny). Results are ordered by crime type. For each crime type, the following results are presented:

- *Frequency of distances between the offender's residence and crime locations*

  This is the relative frequency distribution of distances between the offender's residence and crime locations. Frequencies are shown in 0.25 mile intervals. The frequency distribution is derived from the calibration data set. The calibration data set for a crime series of a particular crime type is a collection of distances between each crime incidence location and actual 'haven' location for each case of that same crime type excluding its own. For example auto theft has 31 serial cases. Thus, for auto theft case 1 (with 3 crime incidents), the calibration data set consists of the distance between the crime incidence location and actual 'haven' location for auto theft cases 2 to 31, excluding the distances for the 3 incidents that belong to auto theft case 1. The distance distribution graph helps to determine the distance of peak crime occurrence and also to study how an offender for a particular crime type travels to and from the crime scene. Such behavior will be defined in the form of distance decay functions described in Chapter 3.

- *Linear regression results for best fitting distance decay function*

  Using SPSS® 15.0, a series of univariate regression functions are executed to model the frequency (or the percentage) as a function of distance (i.e., distance decay function). Here frequency is the relative frequency of a crime type committed at a distance from the 'haven'.

The value of R^2 (coefficient of determination) in the regression result determines the best fitting distance decay function. It ranges between 0 and 1 and indicates how much of the dependent variable could be explained by the independent variable. Smaller values of R^2 indicate that the model does not fit the data well. This indicates that other factors, besides the independent variable, exist, that can explain the nature of the dependent variable, but have not been included in the regression analysis model.

- *Maximum, minimum and median values for each accuracy measure*

  The three accuracy measures for assessing the JTC GP methods are Euclidean distance error, hit score percentage and search area. Descriptive statistics, including the minimum, maximum and median values of each of these accuracy measures for both the calibrated and default value distance decay functions will be calculated and presented.

- *Paired sample t-tests results*

  Paired sample t-tests to compare the results of each accuracy measure, calculated from the calibrated and default distance decay functions, with each other. The paired sample t-test is an inferential statistic that assesses whether means of two related groups are statistically different from each other. A low significant value (typically less than 0.05) indicates that a statistically significant difference exists between the two groups. The t test statistic is compared to the critical t-value from the t- table for particular degrees of freedom. For a two-tailed t test if the t-statistics lies between $\pm$critical t-value, then the null hypothesis (the two groups are not statistically significant different) cannot be rejected. In this research, a two tailed paired sample t-test is used.

4.1 Results for Auto Theft Serial Offenders

Figure 13 illustrates the frequency distributions of auto theft for 0.25-mile bin distance interval created from the calibration group data set (31 cases). The very high frequency near the offender's residences supports environmental criminology research results, which indicate that the majority of human activities are performed within close proximity to the offender's home. The trend follows more or less the truncated negative exponential distance decay function. Therefore the frequency of crime increases to a peak distance close to the offender's residence and then rapidly declines. In terms with environmental criminology research, a buffer zone effect is observed. However, the sudden spike at the distance of 40 miles could be explained by the marauder type behavior of offenders, as the data set does not exclude such cases.

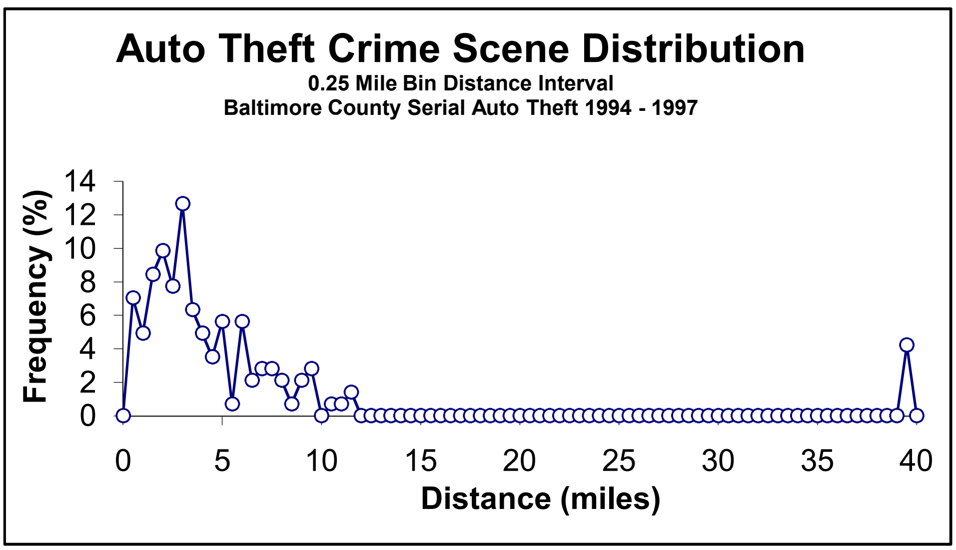


Figure 13. Auto Theft - Crime Scene Distribution using a 0. 25 Mile Bin Distance Interval (Source: Baltimore County Police Department, Maryland)

According to the regression results (Table 9), the best fitting function for auto theft is surprisingly the normal distance decay function since the R^2 value for it is the highest. However, since the R^2 value is not higher than 0.5 for any of the distance decay functions, it could be concluded that the relative frequency of auto thefts cannot be completely explained by the distance from the offender's residence, and that there are other deciding factors. For example, economic, social, or ethnic conditions of the areas near to the ‘haven’, the geography of the area, ease of commuting, road networks, and if the offender’s residence is located in an urban or rural setting may be some of the deciding factors (Quinney, 1966; Clark & Harris, 1992; Rhodes & Conly, 1981).

Table 9. Auto Theft – Summary of Regression Results

Function Type	R square	Df
Linear	0.33 – 0.36	157
Negative Exponential	0.39 – 0.44	157
Truncated Negative Exponential	0.28 – 0.39	157
Normal	0.43 – 0.50	157
Lognormal	0.09 – 0.18	157

It could be concluded from Table 10 that there is no difference in the linear distance decay function for the Euclidean distance error, hit score percentage and top profile area whether the function was individually calibrated or default parameters were used. This conclusion is also supported by the paired sample t-test that resulted in no output / no variance for the linear distance decay function because both the calibrated and the default distance decay functions produced the same results, which made the standard error of the difference zero (Tables 11-13).

Table 10. Auto Theft – Maximum, Minimum and Median Values of Accuracy Measures

Function Type	Euclidean Distance Error (in meters)						Hit Score Percentage (in %)						Search Area (in sq. miles)					
	Calibrated			Default			Calibrated			Default			Calibrated			Default		
	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max
Linear	347	4472	11617	347	4472	11617	3	24	78	3	24	78	0	19	89	0	19	89
Neg Exp	347	4578	12293	347	4472	11968	3	38	88	2	34	78	0	38	168	0	21	104
Trunc Neg Exp	911	4668	11941	438	4025	12545	6	35	97	1	41	98	0	21	155	0	30	161
Normal	911	6693	20687	911	4681	9779	23	78	97	5	31	97	1	78	147	0	27	77
Lognormal	911	6738	15718	578	5003	15374	6	44	98	2	40	90	1	78	147	0	27	77

For the negative exponential, normal and lognormal functions, the minimum and maximum values of the Euclidean distance error, the hit score percentage and the top profile area are lower when using default parameters to estimate the two distance decay functions (Table 10). In contrast, for the truncated negative exponential, the maximum values of the Euclidean distance error, the hit score percentage and the top profile area are higher when using the default parameters. Finally, the normal distance decay function is statistically significantly better for all three accuracy measurements (at $\alpha < 0.05$), when default parameters are used (Tables 11-13).

Table 11. Auto Theft – Paired Sample T-Test Results of Euclidean Distance Error (in meters)

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	4766	4766	30	No variance	N/A
Negative Exponential	5354	5338	30	.544	0.591
Truncated Negative Exponential	5173	4994	30	.375	0.710
Normal	9545	4807	30	4.916	0.000
Lognormal	7181	6508	30	1.353	0.186

Table 12. Auto Theft – Paired Sample T-Test Results of Hit Score Percentage

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	34.34	34.34	14	No variance	N/A
Negative Exponential	43.29	36.31	14	1.772	0.098
Truncated Negative Exponential	43.23	46.69	14	-.0296	0.771
Normal	65.85	43.93	14	2.845	0.013
Lognormal	51.98	39.58	14	1.448	0.17

Table 13. Auto Theft – Paired Sample T-Test Results of Top Profile Area (in sq. miles)

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	34.34	34.34	14	No variance	N/A
Negative Exponential	40.55	30.08	14	2.056	0.0059
Truncated Negative Exponential	32.16	36.98	14	-1.121	0.281
Normal	62.06	30.87	14	3.189	0.007
Lognormal	62.04	30.86	14	3.189	0.007

4.2 Results for Larceny Serial Offenders

Figure 14 illustrates the frequency distributions of larceny for 0.25-mile bin distance interval created from the calibration group data set (59 cases). The significant spike (very high frequency) near the offender's residences corresponds with environmental criminology research results, which indicate that the majority of human activities are performed within close proximity to the home. More than half of all larcenies are committed at a location that is within five miles of the offender's residence. In contrast to environmental criminology research, a buffer zone effect is not observed. As the distances increase, there is a general increase in the number of larcenies committed. This increase continues to a peak distance of about 1- 2 miles from the offender's home. Then the frequency steadily decreases. Thus, the frequency distribution of larceny serial crimes looks more like a truncated negative exponential distance decay function.

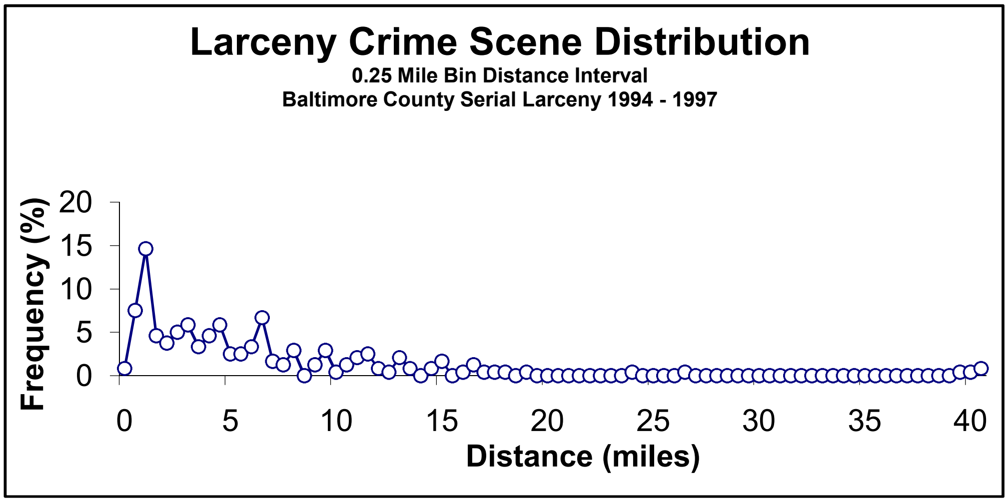


Figure 14. Larceny - Crime Scene Distribution using a 0. 25 Mile Bin Distance Interval (Source: Baltimore County Police Department, Maryland)

According to the regression results (Table 14), the best fitting function for larceny is the negative exponential since the R^2 (the coefficient of determination) value for it is the highest.

Table 14. Larceny – Regression Result Summary

Function Type	R^2	Df
Linear	0.36 – 0.45	160
Negative Exponential	0.45 – 0.73	160
Truncated Negative Exponential	0.42 – 0.47	160
Normal	0.004 – 0.55	160
Lognormal	0.11 – 0.17	160

Thus, 45 to 73 % of the relative frequency of larceny incidents could be attributed to or explained by the distance from the offenders' activity 'haven'. Other influencing factors could be the level of poverty in the area, the degree of tourism, the presence of police, geography and road networks of the area, the unemployment rate and the apprehension rate for larceny (Howsen & Jarrell, 1987; Rhodes & Conly, 1981).

It could be concluded from Table 15 that there is no difference between the Euclidean distance error, hit score percentage and top profile area for linear distance decay function whether it was individually calibrated or default parameters were used   This conclusion is also supported by the paired sample t-tests that resulted in no output / no variance for the linear distance decay function because both the calibrated and the default distance decay functions produced the same results, which made the standard error of the difference zero (Tables 16-18).

Table 15. Larceny – Maximum, Minimum and Median Values of Accuracy Measures

Function Type	Euclidean Distance Error (meters)						Hit Score Percentage (%)						Search Area (sq. miles)					
	Calibrated			Default			Calibrated			Default			Calibrated			Default		
	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max	Min	Med	Max
Linear	11	5288	64394	11	5288	64394	0	29	77	0	29	77	0	6	172	0	6	172
Neg Exp	8	5557	64469	8	5375	64469	0	27	84	0	32	77	0	2	174	0	5	156
Trunc Neg Exp	88	5311	64537	116	5263	63180	2	35	100	0	47	100	0	14	174	0	51	403
Normal	240	10538	62235	240	6397	62235	0.01	20.9	87	0.2	58.3	100	0	3.3	173.7	0.02	12.1	177
Lognormal	18	8242	64494	18	9315	64494	0.01	20.9	87	0.01	21	87	0	3.3	173.7	0.03	3.2	173.7

For the negative exponential function, the minimum and maximum values of the Euclidean distance error, the hit score percentage and the top profile are lower when using default parameters to estimate the distance decay function (Table 15). In contrast, for the truncated negative exponential and normal functions, the minimum and maximum values of the hit score percentage and the top profile area are higher when using default parameters to estimate the two distance decay functions (Table 15).   The lognormal shows little or no difference between the minimum and maximum values of the Euclidean distance error, the hit score percentage and the top profile area. Finally, the normal distance function is statistically significantly better for the Euclidean distance error (at $\alpha < 0.05$) when default parameters are used (Table 16) but it is better for the hit score percentage (at $\alpha < 0.05$) when individually calibrated parameters are used (Table 17). The truncated negative exponential distance decay function is statistically significantly better for the top profile area (at $\alpha < 0.05$) when individually calibrated parameters are used (Table 18).

Table 16. Larceny– Paired Sample T-Test Results of Euclidean Distance Error (in meters)

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	7754	7754	55	No variance	N/A
Negative Exponential	8214	7853	55	0.895	0.375
Truncated Negative Exponential	7885	7777	55	0.574	0.568
Normal	14571	8935	55	5.480	0.000
Lognormal	11251	11236	55	0.157	0.876

Table 17. Larceny – Paired Sample T-Test Results of Hit Score Percentage

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	31	31	27	No variance	N/A
Negative Exponential	30	31	27	-0.090	0.929
Truncated Negative Exponential	45	50	27	-1.447	0.159
Normal	27	55	27	-2.579	0.016
Lognormal	27	28	27	-1.393	0.175

Table 18. Larceny – Paired Sample T-Test Results of Search Area (in sq. miles)

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	93	93	27	No variance	N/A
Negative Exponential	31	27	27	0.428	0.672
Truncated Negative Exponential	31	93	27	-3.999	0.000
Normal	30	34	27	-0.374	0.711
Lognormal	29	30	27	-0.889	0.382

4.3 Results for Residential Burglary Serial Offenders

Figure 15 illustrates the frequency distributions of residential burglary for 0.25-mile bin distance interval created from the calibration group data set (56 cases). The significant spike (very high frequency) near the offender's residences corresponds with environmental criminology research results, which indicate that the majority of human activities are performed within close proximity to the home. In contrast to environmental criminology research, a buffer zone effect is not observed. As the distances increase, there is a general increase in number of residential burglaries committed which continues to a peak distance of about 1- 3 miles from home and then the frequency steadily decreases. Approximately 71% of the residential burglary cases were committed between 0 – 5 miles from the home or haven. The distribution could be categorized as a lognormal distribution, i.e., a normal distribution skewed to the left in this case.

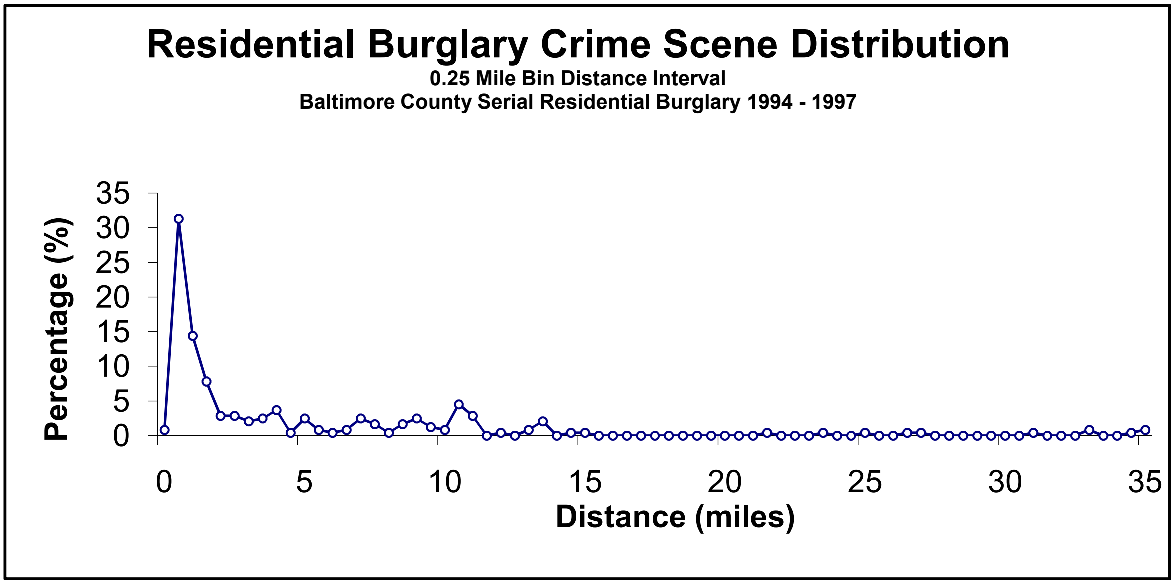


Figure 15. Residential Burglary - Crime Scene Distribution using a 0. 25 Mile Bin Distance Interval (Source: Baltimore County Police Department, Maryland)

According to the regression results (Table 19), the best fitting function for residential burglary is the lognormal since the R^2 (the coefficient of determination) value for it is the highest.

Table 19. Residential Burglary – Regression Result Summary

Function Type	R^2	Df
Linear	0.15 – 0.17	139
Negative Exponential	0.31 – 0.5	139
Truncated Negative Exponential	0.37 – 0.42	139
Normal	0.14 – 0.3	139
Lognormal	0.54 – 0.76	139

Thus, 54 to 76% of variation in the relative frequency of residential burglaries could be explained by the distance from the offenders' activity 'haven'. Other deciding factors could be the economic condition of the areas near to the haven which is related to the amount of booty, geography of the area, social condition of the area – urban or rural population, environmental characteristics like road networks, spatial arrangement of homes or apartments, may be some of the other deciding factors (Quinney, 1966; Clark & Harris, 1992).

It could be concluded from Table 20 that there is no difference between the Euclidean distance error, hit score percentage and top profile area for linear distance decay function whether it was individually calibrated or default parameters were used. This conclusion is also supported by the paired sample t-tests that resulted in no output / no variance for the linear distance decay function because both the calibrated and the default distance decay functions produced the same results, which made the standard error of the difference zero (Tables 21 – 23). For the negative exponential, truncated negative exponential, normal and lognormal functions, there is very little or no difference in the minimum and maximum values of the hit score percentage and the top profile area (Table 20). Finally, only the normal distance decay function is statistically significantly better for the Euclidean distance error (at $\alpha < 0.05$) when default parameters are used (Table 21). For the hit score percentage and the top profile area it could be concluded that the t-test results do not show any statistical difference between the two sets of parameters (Tables 22 & 23). Thus, using the default values for the distance decay function parameters in the Journey-to-Crime routine of CrimeStat® 3.1 will yield the same result as using the individually calibrated values for the same serial crime incidents.

Table 20. Residential Burglary – Maximum, Minimum and Median Values of Accuracy Measures

Function Type	**Euclidean Distance Error (in meters)**						**Hit Score Percentage (in %)**						**Search Area (in sq. miles)**					
	Calibrated			**Default**			**Calibrated**			**Default**			**Calibrated**			**Default**		
	Min	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**	**Min**	**Med**	**Max**
Linear	0	2218	56138	0	2218	56138	0	29	97	0	29	97	0	0	120	0	0	120
Neg Exp	0	2008	12882662	0	2218	56144	0	16	97	0	27	97	0	0	153	0	0	108
Trunc Neg Exp	30	2023	56058	30	2409	55706	2	50	100	2	61	100	0	1	153	0	3	178
Normal	30	4923	54464	30	4079	54464	3	66	100	1	75	100	0	3	293	0	5	169
Lognormal	28	2476	56261	0	2475	56261	0	12	96	0	13	96	0	0	148	0	0	149

Table 21. Residential Burglary - Paired Sample T-Test Results of Euclidean Distance Error

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	7391	7391	55	No variance	N/A
Negative Exponential	237461	7426	55	1.000	0.322
Truncated Negative Exponential	7886	7777	55	0.574	0.568
Normal	10230	7206	55	3.779	0.000
Lognormal	7440	7411	55	0.466	0.643

Table 22. Residential Burglary – Paired Sample T-Test Results of Hit Score Percentage

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	39	39	35	No Variance	N/A
Negative Exponential	33	37	35	-1.571	0.125
Truncated Negative Exponential	48	55	35	-1.354	0.184
Normal	61	63	35	-0.543	0.591
Lognormal	25	25	35	-1.042	0.305

Table 23. Residential Burglary –Paired Sample T-Test Results of Search Area (in sq. miles)

Distance Decay Function	Calibrated (mean)	Default (mean)	Df	T-Test Statistic	Significance (2-tailed)
Linear	23	23	35	No Variance	N/A
Negative Exponential	17	22	35	-1.105	0.277
Truncated Negative Exponential	21	25	35	-1.375	0.178
Normal	38	32	35	0.725	0.473
Lognormal	16	16	35	-0.684	0.499

CHAPTER 5: CONCLUSIONS

With the advent of technological advances, there has been a steady increase in the use of geographic mapping applications in the criminal investigative process to display, analyze and model criminal activities for the last one hundred years (Kent, 2003). This reformed the traditional way of crime mapping (using wall-sized pin maps to survey the occurrence of crime). Today the use of desktop applications that help to 1) identify patterns and concentrations of crime, 2) explore the relationships between crime and environmental or socio-economic characteristics, and 3) assess the effectiveness of policing and crime reduction programs targeted to geographical areas. GIS is being used by thousands of law enforcement agencies across the US, and around the world, to model real-time trends and predict criminal activity using personal computers and hand-held devices for both serial and non-serial crimes. In a survey of 2004 US police departments, 85% of respondents stated that computer mapping was a valuable tool and reported both increasing interest and implementation (Rossmo, 2000).

These technological advances GIS-based crime mapping and GIS can be best exploited in the criminal investigative process for localized serial offenses. Kent (2003, pp. 96) stated that "serial crime incorporates a complex set of psychological and ecological phenomena that requires specialized investigative tools and strategies that extend beyond the traditional criminal investigative processes". Geographic profiling has thus revolutionized the serial crime investigative process by building upon existing environmental criminology theories and traditional mapping techniques to identify key components of an offender's behavior by analyzing the quantitative and qualitative relationship the criminal and his/her target share with the immediate environment and how the offender behaves within his/her activity space. When coupled with journey-to-crime modeling techniques used to quantitatively describe the travel behavior of criminals, geographic profiling can be used by the criminologist to develop new and enhance existing investigative strategies and potentially predict the offender's residence, or 'haven'. Several geographic profiling models, algorithms and software have been developed in recent years to facilitate in the serial criminal investigative process, such as RIGEL, DRAGNET, and Journey-to-Crime of Crime Stat® 3.1.

This research analyzes the accuracy of the Journey-to-Crime algorithm in predicting the 'haven' of a serial offender for property crimes (auto theft, larceny and residential burglary) using five different distance decay functions / models (linear, negative exponential, truncated negative exponential, normal and lognormal). The same set data is divided into two data groups – the calibration group and test data group. In the calibration group all but one serial crime series for a particular crime type are used to individually calibrate the parameters of each of the five distance decay function. The test data group includes only one serial crime. This serial crime is then used to run the previously individually calibrated JTC model in CrimeStat® 3.1. This process is repeated for each serial crime for each of the three different crime types. The GP maps generated from running the JTC model are then analyzed and three different accuracy measures, including the Euclidean distance error, the hit score percentage and the top profile area are calculated. Finally, paired sample t-test analysis is conducted to compare for statistical differences between the default and individually calibrated distance decay functions for all three accuracy measurements.

For auto theft, surprisingly, the default parameters produced significantly better results for the normal distance decay function than when the parameters are individually calibrated. For the other four distance decay functions no significant differences were found. Larceny serial crimes produced similar results as compared to the auto theft serial crimes. Again, the normal function when run with the default parameters resulted in JTC GP that is more accurate for the Euclidean distance error and hit score percentage. In addition, the default parameters produced more accurate results for the top profile area, but only for the truncated negative exponential function. In case of residential burglary serial crimes, using the default parameters when running the normal function for the Euclidean distance error resulted in a more accurate profile when compared to individually calibrated parameters. In general, these results indicate that spending time and resources to individually calibrate distance decay functions may not be necessary for auto theft and residential burglary. In contrast, for larceny, individually calibrated values gave better results than default values for hit score percentage and top profile area. However, the default values are better than the individually calibrated values, if Euclidean distance error is used as an accuracy measure for any larceny JTC GP. Thus, for both auto theft and larceny the null hypothesis is rejected but for residential burglary the null hypothesis is accepted.

However, this research has some limitations. First, this analysis was conducted with serial offense data from a single study area (Baltimore County). Additional research should compare these results with similar research from other differently structured study areas. Second, the analysis does not differentiate between marauder and commuter type offenders. A marauder type offender moves out from his home or base ('haven') to commit his crimes and then returns to the base, going out on different directions on different occasions (Canter & Gregory, 1994). A commuter type offender, on the other hand, travels from his home or base ('haven') into a selected area from which he moves out when travelling to his offence venue or venues (Canter & Gregory, 1994). Since a marauder exhibits a different activity nature than a commuter, more accurate results could be obtained by redoing the analysis with marauder type offenders only. Third, a better analysis of the data could be achieved by incorporating the road networks for the study area. Thus, a network analysis could be a very interesting extension to this research to understand the influence of the geography of the area on the movement of serial offenders and also on the occurrence of crime in those areas. For future research, this analysis could also be performed with the Bayesian JTC routine recently implemented in CrimeStat® 3.1, which will be briefly explained below.

The Bayesian JTC is an extension of the distance based Journey-to-Crime routine (JTC). Unlike the five JTC functions described above that uses a typical travel distance algorithms to predict the likely residence location of serial offenders, the Bayesian JTC routine is based on the Bayes theorem. The routine involves the use of an origin-destination matrix of an offender for particular origins ('haven') and destinations (where the crime is committed) (Levine, 2007). Bayes theorem states the relationship between the conditional and marginal probability distributions of random variables. The marginal probability or the normal probability of variables, such as A and B are P(A) and P(B) is independent of any other conditions. The conditional probability, on the other hand is the probability of an event which is dependent on the occurrence of some other event. Thus for A, it could be written as P(A|B), i.e., event A given that event B has occurred. In probability theory it is defined as:

$$P(A|B) = \frac{\mathrm{P\ (A\ and\ B)}}{\mathrm{P(B)}} \qquad (3.16)$$

The statistical interpretation of the Bayes theorem where the probabilities are estimates of a random variable would result in the following equation:

$$\mathrm{P}(\emptyset|\mathrm{X}) = \frac{\mathrm{P}(\mathrm{X}|\emptyset) * \mathrm{P}(\emptyset)}{\mathrm{P}(\mathrm{X})} \qquad (3.17)$$

Where $\emptyset$ is the parameter of interest, X is some data and P(X) is the spatial distribution of all crimes. In the JTC framework, P($\emptyset$) is the probability where the offender lives for a particular location $\emptyset$.

The matrix is created by imputing information from a sample of known offenders, where both the crime locations ('destinations') and the residence locations ('origins') are known. These locations are then assigned to a set of zones to produce an origin-destination or a trip distribution matrix. For example, if we have a certain distribution of incidents committed by a particular serial offender, we can use the origin-destination matrix for that particular offender to predict the likely origin zones that the offender lives, independent of any assumptions about travel distance. Thus, it improves the estimate of the likely location of a serial offender by updating the estimate from the JTC methods, $P(\emptyset)$, with information from an empirically-derived likelihood estimate, $P(X|\emptyset)$. Therefore the Bayes JTC extension is an improvement to the existing JTC travel distance methods since the Bayes theorem can be used to create an estimate that combines information both from a travel-distance function and an origin-destination matrix. Rewriting equation in the JTC terms,

$$P(Jtc|O) = \frac{P(O|Jtc) * P(Jtc)}{P(O)} \tag{3.18}$$

Where:

P(JTC|O) = an estimate of the residence location of a single offender based on the distribution of offenders given the distribution of incidents committed by the single offender.

P(JTC) = an estimate of the residence location of a single offender based on the location of the incidents that the offender committed and an assumed travel distance function.

P(O) = an estimate of the resident location of a single offender based on a general distribution of all offenders, irrespective of any particular destinations for incidents.

Geographic profiling is a probability oriented methodology and thus the more number of crime incidents in a series, the better the GP model can predict the likely location of the offender's 'haven'. If applied correctly, GP can immensely contribute to the acquisition of a serial offender, however geographic profiling alone cannot solve a crime.

REFERENCES

Anselin, L., Cohen, J., Cook, D., Gorr, W., & Tita, G., (2000). Spatial Analysis of Crimes. Measurement and Analysis of Crime and Justice, United States Department of Justice, Office of Justice Programs, Washington D.C., 4: 213-263.

Baldwin, J. & Bottoms, A. E., (1976). The urban criminal: a study in Sheffield. Tavistock Publications, London.

Bernasco, W., (2006). Co-offending and the Choice of Target Areas in Burglary. Journal of Investigative Psychology and Offender Profiling, Wiley InterScience, 3: 139 – 155.

Brantingham, P.J., & Brantingham, P.L., (1981). Environmental Criminology. Sage Publications, Beverly Hills, CA.

Brantingham, P.J. & Brantingham, P.L., (1984). Patterns in Crime. Macmillan Press, New York, NY.

Brown, B. B. & Altman, I., (1981). Territoriality and residential crime. In P. J. Brantingham & P. L. Brantingham (Eds.), Environmental Criminology: 56-76, Sage Publications, Beverly Hills, CA.

Canter, D. & Larkin, P., (1993). The environmental range of serial rapists. Journal of Environmental Psychology, Elsevier, 13: 63-69.

Canter, D. & Gregory, A., (1994). Identifying the residential location of rapists. Journal of the Forensic Science Society, 34: 169-175.

Canter, D., Coffey, T., Huntley, M., & Missen, C., (2000). Predicting serial killers' home base using a decision support system. Journal of Quantitative Criminology, Plenum Publishing Corporation, 16(4): 457-478.

Capone, D. L. & Nichols, W. W., (1975). Crime and Distance: An Analysis of Offender Behavior in Space. Proceedings of the Association of American Geographers, 7: 45-49.

Chainey, S. & Ratcliffe, J., (2005). GIS and Crime Mapping. John Wiley & Sons Ltd.

Clark, A. N. (1990). The New Penguin Dictionary of Geography. Penguin Books, London.

Clarke, R. & Cornish, D., (1985), Modeling offender's decisions: a framework for policy and research. In Tonry, M. & Morris, N. (Eds),Crime and Justice: An Annual Review of Research (pp 147-185). University of Chicago Press, Chicago, IL.

Cornish, D. B. (1993). Theories of action in criminology: Learning theory and rational choice approaches. In R. V. Clarke & M. Felson (Eds.), Routine activity and rational choice (pp. 351–382). Transaction, New Brunswick, NJ.

Crime Mapping Research Center, (1999). Mapping Crime: Principle and Practice.

Environmental Systems Research Institute, Inc., (2002). ArcMap® version 9.2 for Windows®. Redlands, CA.

Federal Bureau of Investigation (2006). Offense Definitions – Crimes in the United States 2006. Published online: http://www.fbi.gov/ucr/cius2006/about/offense_definitions.html

Felson, M. & Clarke, R., (1998). Opportunity Makes a Thief, Practical theory for crime prevention. Police Research Series, Paper 98. Home Office, Policing and Reducing Crime Unit, Research, Development and Statistics Directorate, London.

Gabor, T. & Gottheil, E. (1984). Offender characteristics and spatial mobility: An empirical study and some policy implications. Canadian Journal of Criminology, 26: 267-281.

Georges**,** D.E. (1978). The geography of crime and violence: a spatial and ecological perspective, Association of American Geographers: Resource papers for college geography, 78 (1).

Godwin, M. & Canter, D., (1997). Encounter and death: The spatial behavior of U. S. serial killers. Policing: An International Journal of Police Strategy and Management. 20(1): 24-38.

Gore, R. & Pattavina, A., (2001). Linking Offender Residence Probability Surfaces to a Specific Incident Location: An Application for Tracking Temporal Shifts in Journey to Crime Relationships and Prioritizing Suspect Lists and Mug Shot Order.. Paper presented at the 5th Annual International Crime Mapping Research Conference. April 2001. Published online: http://www.ojp.usdoj.gov/nij/maps/.

Gore, R. & Tofiluk, N., (2002). The Use and Effectivness of Distance-Decay Curves and Incident Based Offender Residence Probability Surfaces in Ranking an Offender on a Randomly Generated Suspect List. Annual International Crime Mapping Research Conference, Denver, CO., December 2002.

Harries, K., (1999). Mapping Crime, Principle and Practice. United States Department of Justice, Crime Mapping Research Center (CMRC), Washington D.C.

Hesseling, R. B. P., (1992). Using data on offender mobility in ecological research. Journal of Quantitative Criminology, 34: 95-112.

Howsen, Roy M. & Jarrell, Stephen B., (1987). Some Determinants of Property Crime: Economic Factors Influence Criminal Behavior but Cannot Completely Explain the Syndrome. American Journal of Economics and Sociology, 46(4): 445-457.

Kent, Joshua D., (2003). Using functional distance measures when calibrating journey-to-crime distance decay algorithms. Thesis: Louisiana State University. http://etd.lsu.edu/docs/available/etd-1103103-095701/

Kent, Joshua D., Leitner, M. & Curtis, A., (2006). Evaluating the usefulness of functional distance measures when calibrating journey-to-crime distance decay functions. Computers, Environment and Urban Systems 30(2): 181-200.

LeBeau, J., (1987). The Methods and Measures of Centrography and the Spatial Dynamics of Rape. Journal of Quantitative Criminology, 3: 125-141.

Levine, N., (2007). CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 3.1). Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC, March.

Levine, N., (2007). Crime Stat® 3.1 Manual. Ned Levine and Associates. Published online: http://www.icpsr.umich.edu/CRIMESTAT/files/CrimeStatChapter.10.pdf

Levine, N., (2002). Modeling Metropolitan Criminal Travel Behavior. Crime Mapping Research Conference, Denver, CO, December 10, 2002.

Lottier, S., (1938). Distribution of Criminal Offences in Sectional Regions. Journal of Criminal Law, Criminology, and Political Science, 29(3): 329-344.

McGrew, J. C. & Monroe, C., (1993). An Introduction to Statistical Problem Solving in Geography. Wm C. Brown Publisher, Dubuque, IA.

Paulsen, D., (2004). Geographic Profiling: Hype or Hope? Second UK Crime Mapping Conference, March 2004, London.

Paulsen, D., (2006). Connecting the Dots. Policing: An International Journal of Police Strategies & Management, 29(2): 306 – 334.

Phillips, Phillip D., (1980). Characteristics and topology of the journey to crime. In D. E. Georges-Abeyie & K. D. Harries (Eds.), Crime: A Spatial Perspective, pp. 169–80. Columbia University Press, New York.

Ramsland, K., (2005). Random or Planned? The Zodiac Killer. Crime Library. Published online: http://www.crimelibrary.com/forensics/geog/.

Ratcliffe, Jerry H., (2003). Suburb Boundaries and Residential Burglaries. Trends and Issues in Crime and Criminal Justice, March 2003, 1-6.

Reber, A. S., (1985). The Penguin dictionary of psychology. Penguin Books, London.

Rengert, George F., (1980). Spatial Aspects of Criminal Behavior. In D. E. Georges-Abeyie & K. D. Harries (Eds.), Crime: A Spatial Perspective, pp. 47 – 57. Columbia University Press, New York.

Rengert, George F., (1981). Burglary in Philadelphia: A critique of an opportunity structure model. In Brantingham and Brantingham (Eds.), Environmental Criminology, (pp. 189-201). Sage, Beverly Hills, CA.

Rengert, G., Piquero, A. & Jones, P., (1999). Distance Decay Reexamined. Criminology, 37 (2): 422-445.

Rhodes, W.M. & Conly, C., (1981). Crime and mobility: An empirical study. In Brantingham & Brantingham (Eds.), Environmental Criminology, pp 167-188. Waveland Press, Inc., Prospect Heights, IL.

Rossmo, Kim D., (2000). Geographic Profiling. CRC Press, Boca Raton, FL.

Rossmo, Kim D., (2003). Environmental Criminology Research Inc. Corporate web site, Canada. Published online: http://www.ecricanada.com/

Sherman, L.W., Gartin, P. R. & Buerger, M. E., (1989). Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place. Criminology, 27 (1): 27.

Snook, B. (2004). Individual Differences in Distance Traveled by Serial Burglars. Journal of Investigative Psychology and Offender Profiling, 1:53-66.

Snook, B., Taylor, P. & Bennell, C. (2004). Geographic Profiling: The Fast, Frugal, and Accurate Way. Applied Cognitive Psychology, 18: 105-121.

SPSS, (2001). SPSS for Windows®, Release 11.0.1 (15 Nov-2001). SPSS Inc., 1989- 2001.

Taras, G. (1996). The Geography of Crime. Geographical, 68 (9): 26-27.

Turner, S., (1969). Delinquency and distance. In T. Sellin, M. E. Wolfgang (Eds.), Delinquency: Selected studies, pp. 11-26. John Wiley, New York.

Van Koppen, P. J. & Jansen, R. W. J., (1998). The road to robbery: travel patterns in commercial robberies. British Journal of Criminology, 38: 230-246.

Wiles, P. & Costello, A., (2000). The 'road to nowhere': the evidence for traveling criminals. Home Office, Research, Development and Statistics Directorate, London.

Zipf, G.K., (1949). Human Behavior and the Principle of Least Effort. Cambridge, MA.

VITA

Tania Pal was born in Calcutta, India. She received the Bachelor of Science in Economics from University of Calcutta in 2001. Her interest in information technology motivated her to pursue extensive post graduate courses in software development and applications after undergraduate studies. During 2003 – 2004, she taught computer science to middle and high school students in a renowned school in Calcutta and worked as a database manager for an accounting firm. Tania began her studies in Louisiana State University in Spring of 2005. She also interned as a programmer/database analyst with Environmental Systems Research Institute (ESRI) Inc., Redlands in Summer 2007 and plans to join the company as a full time database analyst after her graduation in December 2007.