12-2013

Spatial Analysis of Fatal Automobile Crashes in Nashville, TN, 2001-2011

Yan Chen
Western Kentucky University, yan.chen137@topper.wku.edu

SPATIAL ANALYSIS OF FATAL AUTOMOBILE CRASHES IN NASHVILLE, TN, 2001-2011

A Thesis
Presented to
The Faculty of the Department of Geography and Geology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Yan Chen

December 2013

SPATIAL ANALYSIS OF FATAL AUTOMOBILE CRASHES
IN NASHVILLE, TN, 2001-2011

Date Recommended 12/10/2013

Jun Yan, Director of Thesis

David Keeling

Kevin Cary

12-11-13

Dean, Graduate Studies and Research   Date

ACKNOWLEDGMENTS

It would have been impossible to complete this thesis without the help and the support that I have received. Foremost, I would like to express my sincere gratitude to my advisor, Dr. Jun Yan, for his patience, motivation, and careful guidance. I would also like to thank the rest of my thesis committee, Dr. David Keeling and Prof. Kevin Cary, for their encouragement and insightful comments. My sincere thanks also belong to all the professors who taught me over the past two years in the Geography and Geology Department, for offering me such specialized knowledge.

Last, but not the least, I would like to thank my family and my friends for always giving me encouragement and helping me get through the tough times during my entire graduate study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SPATIAL ANALYSIS OF FATAL AUTOMOBILE CRASHES
IN NASHVILLE, TN, 2001-2011

Yan Chen December 2013 118 Pages

Directed by: Dr. Jun Yan, Dr. David Keeling, and Kevin Cary

Department of Geography and Geology Western Kentucky University

With increasing levels of motor vehicle ownership, automobile crashes have become a serious public issue in the U.S. and around the world. Knowing when, where, and how traffic accidents happen is critical in order to ensure road safety and to plan for adequate road infrastructure. There is a rich body of literature pertaining to time-related fatal crashes, most of which focuses on non-spatial factors such as a driver's visibility at night, drinking and drug use, and road conditions. These studies provide a theoretical basis for understanding the causes of crashes from a non-spatial perspective, and a number of traffic laws and policies consequently have been enacted to minimize the impacts of non-spatial factors. Over the past a few years, advances in Geographic Information Systems (GIS) have greatly enhanced our ability to analyze traffic accidents from a spatial perspective. This study aims to fill a void in traffic safety studies by comparing and analyzing the differences in the spatial distribution of fatal crashes based on temporal factors, specifically in three periods: 1) day and night; 2) A.M. rush hours and P.M. rush hours; and 3) weekdays and weekends. With the Nashville Metropolitan Area as the study area, the research utilized a number of spatial point-pattern analysis (SPPA) methods, including planar KDE, planar global auto K function, network global cross K functions, and network local cross K functions.

All fatal crashes in the Nashville area were found to be clustered and generally follow the patterns of average daily traffic flow. All time-based subtypes of fatal crashes also were found to be concentrated within the central urban area of Nashville, mostly along major roads, and especially near major road intersections and highway interchanges. No notable spatial differences were detected among the subtypes of fatal crashes when applying network global cross K function. However, with the help of the network local cross K function, some localized spatial differences were identified. Some specific locations of hotspots of nighttime and P.M. rush hour fatal crashes were found not to be at the same locations as those at of daytime and A.M. rush hour fatal crashes, respectively.

The approach adopted in this study not only provides a new way to analyze spatial distribution of spatial point events such as fatal crashes, but it also can be applied readily to real-world applications. A good understanding of where these spatial differences are should help various agencies practice effective measures and policies in order to improve road conditions, reduce traffic accidents, and ensure road safety.

CHAPTER 1. INTRODUCTION

According to the United Nations (UN, 2012), in 2011 more than half of the world's population lived in cities and, by 2050, the world urbanization rate is expected to reach around 70%. Due to the increased level of urbanization and car ownership, many urban traffic problems have became ever more serious in global cities, including those in the U.S. Too much dependency on personal vehicles, plus misguided land use and road planning and construction practices, could easily lead to severe road congestion, unnecessary energy consumption, serious air pollution, and frequent traffic accidents. According to Ewing et al. (2003), urban sprawl is one of the major contributors to the increase in automobile crashes in the U.S. As roads become more widespread, the miles driven much higher, and traffic congestion more frequent, people who live further away from urban centers have to spend increasingly more time to get to places. With more and more people living in cities around the world, a higher demand for mobility in urban areas, longer and congested commuting trips, and increased dependency on personal vehicles inevitably lead to greater exposure to traffic dangers, and thus increase drivers' risk of being involved in automobile crashes. A clearer understanding of the spatial distribution of fatal crashes could guide decision makers in implementing effective measures to improve road safety by adopting proper land use practices, building efficient transport systems, and formulating appropriate traffic policies and laws.

Traffic-related injuries and fatalities have become a serious global public health issue. Each year around the world, it is estimated that over one million people are killed and 50 million people are injured because of automobile crashes (WHO, 2013). In the U.S., automobile crashes claimed more than 32,000 lives in 2011, a number that has since dropped from the peak (about 44,000 annually) in 2005 (FARS, 1994-2011). The rate of automobile fatalities per mile driven has steadily fallen for the past few years as well, due to safer cars and roads, increased safety measures, and stricter laws that prohibit drunk driving and cell phone use. The cost of automobile crashes, however, remains high, as fatal automobile crashes and traffic-related injuries have significant impacts on traffic management, the national economy, public health-service systems, and affected families. When taking into account several consequences of fatal crashes, such as medical expenses, lost wages, and property damage, etc., the National Safety Council (NSC, 2003) estimated that each fatal crash can result in a direct economic loss of $5.4 million.

Many studies found that time is one of the main factors in the occurrence of traffic accidents and fatal crashes (Opiela et al., 2003; Farmer and Williams, 2005; Hu et al., 2011). These studies largely focus on three main areas: 1) the impact of drivers' visibility in dark conditions; 2) the roles of drinking, drug use, and distractions in crashes; and 3) the impacts of certain roads on crashes (e.g., near bars serving alcohol, road intersections, sharp curves, and bad road conditions with few streetlights). Only a handful of studies examine directly the spatial distribution of time-based fatal crashes (Xie and Yan, 2008; Anderson, 2009; Dai, 2012). Not surprisingly, many of them tend to focus on detecting the hotspots of different subgroups of fatal crashes based on a single factor, such as time of day or day of the week. Knowing when, where, and how traffic accidents happen is indeed critical to ensure road safety and to plan for adequate and improved road infrastructure. In general, the spatial distribution of automobile crashes tends to exhibit discernible spatial-temporal patterns, because they are highly related to people's daily activities, and different traffic volume along certain roadway segments contributes to unique temporal and spatial patterns of traffic accidents (Black, 1991). Being able to compare the spatial differences of different types of fatal crashes based on time factors would help us to understand better the causes behind many fatal crashes. This study aims to fill a void in traffic safety studies by comparing and analyzing the differences in the spatial distribution of fatal crashes based on temporal variables, particularly in three periods: 1) day and night; 2) A.M. rush hours and P.M. rush hours; and 3) weekdays and weekends. Using the metropolitan area of Nashville, Tennessee, one of the most sprawled cities in the U.S., as the study area, this research intends to achieve the following three objectives:

1. **Objective I** is to identify the spatial patterns of fatal automobile crashes in Nashville, TN, from 2001 to 2011. It is hypothesized that fatal crashes, no matter when they occurred, tend to cluster in areas with high traffic volume and tend to locate along major roads, particularly near major intersections and highway interchanges. This is because the risk of car crashes and fatalities would increase when drivers need to make complex decisions on which way to turn, when to turn, when to yield, which lane to merge into, when to shift lanes, etc.
2. **Objective II** is to identify if there are any differences in the spatial distribution of fatal crashes by examining their occurrence at different times. It is hypothesized that we should expect more hotspots of daytime fatal crashes than those of nighttime fatal crashes, since there are more activities and traffic flow during the daytime. Second, we should observe some similarities in the clustering of A.M. rush hour and P.M. rush hour fatal crashes because people’s working hours and commuting routes usually exhibit regular patterns in space and time. Hotspots of

weekday fatal crashes should have noticeable spatial differences compared to weekend fatal crashes, simply because two different types of activities (e.g., work-related and leisure-related) occur during weekdays and weekends respectively. For instance, hotspots of weekend fatal crashes are expected to be located near shopping centers and areas such as bars and restaurants where alcohol is served.

3. **Objective III** is to find out where discernible spatial differences between subtypes of fatal crashes are likely to be detected; e.g., daytime vs. nighttime, A.M. vs. P.M. rush hour, and weekday vs. weekend fatalities. To be specific, the goal is to determine statistically whether the hotspots of one subtype of fatal crashes are also the hotspots for its corresponding subtype.

Over the past few years, advances in Geographic Information Systems (GIS) and spatial analysis have greatly enhanced our ability to analyze traffic accidents from a spatial perspective. To achieve these objectives, fatality data from 2001 and 2011 were obtained from the Fatality Analysis Reporting System (FARS, 1994-2011). The research used the Kernel Density Estimation (KDE), one of the most commonly used spatial point-pattern analysis (SPPA) methods, to detect visual hotspots of fatal crashes. However, due to the inherent limitation of KDE, spatial patterns of fatal crashes cannot be compared directly from a statistical perspective. To solve this problem, this study adopted an integrated approach by utilizing several SPPA methods together, including planar KDE, planar global auto K function, and network cross K functions.

CHAPTER 2. RESEARCH BACKGROUND

There is a rich body of literature pertaining to time-related fatal crashes, most of which focuses on non-spatial factors such as drivers' visibility at night, drinking and drug use, and road conditions. This type of research provides a theoretical basis for understanding the causes of traffic accidents and fatal crashes from a non-spatial perspective, and a number of traffic laws and policies thus were established to minimize the impacts of non-spatial factors. With advances in GIS, the analysis of traffic accidents and fatal crashes has been greatly enhanced over the past a few years. Armed with powerful spatial analytical tools and GIS, it is now possible to investigate more accurately the spatial distribution of traffic accidents and fatal crashes. This chapter briefly reviews the studies of direct relevance to this thesis research, namely: 1) the effects of urbanization and suburbanization; 2) traffic safety studies; 3) GIS and spatial analysis; and 4) network extension of GIS and spatial analysis.

2.1. Urbanization, Suburbanization, and Their Effects

Urbanization refers to the increasing number of people residing in urban areas. Currently, about 80% of the U.S. population lives in urban areas. Many cities in the U.S. are experiencing so-called urban sprawl. Urban sprawl, also known as suburbanization and urban decentralization, is often associated with the following characteristics: (1) people are widely dispersed in suburban development areas; (2) homes, shops, and workplaces are rigidly separated in space; (3) there is no distinct, thriving activity center; and (4) the network of roads is highly developed (Ewing et al., 2003).

Urban sprawl has a great influence on mobility and accessibility. Mobility is defined as the ability to move between different activity sites in an urban area, while accessibility stands for the number of activity sites within a certain space-time constraint (Gutiérrez et al., 2010). Overall, urban sprawl leads to improved urban mobility for many urban dwellers by providing well developed street networks and multiple travel mode choices (e.g., cars, light rails, and subways) and by shortening travel time (García-Palomares, 2010). However, certain segments of vulnerable urban populations (e.g., low-income families, minority, women, and the elderly) often are left in disadvantaged positions due to the lack of a reliable means of transportation. In addition, increased traffic flows could result in frequent traffic congestion in urban centers, and therefore accessibility to downtown areas and to other activity centers could be reduced (Vandenbulcke et al., 2009). Moreover, travelers are exposed to increased air pollution and longer waiting times brought about by frequent traffic jams. As a result, more and more people have chosen to move from crowded urban centers to those suburbs with higher accessibility, especially where multiple suburban centers are found (Kotavaara et al., 2011).

Without doubt, the spatial distribution of automobile crashes is directly related to urban land use distribution and urban sprawl. General speaking, living in or near urban centers should provide less exposure to the risk of being involved in traffic accidents compared with living in suburbs and rural areas, for two reasons. First, fewer vehicle miles traveled would be required in urban centers, since there are many different activity centers within a relatively shorter distance. Second, due to traffic congestion, the speed of vehicles is slower, so that the risk of fatal accidents is lower too (Ewing and Dumbaugh, 2009). In contrast, driving on rural roads seems more dangerous. With urban sprawl, there are more roads, the miles driven are longer, and traffic congestion is more severe and more frequent, such that people who live further away from urban centers have to spend more time to get to places. For them, walking to activity centers in some locations is not feasible. This could explain why people have to depend more and more on automobiles. Consequently, a higher demand for travel and mobility and an increased dependency on automobiles lead to greater exposure to traffic dangers, and thus increase the drivers' risk of being involved in automobile crashes.

2.2. Traffic Safety Studies

Driving is a visual task. Even though the screening test is required during the application for a driver's license, passing it does not mean that the driver can control a vehicle effectively at all times, especially under risky conditions such as at night with poor lighting, difficult weather conditions, or trying to navigate through difficult or unfamiliar environments. A study by the Federal Highway Administration (FHA) reported that, in the U.S., the crash rate during nighttime is about 1.6 times of that during daytime, while the fatal crash rate is nearly four times greater at nighttime (Opiela et al., 2003). The lack of daylight makes nighttime driving different from daytime driving, and it increases the risk of being involved in traffic accidents for drivers. Due to the reduced visibility caused by dark conditions, drivers have to pay more attention to road conditions and more reaction time is needed to make decisions (Plainis and Murray, 2002). When faced with special sections of roads, such as road intersections, sharp curves, reduced streetlights, and fuzzy road signs, driving at night becomes more dangerous, especially for novices who lack driving experience (Konstantopoulos et al., 2010).

Many studies have revealed the role of vision in nighttime driving, and mainly focused on three aspects of visual functions: visual field, reaction time, and visual search strategies. Completing test rides in driving simulators is one of the most commonly used methods in the research of visual functions. The findings of this method indicate that loss of visual field (Kim et al., 2010), longer reaction time (Plainis and Murray, 2002), and poor visual search strategies (Konstantopoulos et al., 2010) are the main factors that may influence drivers' nighttime driving performances. However, some researchers argued that large individual differences exist in the research that may influence the reliability of results (Evans, 2004). In order to avoid the limitation of subject selection problems, researchers began to consider the performance of different groups of drivers under the same simulated driving conditions, such as young and old (Chrysler et al., 1996), novice and expert (Konstantopoulos et al., 2010), and disabled and normal vision (Owsley and McGwin, 1999). As a goal of improving driver's nighttime vision, some traffic management approaches, such as enhancing the road light level, changing vehicles' front light systems, and improving route markings, have also been examined (Carlson et al., 2009; Isebrands et al., 2010; Kim et al., 2010).

It is widely agreed by many researchers that a driver's behavior, such as drunk driving and distracted driving, is linked to the occurrence and severity of motor vehicle crashes. Alcohol impairment, which is frequent during nighttime, at weekends, and on certain days of the year, can contribute to a large number of fatal crashes (Farmer and Williams, 2005). Lerner (2011) reported that there were over 10,000 fatalities each year in the U.S. as a direct consequence of drunken driving. The most commonly used methods in alcohol-related crash studies include a self-report survey, logistic regression analysis, and a driving simulator. A self-reported survey is usually in the form of questionnaires. Respondents who have experienced nighttime drunk driving were asked to answer questions like breath-alcohol concentrations, whether they were taking any drugs before driving, and their feelings about driving while intoxicated (DWI) or under the influence (DUI); for example, did they find it difficult to make judgments about road signs and nearby traffic (Calafat et al., 2009). Even though this kind of survey cannot reflect the relationships between drunk driving or taking drugs and fatal crashes, it found that either having drinks or taking drugs before driving would increase the risk of being involved in traffic accidents. Many scholars have questioned the subjectivity of a self-reported survey. As a result, instead of using data from questionnaires, a significant amount of research based on traffic accident data from the federal transportation department has appeared in academic journals. Analyses of traffic accident data based on logistic regression methods are widely used in nighttime crash research. The aim of this kind of study is to explore the relationships behind crash events (for instance, DWI (or DUI)) and the crash rate (Thombs et al., 2010). A driving simulator was used extensively in proving that alcohol or drug use may influence driving performance by recording drivers' eye movement (Brookhuis et al., 2004).

Some studies have indicated that nighttime crashes also show disparity in space. Space here refers to the locations (e.g., intersections, rural areas, etc., especially with bad road conditions) where motor vehicle crashes are more likely to occur. In 2009 alone, more than two million intersection- related motor-vehicle crashes (approximately 40% of all crashes) occurred in the U.S. (Hu et al., 2011). Fatal crashes occurring at intersections may often be due to improper turning, disobeying or being confused by traffic signals, and failing to yield to the right-of-way because of the complexity of traffic movements (Zhang et al., 1998). In order to reduce crash rates, many before-and-after studies have been conducted to evaluate the impacts on traffic by, for instance, installing red-light cameras, providing adequate left and U-turn lanes, and improving the design of bicycle lanes at intersections (Das et al., 2009; Hu et al., 2011; Chen et al., 2012). Intersections in rural areas tend to have a higher rate of fatal crashes than other places because there are more angle turns, change lanes, and dark lights that drivers may have to deal with, especially when the speed limit is high in these places (Kim et al., 2006). It is also interesting to know that there is a significant spatial difference between weekday crashes and weekend crashes: weekday crashes always occurred in congested places, while weekend crashes tend to happen in un-congested places (Yu and Abdel-Aty, 2013).

Fatal crashes have become a concern for many law enforcement agencies. Many laws and policies have been introduced to reduce the occurrences of traffic accidents and fatal crashes, including Graduated Driver Licensing (GDL), Minimum Legal Drinking Age (MLDA), and Legal Blood Alcohol Content Limits (BACL). Consequentially, many studies have been done to assess the effectiveness of GDL programs on traffic accidents. It was found that extending supervised driving periods and limiting novice drivers' exposure to high-risk conditions (e.g., driving at nighttime or driving with passengers) have effectively reduced involvements by novice drivers in crashes (Shope and Molnar, 2003; Masten et al., 2011; Lyon et al., 2012). The impact of the minimum legal drinking age has also been well studied over the past 40 years (Wagenaar and Toomey, 2000). The goal of this policy is to limit alcohol use and reduce its related problems among teenagers. Debates still exist about whether the MLDA is effective in reducing alcohol-associated driving problems (Wagenaar and Toomey, 2002), but it has been widely acknowledged that, after applying the MLDA laws, significantly fewer incidents and fatalities from crashes were indeed reported (Fell et al., 2008). The Blood Alcohol Concentration (BAC) refers to the percentage of alcohol to blood in the body. According to statistics, more than 80% of the drivers involved in fatal crashes who were intoxicated had BACs at or above 0.08%, the legal level of intoxication in many states (NHTSA, 2003). Many studies related to BACL found that decreasing the legal BACL would significantly reduce alcohol-related fatal crash occurrences (Marques et al., 1999; Wagenaar et al., 2007).

Although traffic laws and policies have a certain effect on the incidence of fatal crashes, traffic safety studies need to continue in order to improve road safety. Traditional traffic safety studies tend not to take seriously the spatial locations of fatal crashes. As discussed before, the recent advances in GIS, especially the increased number of spatial analytical tools available, make the examination of spatial patterns of fatal crashes an easier task.

2.3. Spatial Analysis of Traffic Accidents and Fatal Crashes

Knowing when, where, and how traffic accidents happen is critical in order to reduce traffic accidents and ensure road safety. A better understanding of the spatial patterns of traffic accidents can make this goal easier to achieve. It would be more effective for police to take steps, for instance, to strengthen traffic law enforcement or for transportation engineers to improve transport infrastructures by targeting locations where traffic accidents are prevalent. In many cases, traffic accidents did not happen randomly in space and time but were clustered (known as “hotspots”) (Xie and Yan, 2008), largely due to the fact that different traffic volumes along certain roadway segments contribute to unique temporal and spatial patterns of traffic accidents (Black, 1991). The distribution of traffic accidents is also affected by other spatial factors, such as bad highway conditions, poor road design, and locations of related facilities (e.g., bars and restaurants serving alcohol), and ultimately the dynamics of transportation and land-use interactions.

For the convenience of analysis, spatial scientists and researchers often treat traffic accidents as spatial point events. As a result, spatial point pattern analysis (SPPA) methods were widely used to identify traffic accident patterns and particularly “hotspots” of traffic accidents (Xie and Yan, 2008). Most SPPA methods belong to one of two broad categories: 1) methods focusing on first-order effects of spatial point processes; or 2) methods focusing on second-order effects of spatial point processes (Bailey and Gatrell, 1995; O’Sullivan and Unwin, 2003). The former, known as density-based methods (e.g., crude density, quadrat analysis, and density estimation), measures the variation in the average value of a spatial process, while the latter, known as distance-based methods (e.g., nearest neighbor distance (NND) statistics, G function, F function, and K function), looks for patterns on the basis of distances and separations among spatial point events.

Among the density-based methods, Kernel Density Estimation (KDE) is one of the most frequently used techniques (Silverman, 1986; Bailey and Gatrell, 1995). Many studies use KDE for so-called “hotspot” analysis, such as the study of road accident hotspots (Anderson, 2009), clusters identification of pedestrian-vehicle crashes (Dai, 2012), highway crash hotspots (Erdogan et al., 2008), and wildlife-vehicle collisions

(Krisp and Durot, 2007). The goal of KDE is to generate a continuous surface of density estimations of spatial point events, e.g., traffic accidents, by calculating the density of point events within a specific search bandwidth. Thus, it would help researchers to find out in which part of a study area there is a high density of crashes and, thus, which areas require more attention. Besides KDE, Moran's I and Getis-Ord's Gi were also adopted in a number of studies to identify traffic accident hotspots (Khan and Noyce, 2008; Moons et al., 2009; Steenberghen et al., 2010). However, Kuo et al. (2011) argued that the risk of having traffic accidents would be underestimated if researchers only use the Getis-Ord's Gi method. A similar weakness was also found when applying Moran's I method in traffic accident analysis (Moons et al., 2009).

Like KDE, K function analysis is one of the most popular methods among the second-order SPPA methods (Okabe and Sugihara, 2012). Ripley (1976) first introduced the K function in order to distinguish this method from the mathematical K function. It was termed as Ripley's K function when applied to SPPA (Dixon, 2002). Compared with other distance-based methods that often rely on some form of nearest neighbor distances (NND) among point events, K function includes all possible distance information and utilizes all inter-event distances. K function has been widely adopted as the main method to analyze the spatial distribution of point events such as traffic accidents (Yamada and Thill, 2004). In addition to geoscience fields, K function was also widely applied to botany, ecology, biology, epidemiology, and other related disciplines. Examples include studies to investigate the spatial patterns of trees (Duncan, 1993; Peterson and Squiers, 1995), herbaceous plants (Stamp and Lucas, 1990), bird nests (Gaines et al., 2000), and disease cases (Diggle and Chetwynd, 1991).

Due to recent advances in geospatial technologies, more and more researchers now realize the important role of GIS in traffic accident analysis. In fact, the application of GIS in transportation fields is not entirely new. The term "GIS for Transportation (GIS-T)" was first introduced to the public in the 1960s, when the first GIS was used as a computerized map-measuring system (Goodchild, 2000). GIS is defined as "a computer system for capturing, storing, querying, analyzing, and displaying geospatial data" (Chang, 2012, p.2). The introduction of GIS to the transportation domain started a revolution in both geography and transportation research, because GIS provides a platform for researchers to examine spatial issues quantitatively (Thill, 2000). With advances in GIS and other information technologies, traffic accident analysis gradually moved from non-spatial to spatial analysis. For instance, almost all of the important spatial analysis methods (e.g., KDE and K function) are readily available in the world-leading commercial GIS software, ESRI's ArcGIS for Desktop. However, with the development of research techniques, researchers find that some traditional spatial analysis methods have some limitations when applying them directly to those spatial point events often confined within or along a certain type of network; e.g., motor vehicle crashes. Many spatial methods were thus extended to network space in recent years.

2.4. Network Extensions of Spatial Analysis on Traffic Accidents

Motor vehicle crashes are one of the typical network-constrained point events (or network events, for short), as they are typically confined within a roadway network (Yamada and Thill, 2007). Almost every car crash occurred directly on a road, at a road intersection, or even along an aisle in a parking lot. As a result, crashes can easily be conceptualized as on-network events (Okabe and Sugihara, 2012). Traditionally, spatial scientists often apply spatial methods that originate in two-dimensional Euclidean space (known as planar spatial methods) directly to these network events. However, this usage has been confirmed to lead to false conclusions, and planar spatial methods are indeed not suited for analyzing network events (Lu and Chen, 2007; Yamada and Thill, 2007; Xie and Yan, 2008). To overcome the limitations of planar methods, many researchers started to extend many conventional methods to network spaces. Some examples include network autocorrelation (Black and Thomas, 1998), network distance-decay (Kent et al., 2006), space-time clustering (Black, 1991), network K function (Yamada and Thill, 2004), and network KDE (Xie and Yan, 2008). These methods were widely adopted by many recent studies to analyze the spatial distributions of network events (Shiode, 2008; Yamada and Thill, 2004; Yamada and Thill, 2007).

As one of the widely-used distance-based methods in SPPA, K function was extended to network space and applied to studies such as traffic accident analysis (Yamada and Thill, 2004) and landscape analysis (Spooner et al., 2004), to name a few. The statistical and computational methods for network K function can be found in Okabe and Sugihara (2012) and GIS based tools of various types of network K functions are available in freeware called Spatial Analysis along Networks – SANET (Okabe and Satoh, 2009). When compared to planar K function, with traffic accident analysis as case studies, the advantages of the network K function are confirmed and the overestimation of hotspots of traffic accidents (based on planar K function) can be avoided (Yamada and Thill, 2004; Lu and Chen, 2007).

Like K function, KDE was also extended to the network space (Xie and Yan, 2008). Borruso (2005) used a modified KDE method to study distribution of spatial point events by considering kernels as a network distances based density function instead of straight-line distances. However, his approach is not strictly a network KDE as it is still based on 2-D Euclidian space when compared to the true network KDE developed by Xie and Yan (2008). In addition to Xie and Yan's (2008) work, Okabe and Sugihara (2012) provide more statistical details and additional computational implementations of network KDE. Unlike planar KDE, in network KDE density values are based on linear units (e.g., per kilometer) rather than areal units (e.g., per square kilometer), and this makes the network KDE more accurate by avoiding an overestimation of density values.

As stated before, a number of network extensions of commonly used SPPA methods, such as network KDE and network K function, are available in freeware called SANET. SANET is simply "a plug-in program of GIS which statistically analyzes spatial patterns of events that occur along networks" (Okabe et al., 2006, p.2). It can be easily loaded in the leading commercial GIS software, ESRI's ArcGIS. SANET provides a set of network versions of spatial methods, including KDE, nearest neighbor distance statistics, global K function, local K function, global cross K function, and local cross K function. SANET tools are ideal for analyzing the spatial distribution of network point events such as traffic accidents. In fact, SANET tools have been used in a wide range of studies, such as the identification of crash clusters (Dai et al., 2010), vegetation distribution analysis (Kumagai and Mizuma, 2009), and bus transport service based on urban accessibility (Gonda and Okunki, 2008).

CHAPTER 3. STUDY AREA

The Nashville Metropolitan Area (NMA) is located in the north-central region of Tennessee (Figure 3.1). It is formally defined by the Office of Management and Budget (OMB, 2006) as the Nashville-Davidson-Murfreesboro-Franklin, TN Metropolitan Statistical Area. The original purpose of setting up the Metropolitan Statistical Area (MSA) was for the convenience of statistical reporting by the United States Census Bureau and other agencies. At first, only Davidson County was included. As the surrounding counties' population increased, new counties were added. Now, thirteen counties form the NMA, including Cannon, Cheatham, Davidson, Dickson, Hickman, Macon, Robertson, Rutherford, Smith, Sumner, Trousdale, Williamson, and Wilson.

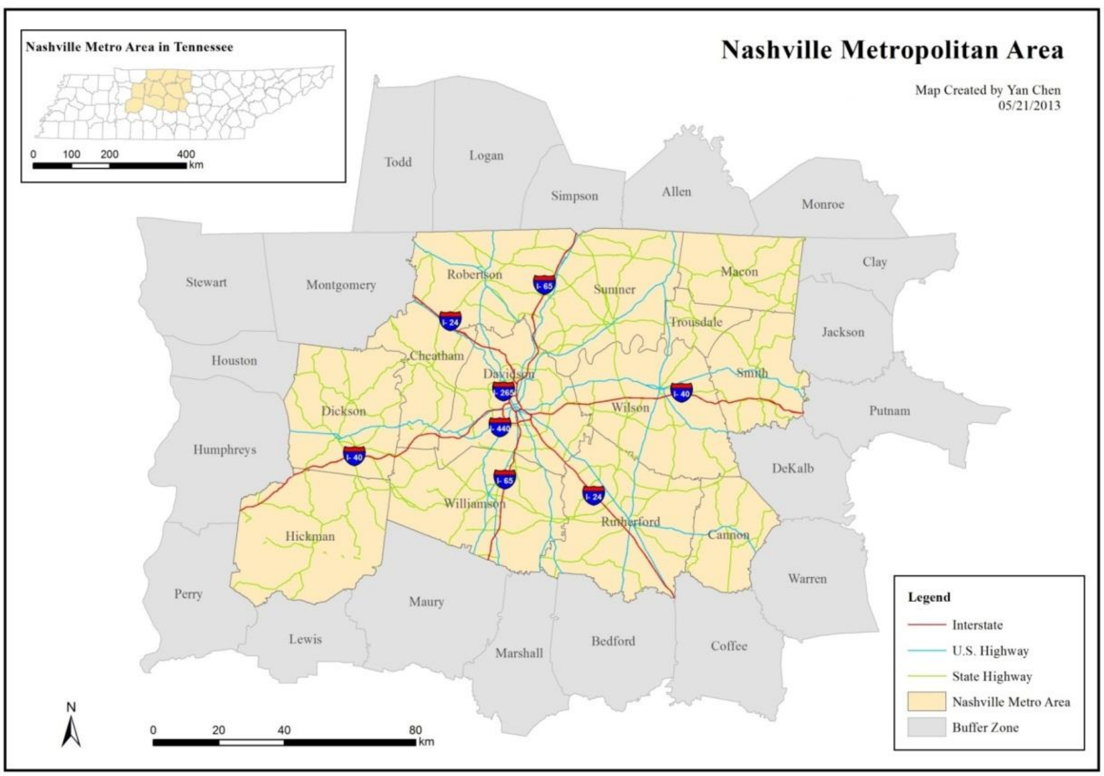


Figure 3.1. Study Area with Primary Roads. Source: U.S. Census Bureau (2012) and TDOT (2012).

As the political capital of Tennessee, Nashville attracts many opportunities in the music, sports, education, health care, publishing, banking and finance, tourism, and transportation industries. Recently, Nashville was ranked by Forbes as the third best city in U.S. positioned to grow in the next ten years (Kotkin, 2011). Different from other cities' development model, Nashville's music industry plays a very important role in economic development. Nashville has been the third largest music production center (after New York and Los Angeles) since the 1970s (Florida and Jackson, 2010). The title of "Music City" for Nashville makes it a famous tourist center for music lovers, especially country music fans. The good education environment also attracts many people. The large number of colleges and universities provides opportunities for students who want to further their study. Besides more employment and education opportunities, the low-cost business environment would be another reason for a large number of migrants choosing to settle down in Nashville (Metropolitan Government of Nashville and Davidson County, 2013). The lower cost-of-living compared with other cities is attractive to job seekers.

According to the U.S. Census Bureau (2012), there are more than 1.6 million people living in the NMA in an area of 6,868 square miles, or 235 people per square mile. The great opportunities for both work and education in Nashville have led to significant population growth in recent years. From 2000 to 2012, population in the NMA increased by more than 330,000, a 25% rise (Table 3.1). The top population increases were in Rutherford, Davidson, and Williamson counties, respectively. Both Rutherford County and Williamson County had a population growth rate of more than 50%. The populations of other counties also increased over the past decade, with relatively lower rates than those of Rutherford and Williamson (Table 3.1. and Figure 3.2.).

No.	County	Pop. in 2000	Pop. in 2012	Pop. Growth	Growth Rate %
1	Cannon	12826	13811	985	7.68
2	Cheatham	35912	39271	3359	9.35
3	Davidson	569891	648295	78404	13.76
4	Dickson	43156	50381	7225	16.74
5	Hickman	22295	24170	1875	8.41
6	Macon	20386	22498	2112	10.36
7	Robertson	54433	66931	12498	22.96
8	Rutherford	182023	274454	92431	50.78
9	Smith	17712	19102	1390	7.85
10	Sumner	130449	166123	35674	27.35
11	Trousdale	7259	7795	536	7.38
12	Williamson	126638	192911	66273	52.33
13	Wilson	88809	118961	30152	33.95
	Total	**1,311,789**	**1,644,703**	332,914	25.38

Table 3.1. Population Change in the NMA from 2000 to 2012. Source: U.S. Census Bureau (2012).

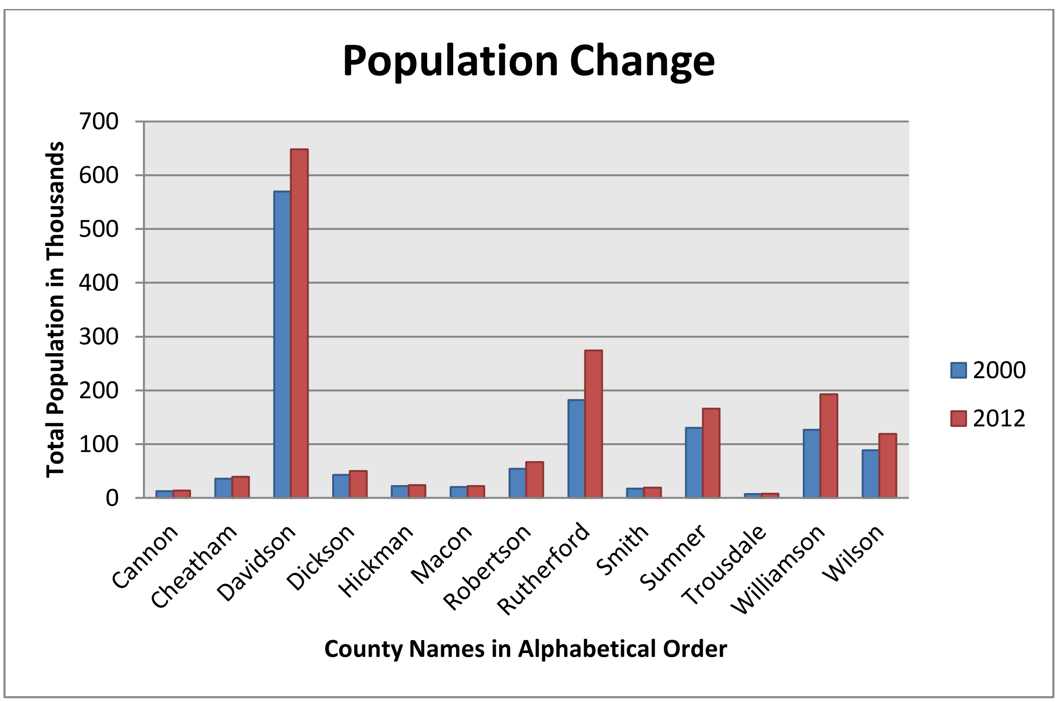


Figure 3.2. Population Change in the NMA from 2000 to 2012. Source: U.S. Census Bureau (2012).

The increased population as well as the growing demand for daily traffic has created stresses on the transportation system. The NMA has a highly developed road network that includes numerous interstate highways, U.S. highways, and state highways. Three important Interstate highways, namely I-24, I-40, and I-65, converge in the central city (Figure 3.3). It is interesting to know that the three counties with the highest population growths, Rutherford, Wilson, and Williamson, are all along I-24, I-40, and I-65 respectively (Figure 3.3). With an increased number of commuters from surrounding suburbs to Nashville, these highways experience heavy traffic each day (Figure 3.3) and the fact that they all converge in the central city makes Nashville one of the most congested North American cities during rush hours. The aviation industry in the NMA is also well developed and supported by the Nashville International Airport. Every day there are flights to and from every corner of the country to transport travelers and cargo. The busy hub also increases the burden on the surface road network.

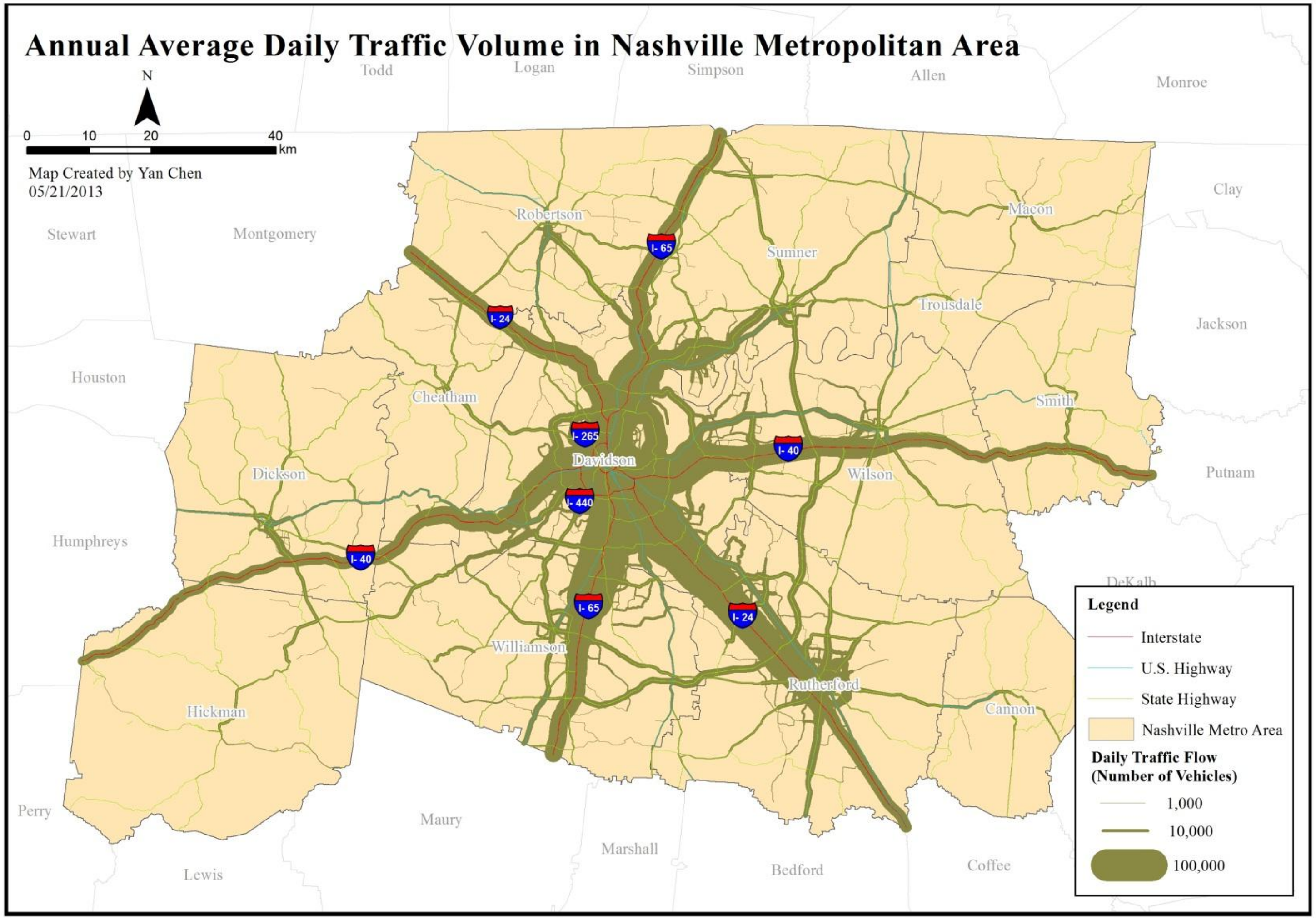


Figure 3.3. Annual Average Daily Traffic. Source: U.S. Census Bureau (2012) and TDOT (2012).

CHAPTER 4. DATA AND METHODOLOGY

In this study, the main data used were acquired from the Fatality Analysis Reporting System (FARS, 2001-2011). FARS fatal crash data were processed and then used in four SPPA methods, namely Planar Kernel Density Estimation (KDE), Planar Global Auto K Function, Network Global Cross K Functions, and Network Local Cross K Functions. The results of planar KDE and global auto K function would help to answer the questions listed in RQI; for instance, where did fatal crashes usually occur and are there any discernible spatial patterns of fatal crashes in the study area? The outputs of network cross K functions would further explain the questions listed in RQII and RQII; for instance, are there any notable spatial differences and where are the differences among subtypes of fatal crashes based on temporal factors, specifically 1) day and night; 2) A.M. rush hours and P.M. rush hours; and 3) weekday and weekend? This chapter is organized as follows: Section 4.1 discusses the main datasets and data preparation tasks, while the detailed discussions of the methods are in Sections 4.2 and 4.3.

4.1. Data and Data Preprocessing

4.1.1. Data Overview

This study utilized data from multiple sources, including FARS (2001-2011), the Tennessee Department of Transportation (TDOT, 2012), the U.S. Census (2012) Tiger/Line database. The National Highway Traffic Safety Administration (NHTSA) maintains the FARS databases, which provide critical information about fatal crashes, such as crash location, time, persons involved, vehicles involved, roadway conditions, etc. The annual traffic flows with roadway descriptions were acquired from TDOT (2012).

Some background data, such as the most updated road-line shapefiles and boundary shapefiles for the state of Tennessee and 13 counties of the study area, were from the U.S. Census (2012) Tiger/Line database. Lastly, photos were taken and used for the ground information in some target areas.

FARS is a nationwide census database providing users with the most updated and complete data regarding motor vehicle fatal crashes. There are two ways to access FARS data: 1) run a query using the FARS web-based encyclopedia; or 2) download all FARS data from 1975 to the present from the FTP site. The data are well prepared in many forms and files have free access by public users. The .dbf files were used for this study since they can be directly opened and viewed in ArcMap for further analysis. The FARS data were organized year by year in files, so it is very convenient for users to search and download. The years of 2001 through 2011 were downloaded from the FTP site for this study. Each set of yearly downloadable files has three .dbf files: accident.dbf, person.dbf, and vehicle.dbf. Each of these three .dbf files records different types of information relating to actual fatal crashes. The accident.dbf file contains information such as coordinates, time (including time of day, day of week, month, and year), number of victims, and road type of fatal crashes. The person.dbf file includes information about victims such as age, gender, and if a DUI was involved or not. The vehicle.dbf provides information about the vehicles involved. For this thesis research, only accident.dbf was used since it includes the information about the location and time of each fatal crash.

4.1.2. Data Preprocessing

Data preprocessing requires good file management skills. All downloaded files need to be unzipped, organized, and named properly according to data types. Three main tasks were carried out, including processing crash data, county boundaries, and road shapefiles.

Processing crash data is the most complex task compared to the other two. First, eleven accident.dbf files from 2001 to 2011 were added to ArcMap. In order to locate the crash events on the map, the .dbf files were first converted to point shapefiles by inputting the coordinates (longitude and latitude) of the events. Since this study aims to look at the overall spatial pattern of fatal crashes from 2001 to 2011, the separated yearly shapefiles were then merged into a single shape file, which contains all fatal crashes during the study period in the entire U.S. However, instead of using fatal crashes nationwide, only data within the NMA were included for this study. Two more steps were then carried out. The fatal crashes were first selected by State Federal Information Processing Standards (FIPS) code (Tennessee's FIPS is '47') so that those fatal crashes that had only occurred in Tennessee were included in the newly exported shapefile. Then, the fatal crashes in the new shapefile were selected again by county FIPS (Table 4.2). Thus, all fatal crashes that occurred in the NMA from 2001 to 2011 were selected. They were then exported to another new shapefile with the UTM Zone 16 North projection. Lastly, this shapefile was subdivided into six shapefiles by day (6 a.m. to 6 p.m.) and night (6 p.m. to 6 a.m.), A.M. rush hours (6 a.m. to 9 a.m.) and P.M. rush hours (4 p.m. to 7 p.m.), as well as weekdays (Monday to Friday) and weekends (Saturday and Sunday).

No.	County Name	FIPS Code	No.	County Name	FIPS Code
1	Cannon	8	8	Rutherford	75
2	Cheatham	11	9	Smith	80
3	Davidson	19	10	Sumner	83
4	Dickson	22	11	Trousdale	85
5	Hickman	41	12	Williamson	94
6	Macon	56	13	Wilson	95
7	Robertson	74	**Note:** State FIPS Code for TN is 47		

Table 4.1. County FIPS. Source: U.S. Census Bureau (2012).

Both the Tennessee state boundary and county boundaries were obtained from the U.S. Census (2012) Tiger/Line database. Since the original downloaded shapefiles contain all the state or county boundaries nationwide, a selection query was used to extract those inside Tennessee. The new shapefiles also were projected with the same coordinate system, UTM Zone 16 North.

Two types of road shapefiles were downloaded from the U.S. Census (2012) Tiger/Line database. One contains only primary and secondary roads, while the other includes all roads. The primary and secondary roads include four types of roadways: interstates, U.S. highways, state highways, and principle county roads. All roads include local streets as well as the primary and secondary roads. All road shapefiles can be searched and downloaded by state and county; however, the primary and secondary road shapefiles are only available as a single shapefile for the entire state. As a result, the primary and secondary roads within the NMA were extracted first. In order to avoid the so-called edge effect, a five-mile buffer was used for the primary and secondary roads within a five-mile radius of the study area. As for the “all roads,” the separate all-road shapefiles for each county were merged into a single shapefile. Again, both roadway shapefiles were projected with the UTM Zone 16 North projection. The primary and secondary roads were later used for the actual analysis, while all roads were mainly used for background information.

4.2. Primary Methods

One density-based method, namely planar KDE, and three distance-based methods, namely planar global auto K function and network cross K functions (both global and local), were used in this study. The hotspots of fatal crashes for all FARS accidents, as well as the hotspots of each subtype fatal crash based on the time factors, were visually interpreted based on their corresponding planar KDE surface. Then the planar global auto K function was used to find out if there are any discernible spatial patterns in all fatal crashes as well as the subtypes of fatal crashes based on the time factors. Finally, two network cross K functions (global and local) were used to detect the differences in the spatial distribution of the subtype fatal crashes during selected periods.

4.2.1. Density-Based Methods

As a density-based method, the planar kernel density estimation (KDE) was used in this study to identify the places with high concentrations of fatal crashes in the NMA from 2001 to 2011. This method can provide researchers with a continuous and smooth surface of spatial density estimations by weighting nearby points more than faraway ones based on a certain kernel function. The kernel function is given below:

$$\lambda(\mathrm{s}) = \sum_{i=1}^{n} \frac{1}{\pi r^2} k(\frac{d_{is}}{r})$$

where $\lambda(s)$ is the density at location s, n is the number of points in study area, r is the search radius (bandwidth) of the KDE; because of this only the points within the distance r will be used. Kernel function k, gives the weight of point i at distance d_{is} to location s; it is usually simplified as a function of the ratio between d_{is} and r.

In this study, the planar KDE was conducted by using directly the Kernel Density Spatial Analyst Tool available in ArcGIS for Desktop10.1. By inputting the point features, this tool generates a density map as a continuous field. This output surface is sensitive to search radius and output cell size. The local variation details are dependent on the cell size of the output raster, while the search radius determines the smoothness of the result density surfaces. In order to look at the patterns at a smaller spatial scale, say street blocks, a search radius of 1,000 meters and a cell size of 100 meters were used for all fatal crashes, as well as for the subsets of fatal crashes (subdivided by day, night, A.M. rush hours, P.M. rush hours, weekdays, and weekends). The hotspots for all these point patterns could be readily interpreted via visual planar KDE surfaces.

Compared with the planar KDE, Network KDE calculates the density of point events based on a network space; for example, streets. The density value of network KDE is based on linear units rather than area units. As a result, the density that is calculated by network KDE is more suitable for the analysis of traffic accidents. However, in this study the planar KDE was chosen mainly because the purpose is not to detect hotspots more accurately but to develop visual interpretations of fatal crashes at relatively large geographic scales. Moreover, the calculation of Network KDE is much more time consuming than that of planar KDE. In fact, the Network KDE tool in SANET was unable to handle the big and complex road network of the study area.

4.2.2. Distance-Based Methods

The distance-based methods rely on the distances among point events to identify spatial patterns. Three distance-based methods were used in this study, including planar global auto K function, network global cross K function, and network local cross K function. Global auto K function was used to determine if there is a clustered, dispersed, or random pattern in FARS crash points over a range of distances. The cross K function methods, instead of examining the distances between a set of points of only a single type as in auto K function, enable users to compare the spatial differences of two types of point events over a range of distances by examining the distances from points of one type to those of the other type.

4.2.2.1. Global Auto K Function

The global auto K function (often called simply the K function) was first introduced by Ripley (1976) and then named Ripley's K function to distinguish it from the term "K function" in mathematics. Ripley's K function makes use of all inter-event distances in a study area, which makes it a complex function to calculate. There are three steps involved in the calculation of a K function. The first step is to determine the number of point events that lie within a specified distance threshold of each event, and iterate this for all events. Then, in the second step, an average number of events at the specified distance is calculated. Lastly, in the third step, the average is divided by the overall event density of the study area to obtain a K value. Hence, the K function is given below:

$$\mathrm{K(d)} = \frac{\sum_{i=1}^{n} no.[S \in C(S_i, d)]}{n\lambda}$$

where λ is overall event intensity, n is the number of total events, S refers to the entire study area, d refers to the distance, and $C(S_i, d)$ is the circle of radius d centered at S_i.

In this study, the global auto K function was conducted by using ArcGIS Spatial Statistics Tool of Multi-Distance Spatial Cluster Analysis (Ripley's K Function) in ArcGIS for Desktop10.1. By inputting a point feature class, this tool outputs a table as well as a summary graph with the expected and observed K values. Three types of spatial patterns, namely clustered (if observed values are above the expected values and the upper envelope), dispersed (if observed values are under the expected values and the lower envelope), or random (if observed values are within the upper and lower envelopes), could be identified from the output graph over a range of distances. The distance band 100 meters and distance increment 1,500 meters were used in this study for all fatal crashes, as well as the subsets of fatal crashes (day, night, A.M. rush hours, P.M. rush hours, weekdays, and weekends). Additionally, the permutation was set to 99 for the calculation of all K functions so that the results can have a 99% confidence level.

Like KDE, the global auto K function has also been extended to the network space. Instead of calculating Euclidian distance among point events in a two-dimensional space, the network global auto K function counts for the distance on a network space (for example, street network), which thus makes network K function an even more complex function than planar auto K. For the same reason that planar KDE was used, planar global auto K function is used in this study even though the network global auto K function is available in SANET.

4.2.2.2. Network Global Cross K Function

The network global cross K function was associated with the mean number of type B points within distance *t* from every point of Type A points. For a given two sets of points, a set of n_A Type A points and n_B Type B points, placed on a given network, let $n(t|p_{B_i})$ be the number of Type A points that are within the distance *t* from the *i*-th Type B point p_{B_i}. The network global cross K function is given by:

$$K_{AB}(t) = \frac{1}{\rho_A} \frac{\sum_{i=1}^{n_B} n(t|p_{B_i})}{n_B}$$

where ρ_A is the density of Type A points. The network global cross K function was available in SANET 4.1 developed by Okabe et al. (2006) at the University of Tokyo, Japan. This tool tested the complete spatial randomness (CSR) hypothesis by using a Monte Carlo simulation. The process includes randomly generating a large number of patterns of *n* point events in a study area, calculating K(t) for each simulation, and then generating two envelopes (lower and upper) and a range of certain confidence levels to assess the observed K(t). The observed values within the upper and lower envelopes that indicate no special pattern exists at the specified confidence level suggest randomness. Values above the upper envelope mean that Type B points tend to cluster around Type A points at that specific distance, while values under the lower envelope mean that Type B points are more likely to be dispersed from Type A points at that specific distance.

In this study, the network global cross K function method was conducted for each pair of FARS subsets by setting Type A points to night, P.M. rush hours, and weekend fatal crash points, and then setting Type B points to day, A.M. rush hours, and weekday fatal crash points, respectively. The bandwidth of 100 meters and significance level of 0.05 were used for this method, as the goal of this research is to examine the spatial differences at localized geographic scales. With the help of the network global cross K function method, the spatial differences of fatal crashes between day and night, A.M. rush hours and P.M. rush hours, week and weekend can be examined over a range of distances for the entire study area. The inputs of Type A points layers and Type B points layers are shown in Table 4.2.

Type A Points Layer	Type B Points Layer
P.M. Accident Points	A.M. Accident Points
4P.M.-7P.M. Accident Points	6A.M.-9A.M. Accident Points
Weekend Accident Points	Weekday Accident Points

Table 4.2. Point Types Used in the Network Global Cross K Function.

4.2.2.3. Network Local Cross K Function

The global cross K function is simply calculated by averaging the local cross K functions. Hence, the averaging step was skipped when calculating the network local cross K function for a specific point location. Similar to the network global cross K function, the network local cross K function is determined by the number of Type B points within distance *t* from every point to each Type A point. Consider a set of n_A Type A points and n_BType B points; let $n(t|p_{B_i})$ be the number of Type A points that are within the distance *t* from the *i*-th Type B point p_{B_i}. The network local cross K function is given as:

$$K_{AB}(t) = \frac{1}{\rho_A} n(t|p_{B_i})$$

where ρ_A is the density of Type A points within the street network. The procedure for testing CSR is the same as that for the network global cross K function method. Different from the network global cross K function (only one single K function for the entire study area and all Type A points), a network local cross K function is calculated for every single Type A point, which means there will be 100 network local cross K functions if there are 100 Type A points. Similar to the interpretation of the network global cross K function, if the observed values fall within the upper and lower envelopes, these indicate that no spatial patterns and differences exist at that location and that the spatial distribution of Type B points appears to be random in space and to have no relationship to the location of a Type A point. Above the upper envelope means that Type B points tend to cluster around a specific Type A point at that specific distance, while under the lower envelope means that Type B points are more likely to be dispersed from that Type A point at that specific distance.

In this study, the network local cross K function method was used to examine the spatial differences between FARS subsets vs. FARS subsets (e.g., night vs. day, P.M. rush hours vs. A.M. rush hours, and weekends vs. weekdays). The bandwidth of 100 meters, cell size of 100 meters, and significant level of 0.05 were used for this method. As a result, the spatial differences of the subsets could be examined from local perspectives. The inputs of Type A points layers and Type B points layers are listed in Table 4.3. In Table 4.3, Type A points for the network local cross K function are all sets of the hotspots of the subtype fatal crashes detected by planar KDE, rather than the actual points of each subtype of fatal crashes. Theoretically, the network local cross K function can be calculated at the location of each Type A point, which will certainly make the calculation of the network local cross K function very time consuming, given a large number of Type A points. Therefore, instead of examining how the spatial distribution of Type B points may differ from that of Type A points, this study only looked at the locations where there are high concentrations of Type A points; that is, the hotspots of

Type A points. In practice, the areas with high concentrations of point events are of more interest in regards to reducing fatal crashes and improving traffic safety.

Type A Points Layer	Type B Points Layer
Day Hotspots	All FARS Accident Points
Night Hotspots	All FARS Accident Points
Week Hotspots	All FARS Accident Points
Weekend Hotspots	All FARS Accident Points
AMR Hotspots	All FARS Accident Points
PMR Hotspots	All FARS Accident Points
Night Hotspots	Day Accident Points
PMR Hotspots	6A.M.-9A.M. Accident Points
Weekend Hotspots	Weekday Accident Points

Table 4.3. Point Types Used in the Network Local Cross K Function

However, two problems have to be resolved. First, the input of the Network Local Cross K Function tool in SANET only accepts point layers, while the hotspots detected by the KDE method are in raster form. Second, only eight Type A points are allowed to run at a time by the tool. To solve the first problem, a new point shapefile was created to represent the regions of high values on the KDE density surface. To achieve these, the centers of the regions of high KDE values were chosen if possible. There are potentially more than eight hotspots for each fatal crash point layer. So, to solve the second problem, these new hotspot centers were labeled with a number starting at zero and then divided into several separate shapefiles named, for instance, as 0-7, 8-15, 16-23, and so on, and each group contains, at most, eight hotspot centers.

CHAPTER 5. RESULTS AND DISCUSSION

The potential hotspots and spatial patterns of all FARS accidents from 2001 to 2011, as well as the subsets (extracted by day, night, A.M. rush hours, P.M. rush hours, weekday, and weekend), were identified by the planar KDE and global auto K function, respectively. Notable spatial differences among the subtypes of fatal crashes based on the above temporal factors were investigated via both network global and local cross K functions. This chapter is organized as follows: Section 5.1 presents the overall spatial patterns of fatal crashes in the Nashville Metropolitan Area (NMA) from 2001 to 2011. Section 5.2 then discusses the spatial patterns of the subtypes of fatal crashes based on temporal factors. Lastly, Section 5.3 reports the findings regarding spatial differences among the subtypes of fatal crashes.

5.1. Overall Spatial Patterns of Fatal Crashes in Nashville

From 2001 to 2011, 3,511 fatal crashes occurred within the NMA. A number of hotspots with high concentrations of fatal crashes can be observed visually on the output planar KDE surface (Figure 5.1). At first glance, the spatial distribution of all FARS accidents during the study period largely reflects the distribution of annual average daily traffic (AADT) flows (Figure 3.3). This is understandable because as the underlying population-at-risk causes, more traffic volume, this could lead to more traffic accidents and fatal crashes. In a broader geographic context, the majority of the visual hotspots tend to cluster around the central city and Davidson County. In addition to Davidson County, Rutherford has the second largest number of hotspots (five). Hotspots can also be found in Robertson, Sumner, and Wilson counties but with smaller numbers. Almost all these hotspots are located along the primary roads such as Interstates, U.S. highways, and state highways (Figure 5.1). Moreover, most of these hotspots are likely to be observed at major road intersections, highway entrances and exits, and highway interchanges (Figure 5.2). These locations all share a common characteristic, at which traffic from many directions converges. The risk of car crashes and fatalities would greatly increase when drivers are forced to make complex decisions on which way to turn, when to turn, when to yield, which lane to merge into, when to shift lanes, etc. Nashville is similar to many major U.S. cities with its convergence of major interstate highways (Figure 3.2). Therefore, it is not surprising that there are several notable hotspots located at the main interchanges of I-24, I-40, and I-65.

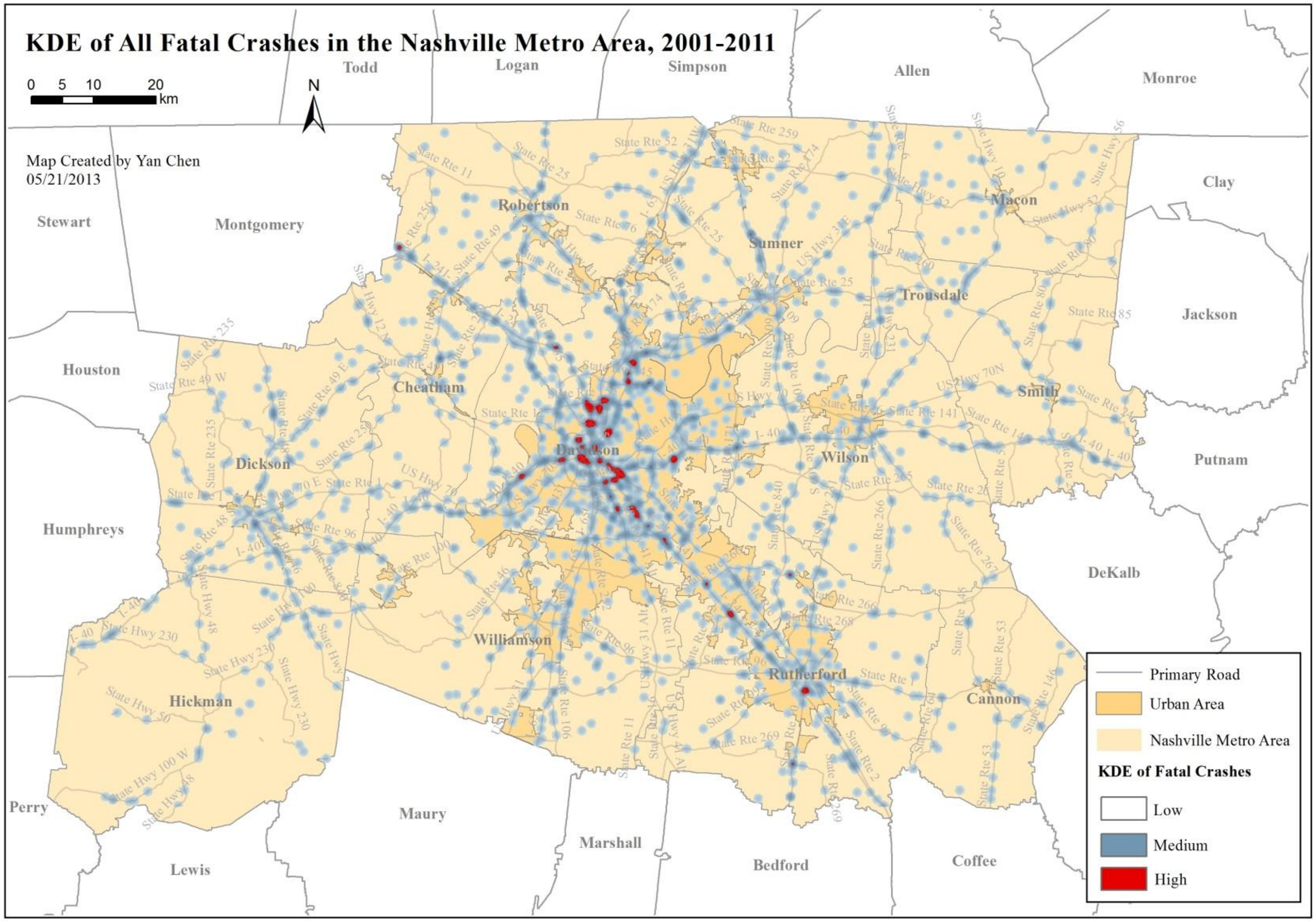


Figure 5.1. Planar KDE of All Fatal Crashes. Source: FARS (2001-2011)

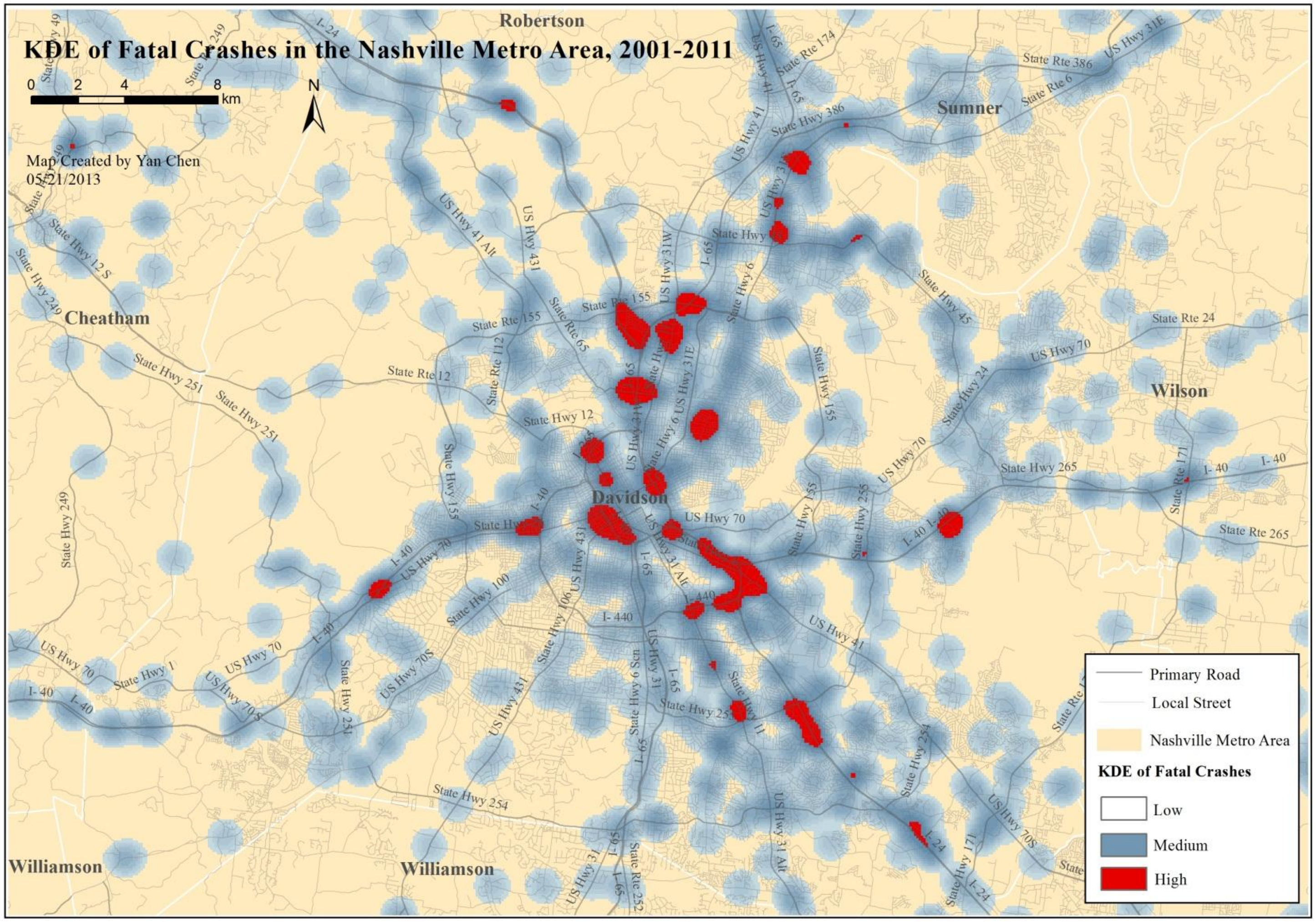


Figure 5.2. A Close-up View of KDE (All Fatal Crashes). Source: FARS (2001-2011)

A cluster pattern of all FARS accidents was detected in the global auto K function, particularly within a distance of about 68 km (Figure 5.3). The spatial pattern of the input point dataset could be identified by interpreting the output graph of Ripley's K Function when comparing the observed K values with expected K values under CSR. When the observed K value is greater than the expected K value (hence **<u>above</u>** it in the graph) for a particular distance, the distribution of fatal crashes would be clustered at that distance as there are more fatal crashes within that distance than otherwise expected randomly. When the observed K value is smaller than the expected K value (hence **<u>below</u>** it in the graph), the distribution is more dispersed at that distance. Moreover, if an observed K value is larger than the upper confidence envelope, for that distance, a cluster pattern could be identified with a high significant level. Similarly, if the observed curve is lower than the lower confidence envelope, spatial dispersion for that distance would be significant. In Figure 5.3, the portion of red curve (the observed K values) above the blue line (the expected K values) for all distances below 67,690 meters, indicates that the FARS accidents in the NMA from 2001 to 2011are indeed clustered below 67,690 meters. Additionally, since this portion of the red curve is also above the upper gray curve (the upper confidence envelope), this clustered pattern below 67,690 meters is statistically significant. The confidence level is dependent on the number of permutations used in the Monte Carlo simulation. The permutation of 99 (roughly corresponding to the 99% confidence level) was used in this study. In sum, all fatal crashes were clustered below approximately a distance of 67.7 kilometers (roughly within a radius of 34 km centered on Nashville downtown area) at a 99% confidence level.

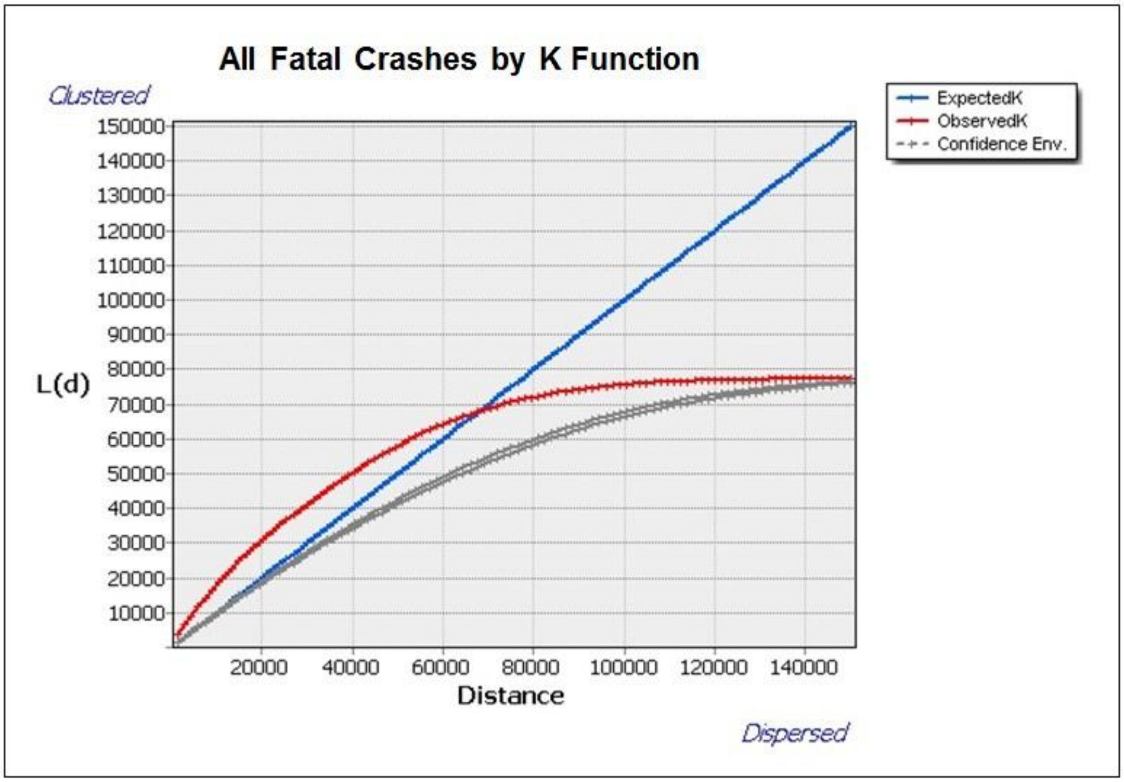


Figure 5.3. Global Auto K Function Graph (All Fatal Crashes). Source: FARS (2001-2011).

5.2. Spatial Patterns of Subtypes of Fatal Crashes by Time Factors

5.2.2. Spatial Patterns of Subtypes of Fatal Crashes Detected by Planar KDE

5.2.2.1. Spatial patterns of daytime and nighttime fatal crashes

The visual spatial patterns of fatal crashes occurring from 6 a.m. to 6 p.m. (1,140 fatal crashes) and from 6 p.m. to 6 a.m. (1,091 fatal crashes) can be readily observed on their planar KDE surfaces, respectively (Figures 5.4 - 5.6). In a broader geographic context, the daytime hotspots are centered in Davidson County and only a few are located in Rutherford, Dickson, Robertson, Sumner, and Smith counties (Figure 5.4). Most of these hotspots were found in urban areas except for a few in the rural areas of Sumner and Dickson counties. Not surprisingly, most of these hotspots are located at the intersection of primary roads. The notable ones include the intersection of I-65 and U.S. Highway 31 in the very center of Davidson County, the intersection of I-65 and State Highway 155 in Davidson County, and the intersection of I-24 and U.S. Highway 231 in the center of Rutherford County.

Similarly, the nighttime hotspots also tend to cluster at the center of urban areas in Davidson County (Figure 5.5). There is only one hotspot found in the rural area of western Robertson County. All these hotspots were found along primary roads. The notable ones include those at the intersection of I-40 and State Highway 24, the interchange of I-40 and I-24, and the intersection of I-65 and U.S. Highway 431.

In sum, the similarity between daytime fatal crashes and nighttime fatal crashes is that the hotspots all tend to cluster in the central urban area and that they are all located along primary roads. However, two differences can be observed (Figure 5.6). First, there are more hotspots detected for daytime fatal crashes. Given that there is roughly the same number of daytime and nighttime fatal crashes, it is safe to say that nighttime fatal crashes are more concentrated geographically than daytime fatal crashes. Second, there are more hotspots found in rural areas for daytime fatal crashes than nighttime fatal crashes. This is not surprising, considering that there is more traffic during daytime than nighttime, particularly in rural areas.

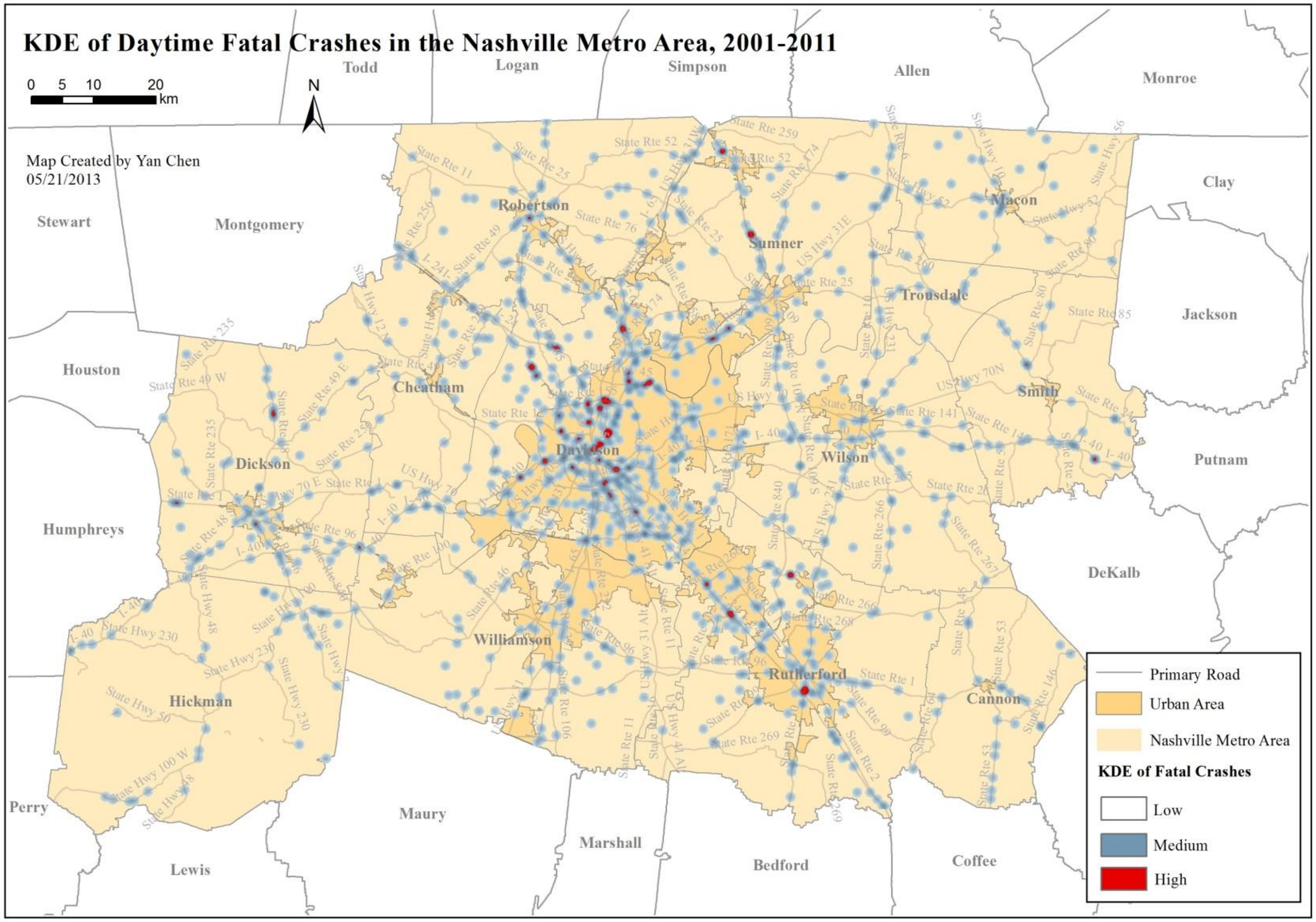


Figure 5.4. KDE of Daytime Fatal Crashes. Source: FARS (2001-2011).

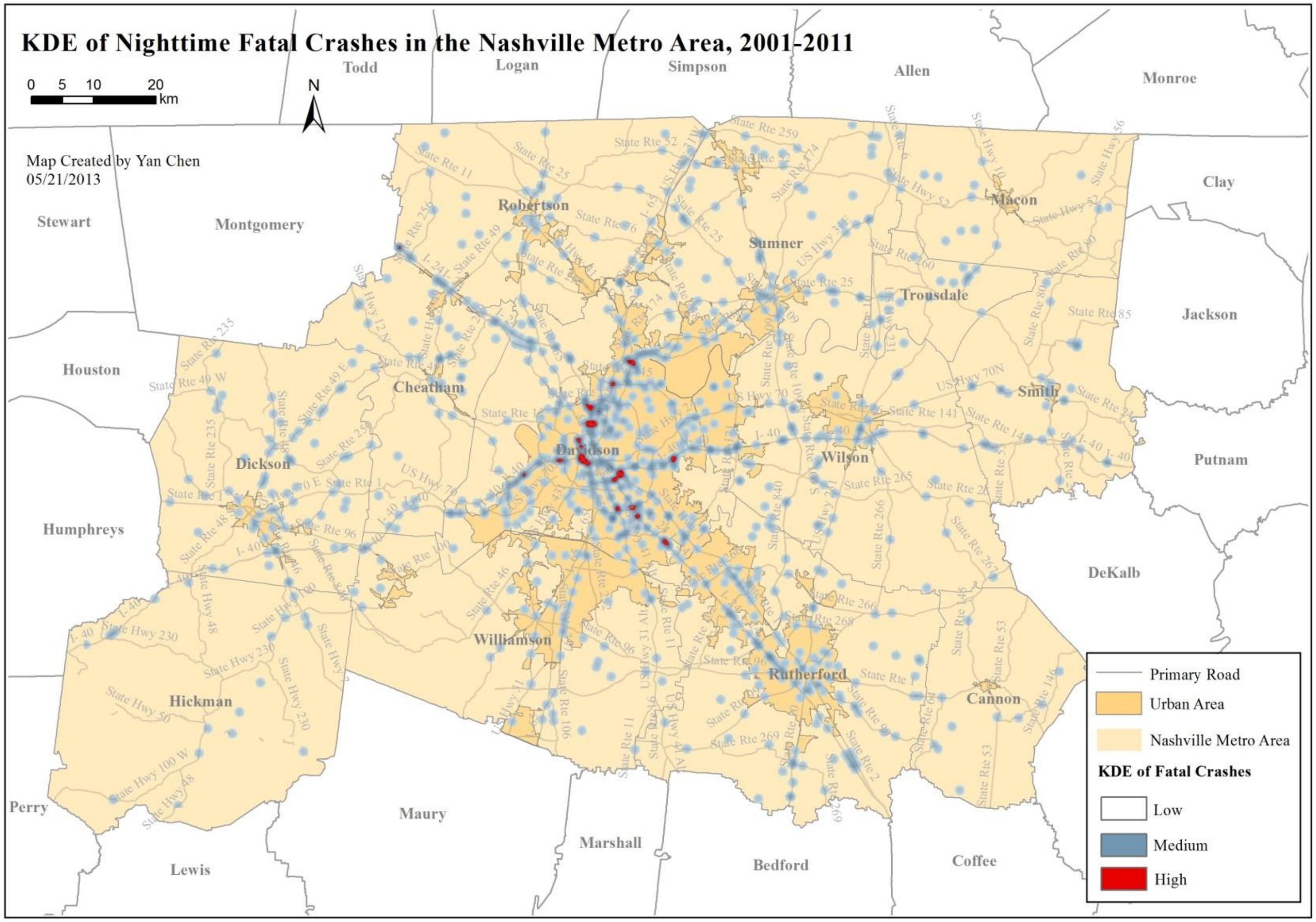
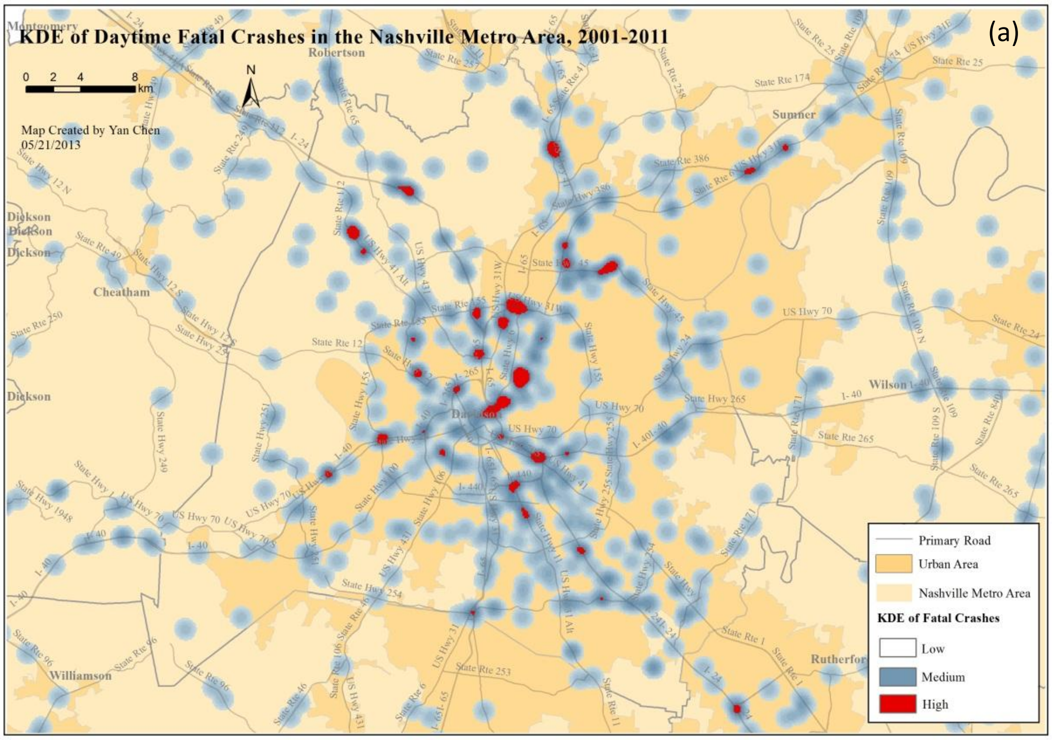


Figure 5.5. KDE of Nighttime Fatal Crashes. Source: FARS (2001-2011).



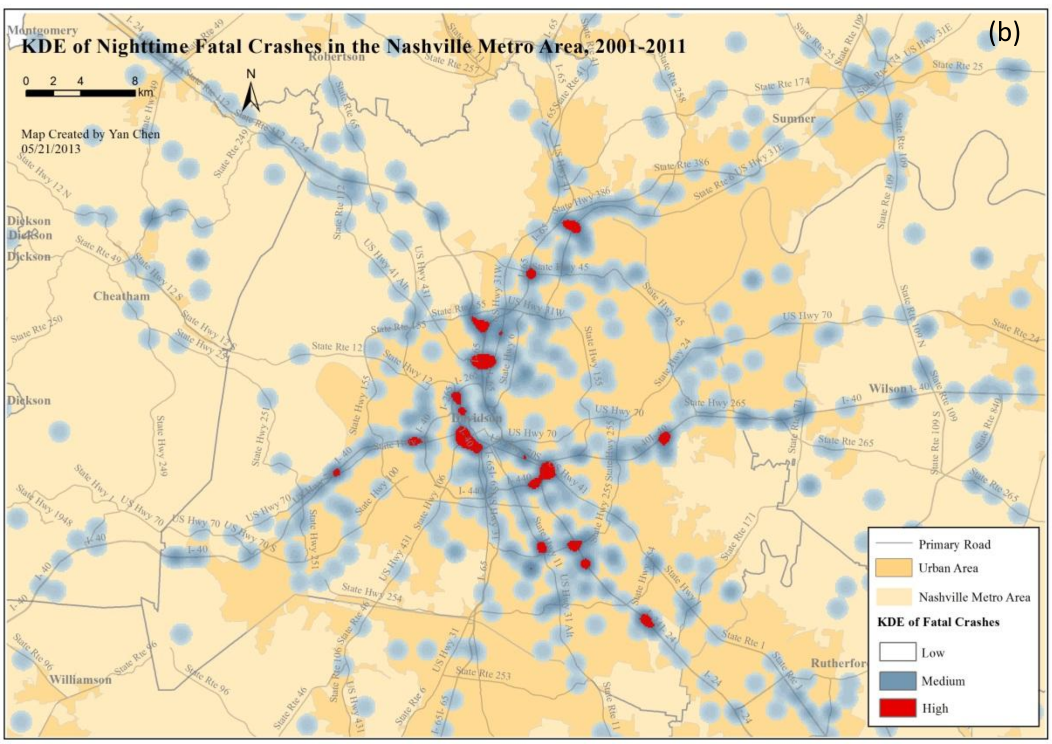


Figure 5.6. Side-by-Side Comparison of KDE of Fatal Crashes: (a) Daytime Fatal Crashes, and (b) Nighttime Fatal Crashes. Source: FARS (2001-2011).

5.2.2.2. Spatial patterns of A.M. rush hours and P.M. rush hours fatal crashes

A total of 252 and 333 fatal crashes were reported during morning rush hours (6 a.m. to 9 a.m.) and night rush hours (4 p.m. to 7 p.m.), respectively within the study period from 2001 to 2011. With fewer data points, the hotspots for these two periods are not as concentrated when compared to those of daytime and nighttime fatal crashes (Figures 5.7 – 5.9). In general, a number of hotspots can be seen in the central city of the NMA during A.M. rush hours, but they tend to stay away from the downtown central business district (CBD) (Figure 5.7). Notable hotspots can also be found outside of Davidson County, such as in Rutherford, Dickson, Robertson, and Sumner counties. Most of these hotspots tend to be located along primary roads, but mainly at the major highway entrance; e.g., the one where U.S. Highway 31W intersects with I-65 and on the edge of Sumner and Davidson counties.

As for P.M. rush hours, a number of hotspots can also be seen in the NMA central city (Figure 5.8). A few hotspots also were found in Rutherford, Robertson, Dickson, and Hickman counties. Most of the hotspots were located in the urban areas of those counties as well. The most notable hotspot was found along I-40 in the western part of Davidson County.

Comparing the A.M. rush hour fatal crashes with the P.M. rush hour fatal crashes, both similarities and differences were found (Figure 5.9). Overall, both of these two subtypes of fatal crashes tend to cluster around the central urban area of the study area as well as along the primary roads. The most obvious difference between them is that more hotspots were found outside of Davidson County during A.M. rush hours than during P.M. rush hours.

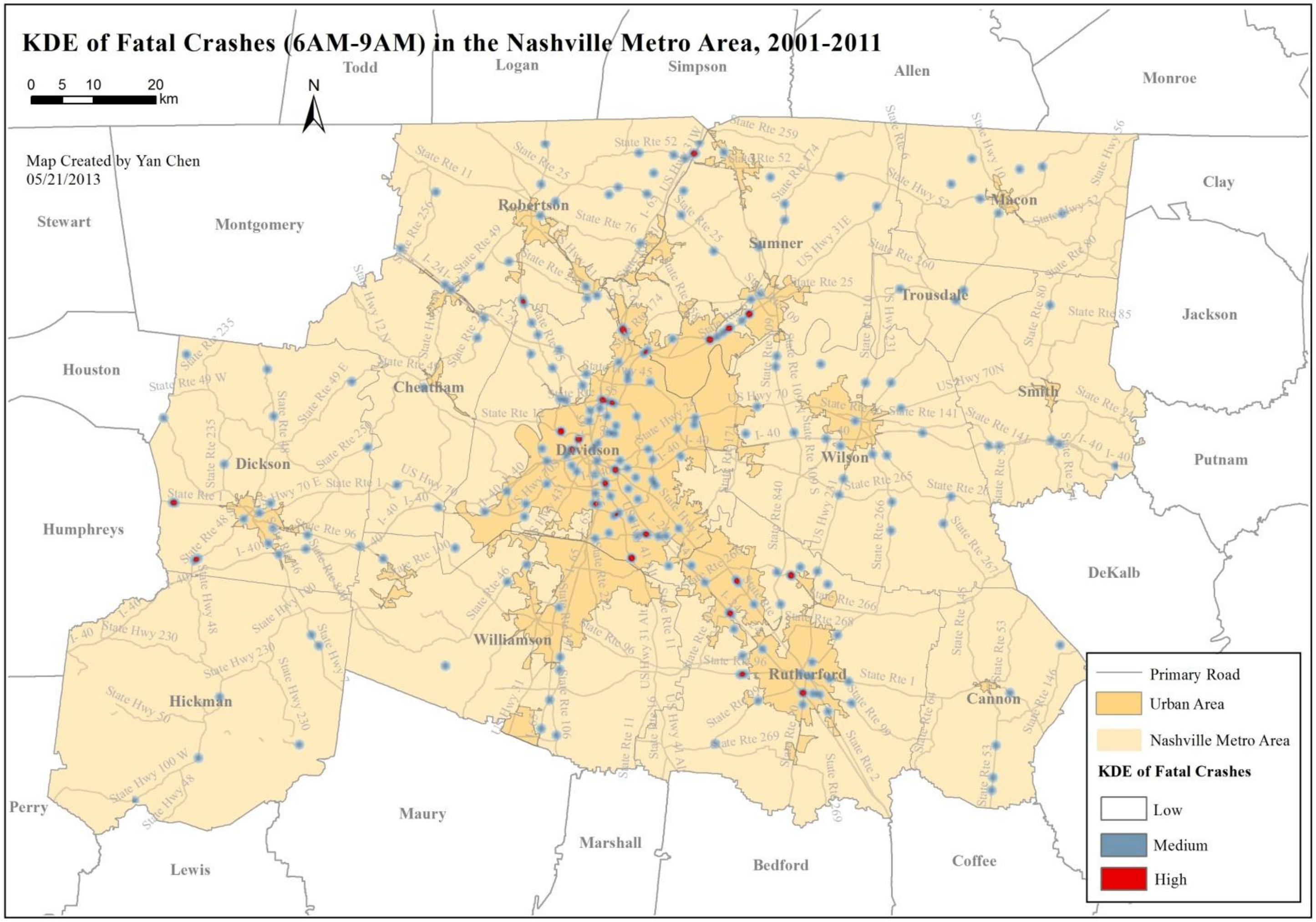


Figure 5.7. KDE of A.M. Rush Hour Fatal Crashes. Source: FARS (2001-2011).

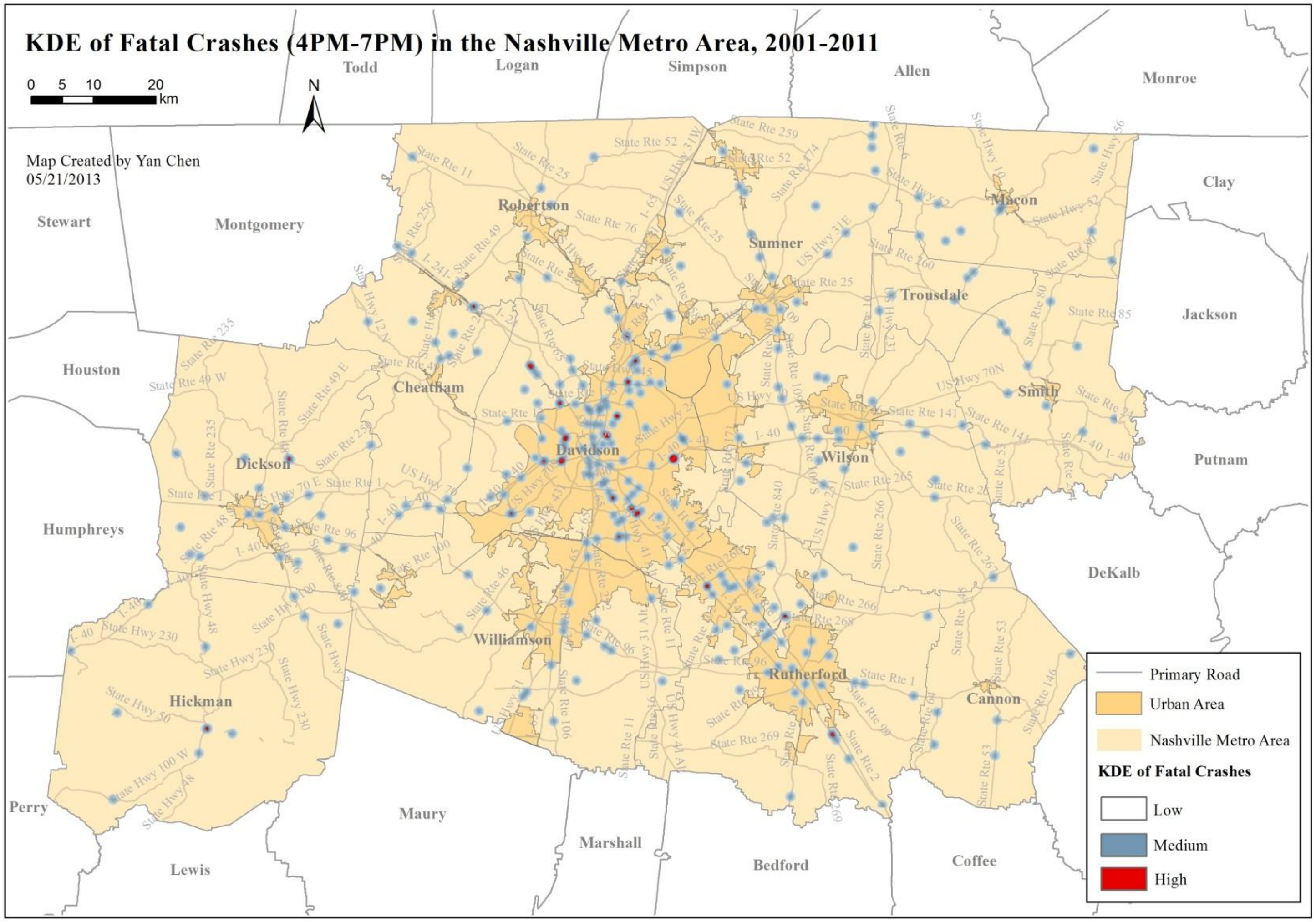


Figure 5.8. KDE of P.M. Rush Hour Fatal Crashes. Source: FARS (2001-2011).

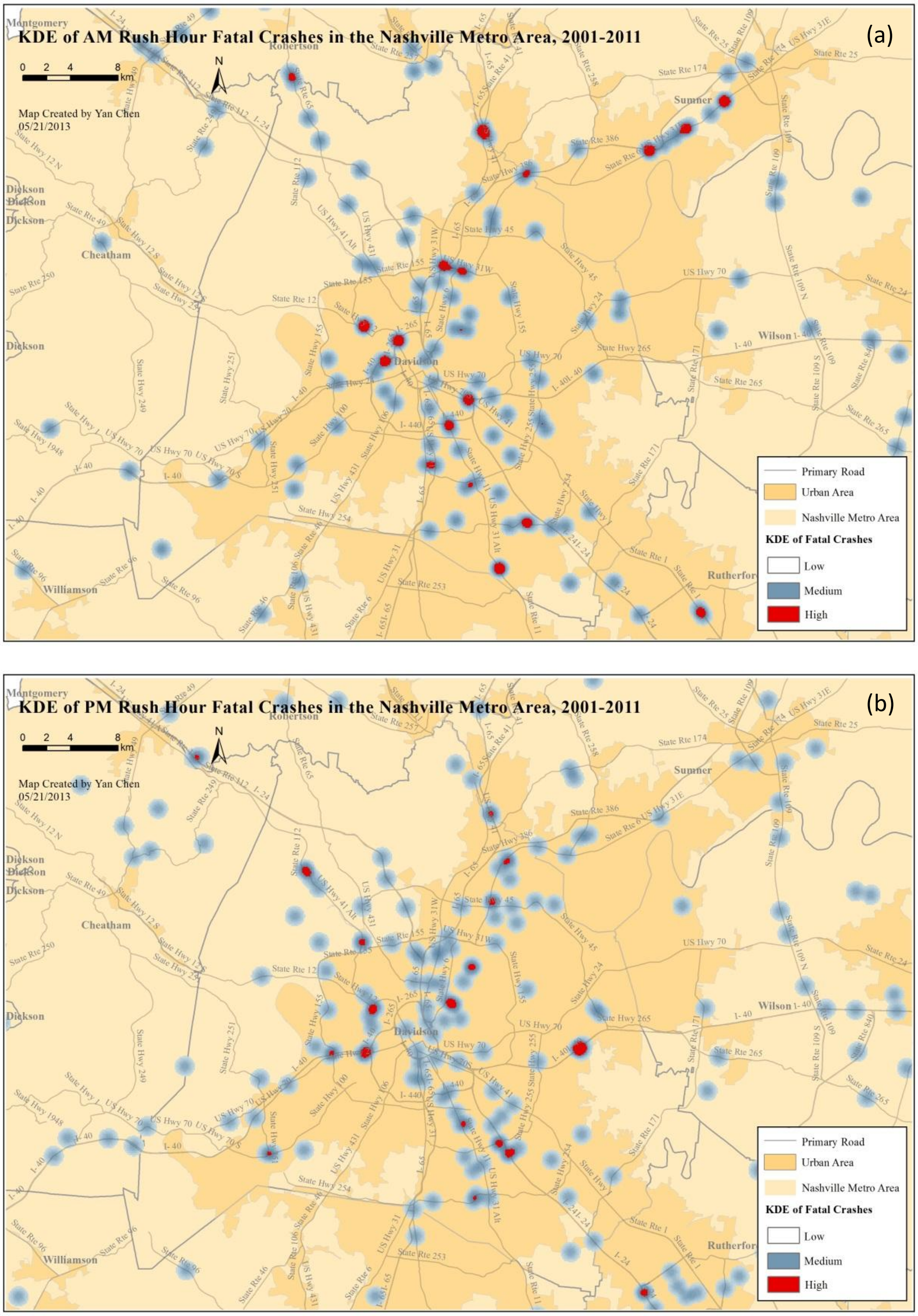


Figure 5.9. Side-by-Side Comparison of KDE of Fatal Crashes: (a) A.M. Rush Hour Fatal Crashes, and (b) P.M. Rush Hour Fatal Crashes. Source: FARS (2001-2011).

5.2.2.3. Spatial patterns of weekday and weekend fatal crashes

There were 1,465 and 821fatal crashes for weekdays (Monday to Friday) and weekends (Saturday to Sunday), respectively. Visual spatial patterns of weekday and weekend fatal crashes can be identified through their planar KDE surfaces, respectively (Figures 5.10 - 5.12). The spatial distributions of the weekday fatal crashes tend to be more conforming to the average daily traffic flows: centered within the urban area of Davidson County, as well as along the major Interstate highways (Figure 5.10). The weekday hotspots form a slant “T” shape in the direction of southeastern Rutherford County. Besides Davidson and Rutherford, weekday hotspots also were found in Wilson, Sumner, Robertson, and Dickson counties, in both urban and rural areas. Two notable hotspots within Davidson County include those at the interchange of I-24 and I-40 and the intersection of I-65 and State highway 155.

As for weekend fatal crashes, most of the hotspots were located in the urban areas of Davidson and Rutherford counties (Figure 5.11). Almost no hotspots were found in the other counties. Three hotspots in Davidson County stand out, including the interchange of I-40 and I-440, the interchange of I-24 and I-65, and the intersection of I-40 and State Highway 24. When comparing weekday and weekend fatal crashes, they all tend to cluster in the urban areas of the NMA and along major highways. The most notable difference between them is that there were fewer hotspots identified during weekends than during weekdays (Figure 5.12). In addition, the weekday fatal crashes tend to spread out more along major thoroughfares that commuters use for work, while the weekend fatal crashes are more concentrated in specific parts of the NMA. This can be explained by the different traffic patterns during weekdays and weekends.

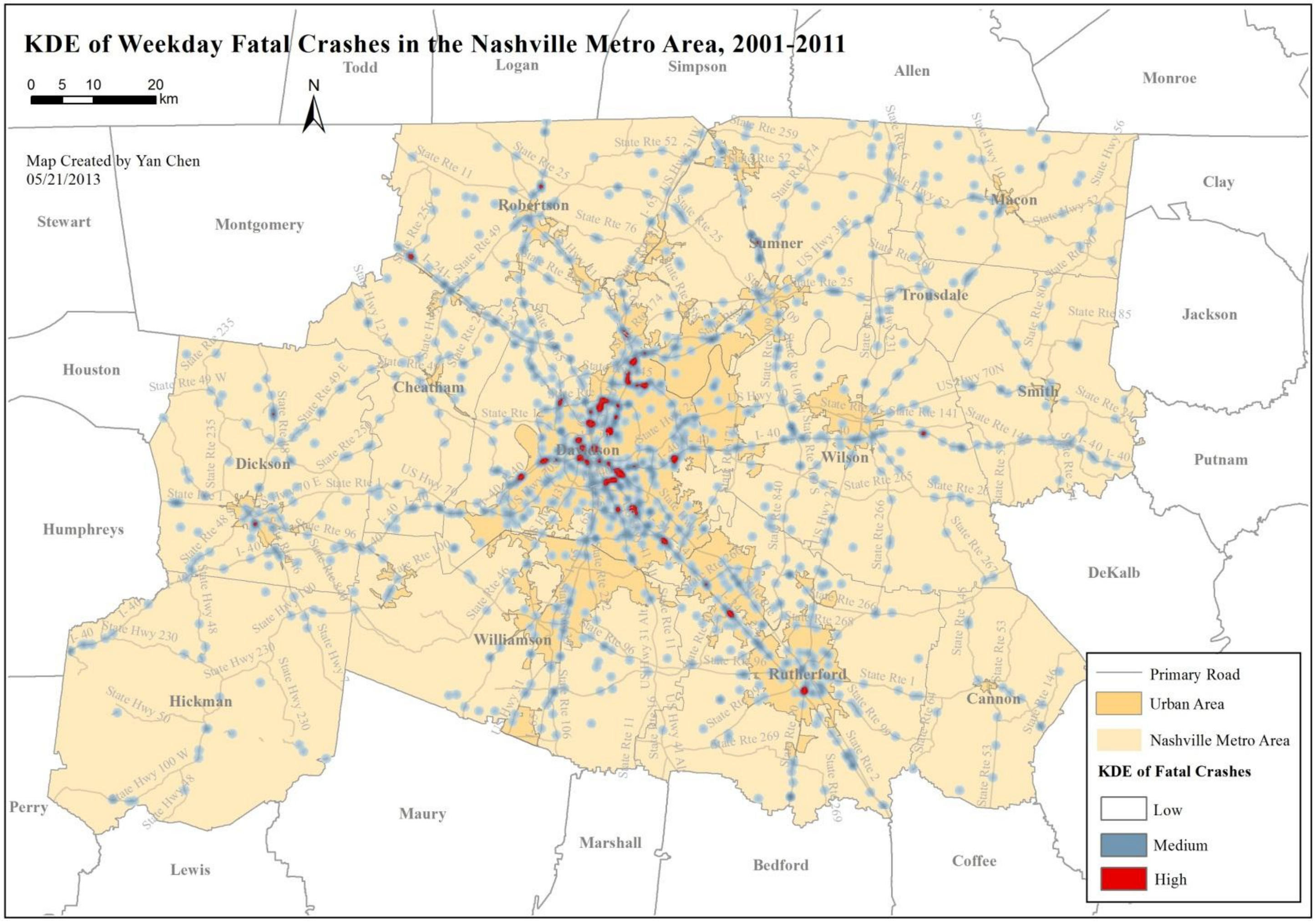


Figure 5.10. KDE of Weekday Fatal Crashes. Source: FARS (2001-2011).

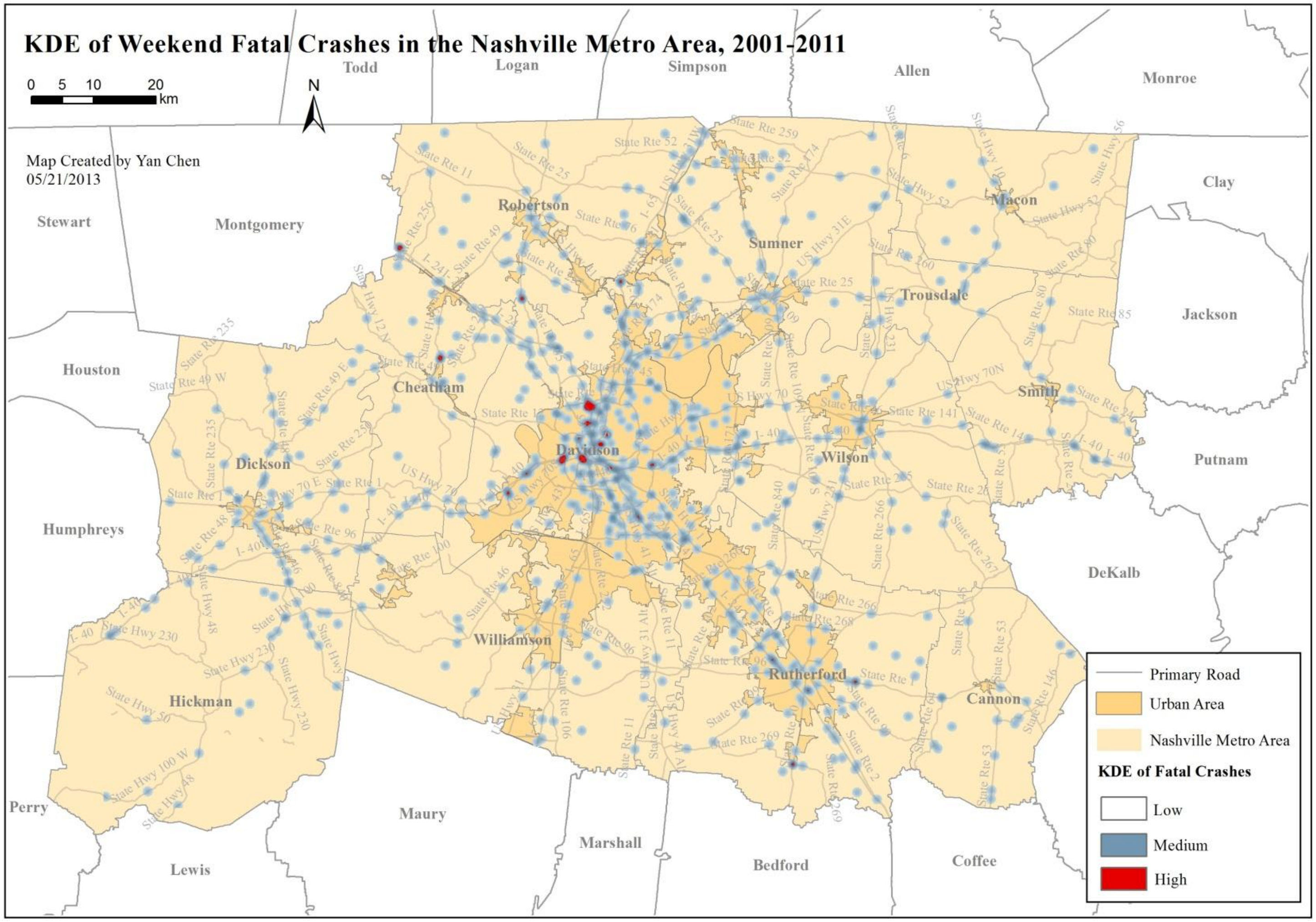
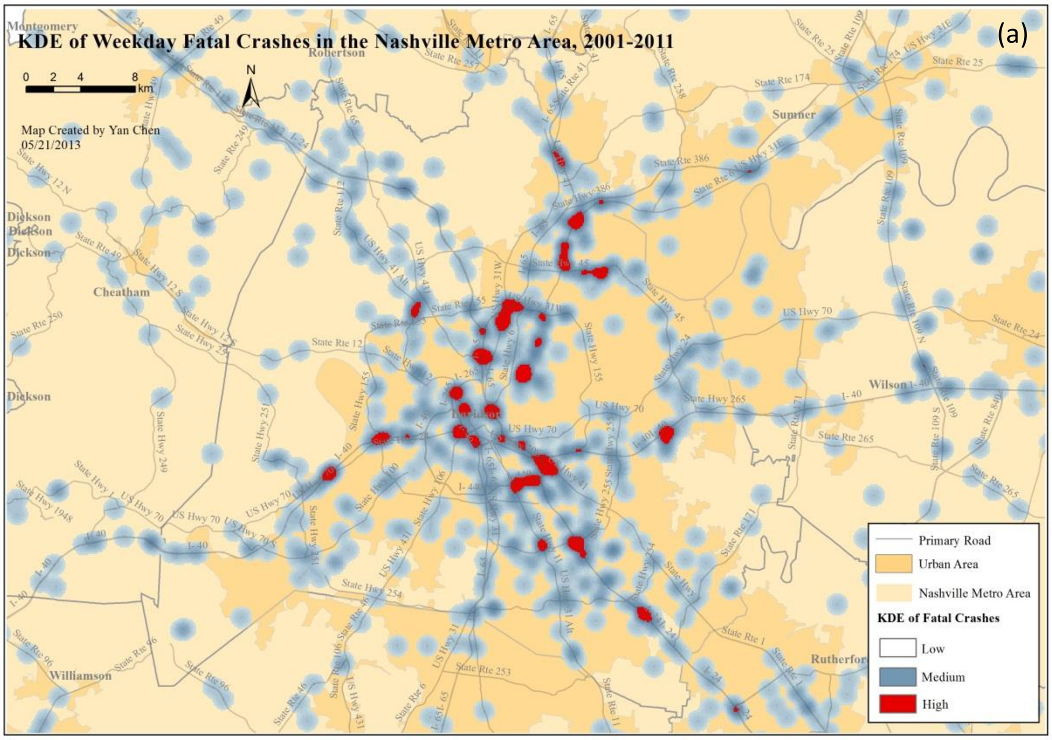


Figure 5.11. KDE of Weekend fatal Crashes. Source: FARS (2001-2011).



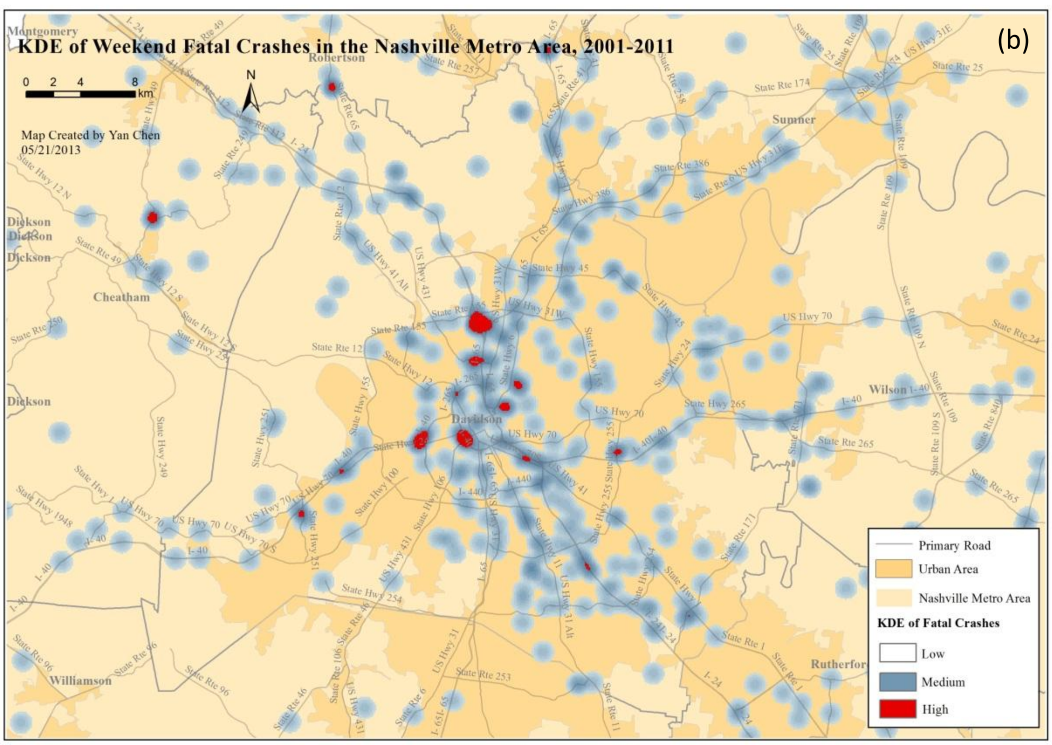


Figure 5.12. Side-by-Side Comparison of the KDE of Fatal Crashes: (a) Weekday Fatal Crashes, and (b) Weekend Fatal Crashes. Source: FARS (2001-2011).

5.2.3. Spatial Patterns of Subtypes of Fatal Crashes Detected by the Global Auto K Function

Spatial patterns for each subtype of fatal crashes based on time factors were observed using the global auto K function method. Detailed discussions about how to interpret the output graphs of Ripley's K Function can be found in Section 5.1.2. In short, all subtypes of fatal crashes were found to be clustered under roughly 60 – 68 km (Table 5.1, Figures 5.13 - 5.18). Since the spatial patterns of fatal crashes are highly confirmed to the daily traffic flow, the cluster pattern are within the radius of 34 km centered with the center of Nashville downtown area. Daytime, A.M. rush hours, and weekend fatal crashes were clustered at a little shorter distance than were their counterparts (nighttime, P.M. rush hours, and weekday, respectively). (Note: all units of output of the global auto K function are meters.)

Subtypes of Fatal Crashes	**Approximated Clustering Distance (km)**
Daytime Fatal Crashes	66
Nighttime Fatal Crashes	68
A.M. Rush Hours Fatal Crashes	60
P.M. Rush Hours Fatal Crashes	65
Weekday Fatal Crashes	67
Weekend Fatal Crashes	64
Note: The clustering distance for all FARS crashes is roughly 68km.	

Table 5.1. Clustering Distance of Subtypes of Fatal Crashes.

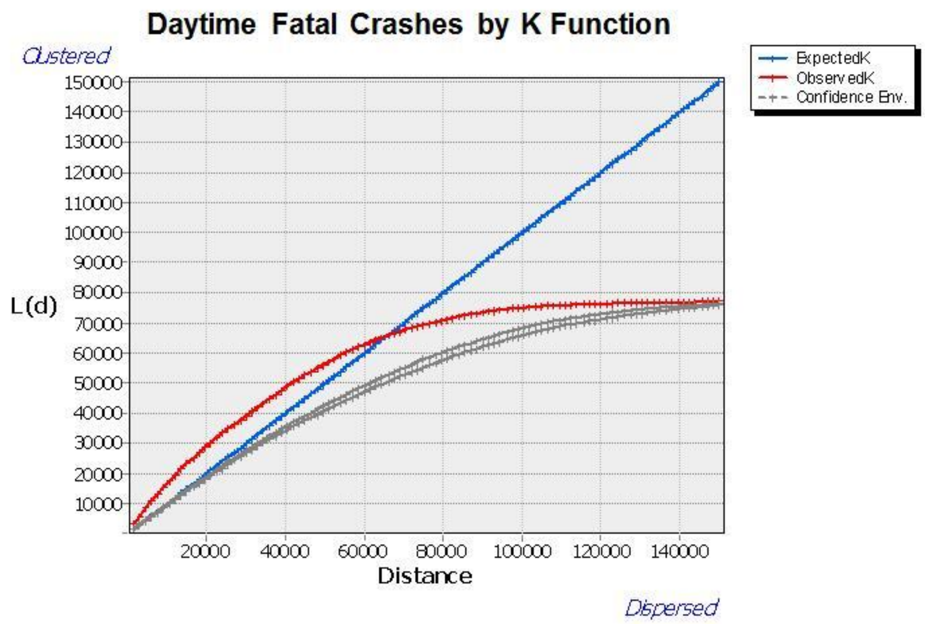


Figure 5.13. K Function for Daytime Fatal Crashes. Source: FARS (2001-2011).

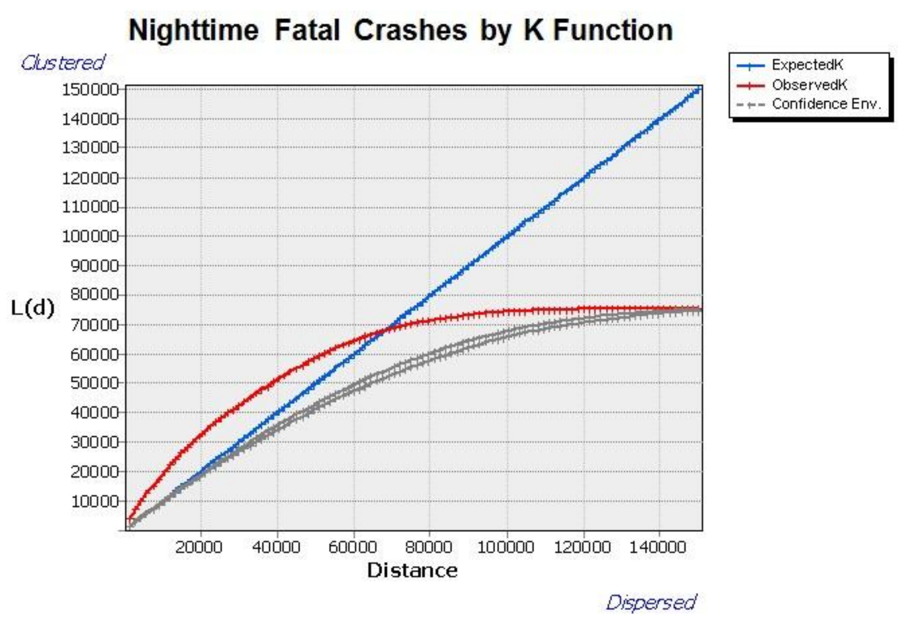


Figure 5.14. K Function for Nighttime Fatal Crashes. Source: FARS (2001-2011).

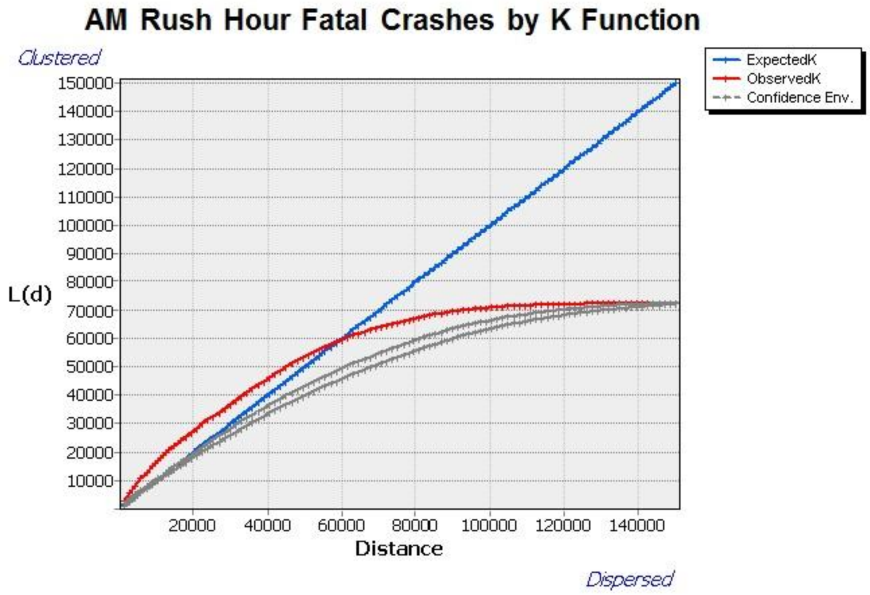


Figure 5.15. K Function for A.M. Rush Hour Fatal Crashes. Source: FARS (2001-2011).

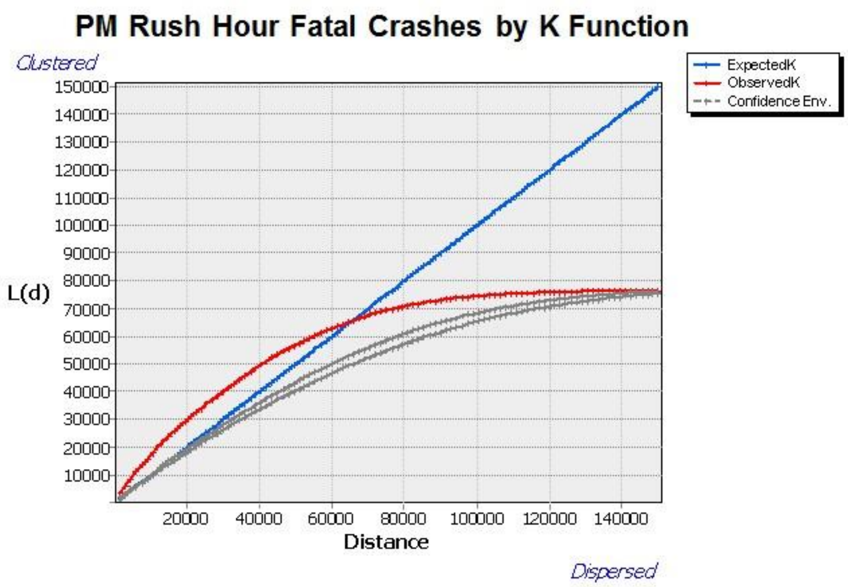


Figure 5.16. K Function for P.M. Rush Hour Fatal Crashes. Source: FARS (2001-2011).

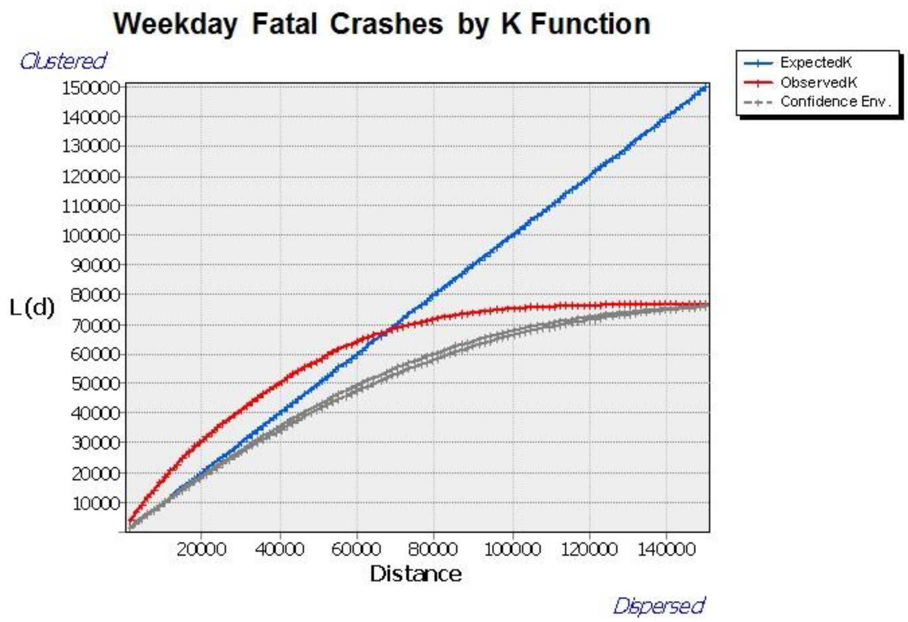


Figure 5.17. K Function for Weekday Fatal Crashes. Source: FARS (2001-2011).

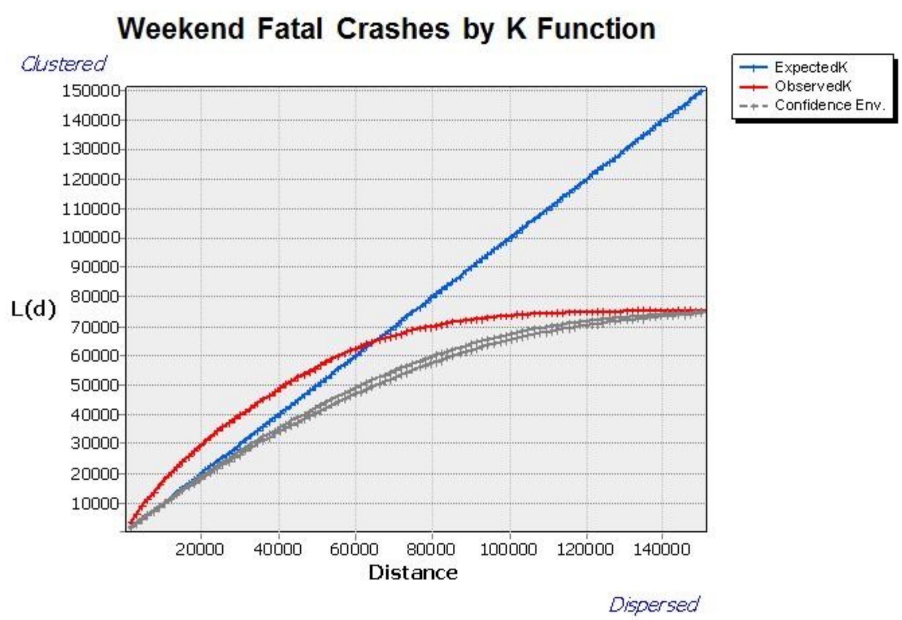


Figure 5.18. K Function for Weekend Fatal Crashes. Source: FARS (2001-2011).

5.3. Spatial Differences Found in Specific Locations

5.3.1. Spatial Differences between Nighttime and Daytime Fatal Crashes

Overall, there is no spatial difference found between daytime and nighttime fatal crashes based on the network global cross K function, as daytime fatal crashes were clustered around the nighttime hotspot locations under all distances at a 95% confidence level (Figure 5.19). However, this does not mean that there may be notable differences at some specific locations. The network local cross K function was adopted to detect the spatial differences between daytime and nighttime fatal crashes at selected hotspots of nighttime fatal crashes. All hotspots for nighttime fatal crashes were then classified into three types of hotspots, namely clustered, random, and dispersed based on the results of the network local cross K function (Figure 5.20). When a hotspot of nighttime fatal crashes was identified as clustered, it meant that there were more daytime fatal crashes than expected surrounding the nighttime hotspot (hence, a hotspot of nighttime fatal crashes is also a hotspot for daytime fatal crashes). When a dispersed type was identified, it meant that fewer daytime fatal crashes than expected were found around that nighttime hotspot and so there would be a significant spatial difference between nighttime and daytime fatal crashes (hence, a hotspot of nighttime fatal crashes is not a hotspot for daytime fatal crashes). Lastly, a random hotspot meant that there was neither a significant clustered nor a dispersed distribution, specifically for daytime fatal crashes, and their occurrence appears to be random at that location. In this study, since the objective is to look for localized spatial differences not at regional scales, all results of local cross K functions were compared only under a short distance of 2 km. In all of the 20 nighttime fatal crash hotspots, two hotspots were found to be clustered (more daytime fatal crashes around it than expected within 2 km of those two locations) and one hotspot was found dispersed (fewer daytime fatal crashes around it than expected within 2 km of that location). The rest were all random (the occurrence of daytime fatal crashes is random around them). A closer look at each representative hotspot type is presented in the following discussion. For the complete list of all network local cross K functions for nighttime hotspots, please refer to Appendix A.

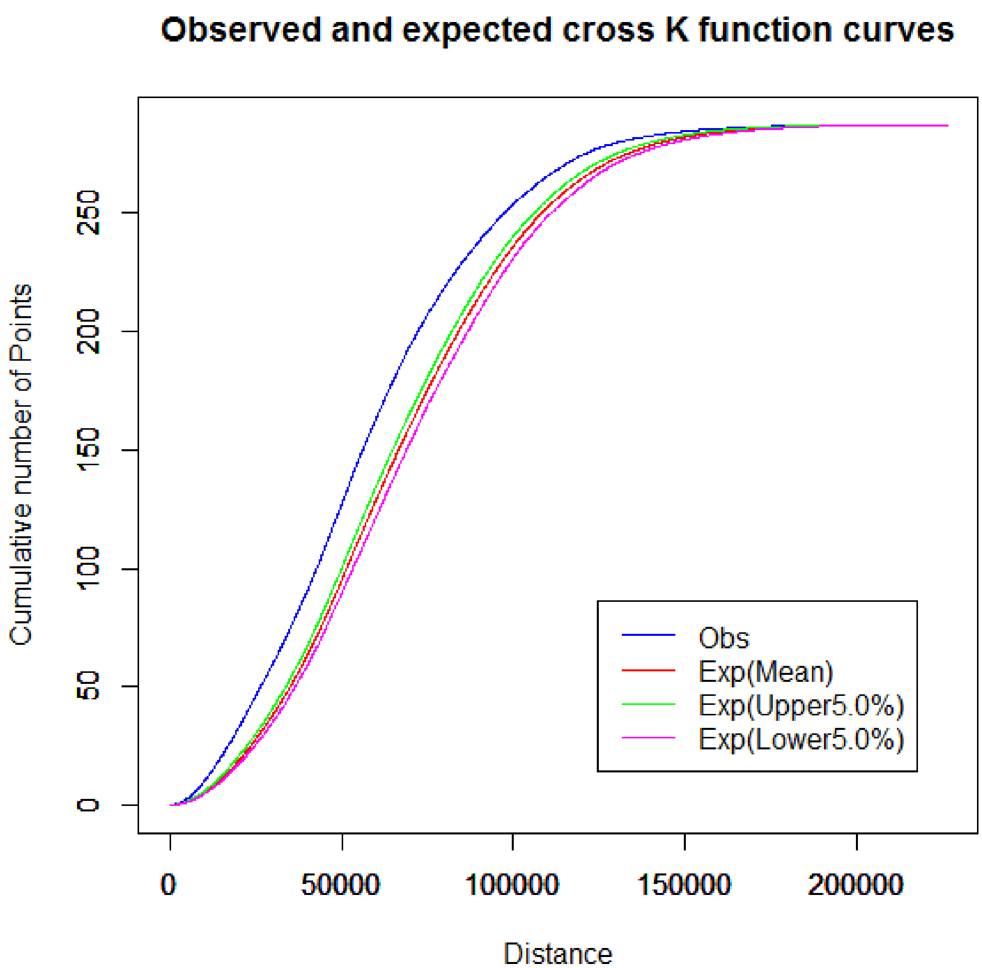


Figure 5.19. Network Global Cross K Function (Daytime Fatal Crashes against Nighttime Hotspots). Source: FARS (2001-2011).

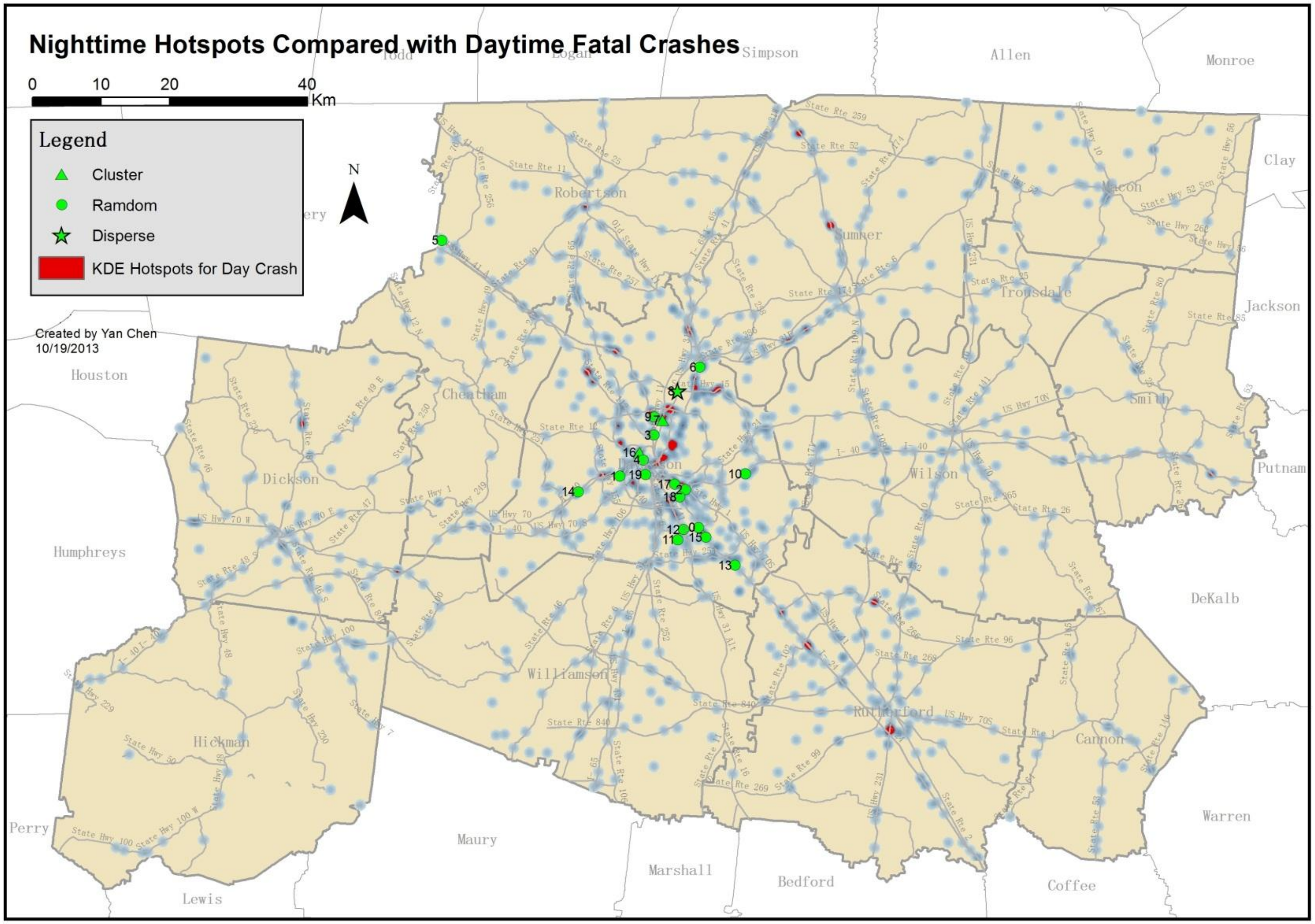


Figure 5.20. Types of Nighttime Hotspots of Fatal Crashes. Source: FARS (2001-2011).

Hotspot 8 of nighttime fatal crashes was identified as a dispersed type of hotspot. When looking at its network local cross K function (Figure 5.21), the blue curve (which stands for observed values) is below the red one (which stands for the expected values) within 2 km, which means the daytime fatal crashes tend to be dispersed from the nighttime hotspot, since there are fewer points found at that distance than expected. Moreover, the blue curve was also below the lower envelope, so this dispersed pattern is statistically significant at a 95% confidence level (the statistical significant level was set to 0.05). At this location, there are high concentrations of nighttime fatal crashes but not for daytime fatal crashes, and hence there is a statistically significant difference between daytime and nighttime fatal crashes at Nighttime Hotspot 8. As shown in Figure 5.22, there are no hotspots of daytime fatal crashes within the proximity of this nighttime hotspot.

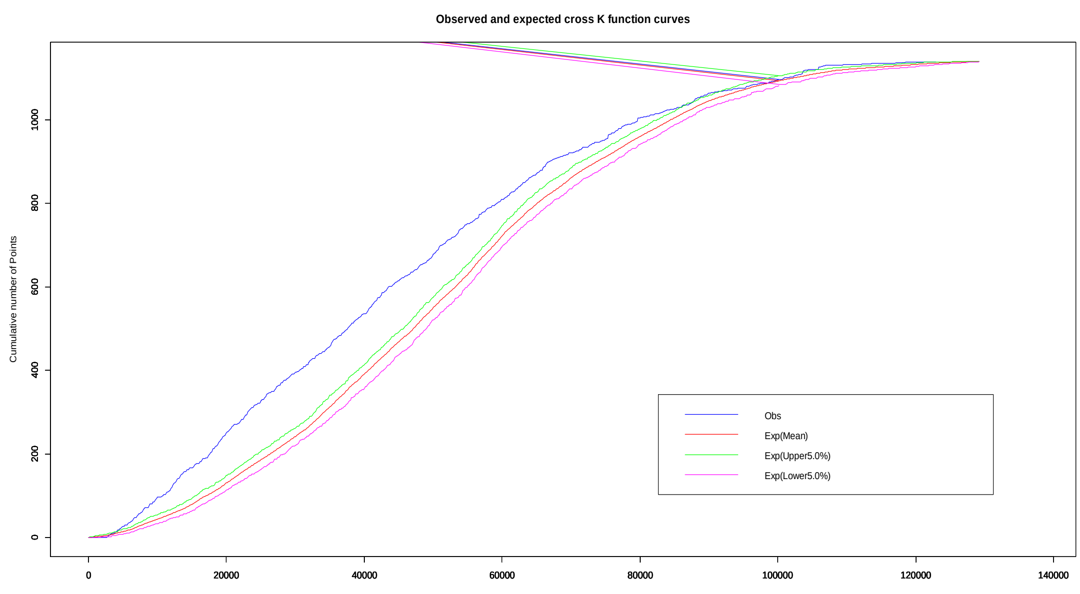


Figure 5.21. Network Local Cross K Function at Nighttime Hotspot 8. Source: FARS (2001-2011).

It would be useful to take a closer look at the location of Nighttime Hotspot 8 and to find out the potential causes behind this difference. Nighttime Hotspot 8 is located at the intersection of I-65 and State Highway 45 in Davidson County (Figure 5.22). In order to find out the potential factors that contribute to fatal crashes, photos were taken along State Highway 45 at Points A, B, C, and D (Figure 5.23). These four points are located roughly within a two-kilometer radius of that intersection.

There are multiple factors that may contribute to fatal crashes during the nighttime (but less likely during daytime) around this intersection. Point A is located at the intersection of US Highway 31W and State Highway 45 (Figure 5.24). It is worth noting that many facilities, such as the Shell gas station, Walgreen pharmacy, and fast food restaurants are located around nearby shopping plazas. The risk of traffic accidents may increase as traffic moves in and out of the highways. Also, turning at this intersection may be risky at night with multiple lanes, particularly during dark conditions.

Point B is located at the intersection of I-65 and State Highway 45 (Figure 5.25). First, this place contains the entrance and exit of I-65. Driving onto or off I-65 could distract drivers and impair their judgment. With a poor field of vision in dark conditions, for example, making decisions at that location could be more difficult. Second, there are two right-angle ramps linking State Highway 45 to I-65 in this location. Due to poor visibility, driving on these sharp curves could be dangerous, especially at night. Moreover, suddenly merging on to the ramps with sharp turns at high speed may make it hard for drivers to control their vehicles. Another possible explanation is that there is significantly heavier traffic during the nighttime at this location.

Point C is located along a local street, S. Graycroft Avenue, and near the intersection of S. Graycroft Avenue and State Highway 45 (Figure 5.26). Two factors may lead to traffic accidents and fatal crashes during the nighttime. First, the street level is lower than the surrounding ground and this could easily block the view field of drivers at this location, especially during nighttime. It could be dangerous for drivers to make turns without a more open vision. Second, drivers could easily drive to an opposite lane if they are not familiar with the roads, especially under dark conditions.

Point D is located at the intersection of a local street, Beverly Drive, and State Highway 45 (Figure 5.27). This is a T-junction. From the photo, it is easy to see that the two streets form a sharp angle, which is difficult for drivers to maneuver when they make a left turn from Beverly Drive to State Highway 45. In addition, driving from a low-speed-limit street to a high-speed-limit street may also be a factor. Under dark conditions and without proper traffic signals, drivers on State Highway 45 (with a higher speed limit) might encounter difficulty avoiding vehicles suddenly trying to merge.

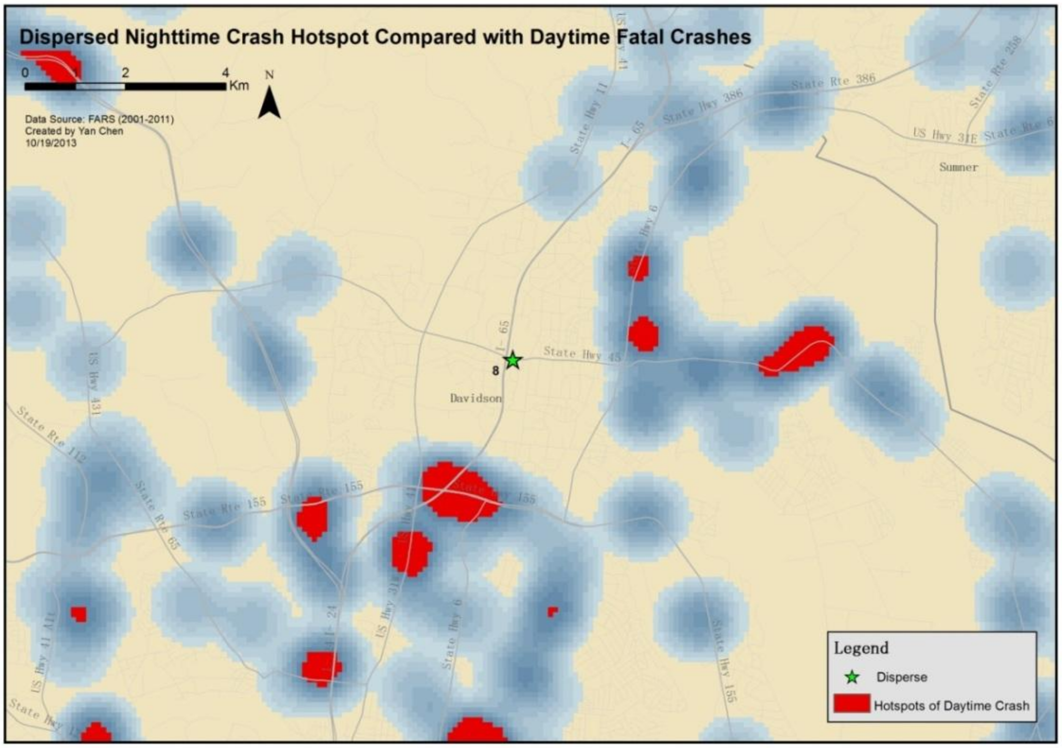


Figure 5.22. A Close-up View of Nighttime Hotspot 8 (Dispersed). Source: FARS (2001-2011).

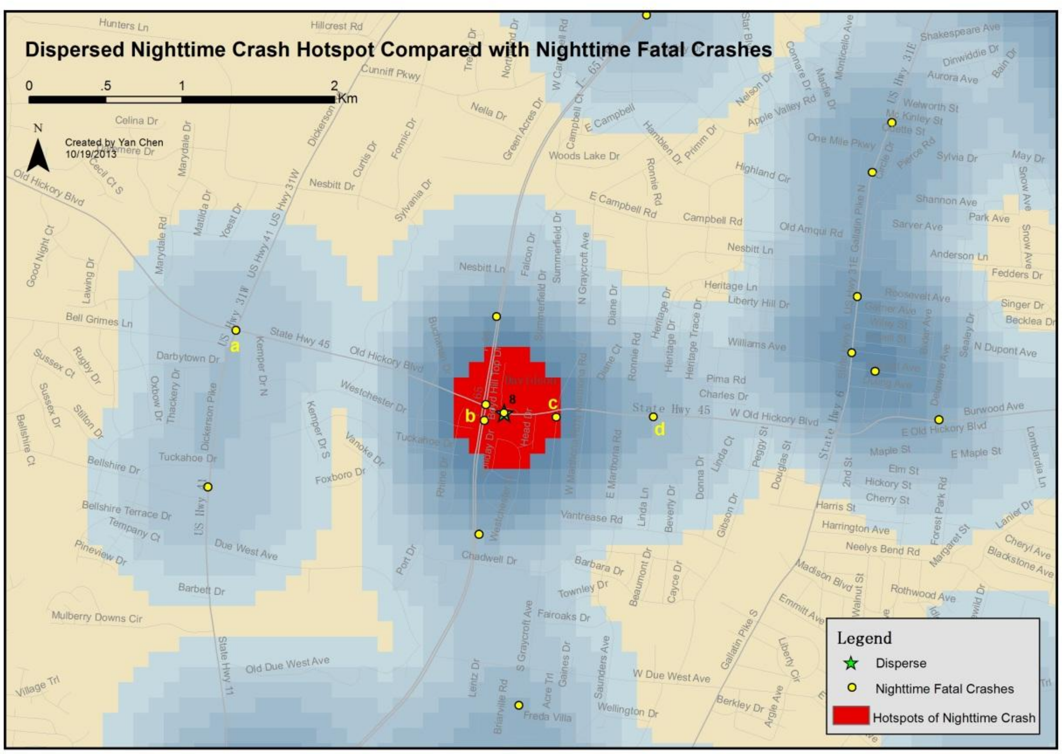


Figure 5.23. A Close-up View of Nighttime Hotspot 8 (Dispersed) and the Surrounding Fatal crashes. Source: FARS (2001-2011).

Figure 5.24. Photo Taken at Point a near Nighttime Hotspot 8.

Figure 5.25. Photo Taken at Point b near Nighttime Hotspot 8.

Figure 5.26. Photo Taken at Point c near Nighttime Hotspot 8.

Figure 5.27. Photo Taken at Point d near Nighttime Hotspot 8.

Let us also take a close look at Nighttime Hotspot 16, a clustered hotspot. In Figure 5.28, the blue curve (observed value) is above the red curve (expected value) below the distances of 2 km. The daytime fatal crashes therefore tend to cluster around the nighttime hotspot, since there were more points found at that distance than expected under CSR. Furthermore, the observed curve is also above the upper envelope (green curve), which means the cluster pattern is statistically significant with a 95% confidence level. Both nighttime fatal crashes and daytime fatal crashes are tending to occur at this location. In other words, Nighttime Hotspot 16 of nighttime fatal crashes is also a hotspot for daytime fatal crashes.

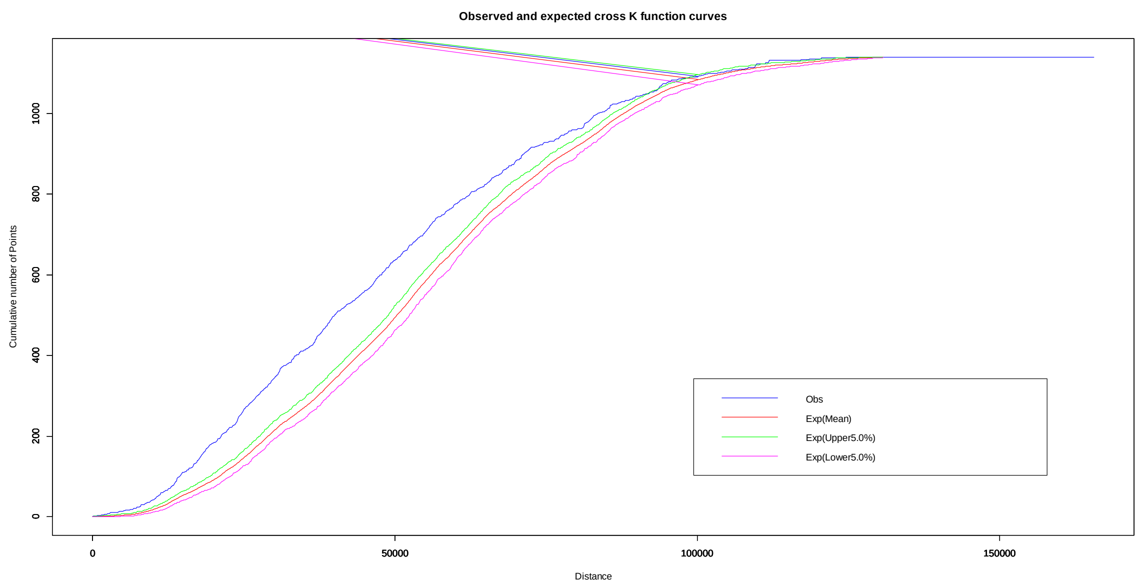


Figure 5.28. Network Local Cross K Function at Nighttime Hotspot 16. Source: FARS (2001-2011).

So what may cause this location to have a high concentration of fatal crashes during the nighttime and daytime? Night Hotspot 16 is located at the intersection of I-65 and U.S. Highway 41 in central Davidson County (Figure 5.29). An overpass, combined with complex road networking, makes this place difficult to navigate. Two main problems, which may cause serious traffic accidents regardless of time, can be found in this location. First, driving off the overpass of U.S. Highway 41 and onto I-65 includes passing through an extremely sharp left turn. Second, cars could be difficult to control when drivers make turns on sharp curves from higher ground to lower ground.

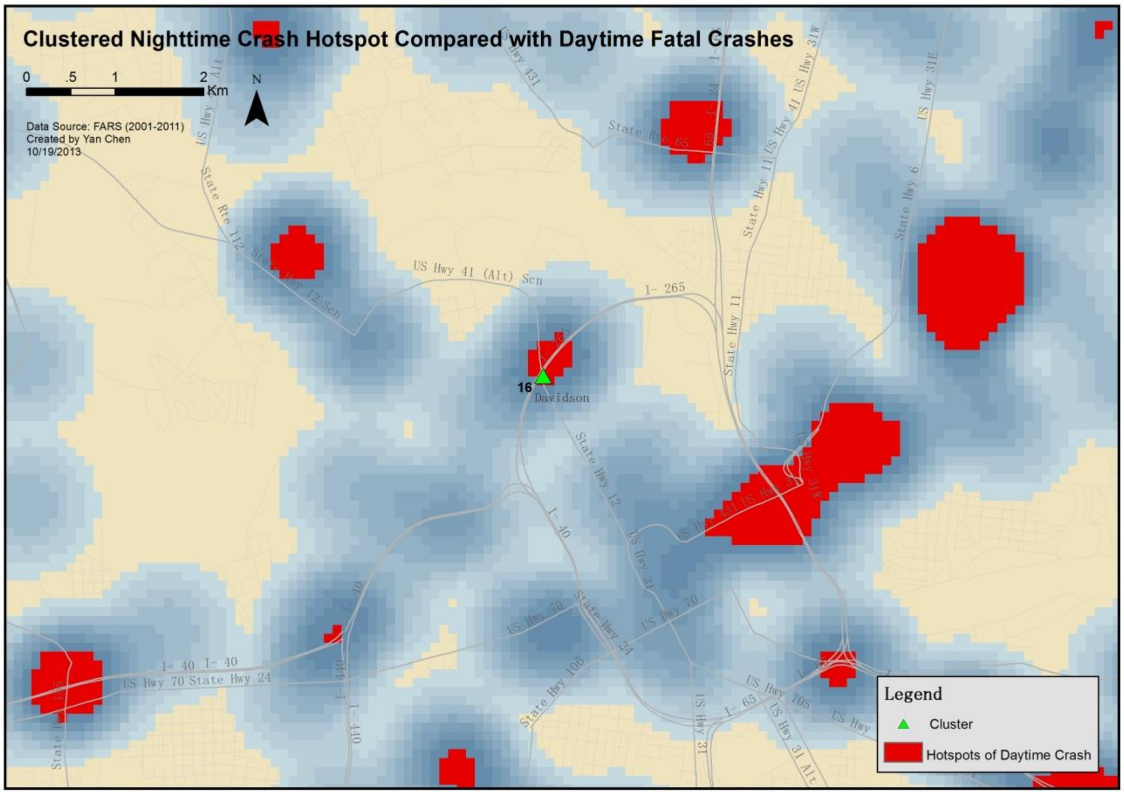


Figure 5.29. A Close-up View of Nighttime Hotspot 16 (Clustered). Source: FARS (2001-2011).

Nighttime Hotspot 19 was found to be random, which means that, statistically speaking, the occurrence of daytime fatal crashes is random at this location. As shown in Figure 5.30, the blue curve is within the upper (green color) and lower envelopes (pink color) under distance of 2,000 meters. Any curve within the envelopes indicates a random distribution. So even though this location can be visually identified as a nighttime hotspot, it cannot be identified as a daytime hotspot statistically.

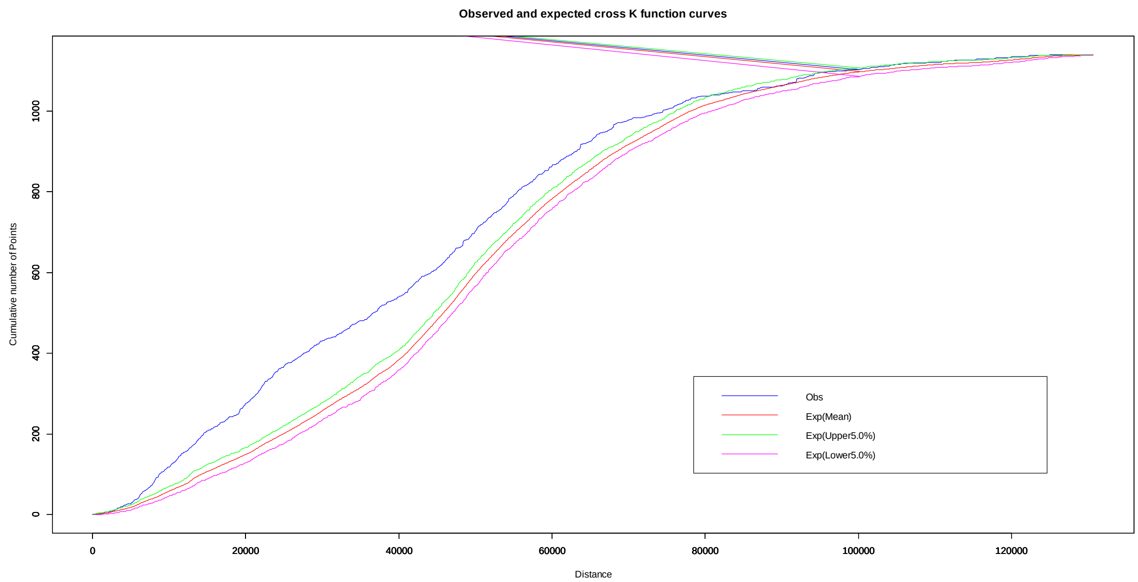


Figure 5.30. Network Local Cross K Function at Nighttime Hotspot 19. Source: FARS (2001-2011).

5.3.2. Spatial Differences between P.M. Rush Hours and A.M. Rush Hours

There is also no overall spatial difference when comparing all P.M. rush hour fatal crashes with all A.M. rush hour fatal crashes through the network global K function (Figure 5.31). To look for spatial differences at the locations of P.M. rush hour hotspots, the network local cross K function method was used. Similarly, all hotspots for P.M. rush hour fatal crashes could be classified into three types of hotspots as well, namely clustered, random, and dispersed. Again, the interpretation of hotspot type was done by examining the network local cross K function under a distance below 2 km. In this case, only two types of hotspots (15 random hotspots and 6 dispersed hotspots) were identified (Figure 5.32). No clustered type of hotspots was found, which indicates that, at the locations of P.M. rush hour hotspots, there was no high concentration of A.M. rush hour fatal crashes. This is a rather surprising finding. It is also worth noticing that all dispersed hotspots of P.M. rush hour fatal crashes are located outside of the CBD. A closer examination of a representative dispersed hotspot is presented below. For the complete list of all network local cross K functions for P.M. rush hour hotspots, please refer to Appendix B.

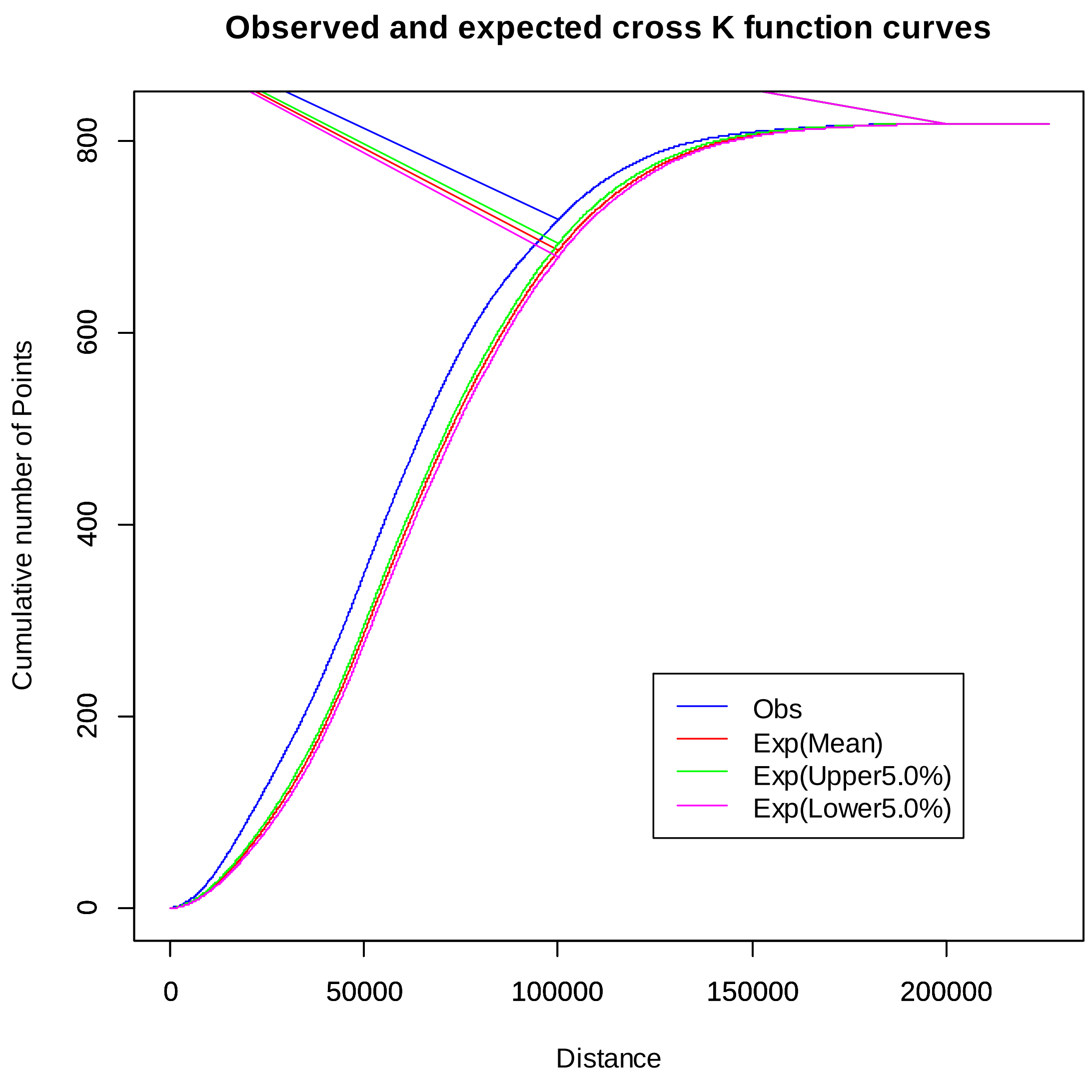


Figure 5.31. Network Global Cross K Function (A.M. Rush Hour Crashes against P.M. Rush Hour Hotspots). Source: FARS (2001-2011).

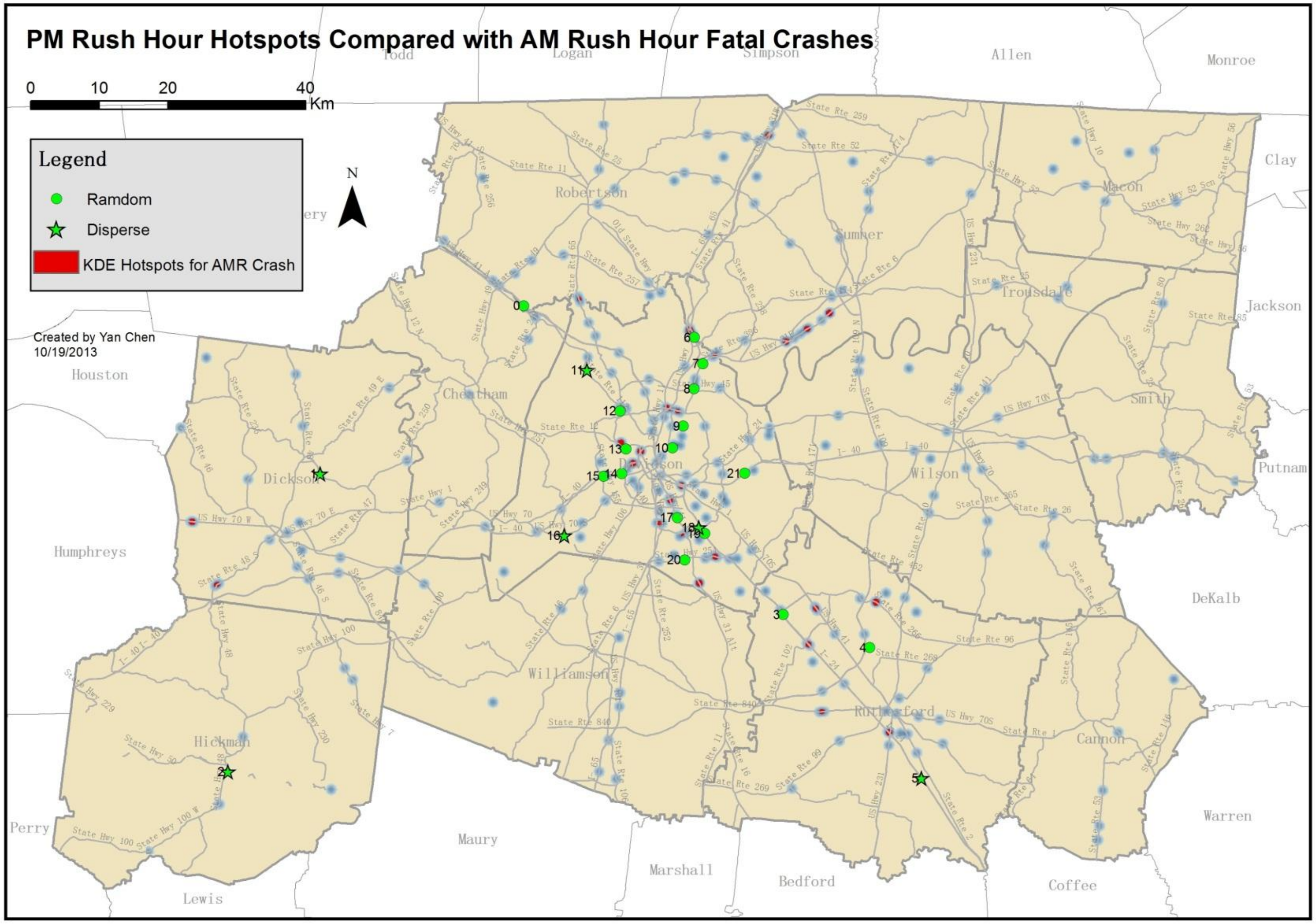


Figure 5.32. Types of P.M. Rush Hour Hotspots of Fatal Crashes. Source: FARS (2001-2011).

Let us take a close look at Hotspot 16 of P.M. rush hour fatal crashes. In Figure 5.33, the blue curve (i.e. observed values) is under the red curve (i.e. expected values) within 2 km. Any blue curve under the red curve indicates a dispersed pattern. In this case, the A.M. rush hour fatal crashes tend to be distributed away from the P.M. rush hour hotspot at this hotspot. The blue curve was also found below the lower envelope, which indicates a statistically significant dispersed pattern. Therefore, at this location, there is a high concentration of fatal crashes during P.M. rush hours but not A.M. rush hours, and there is a statistically significant difference in the spatial distribution of A.M. rush hours and P.M. rush hours fatal crashes at this location.

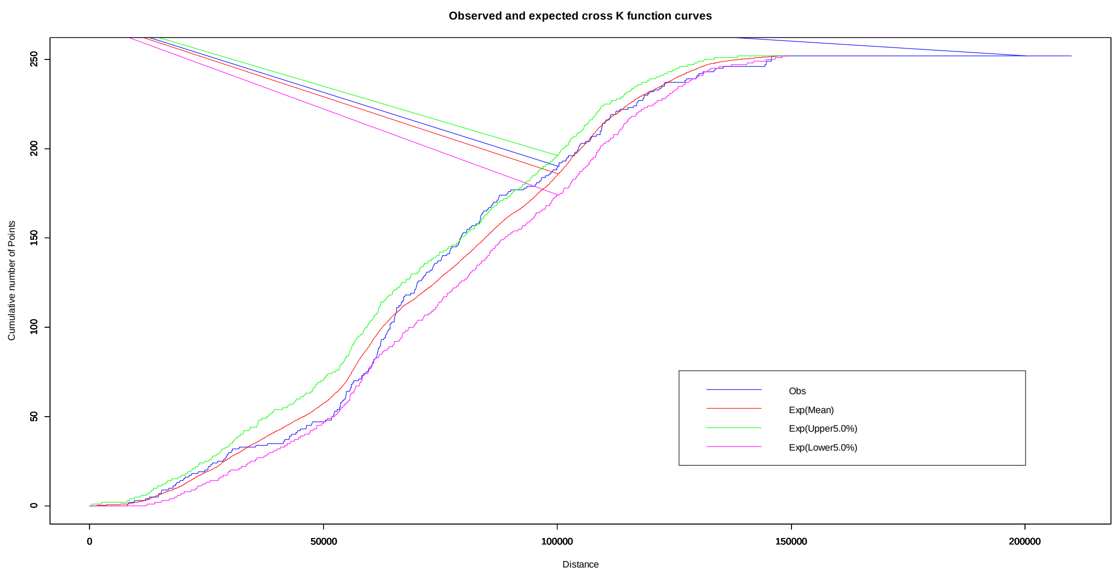


Figure 5.33. Network Local Cross K Function at P.M. Rush Hour Hotspot 16. Source: FARS (2001-2011).

The P.M. Rush Hour Hotspot 16 was located at the intersection of U.S. Highway 70S and State Highway 251 in Davidson County (Figure 5.34). Multi-lane roads from four directions converge at this intersection, which could explain why this intersection becomes a hotspot of fatal crashes. Additionally, a supermarket and several restaurants are located around this area, which may likely contribute to heavy traffic during P.M. rush hour and hence more fatal crashes.

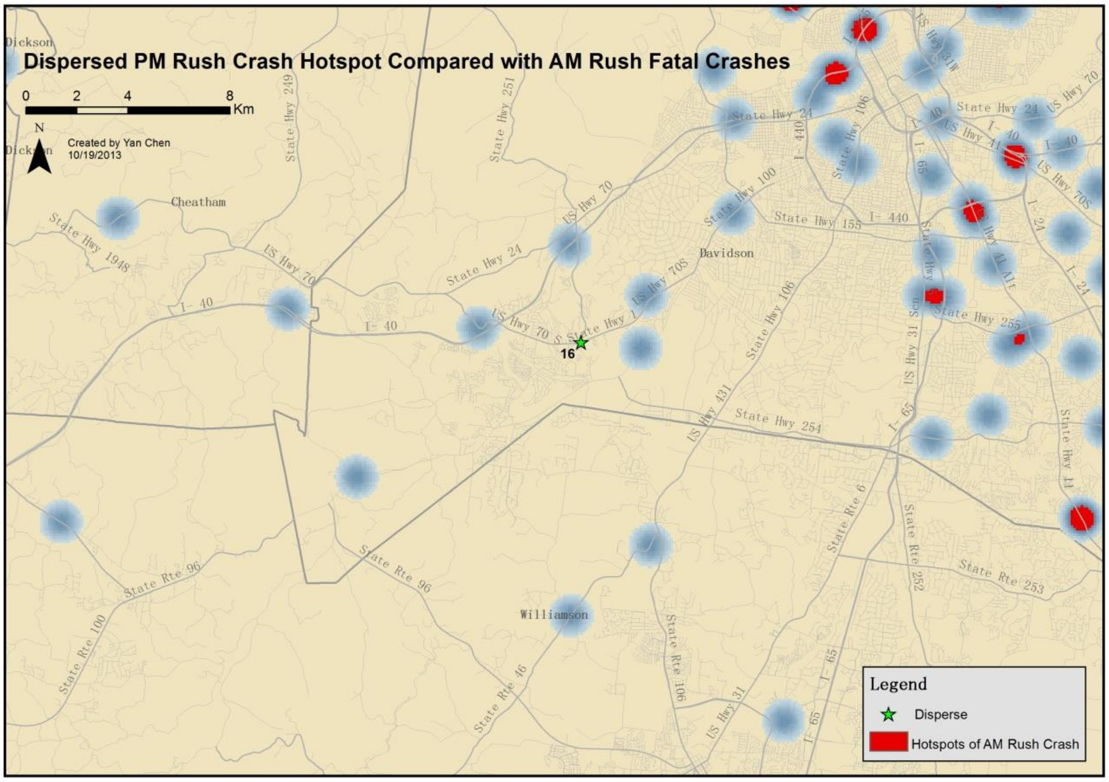


Figure 5.34. A Close-up View of P.M. Rush Hour Hotspot 16 (Dispersed). Source: FARS (2001-2011).

5.3.3. Spatial Differences between Weekend and Weekday Fatal Crashes

Likewise, no overall spatial difference was found between weekend crashes and weekday crashes based on the results of the network global cross K function (Figure 5.35). Spatial differences between weekend fatal crashes and weekday fatal crashes were then analyzed with the network local cross K function at selected hotspot locations of weekend fatal crashes. Similar to the previous analysis, all weekend hotspots of fatal crashes could be classified into three categories, namely clustered, random, and dispersed (Figure 5.36), based upon the interpretation of the network local cross K function within a distance of 2 km. In all of the 21 weekend hotspots, five are clustered hotspots and the rest are all random. A closer look at a representative clustered hotspot is presented below. For the complete list of all network local cross K functions for weekend hotspots, please refer to Appendix C.

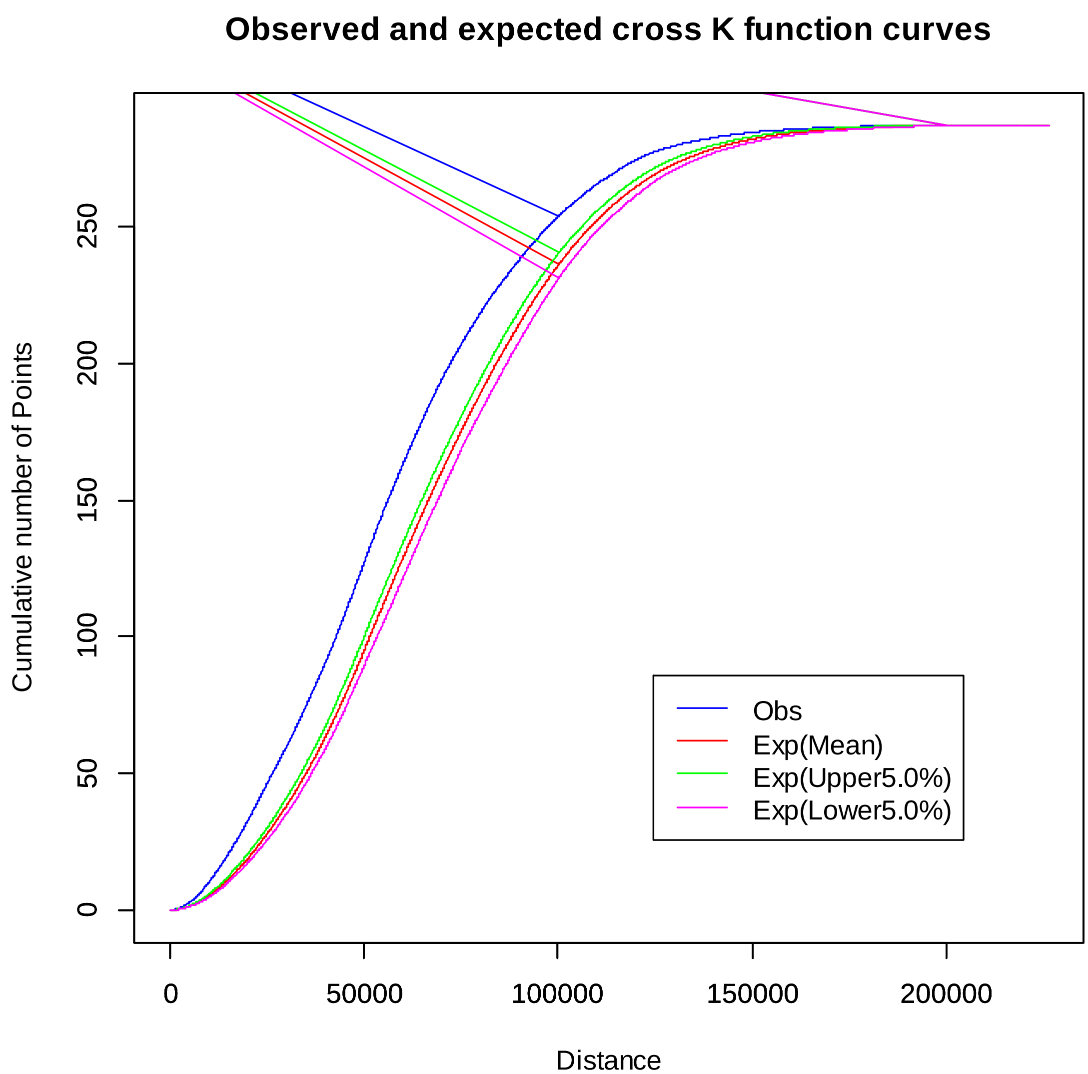


Figure 5.35. Network Global Cross K Function (Weekday Crashes against Weekend Hotspots). Source: FARS (2001-2011).

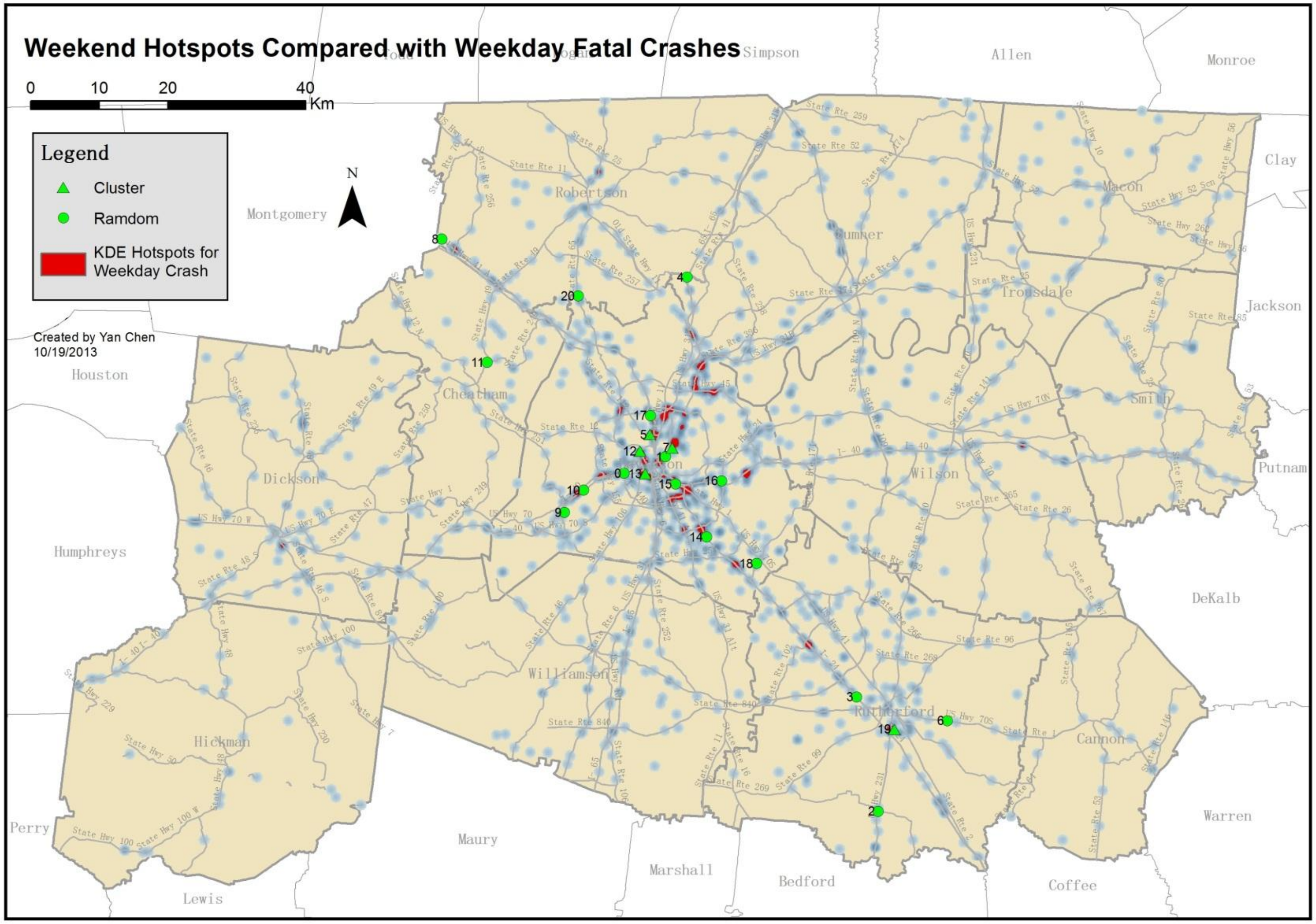


Figure 5.36. Types of Weekend Hotspots of Fatal Crashes. Source: FARS (2001-2011).

Let us look at Hotspot 19 of weekend fatal crashes. In Figure 5.37, the observed curve (blue color) is slightly above the upper envelope curve (green color) within 2 km, indicating that weekday fatal crashes tend to cluster around this weekend hotspot in this distance range. In other words, both weekend fatal crashes and weekday fatal crashes were all likely to occur at this same location. The Weekend Hotspot 19 is located on S Rutherford Blvd between U.S. highway 231 and U.S. highway 41 within the urban area of Rutherford County (Figure 5.38). Driving on this road is dangerous since there are many factors that may result in fatal crashes, including a sharp turn, a T-junction, and a two-lane bi-directional road without safe barriers. Even though the curve of the road seems not to be sharp, driving down from the overpass and continuing with such a curve could make it difficult for drivers to react quickly. Moreover, both the curve and T-junction are linked together, while the poor field of vision could influence drivers' judgment. Additionally, no safety barriers were found in the middle of the bi-directional road. The glaring headlights of other vehicles from the opposite direction could impair drivers' driving performance at night.

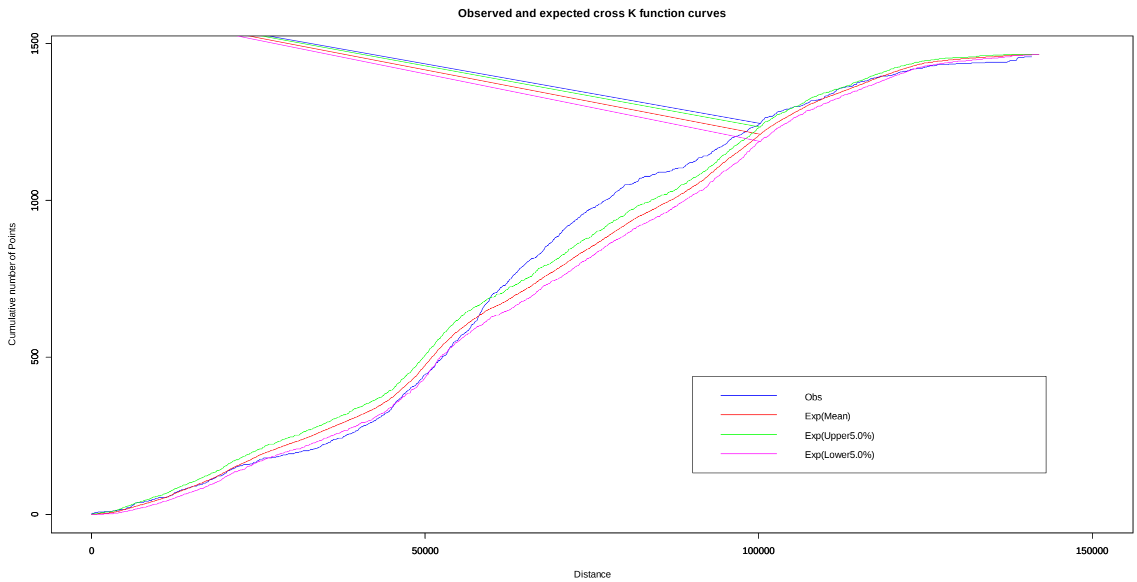


Figure 5.37. Network Local Cross K Function at Weekend Hotspot 19. Source: FARS (2001-2011).

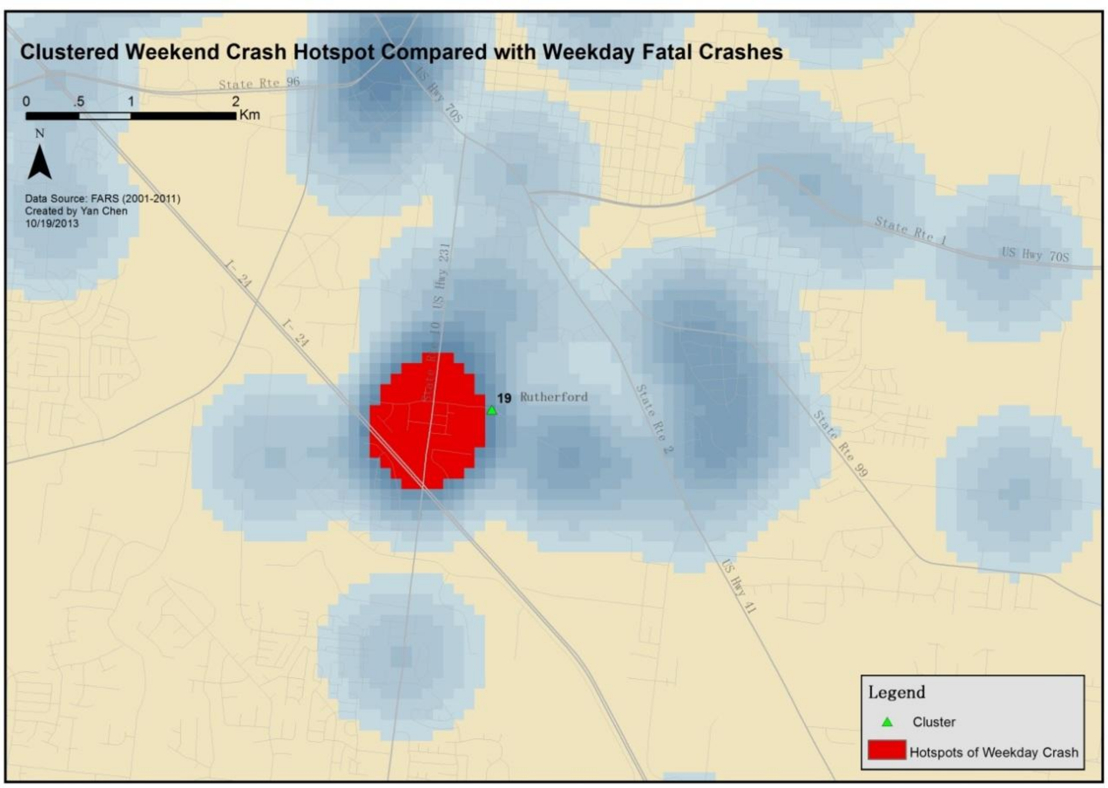


Figure 5.38. A Close-up View of Weekend Hotspot 19 (Clustered). Source: FARS (2001-2011).

CHAPTER 6. CONCLUSIONS AND FUTURE RESEARCH

6.1. Conclusions

The overall spatial patterns of fatal automobile crashes were found to be clustered under the distance of 68 km (within a radius of 34 km centered on the Nashville downtown area) and largely conform to the patterns of average daily traffic flow in Nashville, TN. The spatial distribution of each subtype of fatal crashes based on time of day and day of week have been examined as well. All subtypes of fatal crashes are found to be clustered within the central urban area of Nashville and tend to be located along the major roads, particularly the intersections and interchanges of major highways. These findings agree with the hypothesis of this study that fatal crashes no matter when they occurred are likely to be clustered in locations with high traffic volume.

Moreover, spatial differences among the subtype of fatal crashes based on time factors (daytime vs. nighttime, A.M. rush hours vs. P.M. rush hours, and weekdays vs. weekends) have been compared with both the network global cross K function and network local global cross K function. Based on the output of the network global cross K function, there is no notable overall spatial difference between each pair of subtype fatal crashes as a whole. However, some localized spatial differences between each pair of subtypes of fatal crashes have been observed in the results of the network local cross K function.

All three types of patterns, namely clustered, random, and dispersed, have been identified in the comparison of nighttime hotspots versus daytime fatal crash points. When analyzing A.M. rush hour fatal crashes against P.M. rush hour hotspots, only dispersed and random types were identified, indicating in certain areas that there are statistical differences in the spatial distribution of A.M. and P.M. rush hour fatal crashes. The weekday fatal crashes tend to be more likely to cluster around the weekend fatal crash locations and only clustered random types were identified when comparing weekday fatal crashes to weekend hotspots, indicating that there are no noticeable differences in the spatial distribution of weekday and weekend fatal crashes. This is different from the original hypothesis because weekday fatal crashes tend to cluster around some weekend hotspots.

This study provides a new approach to spatial point-pattern analysis. An integrated two-step approach was adopted, first to identify general hotspots of fatal crashes with KDE, followed by a more localized examination of the network local cross K function. This approach can be very effective because the KDE method is good for the visual detection of the high concentrations of spatial events via estimated density maps but the spatial differences cannot be compared statistically. As for the K functions, they provide a formal means of statistical inference but cannot be directly associated with visual geographic maps. An integration of these two groups of methods combines strengths from both. In this study, the spatial differences among the three subtypes of fatal crashes, day and night, morning rush hours and night rush hours, and weekdays and weekends, could be closely examined at specific localized places. When mapping the hotspot types (e.g., clustered, dispersed, or random) with the KDE surface together, statistical differences in the spatial distribution of subtype fatal crashes can be identified (for instance, no high concentration of daytime fatal crashes were located in the proximity of the dispersed hotspots of nighttime fatal crashes).

The approach adopted in this study not only provides a new way to analyze the spatial distribution of spatial point events such as fatal crashes, but it also can be applied readily to real-world applications. In general, the spatial patterns of fatal automobile crashes are highly influenced by daily human activities. Urban planners could use the findings to evaluate if certain land-use practices in the NMA can reduce the overall regional traffic demand. For instance, could a new shopping center opened in the suburbs relieve the pressure of traffic demand on the central urban area? The findings of this study can also assist transportation planners to make decisions such as which existing roads need to be expanded or where new roads need to be added in order to develop a better coverage of roadway network. Traffic engineers in the NMA could utilize the findings in this study to improve transportation infrastructures as well as roadway designs. For instance, knowing the spatial differences between nighttime and daytime fatal crashes at some locations may suggest that certain localized factors may require close investigation. If the nighttime crash hotspots are detected in locations with bad road conditions such as sharp curves, poorly lit streets, or narrow lanes, then improvement measures such as enhancing the visibility of road signs, adding traffic lights, and widening the lanes could be implemented. Moreover, if restaurants, nightclubs, liquor stores, and supermarkets are found near a nighttime crash hotspot, law enforcement could target this location for blood alcohol content tests or speed tests on nearby roads.

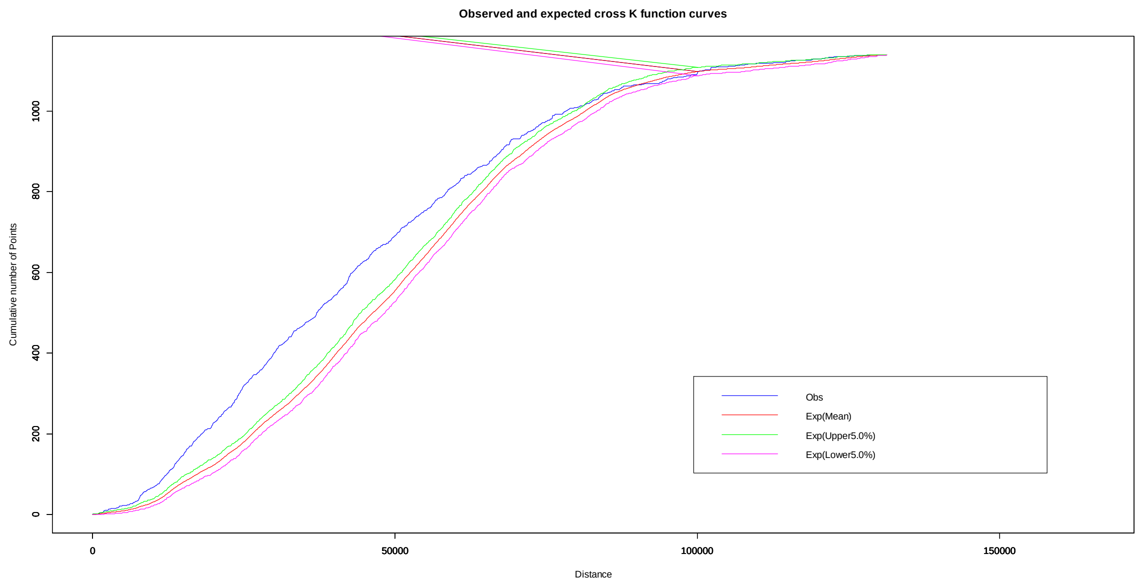
6.2. Future Studies

To understand the locations of fatal crashes in depth, a more detailed statistical analysis of crash rates based on daily traffic counts could be conducted. The unique route

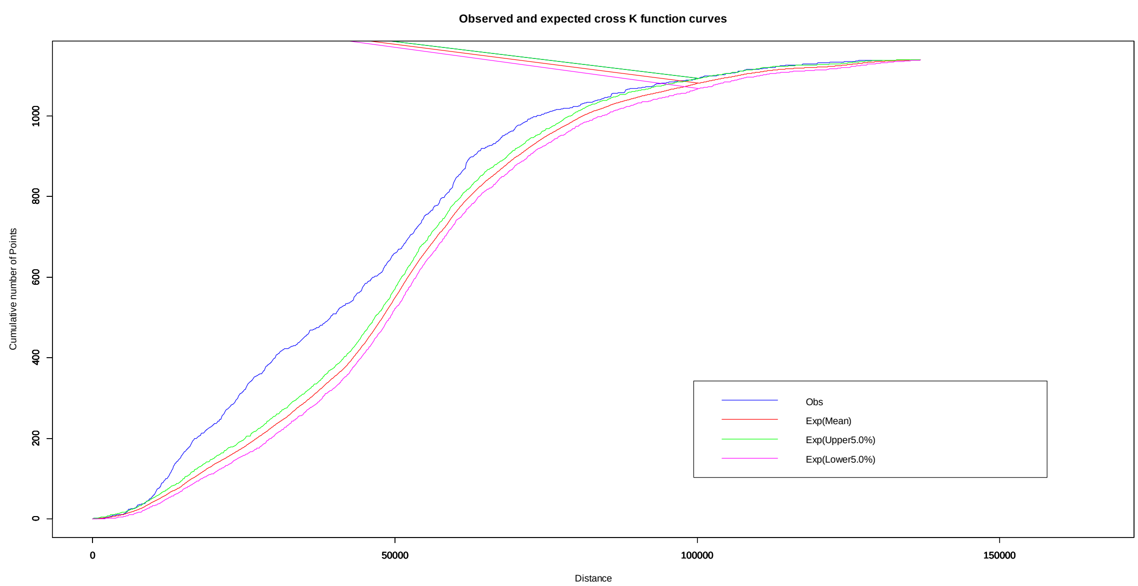
ID of road segments could be connected to every crash-point event to calculate the crash rate at a smaller geographic scale. Due to the limitations of TDOT data, in this study the crash rate cannot be calculated. In addition, the network KDE instead of planar KDE could be applied to all the datasets if the new version of SANET has the ability to reduce the duplicate nodes and segments for the road layers. With advances in computer software and hardware, the K function method could also be applied to all roads in the study area instead of primary roads only. This study focused on the relationship of the spatial distributions among different type of fatal crashes, and it may be a good idea to examine the spatial relationships between fatal crashes and different type of facilities, such as liquor stores, schools, and hospitals. Moreover, this study mainly focused on time factors. In future research, other types of classifications could be studied, such as, drivers' demographic characteristics (age and gender) and weather conditions. A similar study could also be conducted in other places in the U.S. and around the world (if similar data are available). Lastly, the approach used in this study could be applied to other types of spatial point events, such as non-fatal traffic accidents.

APPENDIX A: DAYTIME AND NIGHTTIME FATAL CRASHES BY LOCAL CROSS K FUNCTION (ORDERED WITH HOTSPOT ID)

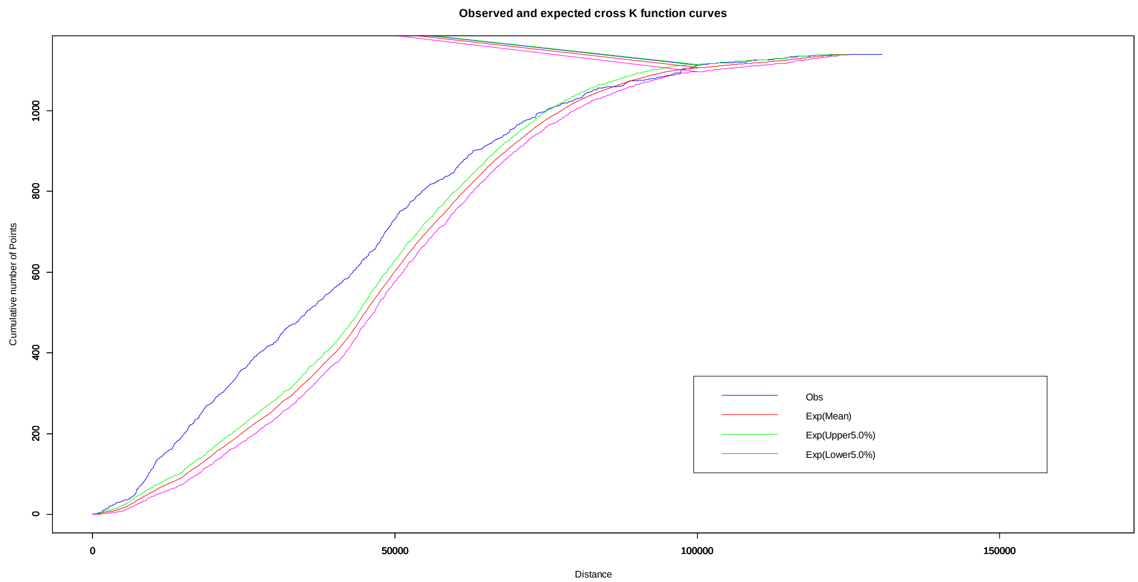
Nighttime Hotspot 0



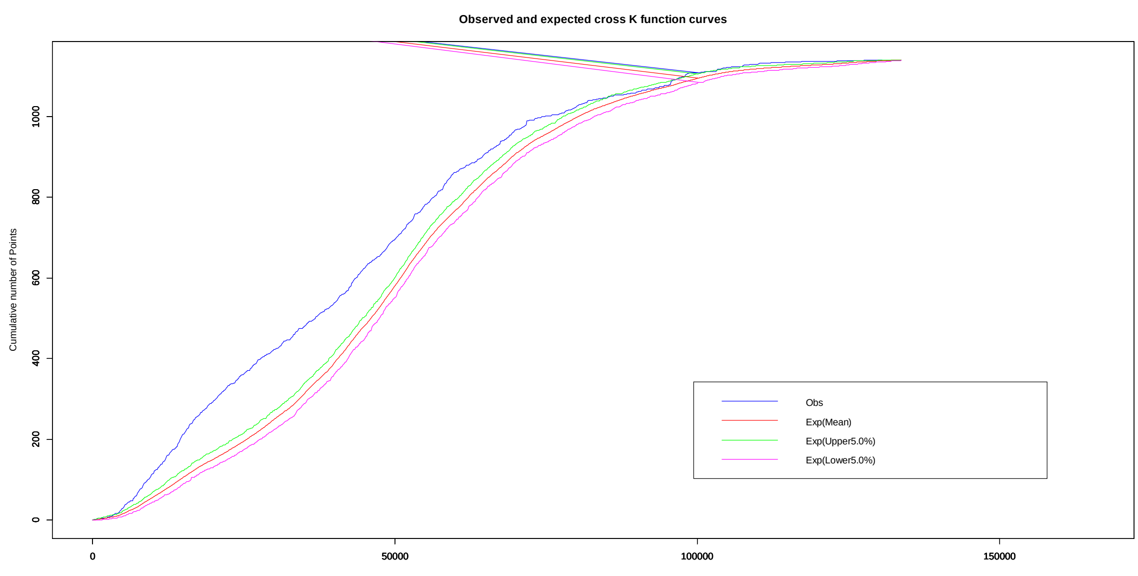
Nighttime Hotspot 1



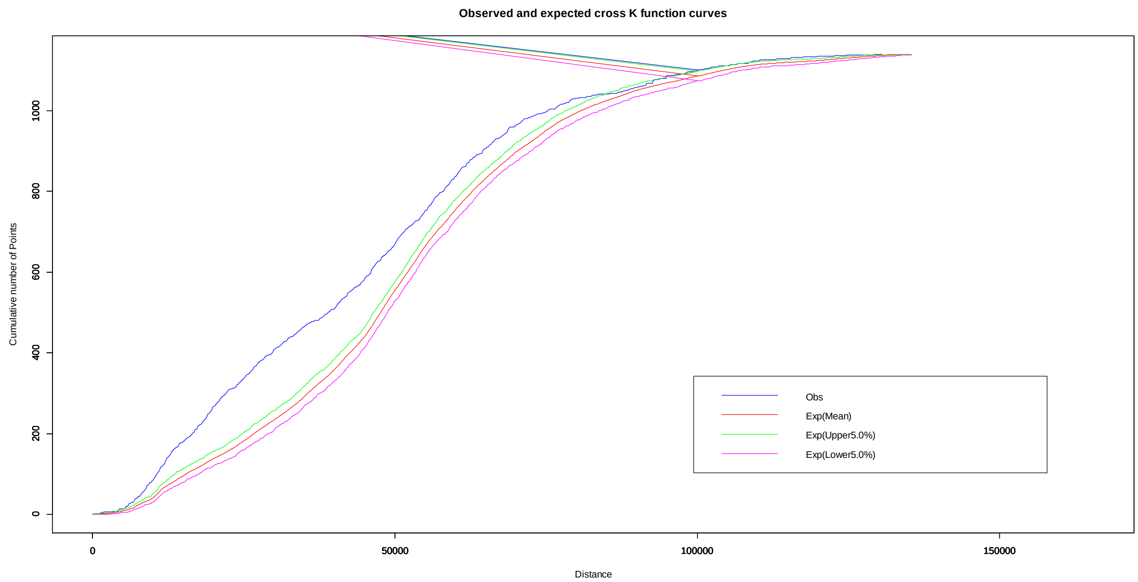
Nighttime Hotspot 2



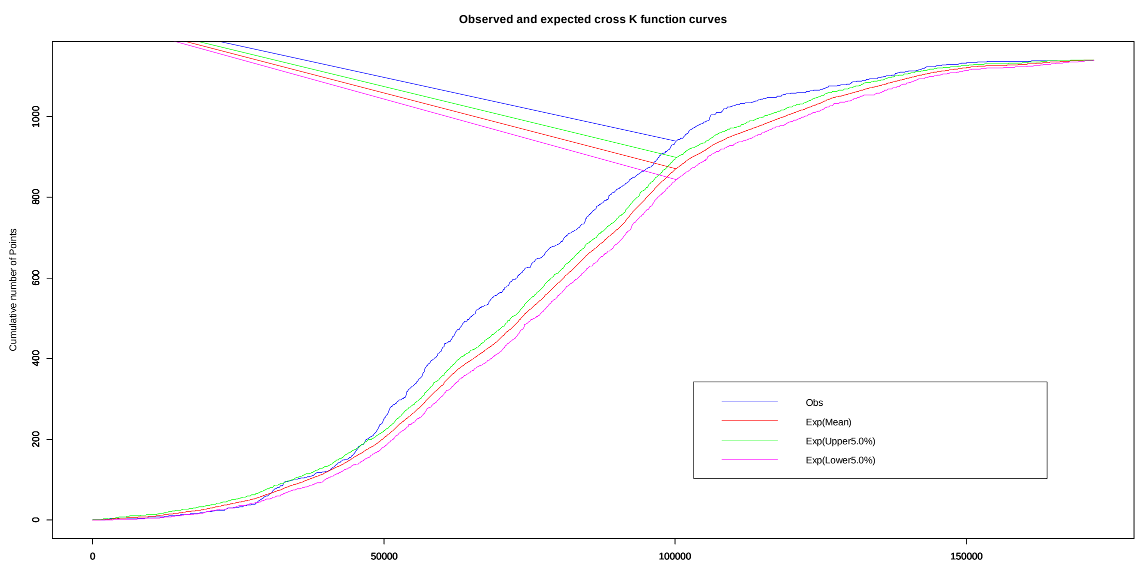
Nighttime Hotspot 3



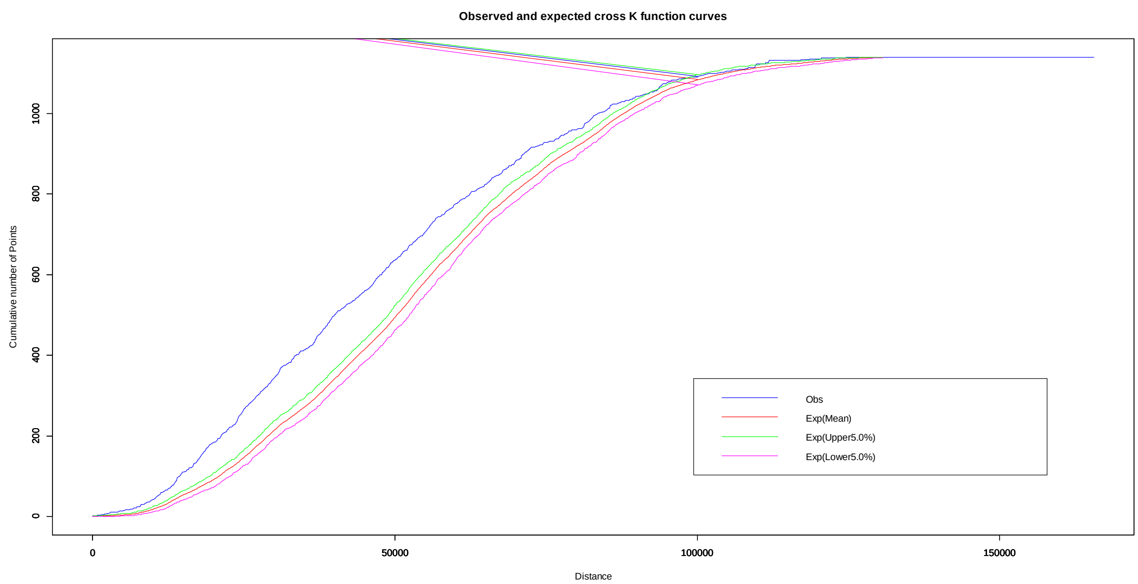
Nighttime Hotspot 4



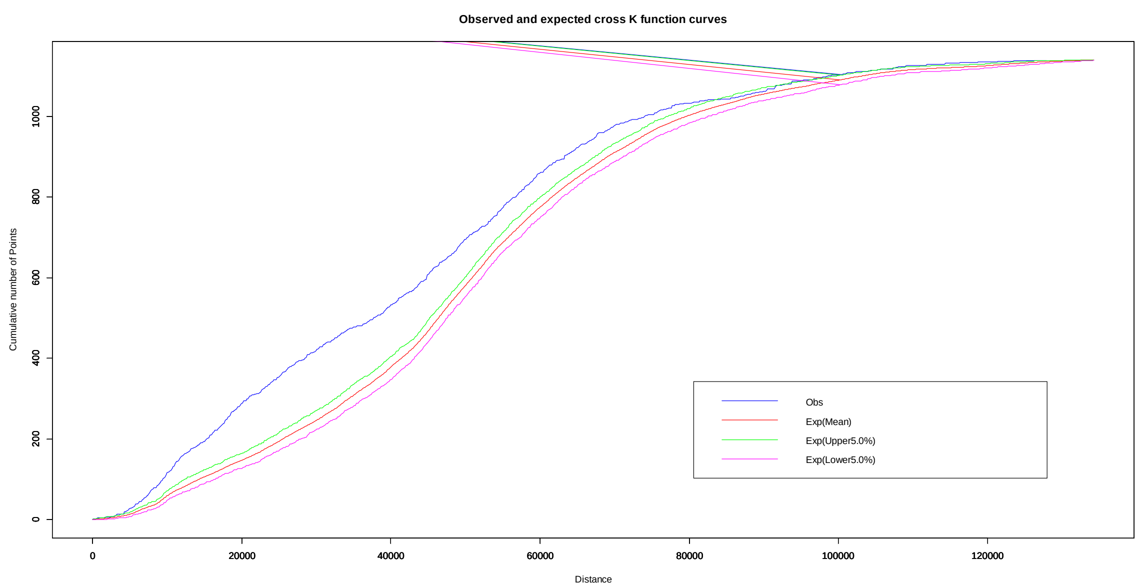
Nighttime Hotspot 5



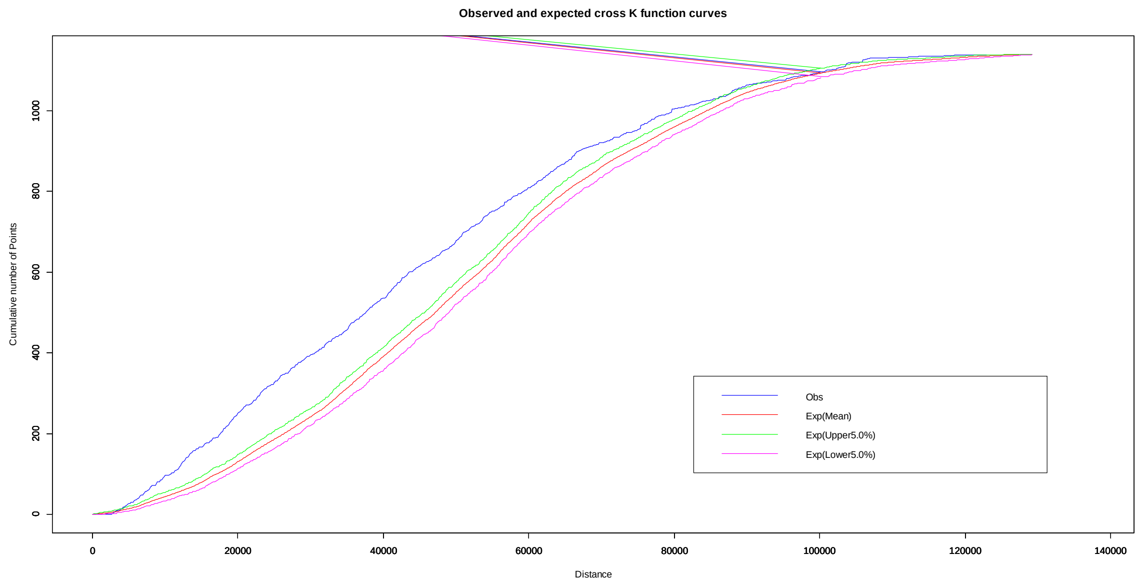
Nighttime Hotspot 6



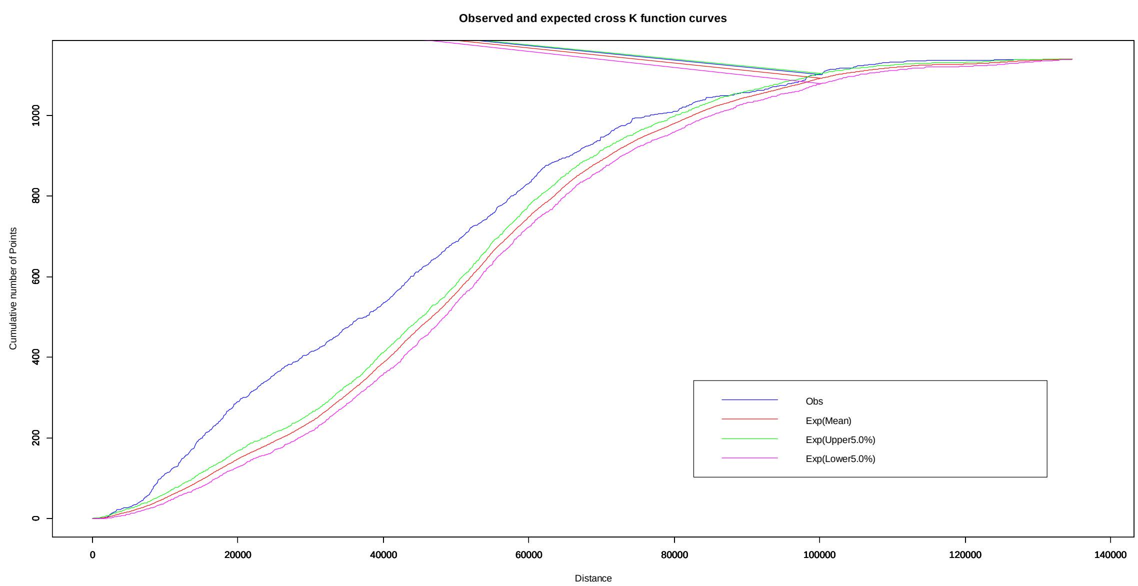
Nighttime Hotspot 7



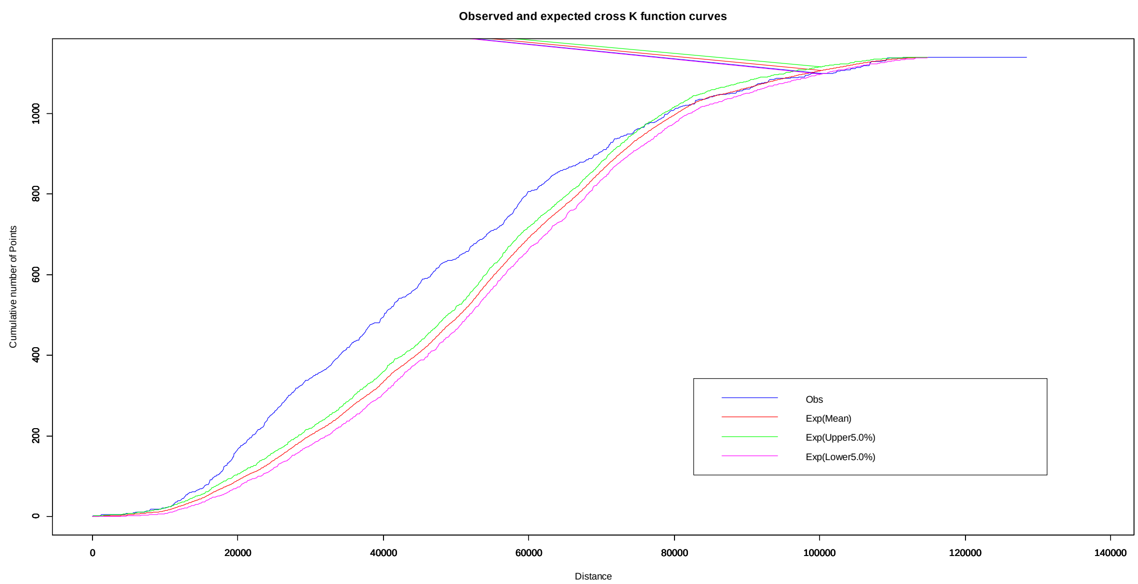
Nighttime Hotspot 8



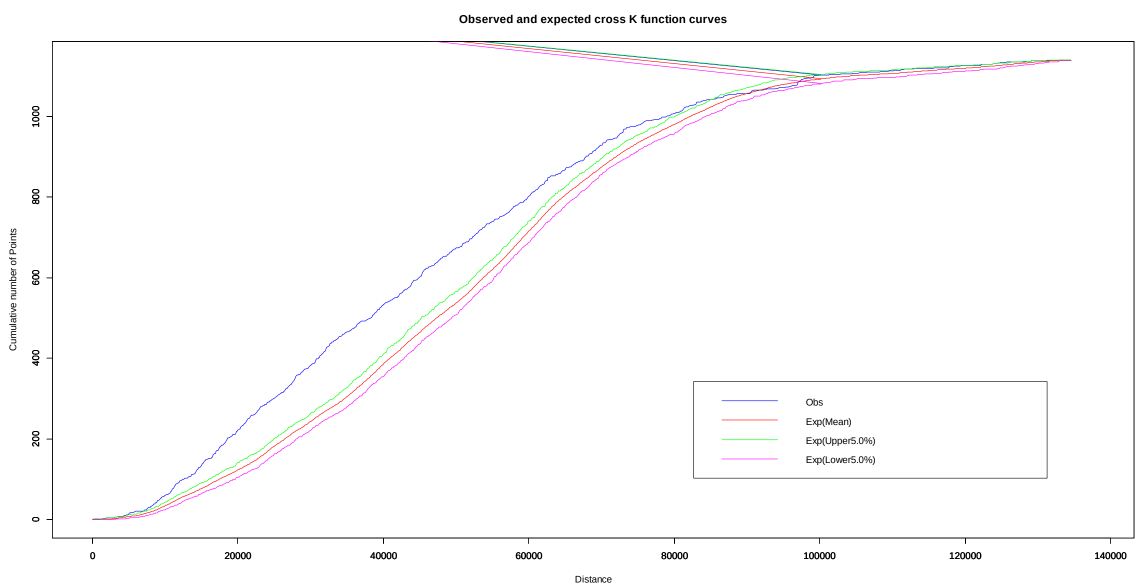
Nighttime Hotspot 9



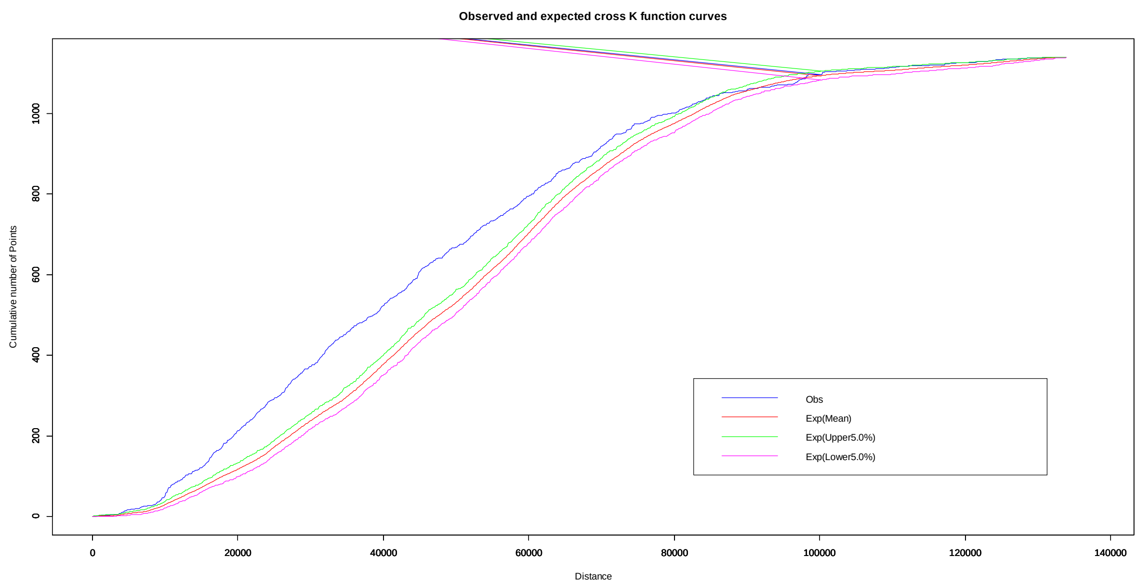
Nighttime Hotspot 10



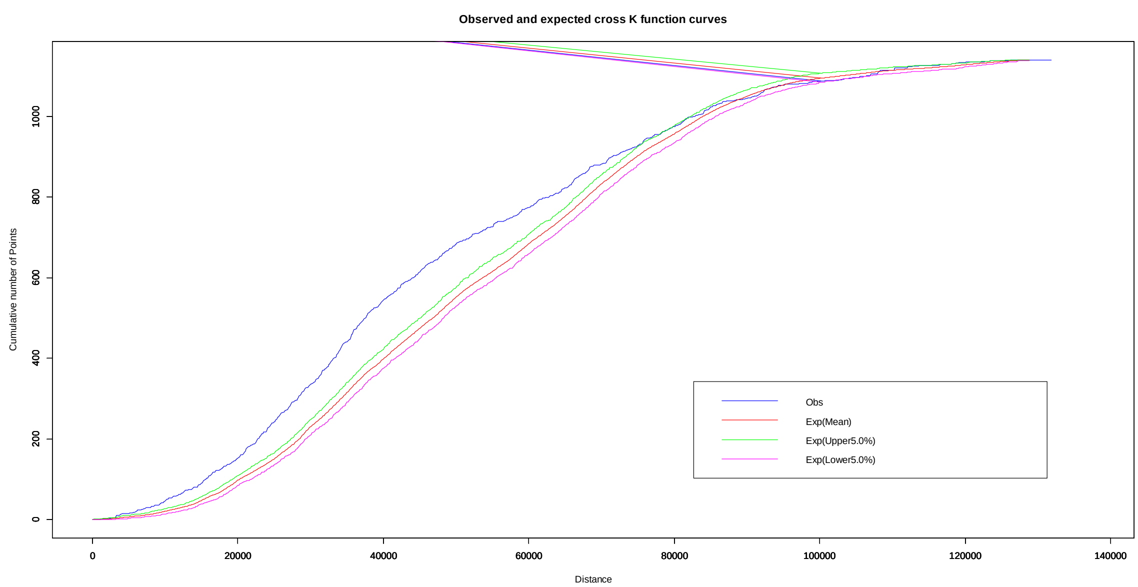
Nighttime Hotspot 11



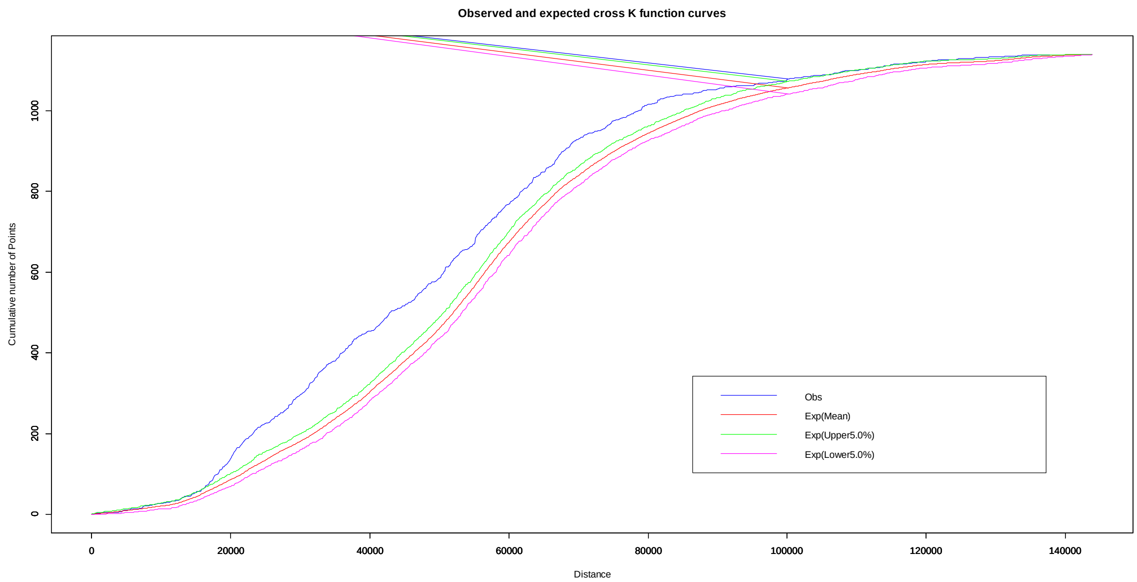
Nighttime Hotspot 12



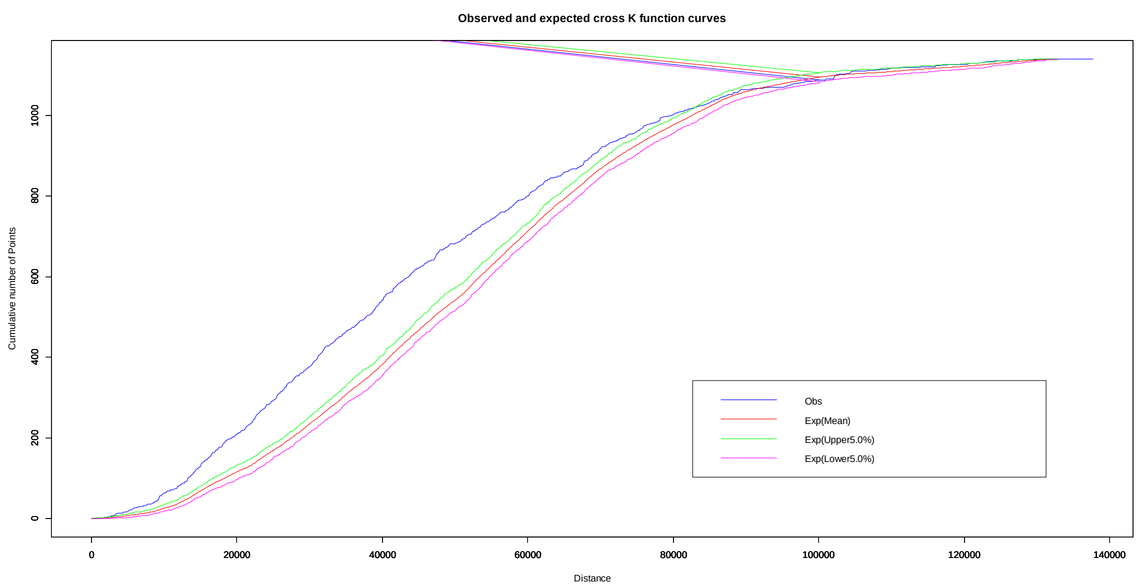
Nighttime Hotspot 13



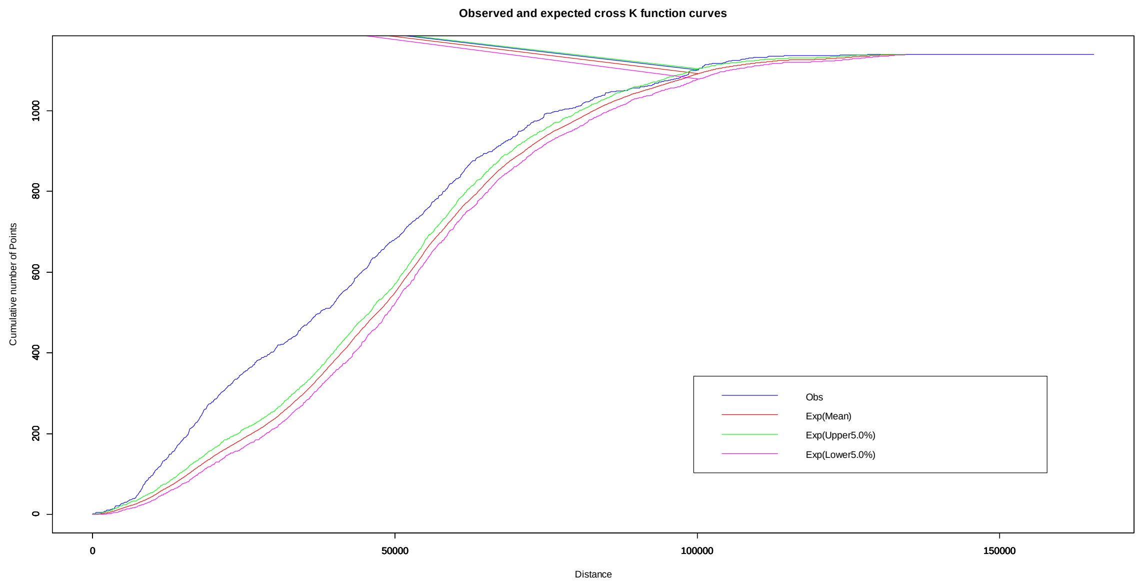
Nighttime Hotspot 14



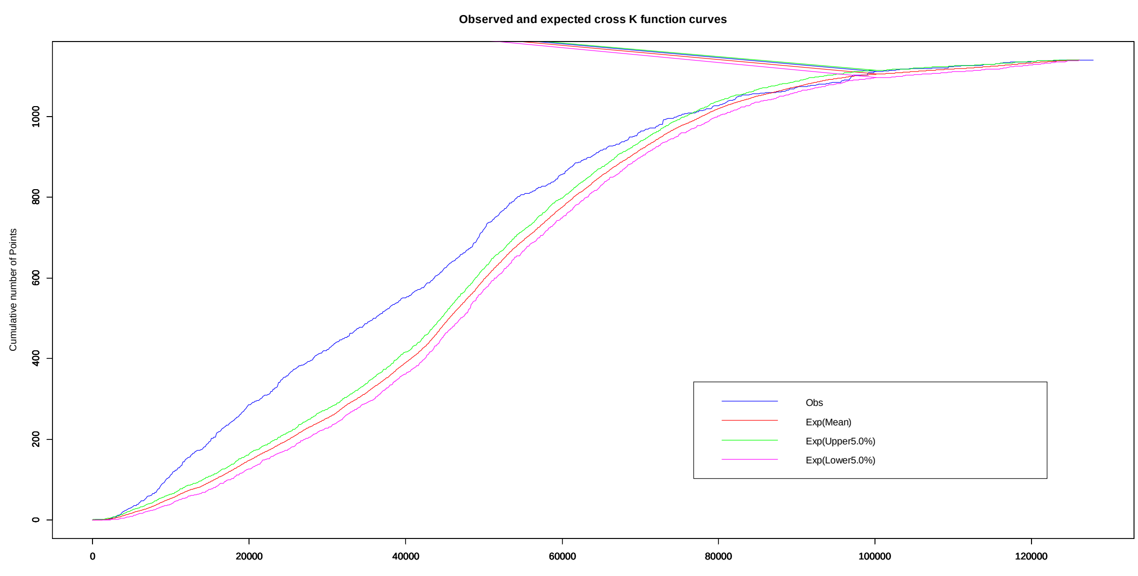
Nighttime Hotspot 15



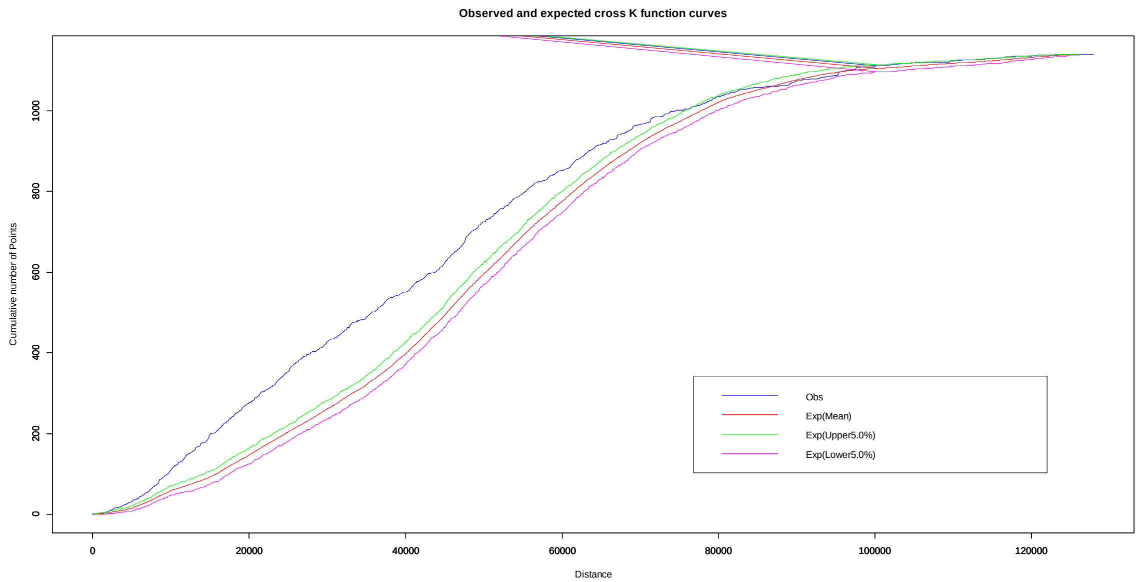
Nighttime Hotspot 16



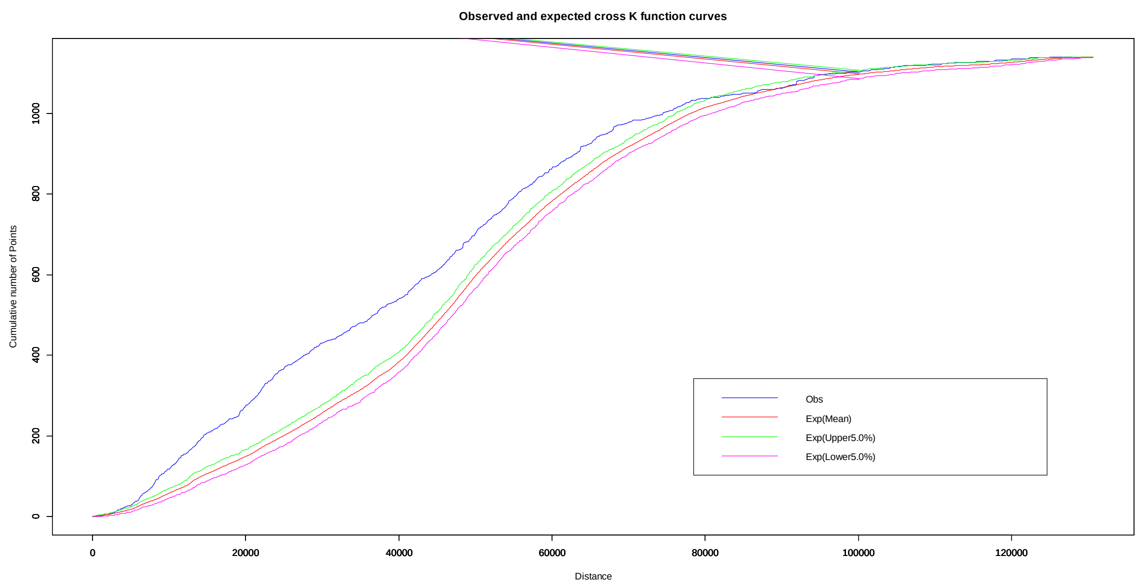
Nighttime Hotspot 17



Nighttime Hotspot 18

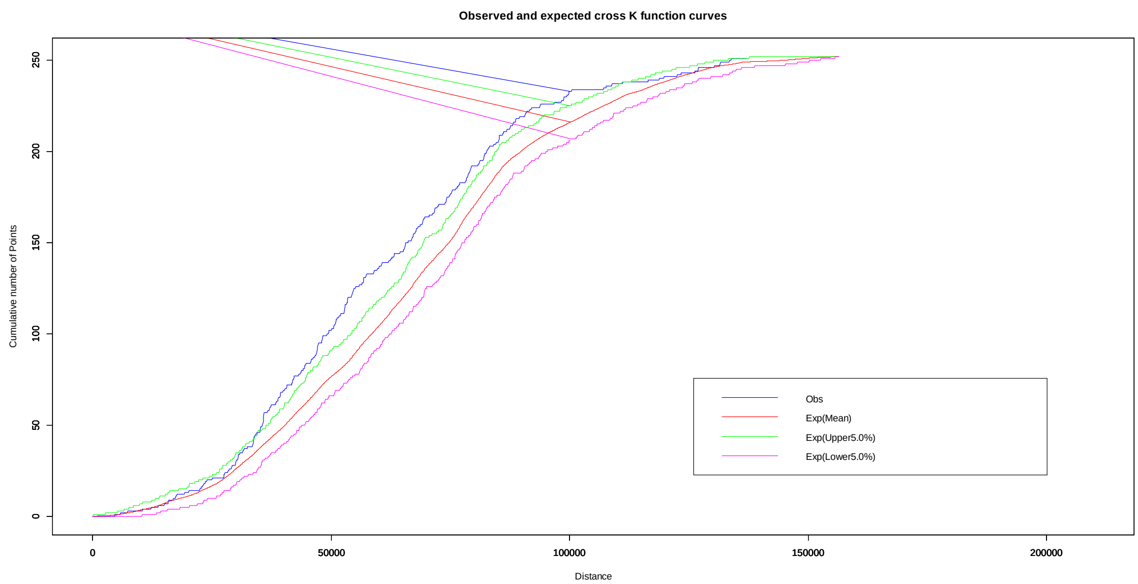


Nighttime Hotspot 19

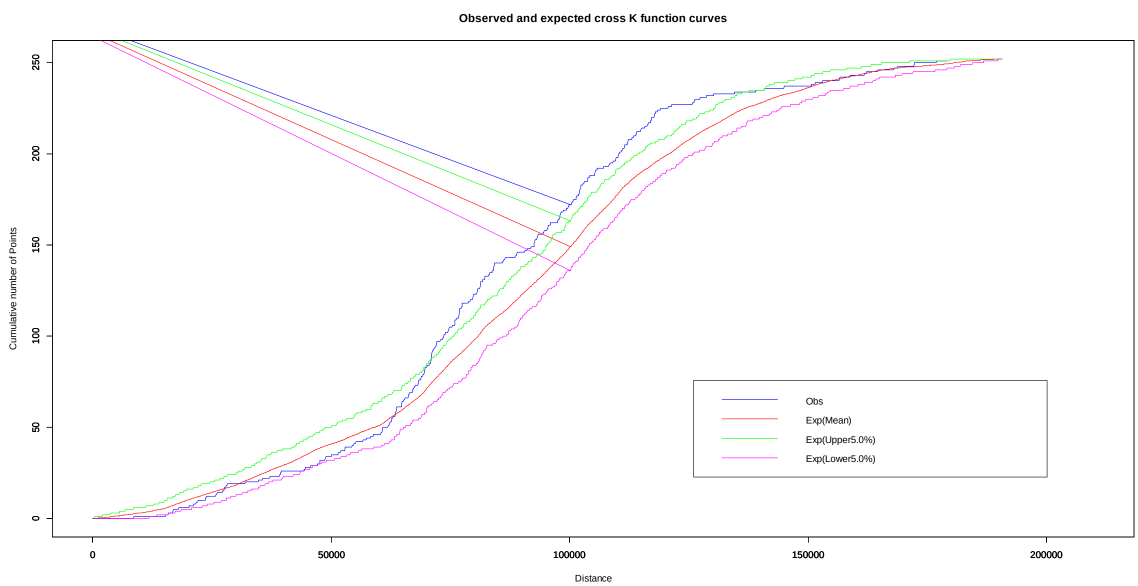


APPENDIX B: A.M. RUSH HOURS AND P.M. RUSH HOURS FATAL CRASHES BY LOCAL CROSS K FUNCTION (ORDERED WITH HOTSPOT ID)

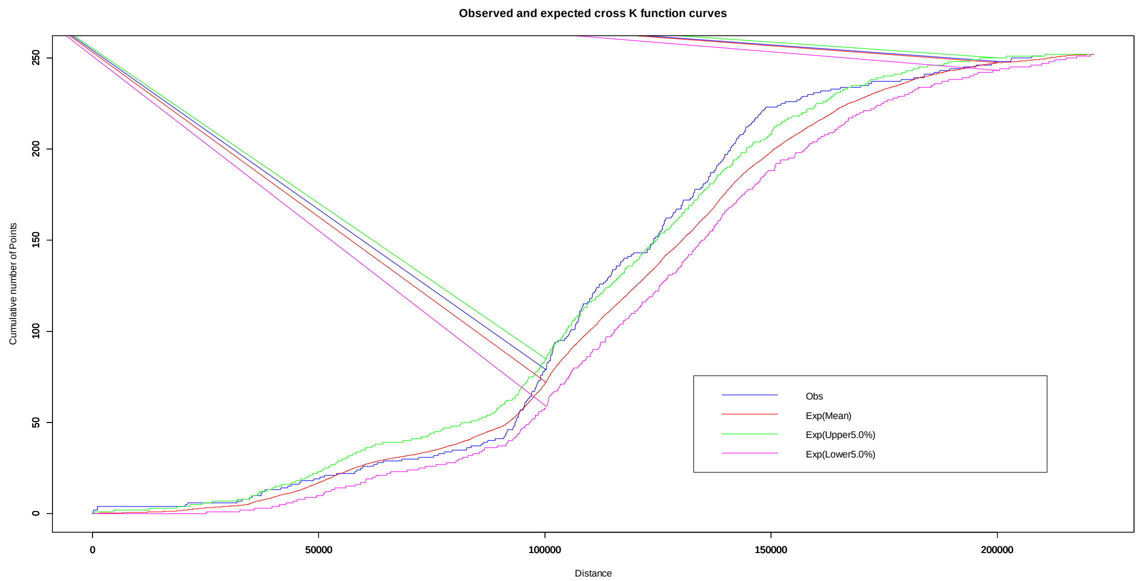
P.M. Rush Hour Hotspot 0



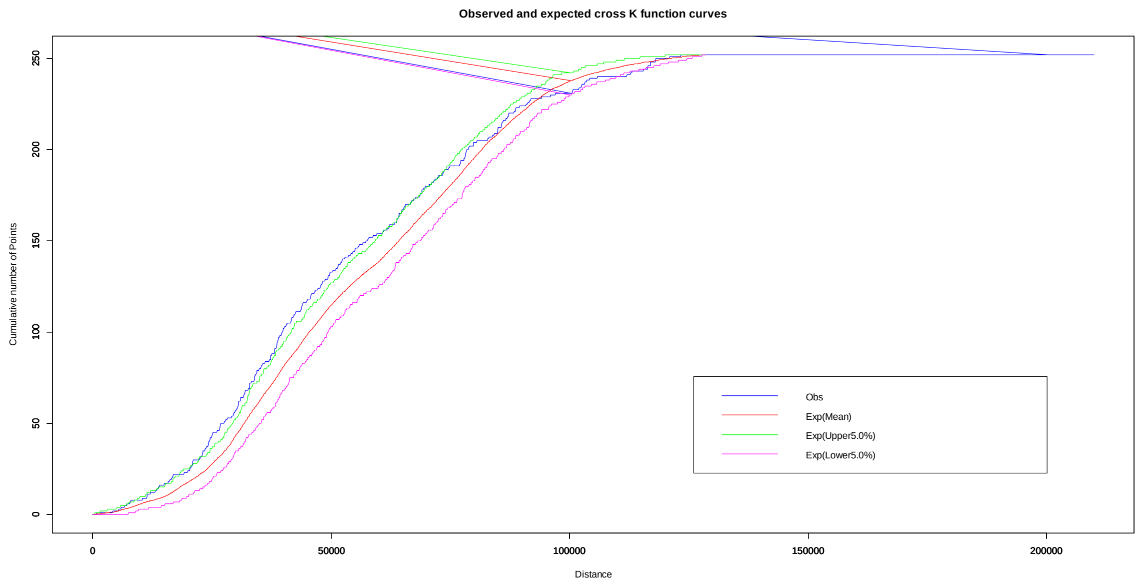
P.M. Rush Hour Hotspot 1



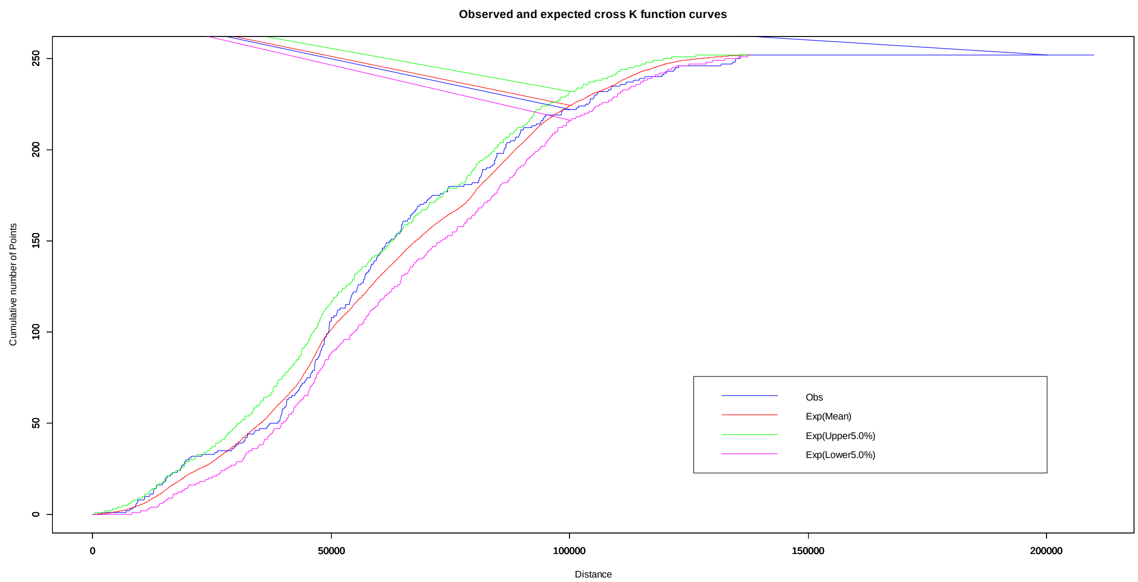
P.M. Rush Hour Hotspot 2



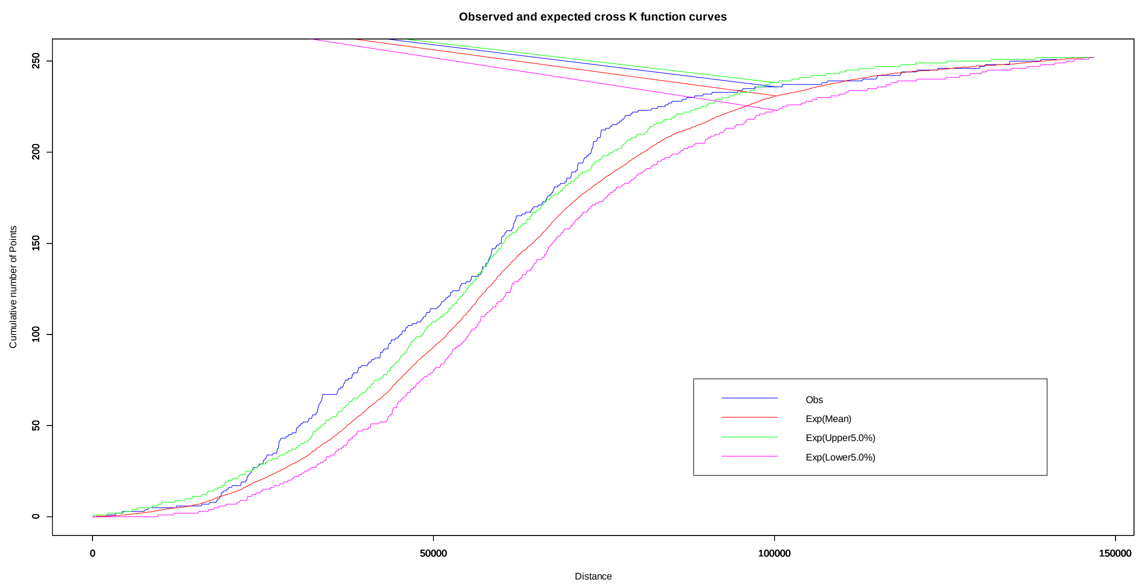
P.M. Rush Hour Hotspot 3



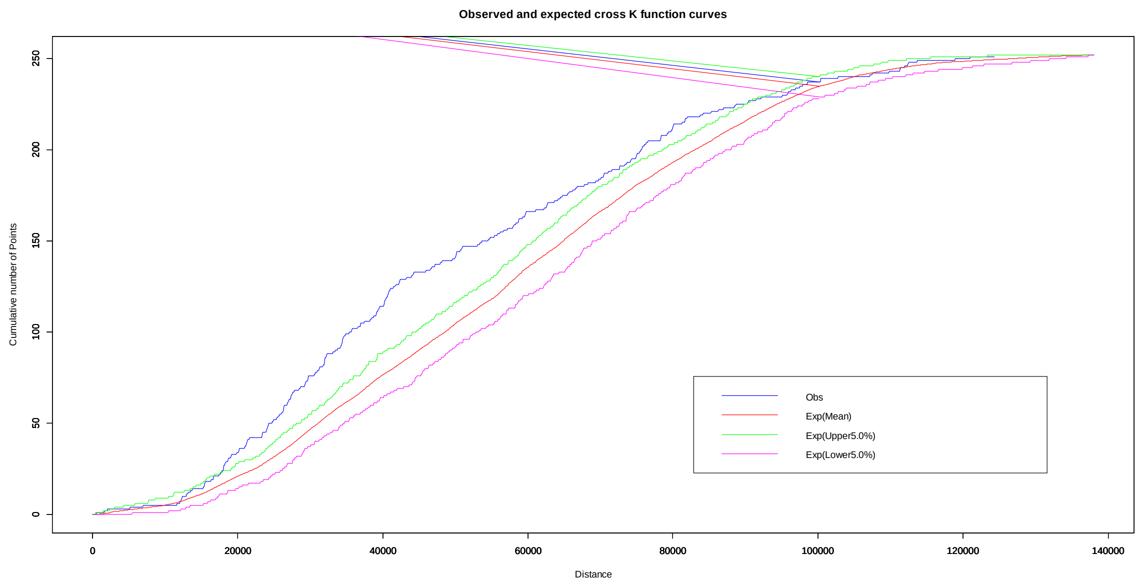
P.M. Rush Hour Hotspot 4



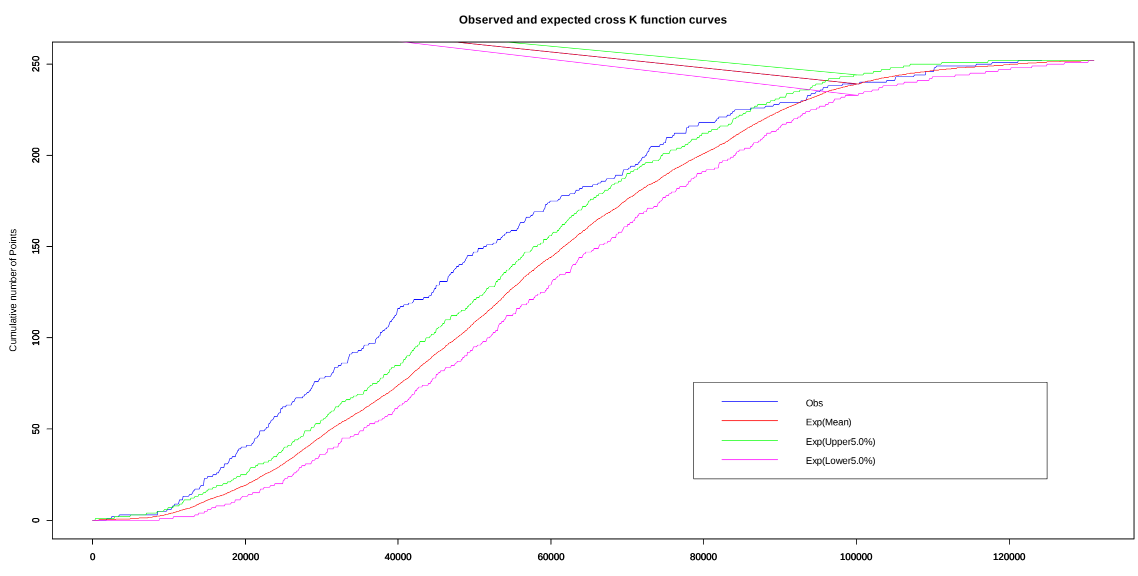
P.M. Rush Hour Hotspot 5



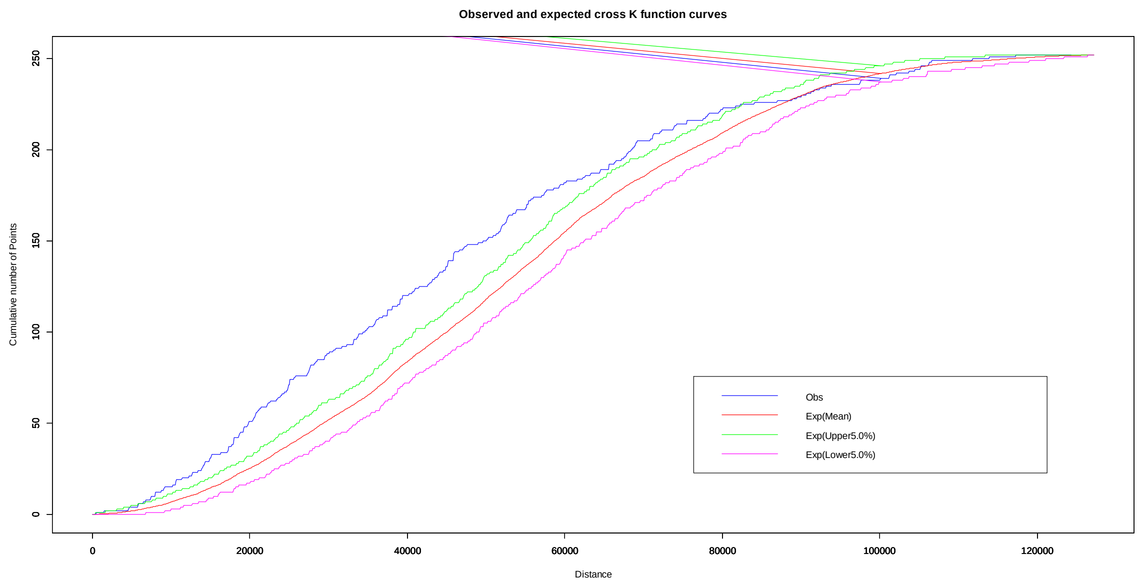
P.M. Rush Hour Hotspot 6



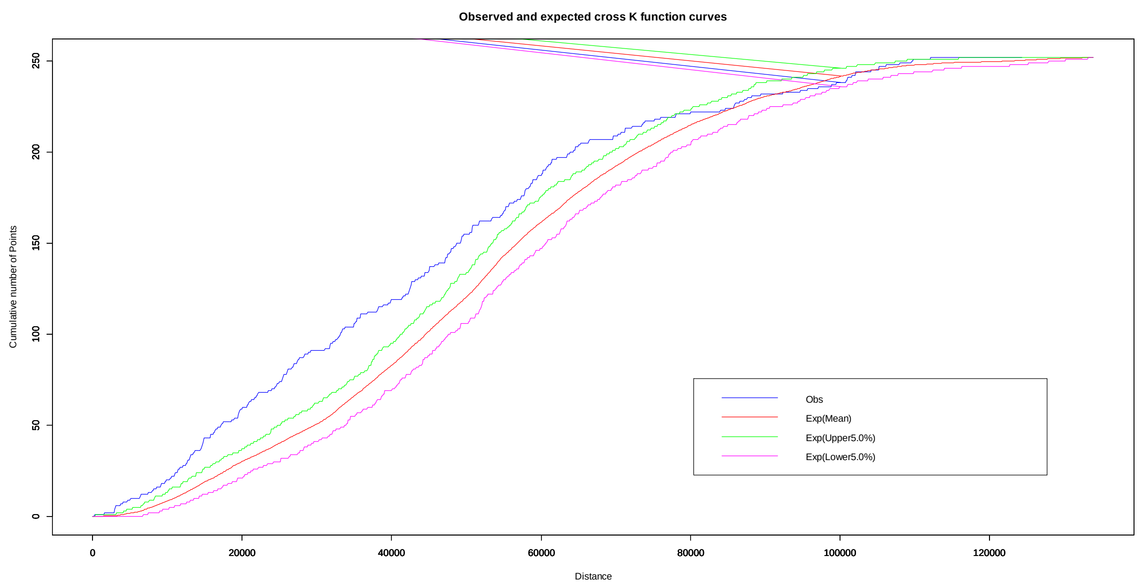
P.M. Rush Hour Hotspot 7



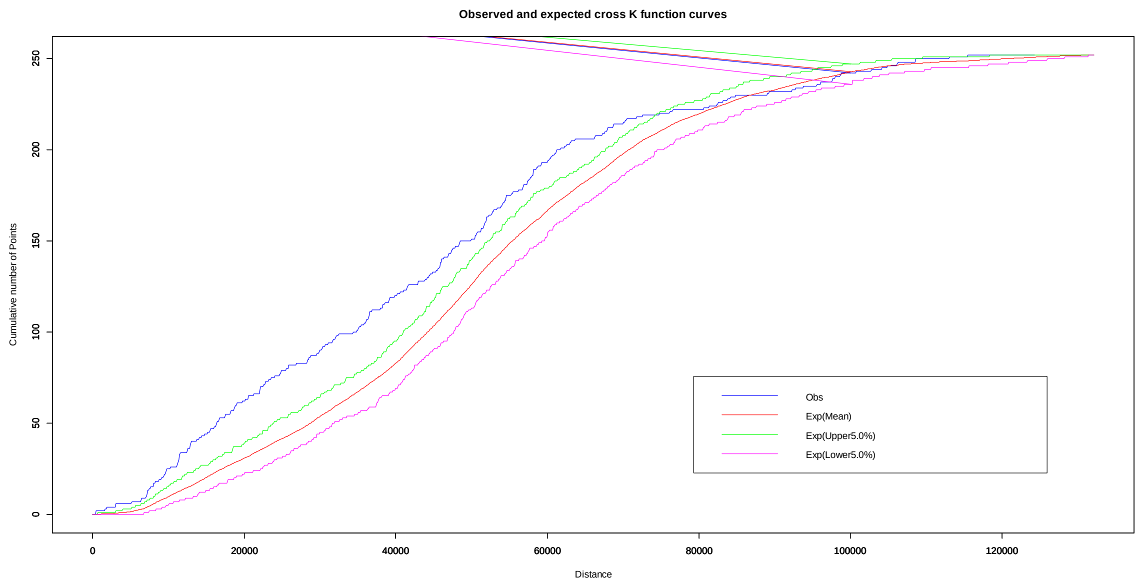
P.M. Rush Hour Hotspot 8



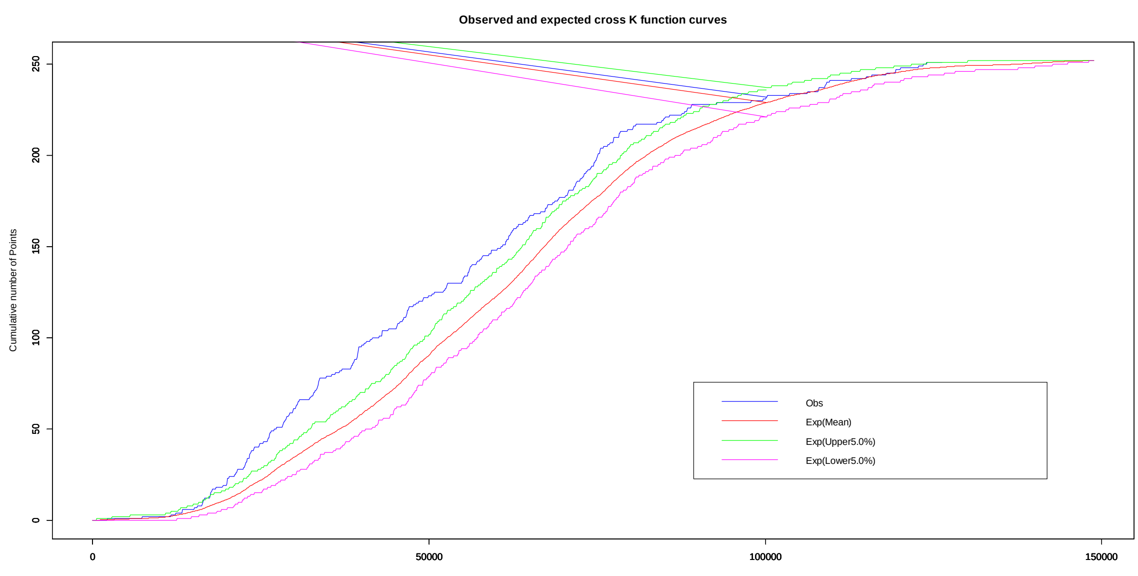
P.M. Rush Hour Hotspot 9



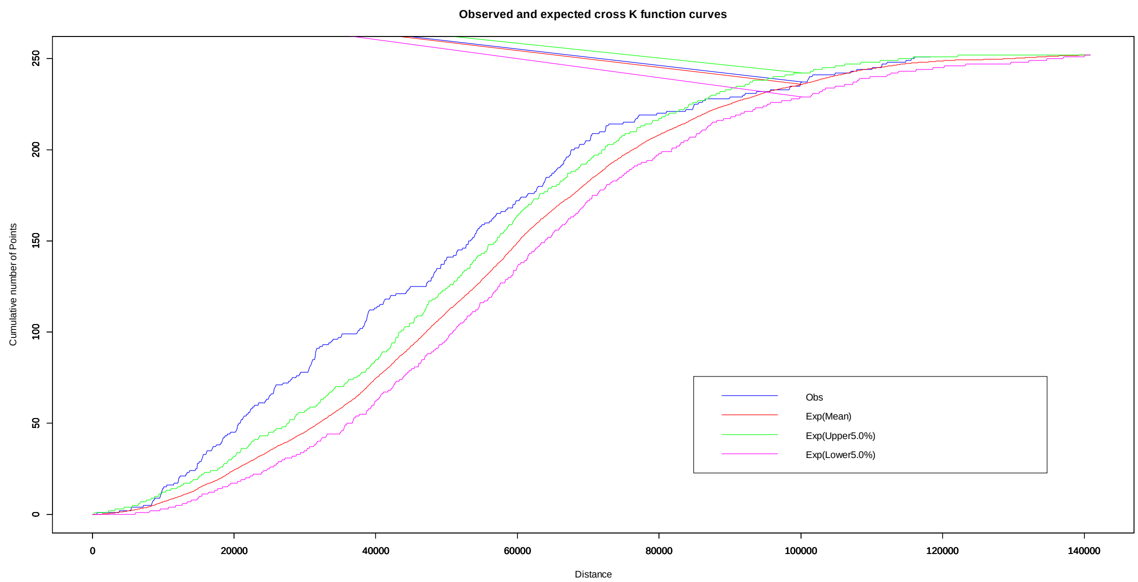
P.M. Rush Hour Hotspot 10



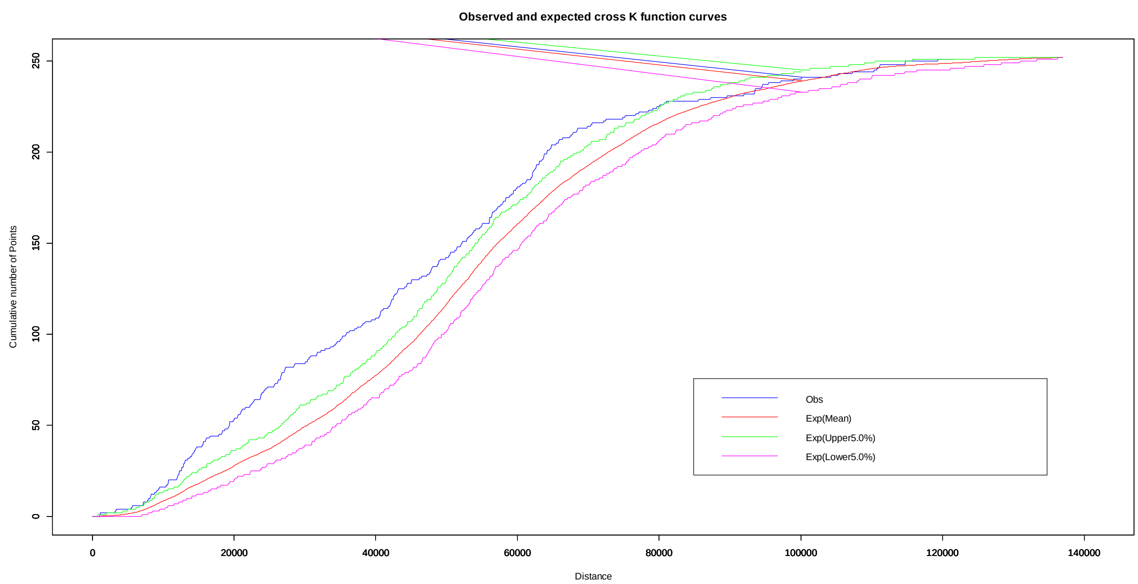
P.M. Rush Hour Hotspot 11



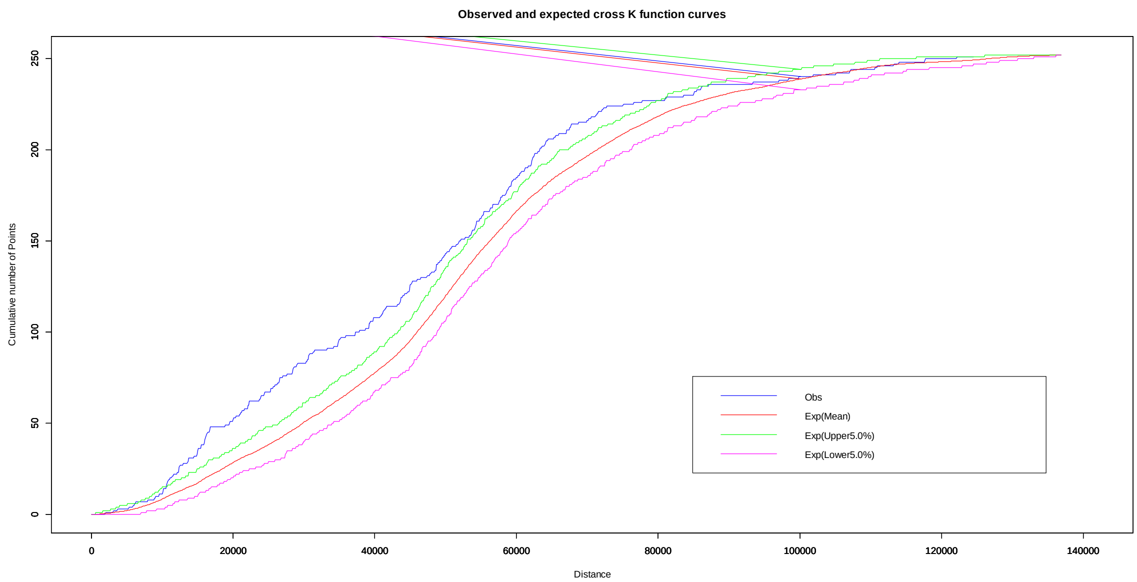
P.M. Rush Hour Hotspot 12



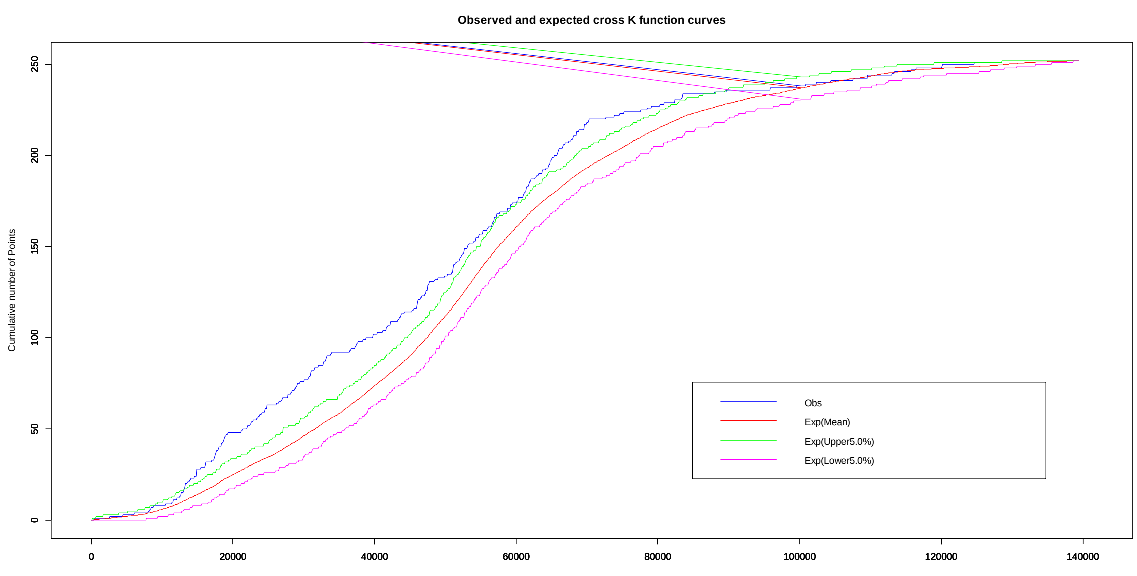
P.M. Rush Hour Hotspot 13



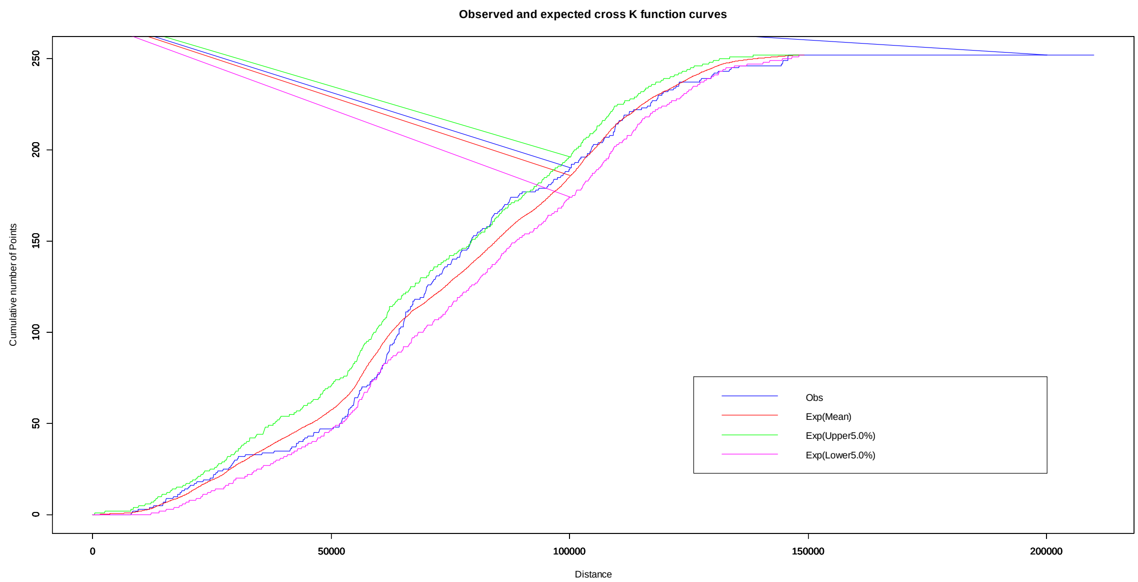
P.M. Rush Hour Hotspot 14



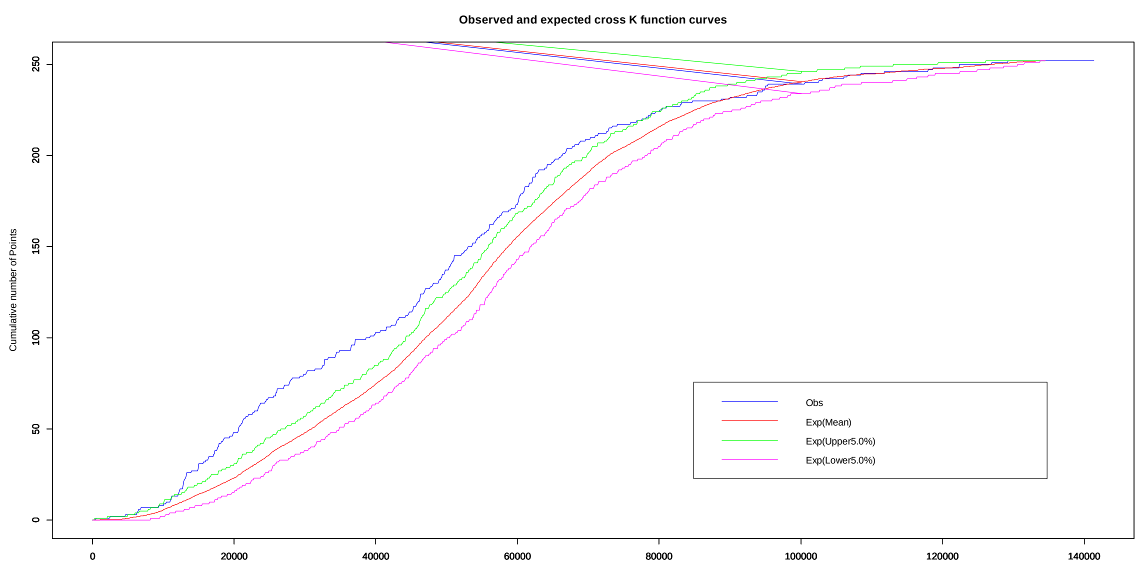
P.M. Rush Hour Hotspot 15



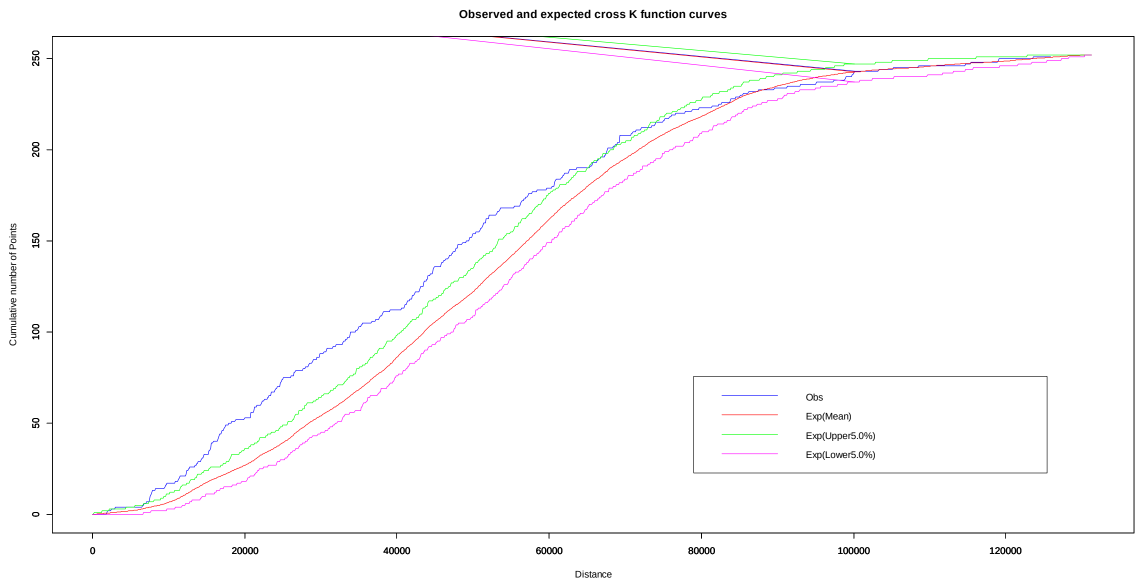
P.M. Rush Hour Hotspot 16



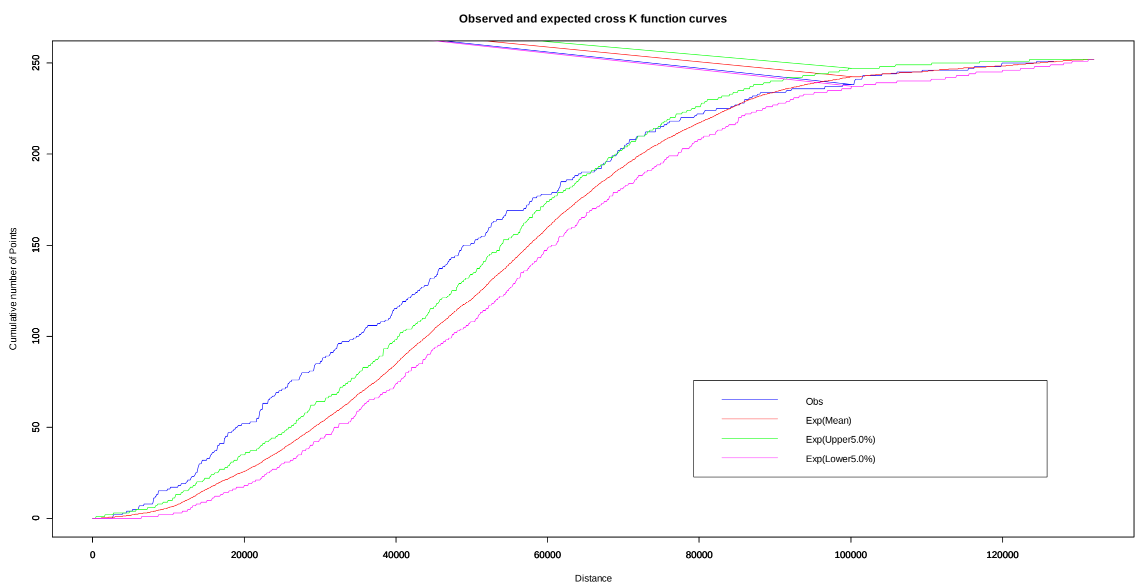
P.M. Rush Hour Hotspot 17



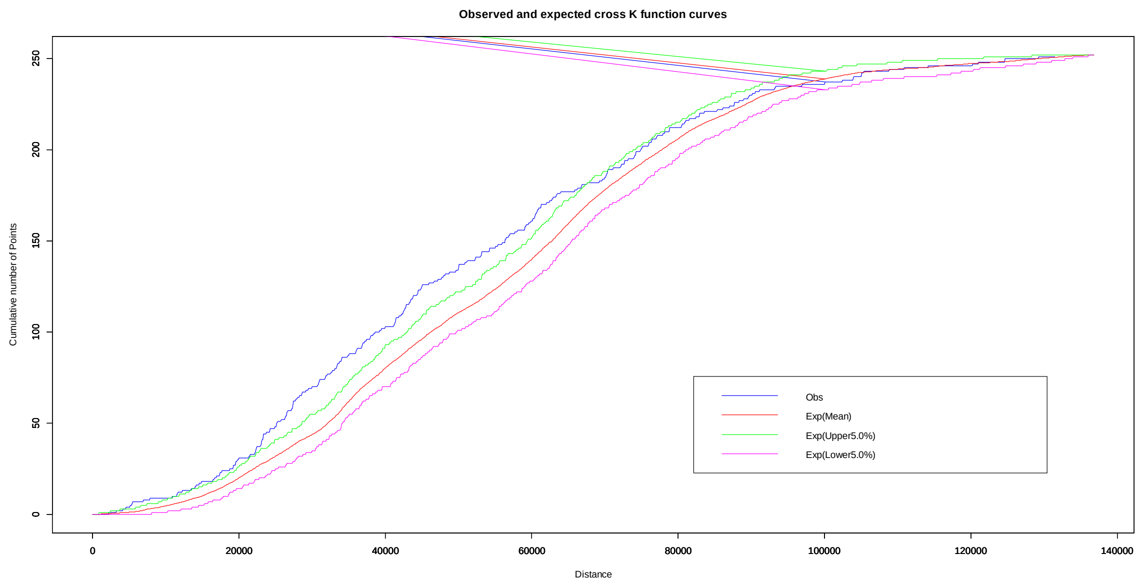
P.M. Rush Hour Hotspot 18



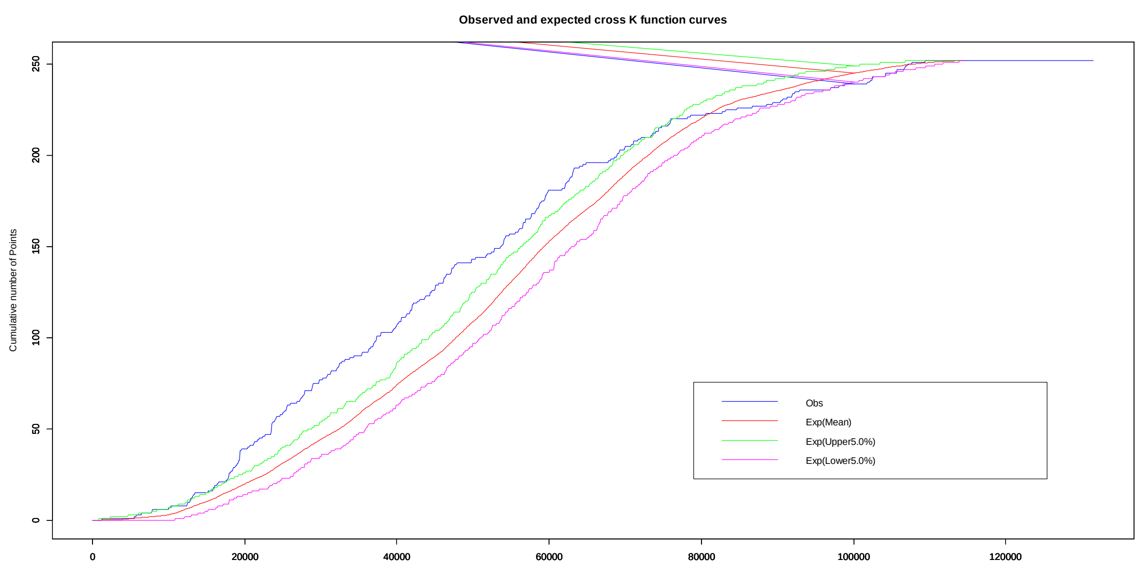
P.M. Rush Hour Hotspot 19



P.M. Rush Hour Hotspot 20

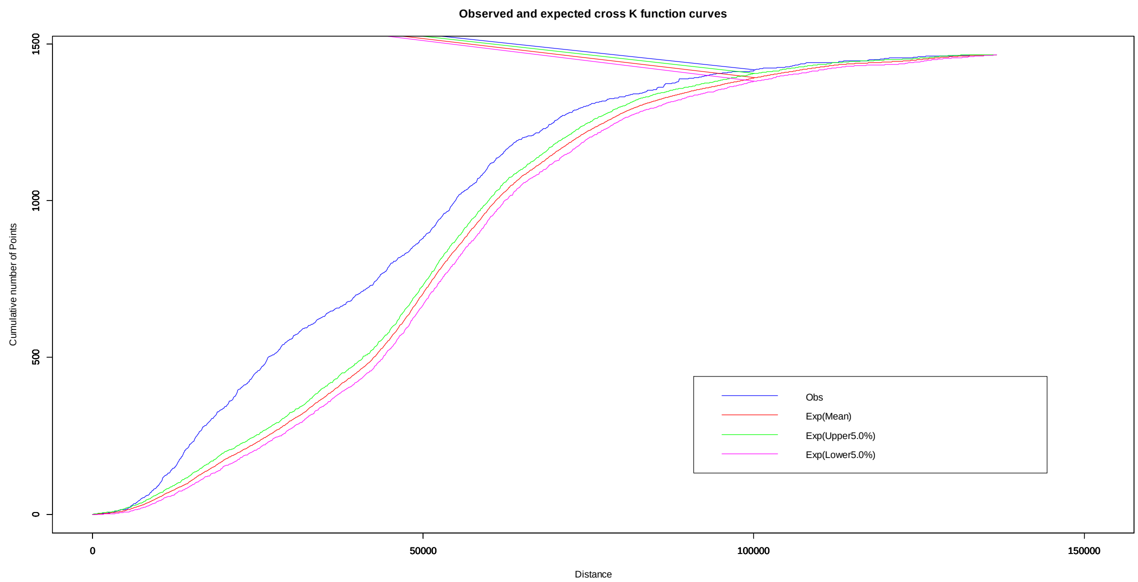


P.M. Rush Hour Hotspot 21

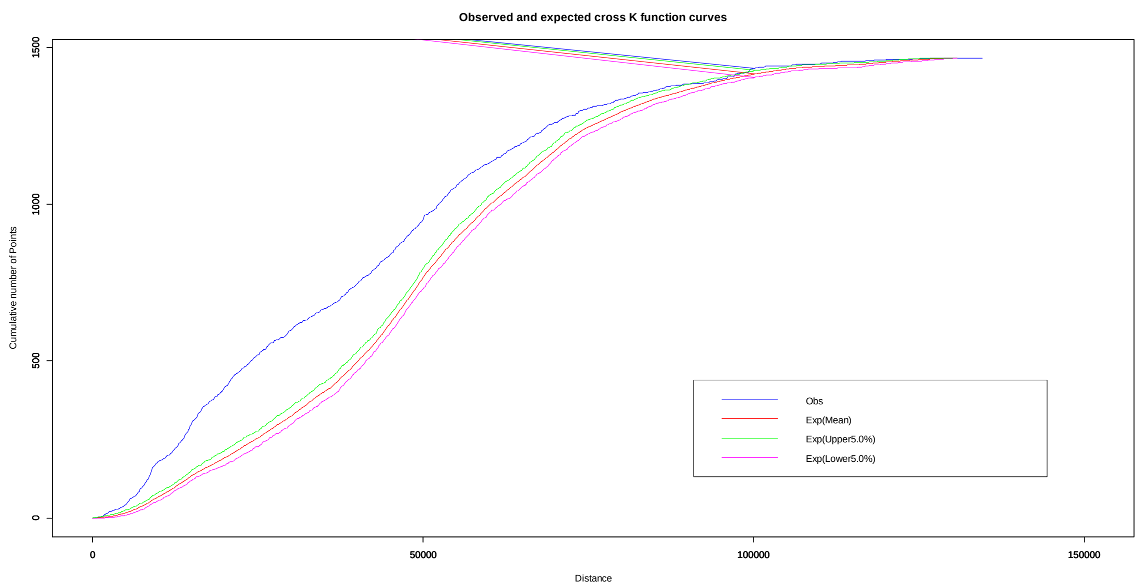


APPENDIX C: WEEKDAY AND WEEKEND FATAL CRASHES BY LOCAL CROSS K FUNCTION (ORDERED WITH HOTSPOT ID)

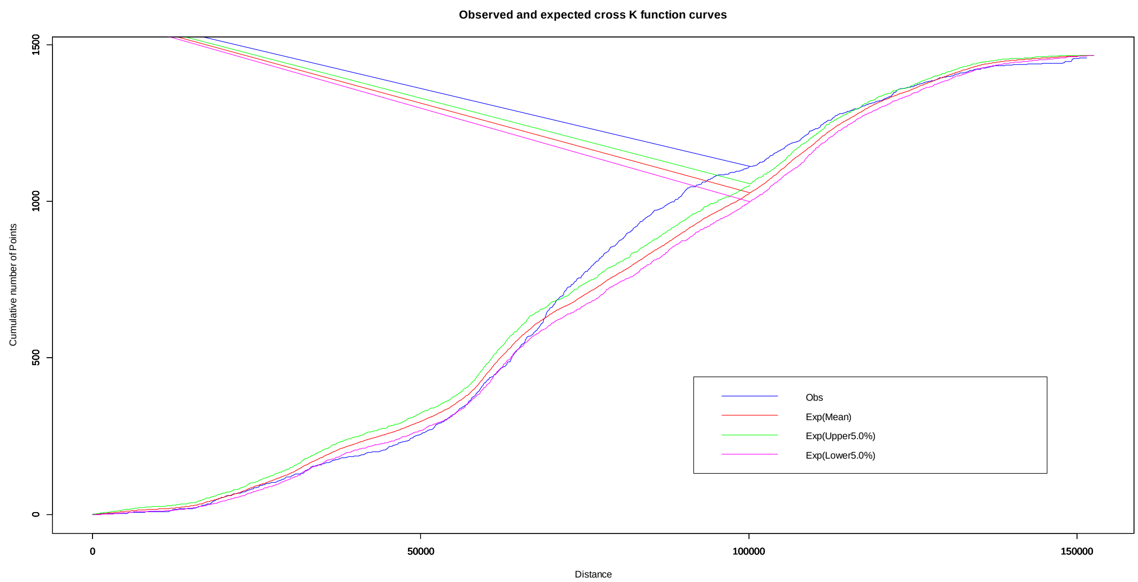
Weekend Hotspot 0



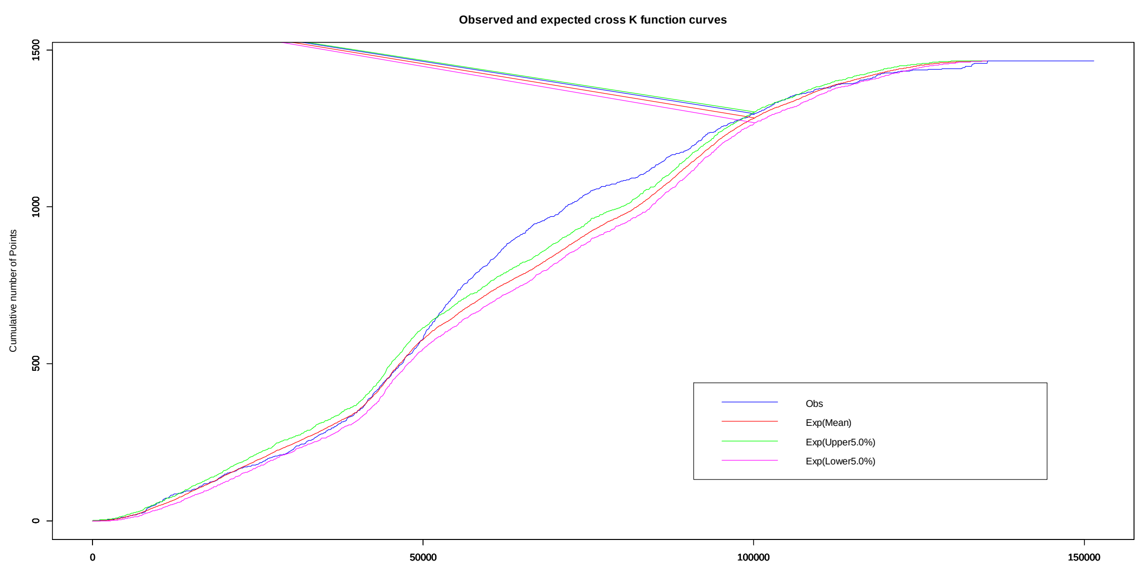
Weekend Hotspot 1



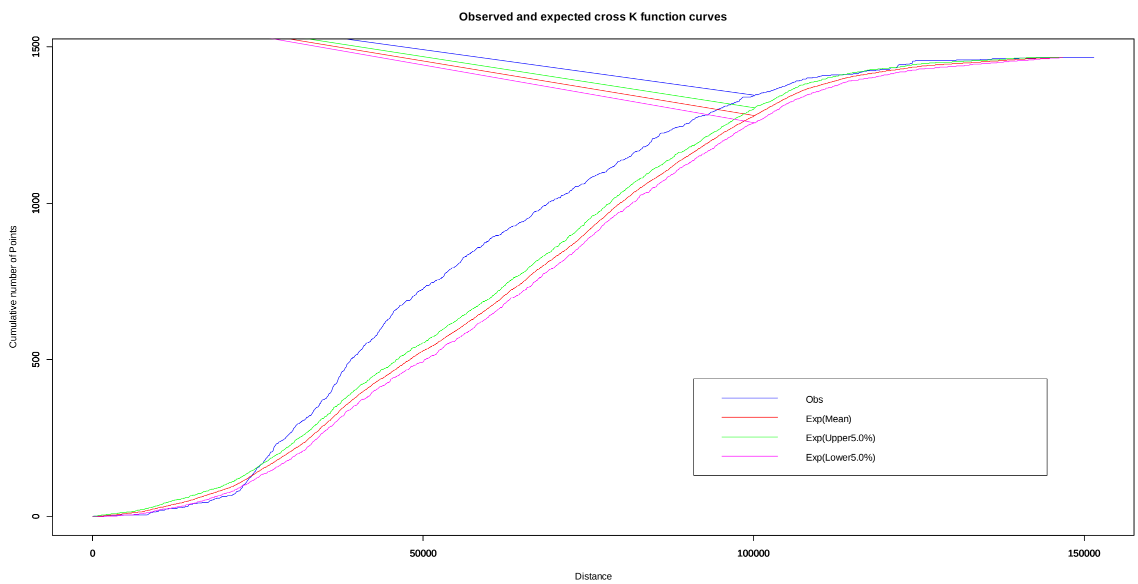
Weekend Hotspot 2



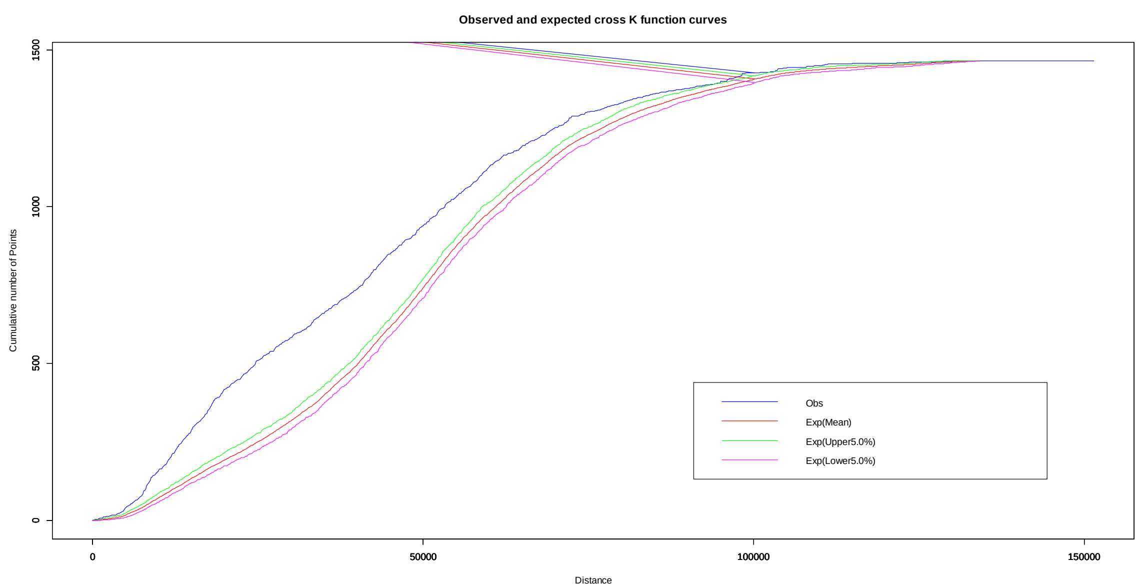
Weekend Hotspot 3



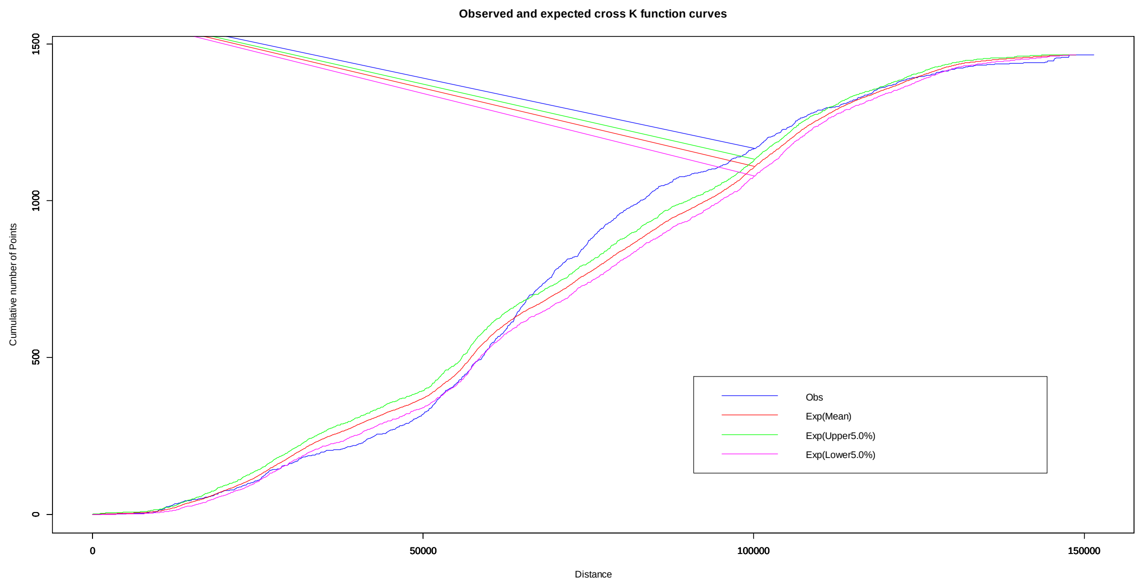
Weekend Hotspot 4



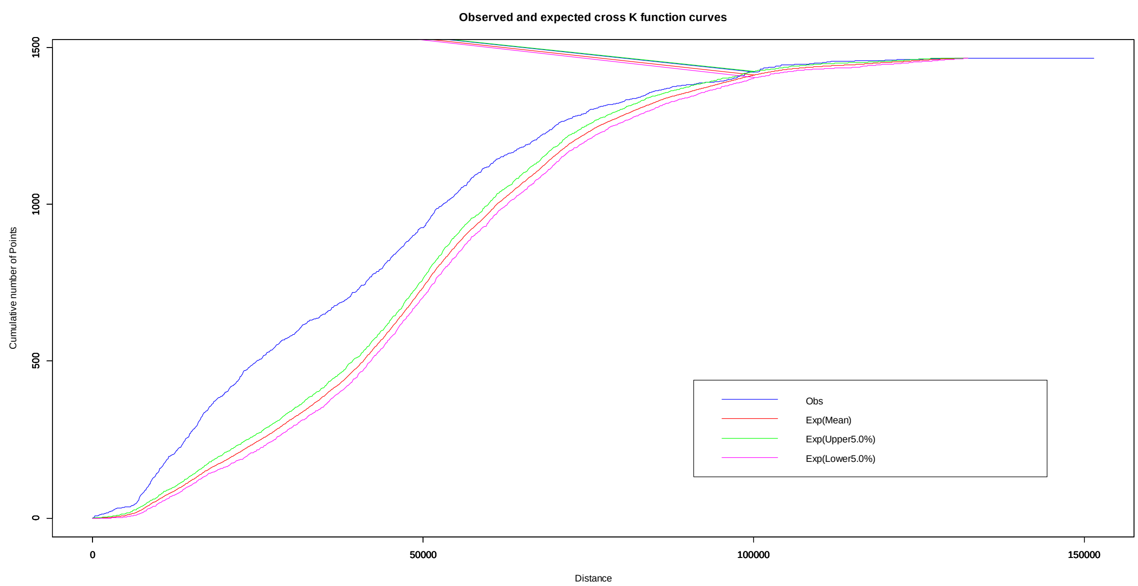
Weekend Hotspot 5



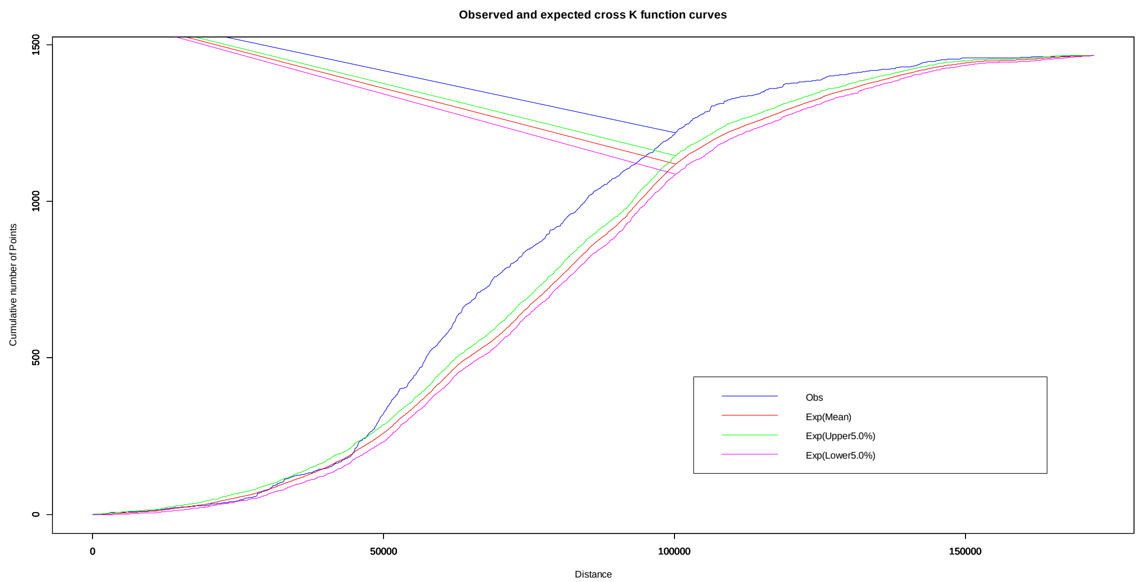
Weekend Hotspot 6



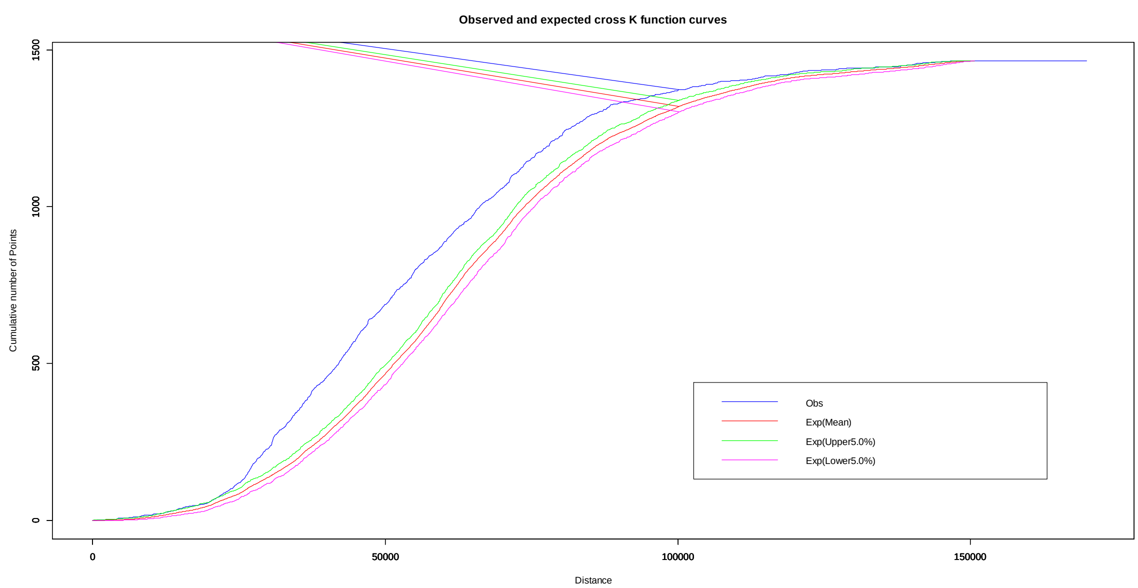
Weekend Hotspot 7



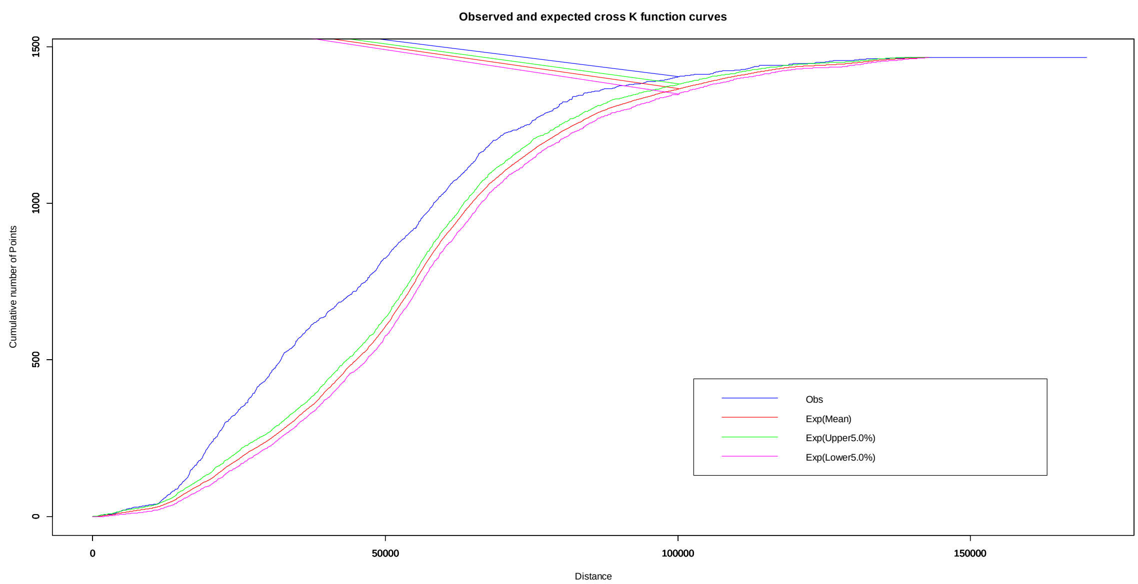
Weekend Hotspot 8



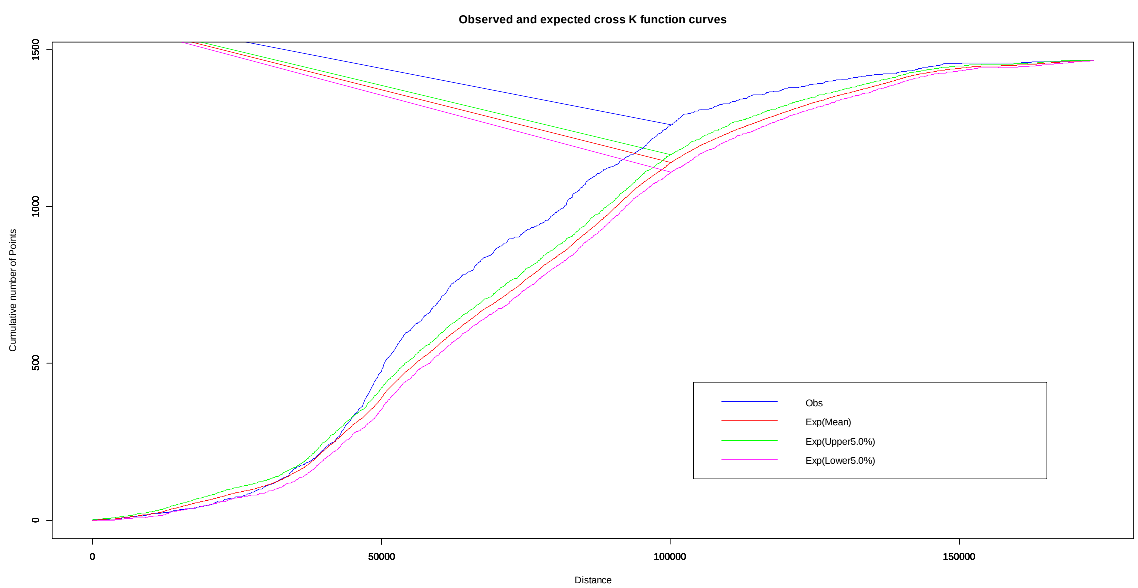
Weekend Hotspot 9



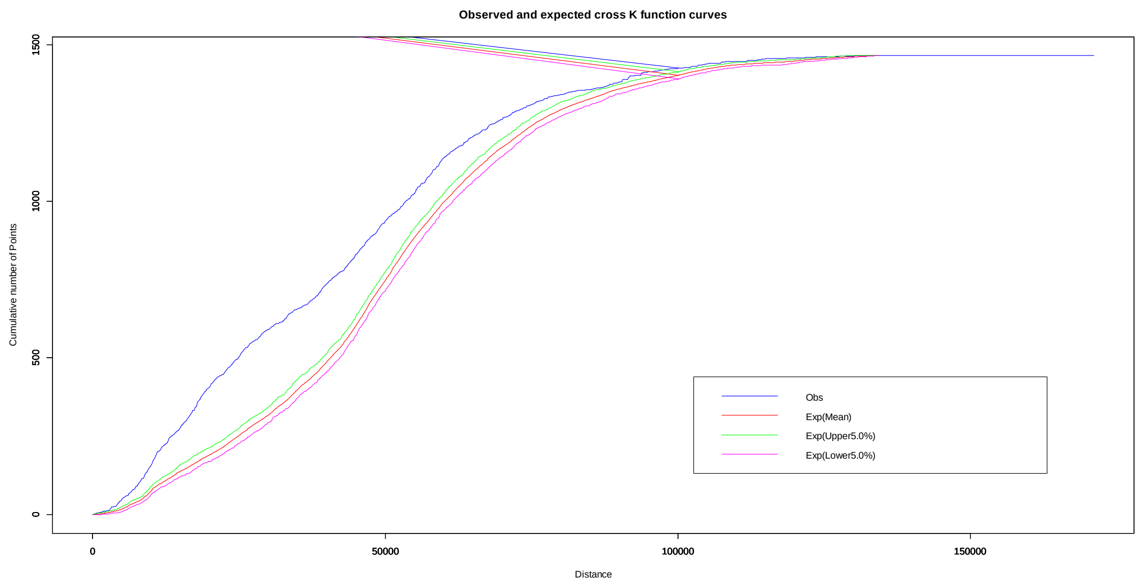
Weekend Hotspot 10



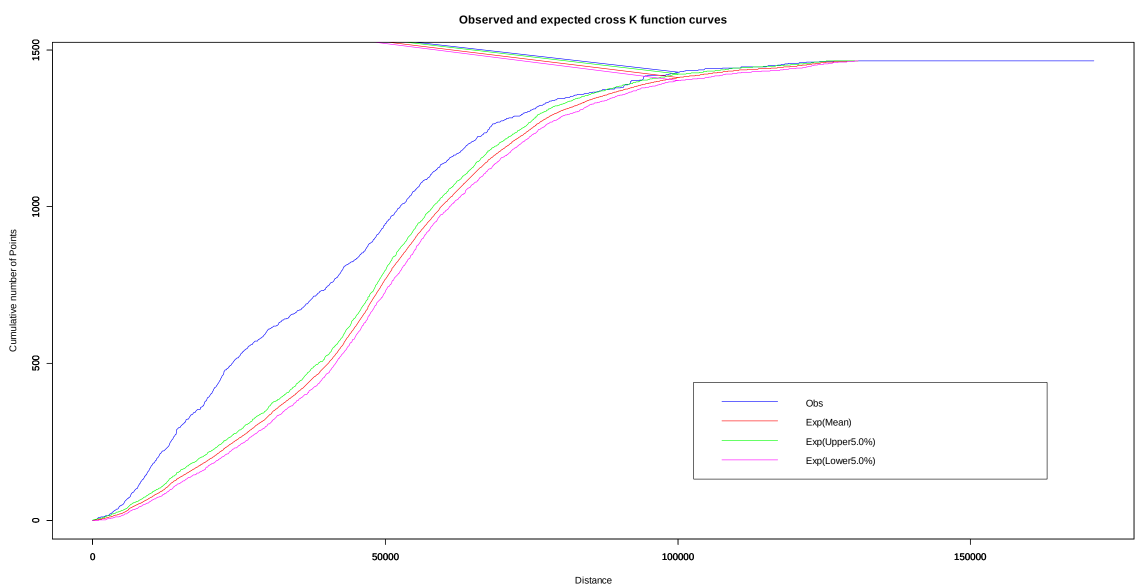
Weekend Hotspot 11



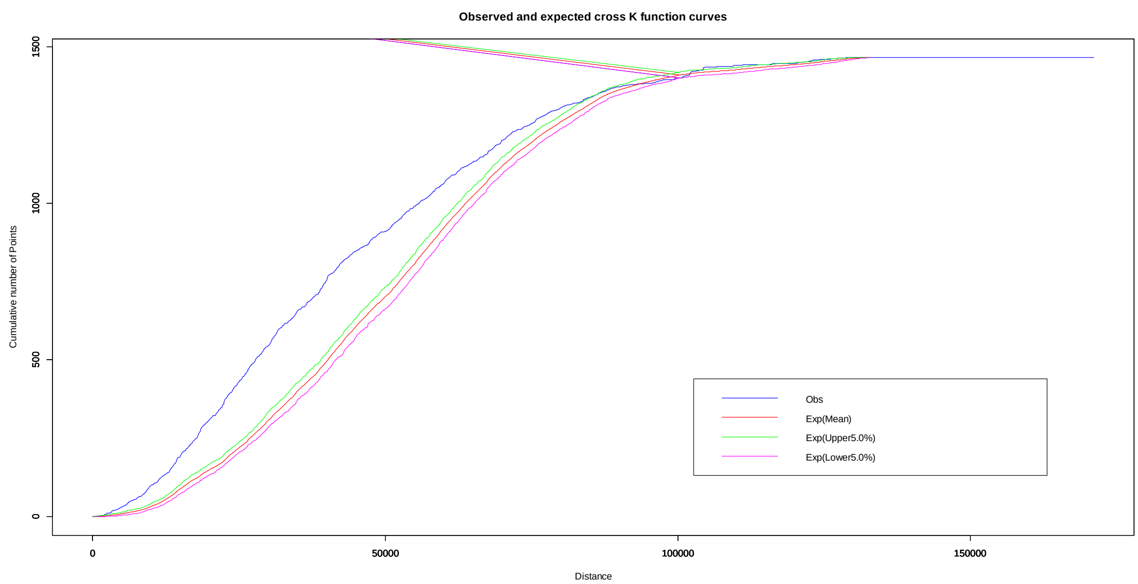
Weekend Hotspot 12



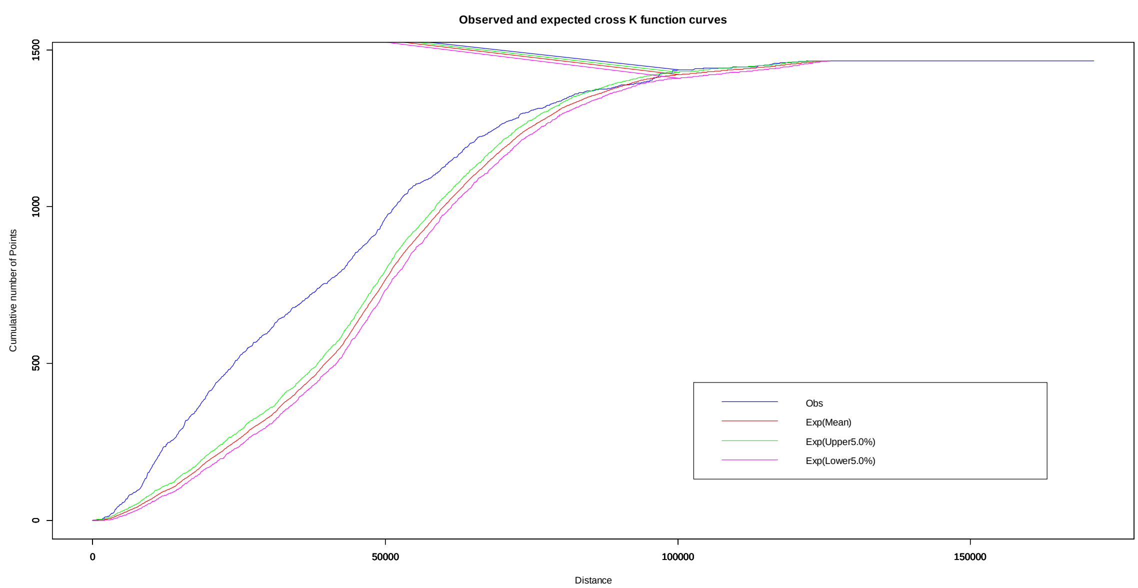
Weekend Hotspot 13



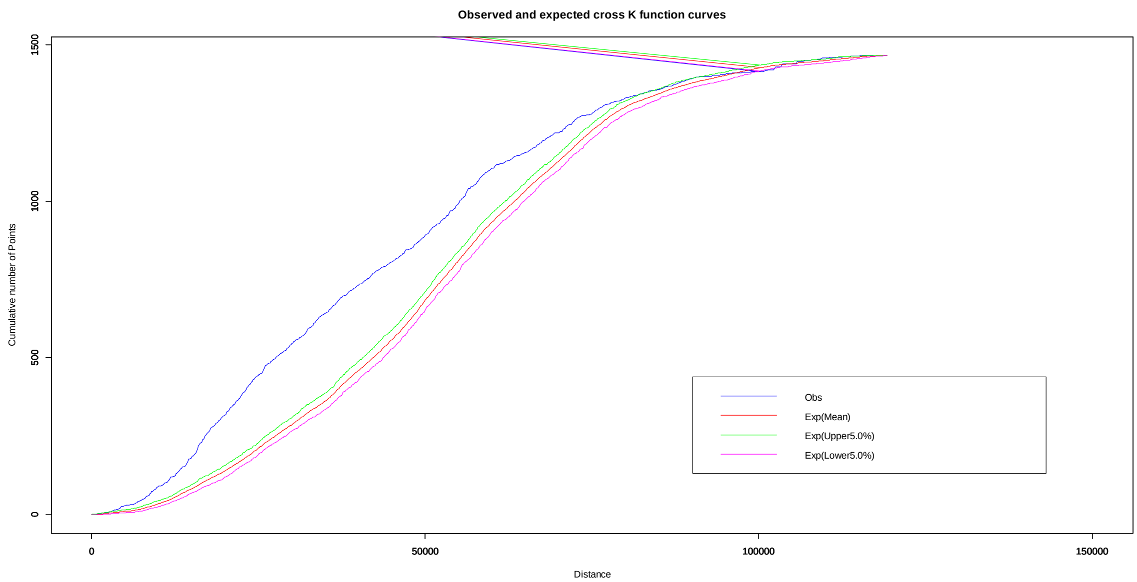
Weekend Hotspot 14



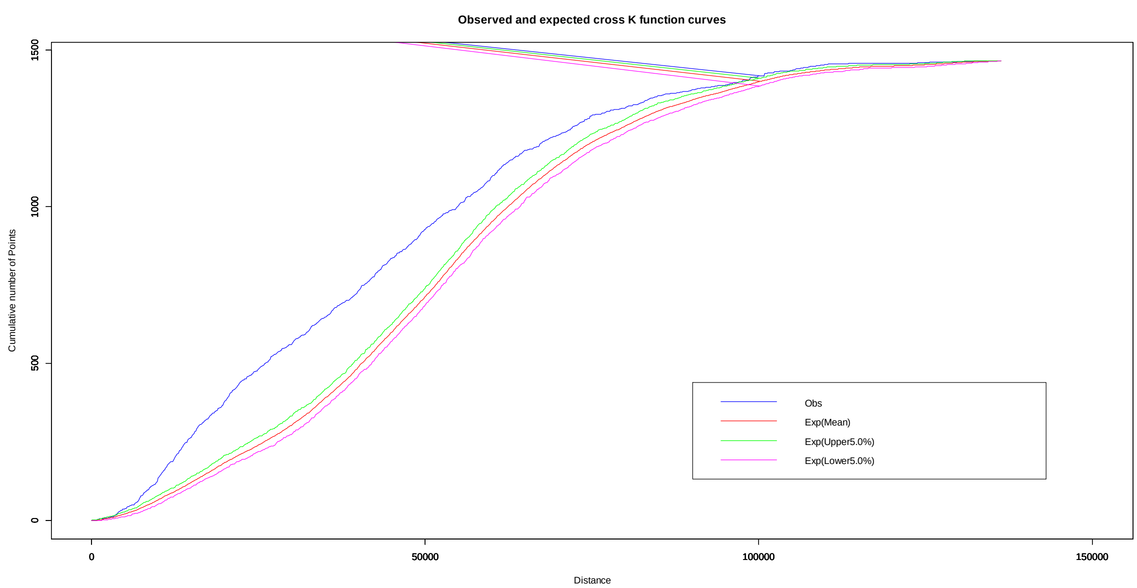
Weekend Hotspot 15



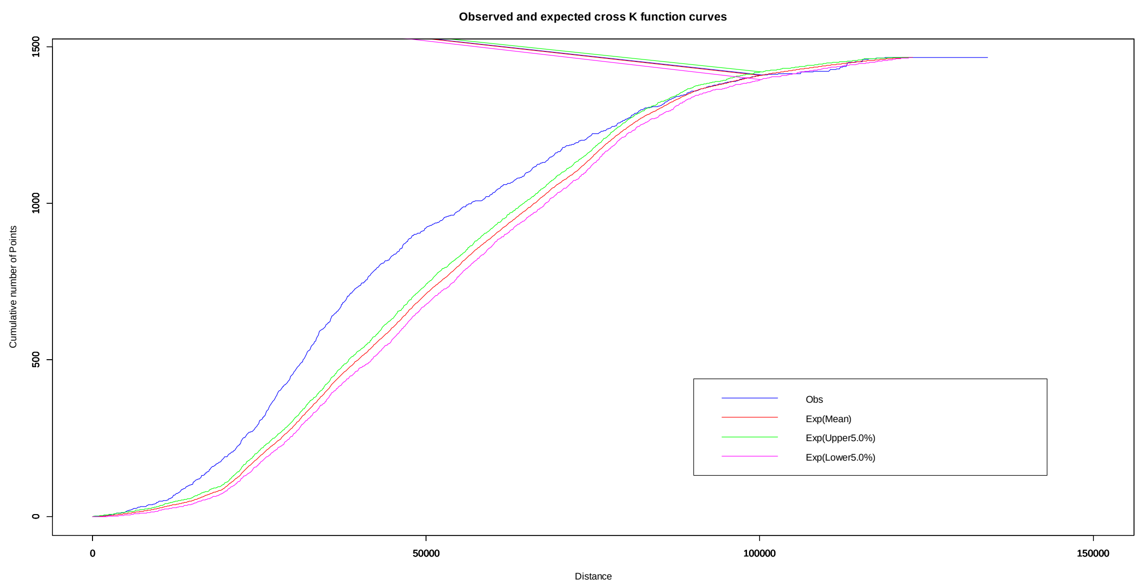
Weekend Hotspot 16



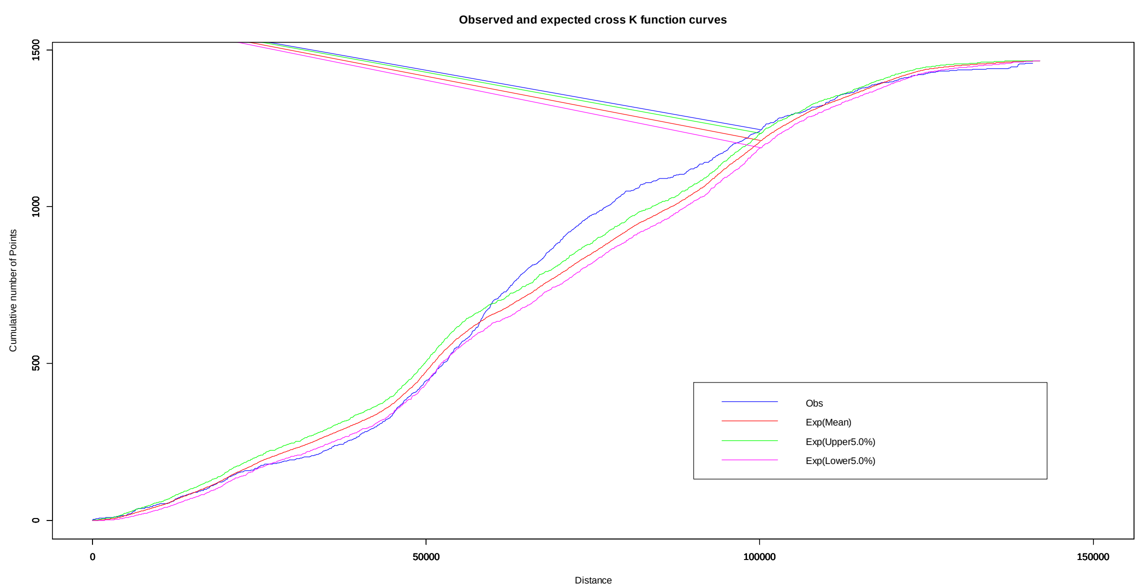
Weekend Hotspot 17



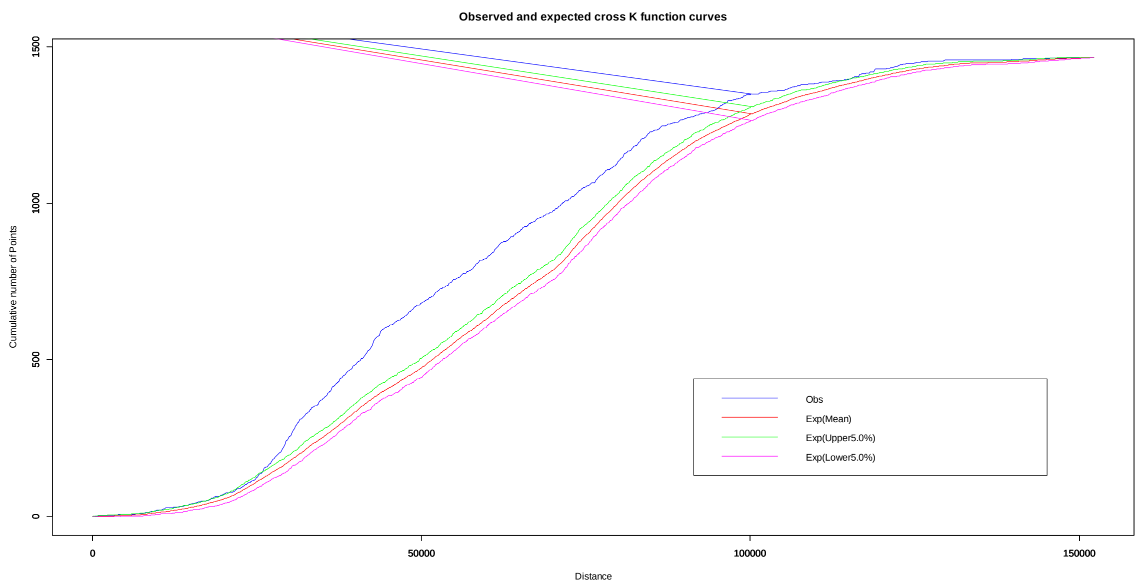
Weekend Hotspot 18



Weekend Hotspot 19



Weekend Hotspot 20



REFERENCES

Anderson, T.K. (2009). Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*, 41(3), 359-364.

Bailey, T.C., and Gatrell, A.C. (1995). Interactive spatial data analysis. London, UK: Longman.

Black, W.R. (1991). Highway accidents: a spatial and temporal analysis. *Transportation Research Record*, 1318, 75-82.

Black, W.R., and Thomas, I. (1998). Accidents on Belgium's motorways: a network autocorrelation analysis. *Journal of Transport Geography*, 6(1), 23-31.

Borruso, G. (2005). Network density estimation: analysis of point patterns over a network. *Lecture Notes in Computer Science: Computational Science and Its Applications*, 3482, 126-132.

Brookhuis, K.A., Waard, D., and Samyn, N. (2004). Effects of MDMA (ecstasy), and multiple drugs use on (simulated) driving performance and traffic safety. *Psychopharmacology*, 173(3-4), 440-445.

Calafat, A., Blay, N., Juan, M., Adrover, D., Bellis, M.A., Hughes, K., Stocco, P., Siamou, I., Mendes, F., and Bohrn, K. (2009). Traffic risk behaviors at nightlife: drinking, taking drugs, driving, and use of public transport by young people. *Traffic Injury Prevention*, 10(2), 162-169.

Carlson, P.J., Park, E.S., and Andersen, C.K. (2009). Benefits of pavement markings. A renewed perspective based on recent and ongoing research. *Transportation Research Record*, 2107, 59-68.

Chang, K. (2012). Introduction to geographic information systems (6th ed.). New York, NY: McGraw-Hill.

Chen, L., Chen, C., Srinivasan, R., McKnight, C.E., Ewing, R., and Roe, M. (2012). Evaluating the safety effects of bicycle lanes in New York City. *American Journal of Public Health*, 102(6), 1120-1127.

Chrysler, S.T., Danielson, S.M., and Kirby, V.M. (1996). Age differences in visual abilities in nighttime driving field conditions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 40(18), 923-927.

Dai, D., Taquechel, E., Steward, J., and Strasser, S. (2010). The impact of built environment on pedestrian crashes and the identification of crash clusters on an urban university campus. *Western Journal of Emergency Medicine*, 11(3), 294-301.

Das, A., Abdel-Aty, M., and Pande, A. (2009). Using conditional inference forests to identify the factors affecting crash severity on arterial corridors. *Journal of Safety Research*, 40(4), 317-327.

Diggle, P.J. and Chetwynd, A.G. (1991). Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics*, 47, 1155-1163.

Dixon, P.M. (2002). Ripley's K function. *Encyclopedia of Environmetrics*, 3, 1796-1803.

Duncan, R.P. (1993). Testing for life historical changes in spatial patterns of four tropical tree species in Westland, New Zealand. *Journal of Ecology*, 81(3), 403-416.

Erdogan, S., Yilmaz, I., Baybura, T., and Gullu, M. (2008). Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar. *Accident Analysis & Prevention*, 40(1), 174-181.

Evans, L. (2004). Traffic safety. Bloomfield Hills, MI: Science Serving Society.

Ewing, R., and Dumbaugh, E. (2009). The built environment and traffic safety: A review of empirical evidence. Journal of Planning Literature, 23, 347-367.

Ewing, R., Schieber, R.A., and Zegeer, C.V. (2003). Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities. *American Journal of Public Health*, 93(9), 1541-1545.

Farmer, C.M. and Williams, A.F. (2005). Temporal factors in motor vehicle crash deaths. *Injury Prevention*, 11, 18-23.

FARS (Fatal Analysis Reporting System Encyclopedia) (1994-2011). National Highway Traffic Safety Administration: Summary of FARS Data Tables. http://www-fars.nhtsa.dot.gov/Main/index.aspx

FARS (Fatal Analysis Reporting System) (2001-2011). Washington, D.C.: National Highway Traffic Safety Administration. ftp://ftp.nhtsa.dot.gov/fars/

Fell, J.C., Fisher, D.A., Voas, R.B., Blackman, K., and Tippetts, A.S. (2008). The relationship of underage drinking laws to reductions in drinking drivers in fatal crashes in the United States. *Accident Analysis & Prevention*, 40(4), 1430-1440.

Florida, R. and Jackson, S. (2010). Sonic city: the evolving economic geography of the music industry. *Journal of Planning Education and Research*, 29(3), 310-321.

Gaines, K.F., Bryan, A.L., and Dixon, P.M. (2000). The effects of drought on foraging habitat selection in breeding wood storks in coastal Georgia. *Waterbirds: The International Journal of Waterbird Biology*, 23, 64-73.

García-Palomares, J.C. (2010). Urban sprawl and travel to work: the case of the metropolitan area of Madrid. Journal of Transport Geography, 18(2), 197-213.

Gonda, Y., and Okunki, K. (2008). A study of evaluating bus transport service in an urban area based on concepts of accessibility. *Papers and Proceedings of the Geographic Information Systems Association*, 17, 363-368.

Goodchild, M.F. (2000). Communicating geographic information in a digital age. *Annals of the Association of American Geographers*, 90(2), 344-355.

Gutiérrez, J., Condeço-Melhorado, A., and Martín, J.C. (2010). Using accessibility indicators and GIS to assess spatial spillovers of transport infrastructure investment. Journal of Transport Geography, 18(1), 141-152.

Hu, W., McCartt, A.T., and Teoh, E.R. (2011). Effects of red light camera enforcement on fatal crashes in large US cities. *Journal of Safety Research*, 42(4), 277-282.

Isebrands, H., Hallmark, S., Li, W., McDonald, T., Storm, R., and Preston, H. (2010). Roadway lighting shows safety benefits at rural intersections. *Journal of Transportation Engineering*, 136(11), 949-955.

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

Khan, X.Q., and Noyce, D.A. (2008). Spatial analysis of weather crash patterns. *Journal of Transportation Engineering*, 134(5), 191-202.

Kim, D., Washington, S., and Oh, J. (2006). Modeling crash types: new insights into the effects of covariates on crashes at rural intersections. *Journal of Transportation Engineering*, 132(4), 282-292.

Kim, K.H., Yum, D.H., Byeon, D.K., Kim, D.Y., and Lee, D. (2010). Improving driver's visual field using estimation of curvature. *Control Automation and Systems, 2010 International Conference on 27-30 Oct*, pp. 728-731.

Konstantopoulos, P., Chapman, P., and Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis & Prevention*, 42(3), 827-834.

Kotavaara, O., Antikainen, H., and Rusanen, J. (2011). Population change and accessibility by road and rail networks: GIS and statistical approach to Finland 1970-2007. Journal of Transport Geography, 19(4), 926-935.

Kotkin, J. (2011). The next big boom towns in the U.S. http://www.forbes.com/sites/joelkotkin/2011/07/06/the-next-big-boom-towns-in-the-u-s/

Krisp, J.M., and Durot, S. (2007). Segmentation of lines based on point densities: an optimization of wildlife warning sign placement in southern Finland. *Accident Analysis & Prevention*, 39(1), 38-46.

Kumagai, K., and Mizuma, Y. (2009). Applicability of the spatial analysis to vegetation distribution in terms of prevention of the fire spreading. *Papers and Proceedings of the Geographic Information Systems Association*, 18, 247-250.

Kuo, P.F., Zeng, X.S., and Lord, D. (2011). Guidelines for choosing hot-spot analysis tools based on data characteristics, network restrictions, and time distributions. *The 91st Annual Meeting of the Transportation Research Board*, Washington, D.C., 12-3788.

Lerner, B.H. (2011). Drunk driving, distracted driving, moralism, and public health. *New England Journal of Medicine*, 365, 879-881.

Lu, Y., and Chen, X. (2007). On the false alarm of planar K-function when analyzing urban crime distributed along streets. *Social Science Research*, 36(2), 611-632.

Lyon, J.D., Pan, R., and Li, J. (2012). National evaluation of the effect of graduated driver licensing laws on teenager fatality and injury crashes. *Journal of Safety Research*, 43(1), 29-37.

Marques, P.R., Voas, R.B., Tippetts, A.S., and Beirness, D.J. (1999). Behavioral monitoring of DUI offenders with the Alcohol Ignition Interlock Recorder. *Addiction*, 94(12), 1861-1870.

Masten, S.V., Foss, R.D., and Marshall, S.W. (2011). Graduated driver licensing and fatal crashes involving 16- to 19- year-old drivers. *Journal of the American Medical Association*, 306(10), 1098-1103.

Metropolitan Government of Nashville and Davidson County. (2013). Working in Nashville. http://www.nashville.gov/Work/About-Nashville.aspx

Moons, E., Brijs, T., and Wets, G. (2009). Improving Moran's Index to identify hotspots in traffic safety. *Geocomputation and Urban Planning*, 176, 117-132.

NHTSA (National Highway Traffic Safety Administration) (2003). Traffic safety facts, 2002. http://www-nrd.nhtsa.dot.gov/Pubs/2002pcyfacts.pdf

NSC (National Safety Council) (2003). Estimating the costs of unintentional injuries. http://www.nsc.org/news_resources/injury_and_death_statistics/Pages/EstimatingtheCostsofUnintentionalInjuries.aspx

Okabe, A., Okunuki, K., and Shiode, S. (2006). SANET: a toolbox for spatial analysis on a network. *Geographical Analysis*, 38(1), 57-66.

Okabe, A. and Satoh, T. (2009). Spatial analysis on a network, in Fotheringham, A.S. and Rogers, P.A. (eds.). *The SAGE Handbook of Spatial Analysis*, Thousand Oaks, CA: Sage Publications, pp. 443-464.

Okabe, A. and Sugihara, K. (2012). Spatial analysis along networks: statistical and computational methods. Hoboken, NJ: John Wiley and Sons.

OMB (Office of Management and Budget) (2006). Update of statistical area definitions and guidance on their uses. http://www.whitehouse.gov/sites/default/files/omb/bulletins/fy2007/b07-01.pdf

Opiela, K.S., Andersen, C.K., and Schertz, G. (2003). Driving after dark. *Public Roads*, 66(4), 22-25.

O'Sullivan, D. and Unwin, D.J. (2003). Geographic information analysis. Hoboken, NJ: John Wiley and Sons.

Owsley, C., and McGwin, G.Jr. (1999). Vision impairment and driving. *Survey of Ophthalmology*, 43(6), 535-550.

Peterson, C.J., and Squiers, E.R. (1995). An unexpected change in spatial pattern across 10 years in an Aspen White Pine forest. *Journal of Ecology*, 83(5), 847-855.

Plainis, S., and Murray, I.J. (2002). Reaction times as an index of visual conspicuity when driving at night. *Ophthalmic and Physiological Optics*, 22(5), 409-415.

Ripley, B.D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability*, 13, 255-266.

Shiode, S. (2008). Analysis of a distribution of point events using the network-based quadrat method. *Geographical Analysis*, 40(4), 401-422.

Shope, J.T., and Molnar, L.J. (2003). Graduated driver licensing in the United States: evaluation results from the early programs. *Journal of Safety Research*, 34(1), 63-69.

Silverman, B.W. (1986). Density estimation for statistics and data analysis. London, UK: Chapman and Hall.

Spooner, P.G., Lunt, I.D., Okabe, A., and Shiode, S. (2004). Spatial analysis of roadside Acacia populations on a road network using the network K-function. *Landscape Ecology*, 19(5), 491-499.

Stamp, N.E., and Lucas, J.R. (1990). Spatial patterns and dispersal distance of explosively dispersing plants in Florida sandhill vegetation. *Journal of Ecology*, 78(3), 589-600.

Steenberghen, T., Aerts, K., and Thomas, I. (2010). Spatial clustering of events on a network. *Journal of Transport Geography*, 18(3), 411-418.

TDOT (Tennessee Department of Transportation) (2012). Project planning division – AADT (Annual Average Daily Traffic). http://www.tdot.state.tn.us/projectplanning/adt.asp

Thill, J.C. (2000). Geographic information systems for transportation in perspective. *Transportation Research Part C: Emerging Technologies*, 8(1-6), 3-12.

Thombs, D.L., O'Mara, R.J., Tsukamoto, M., Rossheim, M.E., Weiler, R.M., Merves, M.L., and Goldberger, B.A. (2010). Event-level analyses of energy drink consumption and alcohol intoxication in bar patrons. *Addictive Behaviors*, 35(4), 325-330.

UN (United Nations), Department of Economic and Social Affairs, Population Division. (2012) *World Urbanization Prospects, the 2011 Revision: Highlights*. http://esa.un.org/unup/pdf/WUP2011_Highlights.pdf

United States Census Bureau – American FactFinder. http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml

United States Census Bureau – (2012) TIGER/Line® Shapefiles. http://www.census.gov/cgi-bin/geo/shapefiles2012/main

Vandenbulcke, G., Steenberghen, T., and Thomas, I. (2009). Mapping accessibility in Belgium: a tool for land-use and transport planning. Journal of Transport Geography, 17(1), 39-53.

Wagenaar, A.C., Maldonado-Molina, M.M., Ma, L., Tobler, A.L., and Komro, K.A. (2007). Effects of legal BAC limits on fatal crash involvement: analyses of 28 States from 1976 through 2002. *Journal of Safety Research*, 38, 493-499.

Wagenaar, A.C., and Toomey, T.L. (2000). Alcohol policy: gaps between legislative action and current research. *Contemporary Drug Problems*, 27, 681-733.

Wagenaar, A.C., and Toomey, T.L. (2002). Effects of minimum drinking age laws: review and analyses of the literature from 1960 to 2000. *Journal of Studies on Alcohol and Drugs*, 14, 206-225.

WHO (World Health Organization) (2013). World health statistics – Indicator compendium. http://www.who.int/gho/publications/world_health_statistics/WHS2013_IndicatorCompendium.pdf

Xie, Z.X., and Yan, J. (2008). Kernel density estimation of traffic accidents in a network space. *Computers, Environment and Urban Systems*, 32(5), 396-406.

Yamada, I., and Thill, J.C. (2004). Comparison of planar and network k-functions in traffic accident analysis. *Journal of Transport Geography*, 12, 149-158.

Yamada, I., and Thill, J.C. (2007). Local indicators of network-constrained clusters in spatial point patterns. *Geographical Analysis*, 39(3), 268-292.

Yu, R., and Abdel-Aty, M. (2013). Investigating the different characteristics of weekday and weekend crashes. *Journal of Safety Research*, 46, 91-97.

Zhang, J., Fraser, S., Lindsay, J., Clarke, K., and Mao, Y. (1998). Age-specific patterns of factors related to fatal motor vehicle traffic crashes: focus on young and elderly drivers. *Public Health*, 112(5), 289-295.