2012 | Faculteit Bedrijfseconomische Wetenschappen

DOCTORAATSPROEFSCHRIFT

Proactive Approaches in Quantitative Micro- and Macroscopic Crash Analysis

Focus on Model-Based Safety Evaluation of Traffic Policy Measures

Proefschrift voorgelegd tot het behalen van de graad van doctor in de verkeerskunde, te verdedigen door:

Ali Pirdavani

Promotor: prof. dr. Tom Brijs
Copromotoren: prof. dr. Tom Bellemans
prof. dr. Geert Wets

D/2012/2451/44

universiteit hasselt
KNOWLEDGE IN ACTION

Preface

Starting my career as an academic researcher was triggered when I wrote a paper for the "Road Safety on Four Continents" conference in 2007. Despite getting my paper accepted, I had a little hope to be able to attend the conference. Hopes came back when I got a fund from the *Global Road Safety Partnership* which was cosponsoring that conference. I will be always grateful for the chance they gave me to meet active researchers in my domain, make contact with them and finally find the opportunity that in a way changed the direction of my professional life. Following that conference, I kept in contact with some of participants and shortly found an available vacancy at IMOB, Transportation Research Institute of Hasselt University. I am indebted to my promoter, Prof. Dr. Tom Brijs who trusted me and gave me the opportunity to pursue my PhD.

The start of my PhD was coupled with the start of a project entitled "A Model-based Approach for Evaluating the Safety and Environmental Effects of Traffic Policy Measures (SBO:MASE)" that I was appointed to. It was a large project including several external partner and was coordinated by IMOB. Although the project was considered to be a safety and environmental oriented research, IMOB's main task was set to provide external partners with input data (a relatively transportation related task). This compelled me to seek new ideas that would later on constitute my PhD dissertation as a road safety oriented research. Prior to the start of my PhD, due to the lack of good quality crash data in my origin country, Iran, my experience was limited to the methodologies which are appropriate in the context of absence or a lack of good quality data. Working on the SBO:MASE project gave me the opportunity to become more familiar with the existing data and to explore new possibilities. Since then, however, my PhD research evolved when new data became available and new ideas were investigated. Dissimilarities between applicable methodologies to study crashes in different data availability situations motivated me to demonstrate advantages and limitations of each methodology that forms the structure of this dissertation. This could be the reason that at first glance the configuration of this dissertation might seem to be less connected.

What is presented in this dissertation has been made possible through the generous support of many others who have contributed to the progression of my work in different ways. First of all, I would like to express my heartfelt appreciation to my promoter Prof. Dr. Tom Brijs. It was indeed a great privilege for me to work under your supervision. I have benefitted very much your experience in doing research and from your wide range of knowledge. Although you are among the busiest people at IMOB, you were always willing to spend time to review my papers and discuss the matters, which is greatly appreciated.

I also owe a debt of gratitude to my co-promoter, Prof. Dr. Tom Bellemans, whose critiques were always relevant and helpful. I think back with great pleasure to our very constructive discussions, not forgetting our inspirational meetings while traveling to the Netherlands for our regular SBO:MAES related meetings. I'm truly grateful for your immeasurable contribution to my research.

My profound gratitude further goes to my co-promoter, Prof. Dr. Geert Wets, and the member of my PhD commission, Prof. Dr. Koen Vanhoof for their thorough guidance during the last four years. I would like to thank you for the added value you brought into this research.

Next, I would like to express my sincere thankfulness to the external members of the jury, Prof. Dr. Elke Hermans, Prof. Dr. Mohamed Abdel-Aty and Dr. Heike Martensen for their thoughtful and meticulous reading of the text. Your constructive suggestions and remarks have significantly improved both the content and style of this dissertation.

Many of my IMOB colleagues have also contributed a great deal to my PhD research. In particular, I would like to thank Konstantinos Perrakis, George Sammour, Bruno Kochan, Dirk Roox, Katrien Declercq, Benoit Depaire, An Neven, Kurt Van Hout, Yongjun Shen and Amin Azmoon. I would also like to extend my thanks to some friends who left IMOB but will always be my dear friends, Hans Tormans, Marlies Vanhulsel, Els Hannes and Diana Kusumastuti.

I want to also thank Kristel Hertogs who has kindly helped me with all administrative matters, from the very first day I arrived in Belgium until now. Moreover, I want to thank Nadine Smeyers for translating summary of this

thesis into Dutch. I would also like to thank all IMOB staff for creating a positive, friendly and supportive work environment.

It would have been unbearable to accomplish my PhD without endless support of my beloved parents. You know how hard was being away from you but I can assure you at this moment that hard work and suffering pay off. My heartfelt gratitude goes to my dear parents, siblings and in-laws for their constant support and encouragement.

I cannot adequately express my gratitude to my beloved wife, Elhaam. You have always been a faithful companion, in difficult and joyful moments, throughout my PhD period. Your never-ending patience, inspiration and love have been and will be the source of success in my life. I can no other answer make, but, thanks, and thanks.

Ali Pirdavani

November 7th, 2012

Summary

Road traffic casualties cause huge social and economic costs that are known to be among the major public health issues for many decades. The current mortality rate is estimated to increase dramatically over the next two decades, unless crashes are properly analyzed and effective safety interventions are appropriately adopted to diminish this undesirable increase. The very first stage of performing any road crash analysis is comprised of collecting relevant data. These data are used within several applications to comprehend the causality of crash occurrence, identify crash hotspots, develop crash prediction models (CMPs), evaluate different traffic policy measures and etc. With this regard, crash data are the most important source of information that is employed to carry out road crash analysis. These data should be accurate, timely, comprehensive and reliable to ensure a successful crash analysis. Although presence of reliable crash data is a key element in the procedure of crash analysis, there are several situations where good quality crash data are not available. Under these circumstances, road safety can be assessed indirectly by means of methodologies that are different from the conventional road safety analyses since they are independent of historical crash data. In contrast and in the case of most of the developed countries, numerous data collection systems are available which facilitate more complex crash analysis.

The first major contribution of this dissertation is therefore set to provide a collection of different methodologies for safety prediction and evaluation, depending on the amount and type of data available. With this regard, two chapters (i.e. Chapters 2 and 3) are devoted to demonstrate two methodologies (gathering expert knowledge and conflict observation by means of microsimulation modeling) that are applied to carry out crash analysis in the absence of crash data. In addition, the second part of this dissertation (i.e. Chapters 4, 5, 6 and 7) is set to develop macro-level CPMs so as to evaluate the road safety impacts of two travel demand management (TDM) strategies (i.e. a fuel cost increase by 20% and a teleworking scenario in which 5% of the working population are assumed to telework). This is considered to be the second major contribution of this dissertation.

Road safety analysis can be performed via different approaches. One of the most followed approaches is the development of predictive models. This approach aims to justify and predict road safety problems by means of analytical modeling. Predictive models can be developed at different levels of aggregation. At a microscopic level, CPMs usually concern an infrastructural element such as a road segment or an intersection. On the contrary, macroscopic crash analysis concerns a relatively large area and associates the probability of crash occurrence with a set of variables that have macro-level characteristics. Limitations in employing micro-level CPMs to perform proactive planning level road safety evaluation, impose pursuing macro-level CPMs. In this dissertation different efforts have been made to develop macro-level CPMs by associating crashes with a number of predictive macro-level characterized variables, such as exposure, network and socio-demographic variables. These models are subsequently utilized to proactively evaluate the safety impacts of the two TDM strategies.

Controlling exposure is known to be a straightforward and implementable approach as a crash reduction strategy. Among the measures to control exposure, reducing travel demand is considered to be an effective measure. Travel demand reduction can be achieved by adopting TDM strategies like pricing (e.g. fuel price increase), telecommunications (teleworking), and so on. TDM strategies are generally implemented to improve transportation systems' efficiency; however, their potential traffic safety impacts should not be ignored. Moreover, identifying the positive traffic safety impacts of TDM strategies and highlighting their benefits can also strengthen the implementation of those strategies among policy makers. In this regard, the study carried out in the second part of this dissertation (i.e. Chapters 6 and 7), makes a contribution to the necessity for safety assessment of TDM strategies.

In order to assess TDM's traffic safety implications, it is necessary to have TDM sensitive exposure measures. Therefore, exposure metrics produced by an activity-based transportation model, called "FEATHERS", are utilized for this purpose. The motivation behind using an activity-based transportation model relies on the fact that it realistically simulates individuals' travel behavior, so as to produce TDM sensitive exposure measure and as such, enables practitioners and researchers to evaluate TDM strategies as accurately as possible. In addition, TDM strategies are often adopted and analyzed at a macro-level rather than on an individual intersection or road segment. Moreover, TDM strategies aim to change travel behavior of road users individually, however the impact of changed behavior on the transportation system or on road safety should be addressed collectively and at an aggregate level. Therefore, a macroscopic approach at the level of traffic analysis zone (TAZ) is conducted to develop the prediction models. These models are referred to as zonal crash prediction models (ZCPMs) as they are constructed on the TAZ based information.

Thus far, we explained the general research motivation and framework of this dissertation. Since this dissertation is considered as a paper-based thesis, we would like to individually recapitulate employed techniques and derived results of each chapter in the next section of this summary.

Chapter 2. In this chapter a framework is presented to identify and prioritize crash hotspots in the absence of crash data. To this intent, two main tasks are performed. In the first task, relevant hotspot criteria and their relative importance are determined by means of an expert knowledge collecting method, namely the Delphi technique. The objective of most Delphi applications is the reliable and creative exploration of ideas or the production of suitable information for decision-making. Once the final criteria weights are obtained, they are applied in a multiple criteria decision making (MCDM) context to develop the crash prioritization model. For this research, we have adopted the Technique for Order Preference Similarity to Ideal Solution (TOPSIS) for prioritizing crash hotspots. The results of the presented framework reveal the usefulness of adopting qualitative methodologies in road safety analysis, specifically in the context of absence or a lack of good quality crash data. Qualitative approaches provides detailed insights into behaviors and values, to a depth of understanding which is often not possible using quantitative methods.

Chapter 3. In this chapter a methodology is established to investigate the relationship between speed, traffic volume and safety levels at intersections that can be employed in case of a lack of crash data. For this purpose, a microsimulation model, namely S-Paramics, is utilized to simulate interactions between cars so as to determine the safety levels of intersections by means of calculating proximal safety indicators. These indicators have a near-crash attribute and are suggested to serve as alternative to the use of crash data. To perform the safety analysis, several sets of scenarios based on different traffic volume and speed limit categories are defined and safety indicators are measured for each of these scenarios. The analyses results reveal that increasing the speed limit on both major and minor approaches of an intersection will deteriorate safety level whilst its magnitude will be larger for higher ranges of traffic volume.

Chapter 4. In this chapter, different statistical models are developed to construct the association between observed crashes and a set of predictor variables. Given the observed overdispersion in the crash data, the Negative Binomial (NB) model within the generalized linear modeling (GLM) framework is adopted to enable the modeling of overdispersed data. The analyses in Chapter 4 reveals that sole use of number of trips (NOTs) - representing trip production/attraction of a TAZ – in construction of CPMs, results in missing important information about the characteristics of travel demand, since NOTs do not contain information on trip time, trip length and route choice. Moreover, transit traffic which just passes through a TAZ is left out when only using the NOTs as the exposure variable. This part of exposure can have a significant share of the total exposure observed in a TAZ. Thus, other exposure variables which are sensitive to the impacts of trip assignment should be taken into account (e.g. vehicle kilometer traveled (VKT) or vehicle hour traveled (VHT)). Comparing the models' performance confirms that models which comprise both trip-based and flow-based exposure variables outperform models which only have one of the exposure variables in their formulation.

Chapter 5. The results of analyses indicate the presence of spatial non-stationarity in the data employed in developing ZCPMs. One of the solutions for taking the spatial variation into account is developing a set of local models, so-called Geographically Weighted Regression (GWR) models. These models rely on the calibration of multiple regression models for different geographical entities. The GWR technique is therefore adapted to the GLM models and form Geographically Weighted Generalized Linear Models (GWGLMs). These models are extensions of models developed in Chapter 4 and are able to model count data (such as number of crashes (NOCs)) while simultaneously accounting for the spatial non-stationarity and are, therefore, particularly useful in the context of this study. Comparing all developed models shows that the GWGLM models always outperform the conventional GLM models.

Chapter 6. As mentioned earlier, the main contribution of this dissertation is set to evaluate the safety impacts of two TDM scenarios. The major attributes of this evaluation task are that the safety evaluation is coupled with the output of an activity-based transportation model and is carried out within a proactive framework. In this chapter the models developed in Chapter 4 are utilized to carry out the first assessment exercise in which fuel price is assumed to increase by 20%. The results of the comparison analysis reveal that this fuel-cost increase scenario has an impact on total travel demand, total crash occurrence, VKT and VHT and mode shift. Despite the improvement of the overall safety situation, changes in the NOCs for different crash-type/severity-level are not identical. After the fuel-cost scenario implementation, many trips are shifted from private cars towards other transportation modes, such as biking and walking (i.e. Slowmode). As a result of an increase in the NOTs for the Slowmode category, crashes which involve cyclists and pedestrians (i.e. vulnerable road users) are generally predicted to increase unlike car only crashes that are expected to decrease. This reveals that the fuel-cost increase scenario affects different road users differently.

Chapter 7. In this chapter the models developed in Chapter 5 are employed to perform the second assessment exercise in which 5% of the working population is assumed to telework in any given day of a week. The results of the comparison analysis confirm that the teleworking scenario has an impact on total travel demand, VKT and VHT and total crash occurrence. The teleworking scenario positively affects the safety levels of different road users. Due to the nature of this scenario and an observed rebound effect, Slowmode trips are expected to reduce slightly less than car trips. Therefore, the positive safety impacts are slightly lower for “Car-Slowmode” crashes compared with “Car-Car” crashes.

Glossary of Acronyms

AADT	Average Annual Daily Traffic
AIC	Akaike Information Criterion
AICc	Corrected AIC
APM	Accident Prediction Models
CAR	Conditional Auto-Regression
CART	Classification and Regression Trees
CCFS	Car-Car/Fatal and Severe Injury
CCSL	Car-Car/Slight Injury
CHAID	Chi-Squared Automatic Interaction Detector
CPMs	Crash Prediction Models
CR	Consistency Ratio
CSFS	Car-Slow Mode/Fatal and Severe Injury
CSSL	Car-Slow Mode/Slight Injury
DF	Degree of Freedom
DM	Decision Maker
DR	Deceleration Rate
DST	Deceleration-to-Safety Time
EPDO	Equivalent Property Damage Only
ETSC	European Transport Safety Council
FEATHERS	Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS
FHWA	Federal Highway Administration
GEE	Generalized Estimating Equation
GLM	Generalized Linear Modeling
GVRD	Greater Vancouver Regional District
GWGLMs	Geographically Weighted Generalized Linear Models
GWPR	Geographically Weighted Poisson Regression
GWR	Geographically Weighted Regression
ITE	Institute of Transportation Engineers
Km/hr	Kilometer per Hour
LISA	Local Indicators of Spatial Association
MCDM	Multiple Criteria Decision Making
MSPE	Mean Squared Prediction Error
NB	Negative Binomial
NOCs	Number of Crashes

NOICs	Number of Injury Crashes
NOTs	Number of Trips
OD	Origin-Destination
OECD	Organization for Economic Co-Operation and Development
OLS	Ordinary Least Square
PCC	Pearson Correlation Coefficients
PET	Post-Encroachment Time
PIARC	Permanent International Association of Road Congresses
PSD	Proportion of Stopping Distance
RI	Robustness Index
RPI	Relative Proximity Index
RSA	Road Safety Audit
RSAP	Road Safety Action Program
RSG	Road Safety Guidelines
SAR	Simultaneous Auto-Regression
SAS	Statistical Analysis System
SDLP	Standard Deviation Of Lateral Position
SEM	Spatial Error Models
TAZ	Traffic Analysis Zone
TDM	Travel Demand Management
TET	Time Extended TTC
TIT	Time Integrated TTC
TLC	Time-to-Line Crossing
TOPSIS	Technique for Order Preference Similarity to Ideal Solution
TRB	Transportation Research Board
TTC	Time-to-Collision
TTZ	Time-to-Zebra
UD	Unsafety Density
UN	United Nations
V/C	Volume over Capacity
VHT	Vehicle Hours Traveled
VIF	Variance Inflation Factor
VKT	Vehicle Kilometers Traveled
VPH	Vehicle per Hour
VTPI	Victoria Transport Policy Institute
WHO	World Health Organization
ZCPMs	Zonal Crash Prediction Models

Table of contents

List of Tables

List of Figures

1. Introduction

1.1. Background

Road traffic casualties and their tremendous social and economic costs are known to be among the major public health issues for many decades. Each year roughly 1.3 million people die and as many as 50 million others are injured as a result of road crashes (*WHO*, 2009). As a result of urbanization, population growth and motorization, the current trends are expected to continue and consequently, the global road safety situation will deteriorate. The current mortality figures are estimated to increase by nearly 65% over the next two decades, unless effective safety countermeasures are implemented to prevent this dramatic increase (*WHO*, 2004).

1.1.1. The Importance of Accurate and Reliable Data

Road crashes occur as a result of different factors, such as driver, road infrastructure and vehicle. To perform any kind of road crash analysis, different types of information at different levels of detail are required. These data are used within numerous applications to comprehend the causality of crash occurrence, identify crash hotspots, develop crash prediction models, implement safety interventions and evaluate different policy measures. In general, there are three major types of data required for road crash analysis. These are crash data, socio-demographic and environmental data and traffic or exposure data (Golembiewski and Chandler, 2011).

Historical crash data are the most important source of information that has been employed to carry out different types of road crash analysis. Crash data should be accessible, accurate, timely, complete and most importantly reliable. Having good quality crash data enables researchers to properly study crashes, accurately identify problems, define suitable safety interventions and appropriately evaluate them.

Another important type of data is the road infrastructure, environment and socio-demographic data. This type of data consists of information about the number and the width of lanes, presence and type of shoulders, presence and type of median, road alignments, pavement condition, drainage condition, road marking condition, road signing condition, presence and configuration of intersections, composition of households, income level, car ownership and etc. Combining this information with historical crash data facilitates a more reliable road safety program evaluation, as it may provide a better understanding of the causality of crash occurrence at specific locations.

The third type of data that is required to carry out crash analysis are traffic or exposure data. It is well documented in the literature that travel demand and exposure are the most important predictors for the number of crashes. Therefore, having comprehensive and informative measures of exposure are indispensable to properly perform crash analysis. In literature, trip production/attraction (e.g. in Pirdavani et al., in press; Naderan and Shahi, 2010; Abdel-Aty et al., 2011a, 2011b), vehicle kilometers traveled (VKT) (e.g. in Pirdavani et al., in press; Lovegrove, 2005; Hadayeghi et al., 2010a, 2010b) and vehicle hours traveled (VHT) (e.g. in An et al., 2011; Pirdavani et al., 2012) have been the most commonly used types of exposure measures.

Although presence of reliable information is required to carry out crash analysis, there are numerous situations where good quality crash data are not available. This can occur for instance when there is no accurate crash data collection system available (typically in developing countries) or where at the planning level of transportation projects, road safety evaluation of newly proposed projects or policies is required while no historical crash data are yet available. Under these circumstances, road safety can be assessed indirectly by different means, such as gathering expert knowledge (e.g. in Donnell et al., 2002; Jamson et al., 2008; Kim et al., 2008; Pirdavani, Brijs, and Wets, 2010), conflict observation (e.g. in Chin and Quek, 1997; Tiwari et al., 1998; Cunto and Saccomanno, 2007, 2008; Laureshyn et al., 2010; Oh et al., 2010; Davis et al., 2011; Autey et al., 2012; Lu et al., 2012), simulation (e.g. in Abdel-Aty et al., 2006; Cunto and Saccomanno, 2008; Pirdavani et al., 2010; Bell et al., 2012) or carrying out a road safety audit (e.g. in *RSA*, 2004; Wang et al., 2011; Mitchell et al., 2012). These methodologies differ from the conventional road safety analyses since they have a proactive characteristic and are independent of historical crash data.

1.1.2. Approaches Towards Road Safety Research

Road safety and crash analysis can be performed via different approaches. According to OECD reports (*OECD*, 1997a, *OECD*, 1997b), there are four approaches available towards road safety research, namely descriptive-, predictive-, risk- and accident consequence models. In the first approach, the main objective is to illustrate the magnitude of road safety problem. In this approach, the main focus is on what causes crashes rather than focusing exclusively on the number of crashes (e.g. in Kam, 2003; Mabunda et al., 2008).

The second approach aims to justify and predict road safety problems by means of analytical macro-level modeling. This approach can be performed by application of predictive models on cross-sectional data or by evaluating road safety in before-and-after studies. Accident hotspot analysis is one of the most well-known applications of cross-sectional studies in which specific locations are identified and prioritized to implement road safety treatments (e.g. in Van den Bossche, 2006; Caliendo et al., 2007; Ma et al., 2008; Abdel-Aty and Haleem, 2011; Abdel-Aty et al., 2012; Pirdavani et al., 2012, in press).

The third approach relies on risk factor models at the micro-level. Risk factor models can be categorized into human factor models and technical models. Human factor models deal with road users' characteristics while technical models explore the relationship between road safety and traffic, road infrastructure or vehicle characteristics (e.g. in Vlahogianni et al., in press; Yau, 2004; Huang et al., 2010; Meng and Weng, 2011; Xu et al., 2012).

In the fourth approach, it is aimed to look into the crash consequences by means of a modeling framework. These crash consequence models can be divided into two major categories of models, namely public health and economy related models (e.g. in de Rome et al., in press; Zaloshnja and Miller, 2004; Czech et al., 2010; Chitturi et al., 2011; Dhondt et al., 2012).

From what has been explained above, it is clear that each single road safety study cannot be exclusively performed within only one of the categorized approaches. In this dissertation, the majority of the work follows the second approach, although some models necessitate including infrastructure and road users characteristics, vehicle information, and so forth.

1.1.3. Reactive vs. Proactive Crash Analysis

From another point of view, road crash analysis can be performed by means of two main approaches; reactive and proactive. Each approach has its own merits and drawbacks in different situations. When crash data are available and safety interventions are immediately required, a reactive approach is often adopted. This approach consists of a set of safety interventions to improve the safety situation of existing crash hotspots. The reactive approach reacts to crashes after they occur, therefore, it requires collecting a substantial amount of crash data before any safety program can be implemented. These data are however often outdated, incomplete and inadequate to be utilized for accurate diagnosis and intervention. Moreover, reactive safety programs are usually more costly as they aren't implemented until crashes have occurred and roads are already built. Despite these shortcomings, reactive approaches are still considered as a valid approach to cope with existing road safety problems. Traditional reactive road safety processes comprise activities such as data collection and management (crash information systems), identification of accident hotspots and analysis, development and realization of countermeasures. The Hazard Elimination Program or a crash hotspot identification and prioritization program are examples of reactive approaches (*FHWA*, 2006).

On the other hand, in the absence of crash data or at the planning level of transportation projects when no crash has occurred yet, a proactive approach seems to be more appropriate. Furthermore, a proactive approach has a preventive attribute and tries to confront safety problems before crashes have occurred. This approach is more ethical and humane and less costly since it doesn’t wait for crashes to occur and lives to be threatened. A proactive approach aims to minimize the risk by evaluating road safety in each stage of the planning process of a transportation project (de Leur and Sayed, 2003).

Road safety audits (RSA), conflict observations and policy impact assessment (see Chapters 6 and 7) are examples of proactive approaches.

Despite the opportunities that proactive road safety programs may provide, there are some obstacles to be overcome. Traditionally, the road planning process does not usually allow planners to consider road safety in their planning decisions. Planners assume that designers will take care of road safety during the design stage. Road designers also think that by applying road design standards, road safety will be accounted for automatically (de Leur, 2001).

1.1.4. Micro- vs. Macro-Level Crash Analysis

Microscopic crash analysis usually concerns an infrastructural element such as a road segment or an intersection. On the contrary, macroscopic crash analysis concerns a relatively large area and associates the probability of crash occurrence with a set of variables that are generally macroscopically characterized. In general, micro-level crash analysis is carried out in a reactive way and by utilizing micro-level crash prediction models (CPMs). Micro-level CPMs have been studied by many researchers in the past. Despite successful implementation of micro-level CPMs and the promising results they may provide at the local level, several shortcomings have been recognized when trying to perform crash analysis by applying micro-level models at the planning level of transportation projects (a proactive approach). For instance, micro-level CPMs are very sensitive to traffic volume values (Lovegrove, 2005) and, therefore, reliability of micro-level CPMs' results greatly depend on very accurate traffic volume forecasts. Accurate traffic volume forecasts of single facilities (e.g. a road section or intersection) can only be carried out by means of short term projections, while predictions at any single location derived from long-term planning-level analyses are known to be generally inaccurate (Lovegrove, 2005). Therefore, application of micro-level CPMs to do planning level analyses is sometimes considered to be less appropriate. However, there are several studies that employed proactive approaches at the micro-level. Gathering expert knowledge (e.g. Donnell et al., 2002; Jamson et al., 2008; Kim et al., 2008; Pirdavani, Brijs, and Wets, 2010), carrying out road safety audits (e.g. *RSA*, 2004; Wang et al., 2011; Mitchell et al., 2012) and the use of proximal safety indicators (e.g. Archer, 2005; Cunto and Saccomanno, 2008; Young and Archer, 2009; Pirdavani et al., 2010) are different domains in application of micro-level proactive approaches. In Chapters 2 and 3 of this dissertation, micro-level proactive crash analysis is demonstrated by means of gathering expert knowledge and simulation respectively.

Limitations in employing micro-level CPMs to perform planning stage road safety evaluation, impose pursuing macro-level CPMs. Several efforts have been made to develop macro-level CPMs (e.g. in Pirdavani et al., in press; Levine et al., 1995a; Amoros et al., 2003; Hadayeghi et al., 2003, 2006, 2007, 2010a, 2010b; Noland and Oh, 2004; Noland and Quddus, 2004; Lovegrove and Sayed, 2006, 2007; Aguero-Valverde and Jovanis, 2006; Lovegrove and Litman, 2008; Quddus, 2008; Huang et al., 2010; Abdel-Aty et al., 2011a, 2011b) by associating crashes with a number of predictive variables which have macro-level characteristics, like exposure, network and socio-demographic variables. These models are generally utilized to deal with traffic safety at the planning level and in a proactive manner. In Chapters 4 and 5 of this dissertation, several macro-level CPMs are developed and further employed in Chapters 6 and 7 to evaluate traffic safety impacts of two travel demand management (TDM) strategies.

1.1.5. Travel Demand Management, Mobility and Crashes

TDM strategies are generally implemented to improve transportation systems' efficiency. Although the primary intention of introducing TDM strategies is to optimize transportation systems, their potential traffic safety impacts should not be ignored. These strategies target and change people's travel behavior at different levels. Changes in travel behavior can be manifested in different ways, such as a mode shift (e.g. using public transportation instead of cars, biking for short distance trips or carpooling), a travel time shift (e.g. avoiding traffic peak-hours by leaving home/the work place earlier or later), or travel demand reduction (e.g. teleworking). Employing strategies that cause such impacts can potentially improve or compromise traffic safety, depending on the type of road users. As described in the PIARC's road safety manual (*PIARC Road Safety Manual*, 2003), there are different crash reduction strategies, such as to control exposure, or to reduce crash and injury risk. Among the measures to control exposure, reducing travel demand is considered to be an effective measure.

Travel demand reduction can be achieved by adopting TDM strategies like pricing and regulation (e.g. fuel price increase), telecommunications (teleworking), and so on. Moreover, identifying the traffic safety impacts of TDM strategies and highlighting their benefits can also strengthen the implementation of those strategies among policy makers. What is carried out in the second part of this dissertation (i.e. Chapters 6 and 7), makes a contribution to the necessity of TDM strategies safety assessment.

A fairly large body of literature has been devoted to investigate the relationships between TDM strategies, exposure metrics and crash rates. In particular, several studies have demonstrated the effectiveness of adopting a fuel cost increase (e.g. in Pirdavani et al., in press; Sivak, 2009; Chi et al., 2010, 2011; Litman and Fitzroy, 2012) and teleworking scenarios (e.g. in Choo et al., 2005; Choo and Mokhtarian, 2007; Vu and Vandebona, 2007; Dissanayake and Morikawa, 2008) in reducing exposure. It is noteworthy to indicate that road safety impacts of different fuel cost increase scenarios were addressed in different studies (e.g. in Pirdavani et al., in press; Grabowski and Morrisey, 2004, 2006; Leigh and Geraghty, 2008; Sivak, 2009; Chi et al., 2010, 2011; Dhondt et al., 2012), while to the best of our knowledge road safety assessment of a teleworking scenario has not been studied before.

Traffic policy measures, TDM strategies or traffic forecasts for transportation projects are often adopted and analyzed at a macro-level rather than on an individual intersection or road segment. Moreover, TDM strategies aim at changing travel behavior of road users individually, however the impact of changed behavior on the transportation system or on road safety should be addressed collectively and at an aggregate level. Therefore, micro-level crash analysis becomes less appropriate to predict safety levels of transportation projects or TDM strategies (*WHO*, 2004; Lovegrove, 2005). When new transportation projects are proposed, macro-level safety impact evaluation is therefore required (*WHO*, 2004). This is to ensure that new projects do not have an opposing safety impact on the whole transportation network. As such, to successfully perform a planning level road safety evaluation, it is necessary to utilize macro-level CPMs. However, a drawback in application of macro-level crash analysis approaches is that they have a tendency to be data intensive.

Therefore, it is less straightforward to perform macro-level crash analysis where there is a lack of reliable information; either historical crash data, exposure data or road infrastructure data.

Reviewing the literature shows that policy makers and researchers are encouraged to consider road safety well in advance of the start of transportation projects (e.g. in ETSC, 1997, PIARC Road Safety Manual, 2003, WHO, 2004, Ex-post evaluation of the RSAP, 2009, ITE, 2009, TDM study, 2012). As stated by ETSC (ETSC, 1997), policy makers should be able to estimate the impact on road safety that results from adopting new projects or substantial changes to the existing infrastructure. Adopting proactive road safety actions is also addressed in PIARC's road safety manual (*PIARC Road Safety Manual*, 2003). Among various types of proactive actions, it is suggested to conduct impact studies before and during the development stage of transportation projects to evaluate their impacts on road safety (*PIARC Road Safety Manual*, 2003). In another before-and-after study (*ITE*, 2009) the necessity of providing better assessment methods was emphasized in order to let policy makers explicitly take into account road safety impacts of transportation projects and policies beforehand and in their decision-making process.

1.2. Aims and Research Contributions

Road safety has become a critical worldwide public policy issue. Therefore, it is essential for practitioners and policy makers to carry out road safety analysis. In this regard, good quality data are needed to increase awareness about the magnitude of road traffic casualties and to convince policy makers to take action. Moreover, effective road safety management programs require type of data that can be relied on their accuracy and completeness (Cherry et al., 2006; *WHO*, 2010). This can sometimes be unachievable in certain circumstances. For instance, in developing countries where there is a lack of reliable, accurate and timely crash data, coping with road safety problems becomes challenging. In such circumstances, a crash database that includes accident reporting and recording and an analysis system is essentially required to enable a reliable assessment of the road safety situation (*RSG*, 2003, *WHO*, 2010). On the contrary and in the case of most of the developed countries, numerous data collection systems have been set up to better understand transportation systems and the road safety situation. These data enable researchers and practitioners to gain a deeper understanding of crashes and to model complex policy questions, such as the impacts of TDM strategies on road safety.

The first contribution of this dissertation is set to provide a collection of different methodologies for safety prediction and evaluation, depending on the amount and type of data available. The first part of this dissertation (i.e. Chapters 2 and 3) is therefore devoted to methods which are appropriate in the context of absence or a lack of good quality data. These methods illustrate what can be achieved under difficult circumstances (i.e. when hardly any crash data are available). On the other hand, richness in possibilities for evaluation purposes that can be achieved in circumstances where a good data collection system is available, is illustrated in the second part of this dissertation (Chapters 4, 5, 6 and 7) by developing predictive models and applying them in road safety impact evaluation for a set of TDM scenarios. Discovering the positive impacts of TDM strategies can be useful not only for developed countries, but also for developing countries (Litman, 2004; Broaddus et al., 2009). Many developed countries are now implementing TDM strategies to create less automobile-dependent transportation systems. This is a great opportunity for developing countries to learn more about the advantages of adopting TDM measures and to avoid future problems by implementing such strategies before they become highly automobile dependent (Litman, 2004; Broaddus et al., 2009; VTPI, 2012). In the first part of this dissertation and in the first study, a methodology is developed in case of a lack of crash data. In fact, lack of reliable data is a well known problem in many developing countries. Yet, it is essential to evaluate potential road safety problems before they emerge. A proactive approach is developed in Chapter 2 by means of collecting expert knowledge in combination with a decision making technique (i.e. multi-criteria decision making). Another alternative and/or complementary approach to safety prediction, in the absence of crash data, is to simulate the more frequent occurrence of near-crashes by means of proximal safety indicators since these are believed to have an established relationship to crash occurrence (Archer, 2005; Young and Archer, 2009; Dijkstra et al., 2010). These indicators are defined as measures of crash proximity, based on the temporal and/or spatial measures that reflect the "closeness" of road-users (or their vehicles), in relation to projected points of collision. Chapter 3 of this dissertation is devoted to assess traffic safety of intersections by means of proximal safety indicators within a microscopic simulation framework.

The main objective of the second part of this dissertation (i.e. second main contribution of this dissertation) is set to develop macro-level CPMs so as to evaluate the road safety impacts of two TDM strategies (i.e. a fuel cost increase by 20% and a teleworking scenario in which 5% of the working population are assumed to telework, respectively in Chapters 6 and 7) at the planning level and within a proactive framework. The main motivation for selecting these two TDM strategies comes from the fact that they are well effective towards reducing travel demand and are easily implementable (VTPI, 2012). These two strategies can influence and reduce travel demand differently; by making vehicle travel more expensive (fuel cost increase) and reducing commute trip (promoting teleworking). Such strategies can be potential answers to the challenges in controlling environmental, traffic safety and energy issues that many of the developed and also developing countries are facing with (Witlox et al., 2009; VTPI, 2012).

To accomplish the evaluation task, detailed geo-coded historical crash data together with disaggregate exposure information, produced by an activity-based transportation model for Flanders (Bellemans et al., 2010), are adopted. The reason for using an activity-based transportation model is because it provides an adequate range of in-depth information about individuals' travel behavior to realistically simulate and evaluate TDM strategies. The main advantage of these models is that the impact of applying a TDM strategy will be accounted for, for each individual, throughout a decision making process. Activity-based models provide more reliable information since, unlike traditional models, TDM strategies are inherently accounted for in these models (Arentze and Timmermans, 2004; Bellemans et al., 2010).

1.3. Scope

The scope of this dissertation, which addresses different proactive approaches towards road safety impact assessment, is defined and further discussed in this section. The present research can be broken down into two parts. In the first part a proactive crash analysis is performed to identify and prioritize crash hotspots in the absence of crash data. This is carried out at a micro-level consisting of road sections and intersections. Furthermore, proximal safety indicators are utilized to evaluate road safety of intersections, by means of a microsimulation model. The second, and also major part of this thesis, consists of the development of CPMs and their application to evaluate the safety levels of two TDM strategies; namely a fuel cost increase and an increase in the level of teleworking.

1.3.1. Study Areas

The proposed procedure for crash hotspot identification and prioritization was developed based on Iranian road safety experts and was applied and illustrated on a sample of 20 road sections between the two main cities of Tehran and Semnan in Iran.

The microscopic crash analysis of intersections can be considered as a primarily theoretical study and, therefore, was not applied to any case area. Its main purpose was to investigate the potential of microscopic simulation models for safety analysis.

The study area in the second part of the research is the Dutch-speaking region in northern Belgium, Flanders.

1.3.2. Data

In the first part of this dissertation, data collection for establishing the predictive identification and prioritization model consists of two phases:

(1) finding out the relevant criteria to identify accident hotspots and

(2) to measure their importance.

To determine the relevant crash hotspot criteria, road safety experts were asked to select a set of hotspot-related criteria by means of a series of questionnaires. These criteria correspond to the ones that have been used in classical road safety studies, such as traffic, road characteristics and network variables. By analyzing the responses to the questionnaires, a strong agreement was found on five main criteria; namely geometric conditions, physical conditions, traffic conditions, specific locations and distance from population centers. For each of these main criteria, additional sub-criteria were identified. Detailed description of these criteria and sub-criteria can be found in Chapter 2 of this dissertation. In the second phase of the data collection and after having identified all the relevant criteria, the relative importance of each of the criteria and sub-criteria was determined by means of a pair-wise comparison approach. Knowing the weight of each criterion and sub-criterion enables us to make a quantitative comparison between the criteria and, therefore, using them in establishing a predictive model. This procedure was then applied on a sample of 20 road sections. Required information of these 20 road sections were collected based on field observations and for each of the given criteria.

In the second part of this dissertation, data collection consists of two major parts; required data for developing CPMs and the data to perform the TDM evaluation task. Since a macroscopic approach at traffic analysis zone (TAZ) level is conducted to develop the prediction models, these models are further referred to as zonal crash prediction models (ZCPMs). ZCPMs are considered to present the association between observed crashes in each zone and a set of predictor variables. To this end, an activity-based transportation model framework is applied to produce exposure metrics which will be used in prediction models. This activity-based model produces traffic demand by means of origin-destination (OD) matrices. These OD matrices include the number of trips for each traffic mode at different disaggregation levels (i.e. age, gender, day of the week, time of day and motive). This traffic demand is then assigned to the Flemish road network to obtain detailed exposure metrics at the network level. Exposure metrics are then geographically aggregated to the TAZ level. This has been carried out at the zonal level, comprising 2,200 TAZs in Flanders. In addition, for each TAZ a set of variables including socio-demographic and road network variables were collected for each TAZ to develop ZCPMs. The crash data used in this study consist of a geo-coded set of fatal and injury crashes that occurred during the period 2004 to 2007 and was provided by the Flemish Ministry of Mobility and Public Works. Detailed descriptive information of the exposure metrics, socio-demographic and road network variables are available in Chapter 3 of this dissertation. For TDM evaluation, the fuel cost increase and the teleworking scenarios are simulated within the activity-based model and corresponding travel demand measures are derived by means of OD matrices.

1.3.3. Methods

Except for Chapter 2 which relies on both qualitative and quantitative methodologies, the rest of this dissertation has a quantitative nature.

In Chapter 2 and for prioritizing crash hotspots in the absence of crash data, two main tasks should be performed. In the first task, relevant hotspot criteria and their relative importance are determined. To this end, an expert knowledge collecting method, namely the Delphi technique, is employed. The Delphi technique is defined by Linstone and Turoff (2002) as "*a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem*". The objective of most Delphi applications is the reliable and creative exploration of ideas or the production of suitable information for decision-making. Once the final criteria weights are obtained, they can be applied in a multiple criteria decision making (MCDM) context to develop the crash prioritization model. For this research, we have adopted the Technique for Order Preference Similarity to Ideal Solution (TOPSIS). The TOPSIS method was developed and improved by Yoon and Hwang (1995) and has been used in several recent group decision-making studies (Lin et al., 2008; Shih, 2008). This method is based on the concept that the chosen alternative should have the shortest Euclidean distance from the best possible solution, while simultaneously being furthest away from the worst possible solution.

Furthermore, an alternative proactive method is utilized to predict the safety levels by means of proximal safety indicators. These indicators which are known to have a near-crash attitude, have been suggested by many researchers to serve as alternative to the use of crash data (e.g. in Darzentas et al., 1980; Archer, 2000, 2005; Minderhoud and Bovy, 2001; Gettman and Head, 2003; Huguenin et al., 2005; El Faouzi et al., 2007; Dijkstra et al., 2010). In this study, a microsimulation model, called S-Paramics (SIAS, 2010) is applied to simulate different predefined scenarios. The output of this microsimulation is then manually analyzed to derive appropriate proximal safety indicators.

In the second part of this dissertation, different statistical models are constructed to present the association between observed crashes in each zone and a set of predictor variables. Given the observed overdispersion in the crash data, the Negative Binomial (NB) model was adopted to enable the modeling of overdispersed data. The NB model is one of the most commonly used models in crash data modeling (e.g. in Amoros et al., 2003; Hadayeghi et al., 2003, 2006, 2007; Noland and Oh, 2004; Noland and Quddus, 2004; De Guevara et al., 2004; Lovegrove, 2005; Aguero-Valverde and Jovanis, 2006; Lovegrove and Sayed, 2006, 2007; Lovegrove and Litman, 2008; Hadayeghi, 2009; Naderan and Shahi, 2010; Lord and Mannering, 2010; Abdel-Aty et al., 2011b; An et al., 2011; Pirdavani et al., 2012, in press). Therefore, different NB ZCPMs were constructed within the generalized linear modeling (GLM) framework.

Moreover, the results of the analysis indicate the necessity of considering spatial variation in developing CPMs. There are various spatial modeling techniques that have been applied in the crash prediction field (e.g. in Levine et al., 1995; Miaou et al., 2003; Flahaut, 2004; Aguero-Valverde and Jovanis, 2006, 2008; Wang and Abdel-Aty, 2006; Quddus, 2008; Wang et al., 2009; Guo et al., 2010; Huang et al., 2010; Siddiqui et al., 2012). One of the solutions for taking the spatial variation into account is developing a set of local models, so called geographically weighted regression (GWR) models (Fotheringham et al., 2002). These models rely on the calibration of multiple regression models for different geographical entities. The GWR technique can be adapted to the GLM models and form geographically weighted generalized linear models (GWGLMs) (Fotheringham et al., 2002). These models are able to model count data (such as the number of crashes) while simultaneously accounting for the spatial non-stationarity and are therefore particularly useful in the context of this study.

Model development and spatial analysis are carried out using the statistical software package R (R: A language and environment for statistical computing, 2011) and GWGLMs are developed using a SAS macro (Chen and Yang, 2012).

1.4. Outline

The present dissertation is a collection of several studies. The research contributions are addressed in the following chapters of this dissertation. Each chapter is based on a paper that has either been published by or submitted to an international peer-reviewed journal. Consequently, there are some overlaps since the original papers have parts of the literature review, data description and methodological framework in common. Although the aim was to use the same terms throughout all chapters, some differences might still remain. However, the dissertation is consistent in the use of numbering for chapters, sections, table headings, figure captions, body text style and reference style.

Chapter 2 aims to define the relevant criteria for identifying accident hotspots, then quantifying each criterion's importance in order to develop a model to prioritize accident hotspots. A Delphi method combined with a MCDM procedure was therefore adopted to prioritize accident hotspots based on several criteria that were considered by experts as relevant for the road safety problem, such as geometric characteristics, traffic conditions, physical conditions and location characteristics. More specifically, expert opinions about relevant hotspot identification criteria and their relative importance were obtained from a Delphi experiment. Subsequently, a MCDM procedure was adopted to prioritize road sections based on their performance on each of the selected criteria. This procedure is illustrated on a collection of 20 road sections and the model is validated against an existing database of road sections containing safe locations and hotspots. Finally, a sensitivity analysis is carried out on the proposed method.

In Chapter 3, a proactive safety evaluation procedure is presented to assess the traffic safety of unsignalized intersections under different traffic circumstances. This is carried out by using a microsimulation software (S-

Paramics) and application of proximal safety indicators. For this study, post-encroachment time (PET), as one of the most commonly used proximal safety indicators in the literature, is chosen (Archer, 2005). This measure is used to evaluate situations in which two road users that are on a collision course, pass over a common spatial point or area with a temporal difference that is below a predetermined threshold. Implementing PET as a safety indicator can provide useful information for evaluating safety conditions at unsignalized intersection within different scenarios based on several traffic volume and speed limit categories. This chapter demonstrates the practical merits and limitations of microsimulation and shows how changes in speed limits and also traffic volume will affect road safety.

In Chapter 4, several ZCPMs are constructed to associate the relationship between crash frequency and a number of explanatory variables at the macro-level while the main focus is to compare the impacts of application of different exposure measures on model performance. Number of trips (NOTs), as an exposure variable, does not contain information on trip time, trip length and route choice. Moreover, transit traffic can have a significant share of the exposure observed in a TAZ. This part of the exposure is left out when only using the NOTs. Thus, other exposure variables which are sensitive to the impacts of trip assignment should be taken into account in model construction. The developed models can be categorized into three different groups based on the type of exposure metric that was utilized, i.e.:

(1) flow-based models,

(2) trip-based models and

(3) models based on a combination of the two.

Flow-based models were constructed by associating the number of crashes (NOCs) in each TAZ with VHT or VKT by cars, as the exposure variables, and the road network and socio-demographic variables. Trip-based models use the same road network and socio-demographic variables but only NOTs as the exposure variable. In the third type of model, both flow-based and trip-based variables are included simultaneously as metrics of exposure.

Chapter 5 proceeds by addressing the phenomenon of spatial autocorrelation. We first aim to investigate the presence of spatial variation of dependent and different explanatory variables which are being used in developing CPMs. This is carried out by computing Moran's *I* statistics (Moran, 1950) for dependent and selected explanatory variables. Based on the results of the analysis that emphasizes the necessity of considering spatial correlation, the second objective of this chapter is to develop GWGLMs. These models are further compared with the models developed in Chapter 4 to verify the results of Chapter 4 regarding the selection of best fitted model.

The subsequent two chapters applied the developed models in Chapters 4 and 5 to evaluate two TDM scenarios; the fuel cost increase and the teleworking scenarios. OD matrices for the null and both the fuel cost increase and the teleworking scenarios are derived from the activity-based transportation model for scenario evaluation. After assigning the travel demand to the road network, all required variables become available to set up the evaluation task. The final models are applied and crashes are predicted for each TAZ and for each scenario. The traffic safety evaluation is then conducted by comparing the NOCs, predicted by the prediction models for the null and the TDM scenarios.

In Chapter 6, unlike some prior studies where reactive approaches were generally used to evaluate the road safety impacts of TDM strategies, a proactive methodology is employed to accomplish the evaluation task. This is carried out in an assessment exercise by assuming a 20% increase in fuel price. Chapter 7 demonstrates the impacts of implementing a teleworking scenario in Flanders, in which 5% of the working population will telework. It is noteworthy to mention that in Chapters 6 and 7, different models are fitted to better capture the different impacts of these scenarios on different road users and at different severity levels of crashes (i.e. car only crashes and crashes involving pedestrians and cyclists are modeled separately at two severity levels, namely "fatal and severe injury" and "slight injury" levels). In Chapter 6, prediction models are developed within the GLM framework whereas in Chapter 7, the GWGLM framework is developed to also incorporate the spatial variations in association between NOCs and the explanatory variables.

Chapter 8 summarizes all chapters and presents the overall conclusions derived from this thesis, as well as policy recommendations and further research extensions.

References

Abdel-Aty, M., Dilmore, J., Dhindsa, A., 2006. Evaluation of variable speed limits for real-time freeway safety improvement. Accident Analysis & Prevention 38, 335–345.

Abdel-Aty, M., Haleem, K., 2011. Analyzing angle crashes at unsignalized intersections using machine learning techniques. Accident Analysis & Prevention 43, 461–470.

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011a. Zonal Level Safety Evaluation Incorporating Trip Generation Effects. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011b. Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Abdel-Aty, M.A., Hassan, H.M., Ahmed, M., Al-Ghamdi, A.S., 2012. Real-time prediction of visibility related crashes. Transportation Research Part C: Emerging Technologies 24, 288–298.

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis & Prevention 38, 618–625.

Aguero-Valverde, J., Jovanis, P.P., 2008. Analysis of Road Crash Frequency with Spatial Models. Transportation Research Record: Journal of the Transportation Research Board 2061, 55–63.

Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Accident Analysis & Prevention 35, 537–547.

An, M., Casper, C., Wu, W., 2011. Using Travel Demand Model and Zonal Safety Planning Model for Safety Benefit Estimation in Project Evaluation. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Archer, J., 2005. Indicators for traffic safety assessment and prediction and their application in micro-simulation modelling: A study of urban and suburban intersections, PhD thesis, Royal Institute of Technology Stockholm, Sweden.

Arentze, T.A., Timmermans, H.J.P., 2004. ALBATROSS – Version 2.0 – A learning based transportation oriented simulation system. EIRASS (European Institute of Retailing and Services Studies).

Autey, J., Sayed, T., Zaki, M.H., 2012. Safety evaluation of right-turn smart channels using automated traffic conflict analysis. Accident Analysis & Prevention 45, 120–130.

Bell, M.C., Galatioto, F., Giuffrè, T., Tesoriere, G., 2012. Novel application of red-light runner proneness theory within traffic microsimulation to an actual signal junction. Accident Analysis & Prevention 46, 26–36.

Bellemans, T., Kochan, B., Janssens, D., Wets, G., Arentze, T., Timmermans, H., 2010. Implementation Framework and Development Trajectory of FEATHERS Activity-Based Simulation Platform. Transportation Research Record: Journal of the Transportation Research Board 2175, 111–119.

Broaddus, A., Litman, T., Menon, G., 2009. Travel Demand Management: Training Document. GTZ Transport Policy Advisory Services, Eschborn, Germany.

Caliendo, C., Guida, M., Parisi, A., 2007. A crash-prediction model for multilane roads. Accident Analysis & Prevention 39, 657–670.

Chen, V.Y.-J., Yang, T.-C., 2012. SAS macro programs for geographically weighted generalized linear modeling with spatial point data: Applications to health research. Computer Methods and Programs in Biomedicine 107, 262–273.

Cherry, E., Floyd, R., Graves, T., Martin, S., Ward, D., 2006. Crash Data Collection and Analysis System (Final Report No. FHWA-AZ-06-537). Arizona Department of Transportation.

Chi, G., Cosby, A.G., Quddus, M.A., Gilbert, P.A., Levinson, D., 2010. Gasoline prices and traffic safety in Mississippi. Journal of Safety Research 41, 493–500.

Chi, G., Zhou, X., McClure, T.E., Gilbert, P.A., Cosby, A.G., Zhang, L., Robertson, A.A., Levinson, D., 2011. Gasoline prices and their relationship to drunk-driving crashes. Accident Analysis & Prevention 43, 194–203.

Chin, H.-C., Quek, S.-T., 1997. Measurement of traffic conflicts. Safety Science 26, 169–185.

Chitturi, M.V., Ooms, A.W., Bill, A.R., Noyce, D.A., 2011. Injury outcomes and costs for cross-median and median barrier crashes. Journal of Safety Research 42, 87–92.

Choo, S., Mokhtarian, P.L., 2007. Telecommunications and travel demand and supply: Aggregate structural Equation models for the US. Transportation Research Part A: Policy and Practice 41, 4–18.

Choo, S., Mokhtarian, P.L., Salomon, I., 2005. Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the U.S. Transportation 32, 37–64.

Cunto, F., Saccomanno, F., 2007. Microlevel Traffic Simulation Method for Assessing Crash Potential at Intersections. Presented at the Transportation Research Board 86th Annual Meeting, Washington DC, United States.

Cunto, F., Saccomanno, F.F., 2008. Calibration and validation of simulated vehicle safety performance at signalized intersections. Accident Analysis & Prevention 40, 1171–1179.

Czech, S., Shakeshaft, A.P., Byrnes, J.M., Doran, C.M., 2010. Comparing the cost of alcohol-related traffic crashes in rural and urban environments. Accident Analysis & Prevention 42, 1195–1198.

Davis, G.A., Hourdos, J., Xiong, H., Chatterjee, I., 2011. Outline for a causal model of traffic conflicts and crashes. Accident Analysis & Prevention 43, 1907–1919.

De Guevara, F.L.D., Washington, S., Oh, J., 2004. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. Transportation Research Record: Journal of the Transportation Research Board 1897, 191–199.

de Leur, P., 2001. Improved approaches to manage road safety infrastructure, PhD thesis, University of British Columbia, Canada.

de Leur, P., Sayed, T., 2003. A framework to proactively consider road safety within the road planning process. Canadian Journal of Civil Engineering 30, 711–719.

de Rome, L., Ivers, R., Fitzharris, M., Haworth, N., Heritier, S., Richardson, D., 2011. Effectiveness of motorcycle protective clothing: Riders' health outcomes in the six months following a crash. Injury: international journal of the care of the injured, 1-11.

Dhondt, S., Pirdavani, A., Macharis, C., Bellemans, T., Putman, K., 2012. Translating Road Safety into Health Outcomes Using a Quantitative Impact Assessment Model. Injury Prevention, in press.

Dijkstra, A., Marchesini, P., Bijleveld, F., Kars, V., Drolenga, H., van Maarseveen, M., 2010. Do Calculated Conflicts in Microsimulation Model Predict Number of Crashes? Transportation Research Record: Journal of the Transportation Research Board 2147, 105–112.

Dissanayake, D., Morikawa, T., 2008. Impact assessment of satellite centre-based telecommuting on travel and air quality in developing countries by exploring the link between travel behaviour and urban form. Transportation Research Part A: Policy and Practice 42, 883–894.

Donnell, E., Harwood, D., Bauer, K., Mason, J., Pietrucha, M., 2002. Cross-Median Collisions on Pennsylvania Interstates and Expressways. Transportation Research Record: Journal of the Transportation Research Board 1784, 91–99.

ETSC, 1997. Road Safety Audit and Safety Impact Assessment. European Transport Safety Council, Brussels, Belgium.

Ex-post evaluation of the RSAP, 2009. The preparation of the European Road Safety Action Program 2011-2020. Transport and Mobility Leuven.

FHWA, 2006. U.S. Department of Transportation, FHWA Road Safety Audit Guidelines, final report No. FHWA-SA-06-06). Federal Highway Administration (FHWA).

Flahaut, B., 2004. Impact of infrastructure and local environment on road unsafety: Logistic modeling with spatial autocorrelation. Accident Analysis & Prevention 36, 1055–1066.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression the analysis of spatially varying relationships. John Wiley & Sons Ltd, West Sussex, England.

Golembiewski, G.A., Chandler, B., 2011. Road Safety Information Analysis: A Manual for Local Rural Road Owners (No. FHWA-SA-11-10). Federal Highway Administration (FHWA), Washington D.C.

Grabowski, D.C., Morrisey, M.A., 2004. Gasoline Prices and Motor Vehicle Fatalities. Journal of Policy Analysis and Management 23, 575–593.

Grabowski, D.C., Morrisey, M.A., 2006. Do higher gasoline taxes save lives? Economics Letters 90, 51–55.

Guo, F., Wang, X., Abdel-Aty, M., 2010. Modeling signalized intersection safety with corridor-level spatial correlations. Accident Analysis & Prevention 42, 84–92.

Hadayeghi, A., 2009. Use of Advanced Techniques to Estimate Zonal Level Safety Planning Models and Examine Their Temporal Transferability, PhD thesis, Department of Civil Engineering, University of Toronto.

Hadayeghi, A., Shalaby, A., Persaud, B., 2003. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. Transportation Research Record: Journal of the Transportation Research Board 1840, 87–95.

Hadayeghi, A., Shalaby, A., Persaud, B., 2007. Safety Prediction Models: Proactive Tool for Safety Evaluation in Urban Transportation Planning Applications. Transportation Research Record: Journal of the Transportation Research Board 2019, 225–236.

Hadayeghi, A., Shalaby, A., Persaud, B., 2010a. Development of Planning-Level Transportation Safety Models using Full Bayesian Semiparametric Additive Techniques. Journal of Transportation Safety & Security 2, 45–68.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010b. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. Accident Analysis & Prevention 42, 676–688.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accident Analysis & Prevention 38, 579–589.

Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-Level Crash Risk Analysis in Florida. Transportation Research Record: Journal of the Transportation Research Board 2148, 27–37.

ITE, 2009. Before-and-After Study; Technical Brief. Institute of Transportation Engineers Transportation Safety Council, Washington D.C., USA.

Jamson, S., Wardman, M., Batley, R., Carsten, O., 2008. Developing a driving Safety Index using a Delphi stated preference experiment. Accident Analysis & Prevention 40, 435–442.

Kam, B.H., 2003. A disaggregate approach to crash rate analysis. Accident Analysis & Prevention 35, 693–709.

Kim, T.-G., Donnell, E.T., Lee, D., 2008. Use of cultural consensus analysis to evaluate expert feedback of median safety. Accident Analysis & Prevention 40, 1458–1467.

Laureshyn, A., Svensson, Å., Hydén, C., 2010. Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation. Accident Analysis & Prevention 42, 1637–1646.

Leigh, J.P., Geraghty, E.M., 2008. High gasoline prices and mortality from motor vehicle crashes and air pollution. Journal of Occupational and Environmental Medicine 50, 249–254.

Levine, N., Kim, K.E., Nitz, L.H., 1995a. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. Accident Analysis & Prevention 27, 675–685.

Levine, N., Kim, K.E., Nitz, L.H., 1995b. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. Accident Analysis & Prevention 27, 663–674.

Lin, Y.-H., Lee, P.-C., Chang, T.-P., Ting, H.-I., 2008. Multi-attribute group decision making model under the condition of uncertain information. Automation in Construction 17, 792–797.

Linstone, H.A., Turoff, M., 2002. Delphi Method : Techniques and Applications (digital version). Upper Saddle River, NJ: Addison-Wesley.

Litman, T., 2004. Evaluating Mobility Management Traffic Safety Impacts. Victoria Transport Policy Institute, Victoria, BC, Canada.

Litman, T., Fitzroy, S., 2012. Safe Travels: Evaluating Mobility Management Traffic Safety Impacts. Victoria Transport Policy Institute.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44, 291–305.

Lovegrove, G., Sayed, T., 2007. Macrolevel Collision Prediction Models to Enhance Traditional Reactive Road Safety Improvement Programs.

Transportation Research Record: Journal of the Transportation Research Board 2019, 65–73.

Lovegrove, G.R., 2005. Community-Based, Macro-Level Collision Prediction Models, PhD thesis, University of British Columbia.

Lovegrove, G.R., Litman, T., 2008. Using Macro-Level Collision Prediction Models to Evaluate the Road Safety Effects of Mobility Management Strategies: New Empirical Tools to Promote Sustainable Development. Presented at the Transportation Research Board (TRB) 87th Annual Meeting, Washington D.C. USA.

Lovegrove, G.R., Sayed, T., 2006. Macro-level collision prediction models for evaluating neighbourhood traffic safety. Canadian Journal of Civil Engineering 33, 609–621.

Lu, G., Liu, M., Wang, Y., Yu, G., 2012. Quantifying the Severity of Traffic Conflict by Assuming Moving Elements as Rectangles at Intersection. Procedia - Social and Behavioral Sciences 43, 255–264.

Ma, J., Kockelman, K.M., Damien, P., 2008. A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. Accident Analysis & Prevention 40, 964–975.

Mabunda, M.M., Swart, L.-A., Seedat, M., 2008. Magnitude and categories of pedestrian fatalities in South Africa. Accident Analysis & Prevention 40, 586–593.

Meng, Q., Weng, J., 2011. Evaluation of rear-end crash risk at work zone using work zone traffic data. Accident Analysis & Prevention 43, 1291–1300.

Miaou, S.-P., Song, J.J., Mallick, B.K., 2003. Roadway Traffic Crash Mapping: A Space-Time Modeling Approach. Journal of Transportation and Statistics 6, 33–57.

Mitchell, R., Friswell, R., Mooren, L., 2012. Initial development of a practical safety audit tool to assess fleet safety management practices. Accident Analysis & Prevention 47, 102–118.

Moran, P. A. P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37 (1-2), 17–23.

Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: Introducing crash generation concept. Accident Analysis & Prevention 42, 339–346.

Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention 36, 525–532.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis & Prevention 36, 973–984.

OECD, 1997a. Road safety principles and models (IRRD NO. 888815). Organization for economic co-operation and development, Paris, France.

OECD, 1997b. Road safety principles and models: review of descriptive, predictive, risk and accident consequence models (OECD/GD(97)153). Organization for economic co-operation and development, Paris, France.

Oh, J., Kim, E., Kim, M., Choo, S., 2010. Development of conflict techniques for left-turn and cross-traffic at protected left-turn signalized intersections. Safety Science 48, 460–468.

PIARC Road Safety Manual, 2003. PIARC Technical Committee on Road Safety (C13), France.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., 2012. Developing Zonal Crash Prediction Models with a Focus on Application of Different Exposure Measures. Transportation Research Record: Journal of the Transportation Research Board, in press.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., in press Evaluating the road safety effects of a fuel cost increase measure by means of zonal crash prediction modeling. Accident Analysis & Prevention, in press, doi:10.1016/j.aap.2012.04.008.

Pirdavani, A., Brijs, T., Bellemans, T., Wets, G., 2010. Evaluation of Traffic Safety at Un-signalized Intersections Using Microsimulation: A Utilization of Proximal Safety Indicators. Advances in Transportation Studies an international Journal 22, 43–52.

Pirdavani, A., Brijs, T., Wets, G., 2010. A Multiple Criteria Decision Making Approach for Prioritizing Accident Hotspots in the Absence of Crash Data. Transport Reviews 30, 97–113.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. Accident Analysis & Prevention 40, 1486–1497.

R: A language and environment for statistical computing, 2011. R Development Core Team, Vienna, Austria.

RSA, 2004. Road Safety Audits: A Synthesis of Highway Practice. Transportation Research Board of the National Academies, Washington D.C.

RSG, 2003. Road Safety Guidelines for the Asian and Pacific Region: Guidelines for Decision Makers on Road Safety Policy. Asian Development Bank, Manila, Philippines.

Shih, H.-S., 2008. Incremental analysis for MCDM with an application to group TOPSIS. European Journal of Operational Research 186, 720–734.

SIAS, 2010. S-Paramics Microsimulation, SIAS Limited, Reference Manual S-Paramics, Edinburgh, UK.

Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis & Prevention 45, 382–391.

Sivak, M., 2009. Mechanisms Involved in the Recent Large Reductions in US Road Fatalities. Injury Prevention 15, 205–206.

TDM study, 2012. Integrating Transportation Demand Management into the Planning and Development Process. iCommute, San Diego, CA, USA.

Tiwari, G., Mohan, D., Fazio, J., 1998. Conflict analysis for prediction of fatal crash locations in mixed traffic streams. Accident Analysis & Prevention 30, 207–215.

Van den Bossche, F.A.M., 2006. Road Safety, Risk and Exposure in Belgium: an Econometric Approach, PhD thesis, Hasselt University, Belgium.

Vlahogianni, E.I., Yannis, G., Golias, J.C., 2012. Overview of critical risk factors in Power-Two-Wheeler safety. Accident Analysis & Prevention, 49, 12-22.

VTPI, 2012. VTPI. Online TDM Encyclopedia, Victoria Transport Policy Institute.

Vu, S.T., Vandebona, U., 2007. Telecommuting and Its Impacts on Vehicle-km Travelled. Presented at the International Congress on Modelling and Simulation, University of Canterbury, Christchurch, New Zealand.

Wang, C., Quddus, M.A., Ison, S.G., 2009. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. Accident Analysis & Prevention 41, 798–808.

Wang, X., Abdel-Aty, M., 2006. Temporal and spatial analyses of rear-end crashes at signalized intersections. Accident Analysis & Prevention 38, 1137–1150.

Wang, Y.G., Chen, K.M., Ci, Y.S., Hu, L.W., 2011. Safety performance audit for roadside and median barriers using freeway crash records: Case study in Jiangxi, China. Scientia Iranica 18, 1222–1230.

WHO, 2004. World report on road traffic injury prevention. World Health Organization (WHO), Geneva.

WHO, 2009. Global status report on road safety: time for action. World Health Organization (WHO), Geneva.

WHO, 2010. Data Systems: A Road Safety Manual for Decision-makers and Practitioners. World Health Organization (WHO), Geneva.

Witlox, F., Jourquin, B., Thomas, I., Verhetsel, A., Van Malderen, L., Vanoutrive, T., 2009. Assessing and developing initiatives of companies to control and reduce commuter traffic "ADICCT" (No. SD/ TM/02A). Belgian Science Policy, Brussels, Belgium.

Xu, C., Liu, P., Wang, W., Li, Z., 2012. Evaluation of the impacts of traffic states on crash risks on freeways. Accident Analysis & Prevention 47, 162–171.

Yau, K.K.W., 2004. Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong. Accident Analysis & Prevention 36, 333–340.

Yoon, K., Hwang, C.-L., 1995. Multiple Attribute Decision Making: An Introduction. SAGE, Thousand Oaks, CA.

Young, W., Archer, J., 2009. Traffic Microsimulation Modeling to Study a Traffic Signal Incident Reduction Function. Transportation Research Record: Journal of the Transportation Research Board 2103, 80–87.

Zaloshnja, E., Miller, T.R., 2004. Costs of large truck-involved crashes in the United States. Accident Analysis & Prevention 36, 801–808.

2. A Multiple Criteria Decision Making Approach for Prioritizing Accident Hotspots in the Absence of Crash Data

Pirdavani, A., et al., (2010), Transport Reviews 30 (1), pp 97–113, DOI: 10.1080/01441640903279345.

Pirdavani, A., et al., (2010), DVD Compendium of the 89th Transportation Research Board Conference, Washington D.C., USA.

2.1. ABSTRACT

In an efficient transportation system, traffic safety is an important issue and it is influenced by many factors. In a country like Iran, until now safety improvements are mainly concentrated on road engineering activities, without much attention to vehicle technology or driving behavior. One important aspect of road safety engineering activities is the so-called treatment of hotspots or dangerous accident locations. Until recently, accident hotspots were identified and remedied by the experts' personal judgments and a hand-full of statistics without taking into account other important factors such as geometric or traffic conditions of the road network. Moreover, there are several situations, like in developing countries, where there is a lack of reliable, accurate and timely crash data. This motivates us to pursue a methodology which can be used to analyze crashes in the context of absence or a lack of good quality data. This chapter, therefore, aims to establish a methodology, so as to identify and prioritize accident hotspots when crash data are not available. To this end, the "Delphi" method has been adopted and a prioritization model is produced by the use of a multiple criteria decision making (MCDM) method. The procedure is illustrated on a collection of 20 road sections in Iran. In addition, the model is validated against an existing database of road sections containing safe locations and hotspots. Finally, a sensitivity analysis is carried out on the proposed method. The results of analyses demonstrate the capabilities of application of structured qualitative research methods in the context of road safety.

2.2. Introduction

Because of Iran's degree of urbanization, its extended country surface, long distances between main cities and its privileged location at the crossroad of international trade routes, transportation and mainly traffic safety is one of the most important challenges for Iranian road planners and the general public. For instance, in the past few years the number of vehicles including motorcycles has increased dramatically from 6,380,600 in 2000 to 14,174,400 in 2005 (an increase of 122%) whereas the total length of main and local roads in Iran has increased 'only' from 173,240 km in 2000 to 181,900 km in 2005 (an increase of 5%) (Zekavat, 2006). In other words, the rate of road infrastructure development has not kept up with the rate of motorization causing severe problems of congestion and road safety.

With respect to road safety, the current situation in Iran is indeed among the worst worldwide. For instance, in 1995, around 10,000 road fatalities and more than 50,000 injuries were reported. Only 5 years later, more than 15,000 people died and more than 87,000 were injured in traffic accidents in Iran (Montazeri, 2004) and the problem is still increasing. Whereas most western countries experienced a continuous decline in the number of road fatalities during the past decades, the number of road accident fatalities and injuries in Iran has reached to respectively 30,000 and 285,000 in 2006 (*Iranian Police Statistics*, 2008). Given that the population of Iran was 66,360,000 in 2000 and around 70,000,000 in 2006 ("Globalis - an Interactive World Map," 2008), the rate of road accident fatalities increased from 22.6 to 42.9 per 100,000 inhabitants. With respect to this tremendous rate of traffic accident fatalities, it is strongly vital to establish safety programs to improve the safety performance of the Iranian road network. A well-known road safety action program is the treatment of so-called hotspots or dangerous accident locations (Cheng and Washington, 2005; Brijs et al., 2006, 2007; Elvik, 2008).

In this chapter we will introduce a method for identifying and prioritizing accident hotspots in order to improve traffic safety in Iran. However, our approach will be different from traditional hotspot identification and ranking procedures, which typically depend on the availability of detailed, valid and reliable accident data. Furthermore, some researchers have argued in favor of using existing accident prediction models (APM) constructed in other regions of the world and transferring them to the local context (Sawalha and Sayed, 2006). Although this looks like a promising alternative that needs more research attention in the future, it is far from clear to what extent APM's constructed in developed countries can be easily transferred to countries where the traffic, road infrastructure and road accidents context are strongly different. Moreover, some authors have argued that the current recalibration procedures presented in the traffic safety literature are incomplete and generally, it would seem worthwhile to carry out more research on traffic safety model transferability in order to compensate for its lack of scientific basis (Sawalha and Sayed, 2006; Turner et al., 2007). Alternative methods are, therefore, required to be used for safety assessment and prediction but which do not rely on traffic accident data. The approach presented in this chapter, therefore, relies on a combination of qualitative (Delphi method) and quantitative methods (MCDM) which do not depend on traffic accident data. More specifically, expert choice (from 40 experts) and decision making methods were surveyed and used to identify relevant indicators of accident occurrence, their relative importance and then a model was established to identify and prioritize accident hotspots for a collection of 20 road sections in Iran.

The structure of the chapter is as follows. Initially, we will elaborate on the adopted procedure of the Delphi method and the use of MCDM for prioritizing hotspots. For the purpose of illustration, the procedure is then applied to a collection of 20 road sections in Iran. The procedure is validated against a database of road sections containing some known hotspots and safe sections. Finally, a sensitivity analysis is carried out on the proposed method.

2.3. Methodological Development

2.3.1. The Delphi Method and its Application in Road Safety Studies

The Delphi technique is defined by Linstone and Turoff (Linstone and Turoff, 2002) as "a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem." The objective of most Delphi applications is the reliable and creative exploration of ideas or the production of suitable information for decision making. The Delphi Method is based on a structured process for collecting and distilling knowledge from a group of experts by means of a series of questionnaires interspersed with controlled opinion feedback (Linstone and Turoff, 2002). Consequently Delphi (Helmer, 1977) represents a useful communication device among a group of experts and thus facilitates the formation of a group judgment. This method is an exercise in group communication among a panel of geographically dispersed experts (Linstone and Turoff, 2002). The technique allows experts to deal systematically with a complex problem or task. The essence of the technique is fairly straightforward. It comprises a series of questionnaires sent either by mail or via computerized systems, to a pre-selected group of experts. These questionnaires are designed to elicit and develop individual responses to the problems posed and to enable the experts to refine their views as the group's work progress in accordance with the assigned task. The main point behind the Delphi method is to overcome the disadvantages of conventional committee action. According to Fowles (Fowles, 1978) anonymity, controlled feedback, and statistical response characterize Delphi. The group interaction in Delphi is anonymous in the sense that comments, forecasts, and the like are not identified as to their originator but are presented to the group in such a way as to suppress any identification. In the original Delphi process, the key elements were:

(1) structuring of information flow,

(2) feedback to the participants, and

(3) anonymity for the participants.

Clearly, these characteristics may offer distinct advantages over the conventional face-to-face conference as a communication tool.

The Delphi method has been adopted previously in road safety research. For instance, in order to quantify the potential safety impacts of new in-vehicle technologies (Aittoniemi, 2008) and the prospective assessment of expected safety effects of new road transport informatics (Hydén, 1993). A Delphi-type survey instrument was also used in (Donnell et al., 2002) to gather expert knowledge on median safety issues that provided the impetus for field data collection in order to develop predictive models of cross-median collision crashes. In a similar study (Kim et al., 2008) a cultural consensus methodology was presented and applied to a set of median design and safety survey data that were collected using the Delphi method. To this end, a total of 21 Delphi survey participants were asked to answer research questions related to cross-median crashes. In another study (Jamson et al., 2008), a Delphi study was undertaken to use expert judgment as a way of deriving a first approximation of safety thresholds, i.e. the point at which behavior can be considered unsafe. The aim of the study was to understand the relative weightings that are assigned to a number of driver behaviors and thereby to construct a Safety Index. Finally, in (Pulkkinen and Holmberg, 1997; Simola and Virolainen, 2000; Cojazzi G. et al., 2001), the Delphi method was proposed for several kinds of probabilistic safety assessments, including road safety.

2.3.2. Determining the Set and Importance of Decision Criteria

The Delphi Method suggests 10 to 15 experts, but up to 100 expert opinions can be used for answering the questionnaires (Asgharpoor, 2003). This study was carried out in 2007 and includes the judgments of 40 experts who were selected carefully according to their professional background. All the experts have several years of related experience in the field of traffic safety and all of them have been working in safety departments of the Iranian Ministry of Road and Transportation, the Tehran Municipality or as a university professor. Initially, a general explanation about the problem, the procedure and the purpose of this study was presented to all of the experts separately. They were told how to fill out the set of questionnaires and how to provide their opinions during the Delphi procedure.

The Delphi procedure in this research mainly contains two phases to establish the basic structure of a predictive identification and prioritization model; 1) finding out the relevant criteria to identifying accident hotspots and 2) to measure their importance.

1) Determining Relevant Hotspot Criteria

At this stage, the experts were offered a set of 7 hotspot related criteria and an explanation for each of them in alphabetical order. These criteria correspond to the ones that have been used in classical traffic accident studies. At this stage, experts were asked to choose any criteria out of those 7 criteria and mention any other relevant criteria which they think are important and correlate to the purpose of this research. By analyzing the responses on the first questionnaire, 5 more criteria were mentioned by some experts and consequently added to the first set of criteria. Again, experts were required to choose the most important and relevant criteria out of the new set of 12 criteria. This is one of the advantages of the Delphi Method because it enables experts to improve and modify their opinions in each round with respect to the feedback of other experts in the previous round. Having surveyed the experts' opinions, Table 2.1 shows the response to each criterion and a rank order of perceived importance for each criterion after the second round of survey. Table 2.1 also shows that a strong agreement can be found on the first 5 criteria. This led us to continue our research only with these first 5 criteria, boldfaced in Table 2.1.

Table 2.1 Response to Each Criterion

Criterion	**Number of Responses**	**Criteria's Rank**
Geometric conditions	**32**	**1**
Physical conditions	**25**	**2**
Specific locations	**25**	**3**
Traffic conditions	**24**	**4**
Distance from population centers	**23**	**5**
Weather conditions	12	6
Maintenance costs	12	7
Accidents costs	10	8
Period of the day (day or night)	10	9
Rate of public transportation vehicles	10	10
Topography	9	11
Type of road	7	12

Below follows a short description of these 5 criteria:

- Geometric conditions: This criterion contains curves, slope, width of road and other related parts of geometric characteristics of it.
- Physical conditions: This criterion contains pavement, drainage and other related parts of physical characteristics of a road.
- Traffic conditions: This criterion is related to traffic characteristics of a road like traffic volume, traffic composition and traffic direction.
- Specific locations: This criterion contains some special sections of a road, like roundabouts, tunnels, bridges, etc. The reason for using this criterion is that a speed related safety problem may occur when drivers are approaching these specific sections. The driving maneuver common to all of these situations is deceleration, and it becomes a problem when visual cues induce drivers to underestimate their speed and thus fail to decelerate to an appropriate speed (Charlton and O'Brien, 2002).
- Distance from population centers: In Iran a large amount of accidents occurs near cities, because of combination of more types of road users with different ranges of speed in addition to the fact that transitions between high speed roads and low speed roads are poorly designed, if not completely absent (Golshan, 2005). In general the rural-urban threshold has consistently been identified as a speeding hotspot. Many motorists appear to find it difficult to slow down from open road speeds to a slower speed when entering an urban area. Commonly, the motorist is faced with a situation in which they are required decelerate to a low level of speed after having traveled at a high level of speed for an extended period. In this situation, the motorist tends to underestimate the speed at which they are traveling and, as a result, they find themselves driving too fast (Charlton and O'Brien, 2002). As a conclusion, sections which are located near population centers or rural-urban thresholds need special attention.

According to some experts' suggestions, some of these criteria are more common than the others and they need to be considered in more detail such as geometric, physical and traffic characteristics. So in this step of the procedure, the experts were asked to give their opinion on several sub-criteria for each of the main criteria above. Using the same procedure as for the main criteria, the experts came up with the following set of sub-criteria for each main criterion:

- Geometric conditions (A)
  - Section located in sub-standard horizontal curve (A1)
  - Section located in sub-standard vertical curve (A2)
  - Section located in steep slope (A3)
  - Section located in narrow width of road (A4)
  - Poor visibility (based on the stopping sight distance) (A5)
- Traffic conditions (B)
  - Traffic volume (B1)
  - Traffic composition (% of heavy vehicles) (B2)
  - Traffic direction (one-way or two-way) (B3)
- Physical conditions (C)
  - Poor pavement conditions (C1)
  - Poor drainage conditions (C2)
  - Poor road marking conditions (C3)
  - Poor road signing conditions (C4)
- Specific locations (D)
- Distance from population centers (E)

2) Determining the Importance of the Selected Criteria

After having identified all the relevant criteria, the main purpose is to determine the relative importance of each of the criteria and sub-criteria. Indeed, it is important to know the weight of each criterion and sub-criterion in order to make a quantitative comparison between the criteria to be used in establishing a predictive model.

In literature, several approaches have been proposed to determine weights (Saaty, 1980; Hwang and Yoon, 1981; Hwang and Lin, 1986). The majority of them can be classified into either subjective approaches or objective approaches depending on the information provided. The objective approaches determine weights based on objective information (i.e. a decision matrix) and these weights may be different from one decision matrix to another. In other words, weights which are calculated from two decision matrices with the same criteria but different alternatives will be different (not unique). The subjective approaches select weights based on preference information of criteria given by the decision makers (DM). Amongst others, they include the eigenvector method (Saaty, 1977), the weighted least square method (Chu et al., 1979), and the Delphi method (Hwang and Lin, 1986). This research follows a subjective approach because the purpose of this study is to make one unique weight vector to be used in a comprehensive prediction model. The most important advantage of this unique weight vector is that it can be used for each set of alternatives (sections of a road) to obtain valid results for identifying or prioritizing in the same situation. To do this, a pair-wise comparison matrix using a scale of relative importance should be constructed. More specifically, the fundamental scale of the analytic hierarchy process was used (Saaty, 2000). A criterion which is compared with itself, is always assigned the value 1, so the main diagonal entries of the pair-wise comparison matrix are all 1. The numbers 3, 5, 7, and 9 correspond to the verbal judgments "moderate importance", "strong importance", "very strong importance", and "absolute importance" (with 2, 4, 6, and 8 for compromise between these values).

Table 2.2 Pair-wise Comparison between Criteria

$$W=\begin{array}{c|ccccc} & x_1 & x_2 & x_3 & \cdots & x_n \\ \hline x_1 & 1 & a_{12} & a_{13} & \cdots & a_{1j} \\ x_2 & a_{21} & 1 & a_{23} & \cdots & a_{2j} \\ x_3 & a_{31} & a_{32} & 1 & \cdots & a_{3j} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ x_n & a_{i1} & a_{i2} & a_{i3} & \cdots & 1 \end{array} = \left|\begin{array}{ccccc} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \frac{w_1}{w_3} & \cdots & \frac{w_1}{w_j} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \frac{w_2}{w_3} & \cdots & \frac{w_2}{w_j} \\ \frac{w_3}{w_1} & \frac{w_3}{w_2} & \frac{w_3}{w_3} & \cdots & \frac{w_3}{w_j} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{w_i}{w_1} & \frac{w_i}{w_2} & \frac{w_i}{w_3} & \cdots & \frac{w_i}{w_j} \end{array}\right|$$

At this stage, the experts are asked to fill out 4 pair-wise comparison matrices (with respect to the main criteria and the 3 sets of sub-criteria). We use the following notation:

w_i = weight for criterion i, i=1,..., n where n = number of criteria

a_{ij} = w_i / w_j = the result of a pair-wise comparison between criterion *i* as compared to criterion *j*

W = matrix of pair-wise comparison values, a_{ij}

A set of pair-wise comparisons can thus be represented as Table 2.2

Once pair-wise comparisons have been elicited from the experts' decisions, the next step is to use these matrices to estimate the underlying scale of preferences. Several methods have been proposed in literature to estimate weights from matrices of pair-wise comparisons. The two most common methods for deriving criterion weights are the eigenvector and the logarithmic least squares methods (Saaty, 1977). In fact, it can be shown by algebraic manipulations of the pair-wise definitions that criterion weights can be obtained by finding the eigenvector corresponding to the largest eigenvalue of the W matrix. The eigenvector method was originally proposed by Saaty (Saaty, 1977) and is one of the most popular methods for calculating preferences from inconsistent matrices of pair-wise comparisons. Inconsistency occurs where the pair-wise comparison matrix does not satisfy transitivity for all pair-wise comparisons. Also Saaty's method allows inconsistency, but provides a measure of the inconsistency in each set of judgments. The consistency of the judgmental matrix can be determined by a measure called the consistency ratio (CR). This measure depends on the number of criteria, the maximum eigenvalue of a pair-wise comparison matrix and a putative value which is called the "Random Index". In general, a consistency ratio of 0.1 or less is considered acceptable; this threshold is 0.08 for matrices of size four and 0.05 for matrices of size three. If the value is higher, the judgments may not be reliable and should be elicited again (Saaty, 2000). The CR calculated measures for the 4 matrices of this research are all in acceptable ranges.

The special structure of a square reciprocal matrix means that the eigenvectors can be found and the largest eigenvector can be normalized to form a vector of relative weights (Fichtner, 1986). The weight vector is shown in Equation (1) which contains each criterion's weight value.

Weight vector = $w^t = \{w_1, w_2, ..., w_n\}$ (1)

The normalized weight values for each criterion are shown in Table 2.3.

Table 2.3 Normalized Weight Values for Each Criterion

Normalized Weight Values for the Main criteria					
criterion	(A)	(B)	(C)	(D)	(E)
weight	0.2465	0.2007	0.1989	0.1925	0.1614
Normalized Weight Values for Geometric conditions					
criterion	(A_1)	(A_2)	(A_3)	(A_4)	(A_5)
weight	0.1969	0.1805	0.1678	0.1914	0.2634
Normalized Weight Values for Traffic conditions					
criterion	(B_1)		(B_2)		(B_3)
weight	0.2916		0.2998		0.4086
Normalized Weight Values for Physical conditions					
criterion	(C_1)	(C_2)	(C_3)	(C_4)	
weight	0.2193	0.1818	0.2641	0.3348	

Table 2.4 Final Unique Weight Matrix

Criteria	**Final Weights**
Specific locations (D)	0.1925
Distance from Population Centers (E)	0.1614
Section located in sub-standard horizontal curve (A_1)	0.0486
Section located in sub-standard vertical curve (A_2)	0.0445
Section located in steep slope (A_3)	0.0414
Section located in narrow width of road (A_4)	0.0472
Poor Visibility (A_5)	0.0649
Traffic Volume (B_1)	0.0585
Traffic composition (B_2)	0.0602
Traffic Direction (B_3)	0.0820
Poor Pavement Conditions (C_1)	0.0436
Poor Drainage Conditions (C_2)	0.0362
Poor Road Marking Conditions (C_3)	0.0525
Poor Road Signing Conditions (C_4)	0.0666

According to the assumptions with the structure of the prioritization model, each criterion and sub-criterion must be in the same level (Saaty, 1980). In order to satisfy this condition, the weights of the sub-criteria must be multiplied by the weight of the corresponding main criterion, e.g. for poor visibility ($A*A_5$=0.2465*0.2364=0.0649). As a result, the final weight vector contains 14 criteria weights (Table 2.4).

2.3.3. Establishing the Accident Hotspots Prioritization Model

Once the final criteria weights are obtained, they can be applied in a multiple criteria decision context to rank a set of alternatives for which performance measures on the different criteria are known. More precisely, MCDM refers to making decisions in the presence of multiple, usually conflicting criteria. For this research, we have adopted the Technique for Order Preference Similarity to Ideal Solution (TOPSIS). The TOPSIS method was developed and improved by Hwang and Yoon (Yoon and Hwang, 1995) and has been used in several recent group decision making studies (Lin et al., 2008; Shih, 2008). This method is based on the concept that the chosen alternative should have the shortest Euclidean distance from the positive ideal solution whilst simultaneously being furthest away from the negative ideal solution. The ideal solution is a hypothetical solution for which all criterion values correspond to the maximum criterion values in the database comprising the satisfying solutions. The negative ideal solution is the hypothetical solution for which all criterion values correspond to the minimum criterion values in the database. TOPSIS thus gives a solution that is not only closest to the hypothetically best but that is also the furthest away from the hypothetically worst. The main procedure of the TOPSIS method for the selection of the best alternative from among those available is described below:

Step1: Preparing a Decision Matrix

In order to obtain the performance of a set of alternatives on a given set of criteria, a decision table or matrix is constructed consisting of (a) alternatives A_i *(for* $i = 1, 2, \ldots , n$*)*, (b) criteria B_j *(for* $j = 1, 2, \ldots , m$*)*, and (c) measures of performance M_{ij} *(for* $i= 1, 2, \ldots, n; j=1, 2, \ldots, m$*)* of the alternatives with respect to the criteria (Table 2.5). Given the decision matrix information and a decision-making method, the task of the decision maker is to find the best alternative and/or to rank the entire set of alternatives (Venkata Rao, 2007).

Table 2.5 Decision Matrix

$$D = \begin{array}{c|ccccc} & B_1 & B_2 & B_3 & \cdots & B_m \\ \hline A_1 & M_{11} & M_{12} & M_{13} & \cdots & M_{1m} \\ A_2 & M_{21} & M_{22} & M_{23} & \cdots & M_{2m} \\ A_3 & M_{31} & M_{32} & M_{33} & \cdots & M_{3m} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ A_n & M_{n1} & M_{n2} & M_{n3} & \cdots & M_{nm} \end{array}$$

Step2: Normalizing the Decision Matrix

It should be mentioned that all the elements in the decision matrix must be normalized to the same units, so that all possible criteria in the decision problem can be considered. Conversion of the decision making matrix to a dimensionless matrix is done by using Equation (2) (Saaty, 1980; Venkata Rao, 2007).

$$N_{ij} = \frac{M_{ij}}{\sqrt{\sum_{i=1}^{n} M_{ij}^{2}}} \tag{2}$$

Step3: Obtaining the Weighted Normalized Decision Matrix

Now the weighted normalized decision matrix can be obtained by multiplying the normalized decision matrix N_{ij} with the weight matrix W_j. W_j is a diagonal matrix such that all values are zero except for the major diagonal components which are the weight values of each criterion. Hence, the elements of the weighted normalized matrix V_{ij} are expressed as:

$$V_{ij} = N_{ij} \, . \, W_j = \begin{bmatrix} v_{11} & \cdots & v_{1j} & v_{1m} \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ v_{n1} & \cdots & v_{nj} & v_{nm} \end{bmatrix} \tag{3}$$

Step4: Defining the Positive and Negative Ideal Solutions

The positive ideal (best) and the negative ideal (worst) solutions can now be calculated from the weighted normalized decision matrix using Equation (4) and Equation (5) (Venkata Rao, 2007).

$$\text{Positive ideal solutions} = V^{+} = \left\{ \left(\max v_{ij} \middle| j \in J\right), \left(\min v_{ij} \middle| j \in J'\right) \middle| \; i = 1,2,\ldots,n \right\} = \left\{ v_1^{+}, v_2^{+}, \ldots, v_j^{+}, \ldots, v_m^{+} \right\} \quad (4)$$

$$\text{Negative ideal solutions} = V^{-} = \left\{ \left(\min v_{ij} \middle| j \in J\right), \left(\max v_{ij} \middle| j \in J'\right) \middle| \; i = 1,2,\ldots,n \right\} = \left\{ v_1^{-}, v_2^{-}, \ldots, v_j^{-}, \ldots, v_m^{-} \right\} \quad (5)$$

J and J' stand for subsets of beneficial and non-beneficial criteria, respectively. v_j^+ indicates the positive ideal (best) value of the considered criterion among the values of the criterion for different alternatives. In the case of beneficial criteria (i.e., those for which higher values are desirable for the given application), v_j^+ indicates the higher value of the criterion. In the case of non-beneficial criteria (i.e., those for which lower values are desired for the given application), v_j^+ indicates the lower value of the criterion. v_j^- indicates the negative ideal (worst) value of the considered criterion among the values of the criterion for different alternatives. In the case of beneficial criteria (i.e., those of which higher values are desirable for the given application), v_j^- indicates the lower value of the criterion. In the case of non-beneficial criteria (i.e., those of which lower values are desired for the given application), v_j^- indicates the higher value of the criterion (Venkata Rao, 2007).

Step5: Obtaining the Separation Measures and the relative proximity index (RPI)

The separation of each alternative from the positive ideal and negative ideal solution is given by the Euclidean distance in the Equation (6) and Equation (7) (Venkata Rao, 2007):

$$s_{i+} = \textit{separation from positive ideal solution} = \left\{ \sum_{j=1}^{m} \left(v_{ij} - v_j^{+}\right)^2 \right\}^{0.5}, \quad i = 1,2,\ldots,n \quad (6)$$

$$s_{i-} = separation\ from\ negative\ ideal\ solution = \left\{ \sum_{j=1}^{m} \left(v_{ij} - v_j^- \right)^2 \right\}^{0.5}, \quad i = 1,2,\ldots,n \tag{7}$$

The relative closeness or proximity of a particular alternative to the ideal solution (also called the RPI), P_i, can then be expressed as Equation (8):

$$P_i = \frac{s_{i-}}{s_{i-} + s_{i+}}, \quad 0 \le P_i \le 1, \quad i = 1,2,\ldots,n \tag{8}$$

Step6: Prioritizing Alternatives

According to this proximity value P_i, the set of alternatives can be ranked from the most preferred to the least preferred feasible solutions. P_i may also be called the overall or composite performance score of alternative, A_i.

2.4. Case Study

In what follows, the above described procedure is applied and illustrated on a sample of 20 road sections between the two main cities of Tehran and Semnan in Iran. Firstly, a decision matrix for these 20 road sections was created based on field observations indicating the performance of each section on each of the given criteria (See Table 2.6). To be able to make a decision matrix for a set of road sections, it is needed to answer the questions below and then fill out the database for each alternative (road section). Then it will be easily possible to fill out the decision matrix out of those databases.

- (A) Geometric Conditions
  - (A_1) Section located in sub-standard horizontal curve? If yes put 1, if not 0.
  - (A_2) Section located in sub-standard vertical curve? If yes put 1, if not 0.
  - (A_3) Section located in steep slope? If yes put 1, if not 0.
  - (A_4) Section located in narrow width of road? If yes put 1, if not 0.
  - (A_5) Section has a smaller stopping sight distance compared with standards? If yes put 1, if not 0.

Table 2.6 Initial Decision Matrix of 20 Road Sections

	D	E	A_1	A_2	A_3	A_4	A_5	B_1	B_2	B_3	C_1	C_2	C_3	C_4
1	0	0.33333	1	1	1	0	1	45000	35	0	0	0	1	1
2	0	0.06250	1	0	1	0	0	45000	35	0	0	0	0	1
3	0	0.09091	1	0	0	0	0	45000	35	0	0	0	1	1
4	1	0.16667	0	0	0	0	0	45000	35	0	1	1	0	1
5	1	0.50000	0	0	0	1	0	45000	35	0	0	1	0	1
6	0	0.06250	1	0	0	0	0	45000	35	0	0	1	0	0
7	1	0.16667	0	0	0	0	0	25000	30	0	0	1	1	1
8	1	0.01923	1	0	0	1	1	25000	30	0	0	1	0	1
9	1	0.03846	0	0	0	1	0	25000	30	0	1	1	0	1
10	0	0.06250	1	0	0	1	0	25000	30	0	0	0	0	1
11	0	0.33333	1	0	0	0	1	25000	30	0	0	1	0	0
12	1	0.20000	0	0	0	1	1	15000	25	0	1	0	1	1
13	1	0.16667	0	1	0	1	1	15000	25	0	1	1	0	1
14	1	0.09091	1	0	0	1	1	15000	25	0	1	1	1	0
15	0	0.06250	1	0	0	1	1	15000	25	0	1	0	1	0
16	0	0.04348	1	1	0	1	1	15000	25	0	1	1	1	1
17	1	0.02564	1	0	0	0	1	15000	25	0	0	0	1	0
18	0	0.03226	1	1	0	1	1	15000	25	0	1	1	1	1
19	0	0.06250	1	0	1	1	1	15000	25	0	1	1	1	0
20	0	0.16667	1	0	0	0	0	15000	25	0	0	0	0	1

- (B) Traffic Conditions
  - (B_1) How much is the traffic volume, Average Annual Daily Traffic (AADT)?
  - (B_2) Ratio of heavy vehicles to all vehicles (%)?
  - (B_3) what about traffic direction? If one-way fill out 0, if two-way put 1.
- (C) Physical Condition
  - (C_1) Section located in poor pavement conditions? If yes put 1, if not 0.
  - (C_2) Section located in poor drainage conditions? If yes put 1, if not 0.
  - (C_3) Section located in poor road Marking conditions? If yes put 1, if not 0.
  - (C_4) Section located in poor road Signing conditions? If yes put 1, if not 0.
- (D) Section located in specific places? If yes put 1, if not 0.
- (E) How much is the distance from population centers (Km)? Fill out with the "E" measure using Equation (9); where "d" represents the distance from population centers.

$$E = \frac{1}{1 + 2 \times d} \qquad (9)$$

Distribution of number of occurred traffic accidents against distance from population centers has the best correlation with Equation (9) (Golshan, 2005).

The decision matrix is then normalized according to Equation (2) and weighted by multiplying it by the weight matrix defined in Section 2. Subsequently, the positive and negative ideal solution measures (Table 2.7) are calculated using Equation (4) and Equation (5). In accordance with the principals of the TOPSIS method, along with the purpose of this research which is prioritizing traffic accident hotspots, the positive ideal solutions signify the most dangerous situation and the negative ideal solutions imply the safest condition with respect to each criterion.

Now the separation of each alternative from the positive ideal and negative ideal solution can be calculated along with the relative proximity to the ideal solutions according to the Equation (6), Equation (7) and Equation (8). These measures are shown in columns 2-4 of Table 2.8. The priority of each road section for treatment can now be obtained by ranking the measures of relative proximity in descending order. In other words, the road section ranked first (i.e. with the highest measure of relative proximity) is the most dangerous section. The ranks are shown in column 5 of Table 2.8.

Table 2.7 Measures of Positive Ideal and Negative Ideal Solutions

Criterion	Positive ideal solutions	Negative ideal solutions
(A_1)	0.0135	0.0000
(A_2)	0.0257	0.0000
(A_3)	0.0293	0.0000
(A_4)	0.0142	0.0000
(A_5)	0.0196	0.0000
(B_1)	0.0200	0.0067
(B_2)	0.0159	0.0114
(B_3)	0.0000	0.0000
(C_1)	0.0154	0.0000
(C_2)	0.0105	0.0000
(C_3)	0.0166	0.0000
(C_4)	0.0178	0.0000
(D)	0.0990	0.0038
(E)	0.0642	0.0000

2.5. Validation and Sensitivity Analysis

2.5.1. Validation

In order to validate the hotspot identification procedure presented in this chapter, the results of the proposed procedure are correlated with a database containing 100 road sections from National Road Tehran-Semnan (210 kilometers) in Iran. Forty four road sections on this road were identified as hotspots by the road administration in Iran according to a procedure based on historical crash data (3-years accident database) whereas the others were considered as safe sections. These fifty six safe sections were selected randomly among hundreds of other safe sections on this National Road to balance the validation database and results into a final set of 100 sections for validation.

Table 2.8 Separation from Positive and Negative Ideal Solutions and Relative Proximity

Section	Separation from positive ideal solution	Separation from negative ideal solution	Relative Proximity	Rank
1	0.0409	0.0979	0.7053	1
2	0.0714	0.0422	0.3718	16
3	0.0686	0.0367	0.3483	17
4	0.0516	0.0762	0.5962	6
5	0.0560	0.1200	0.6817	2
6	0.0732	0.0268	0.2681	20
7	0.0502	0.0765	0.6038	5
8	0.0703	0.0735	0.5113	9
9	0.0678	0.0714	0.5131	8
10	0.0720	0.0299	0.2932	19
11	0.0526	0.0686	0.5658	7
12	0.0449	0.0833	0.6497	3
13	0.0473	0.0836	0.6388	4
14	0.0671	0.0407	0.3777	15
15	0.0711	0.0375	0.3452	18
16	0.0719	0.0493	0.4068	12
17	0.0700	0.0706	0.5022	10
18	0.0736	0.0491	0.4001	13
19	0.0694	0.0487	0.4124	11
20	0.0603	0.0382	0.388	14

The method used to identify different road sections as "accident hotspot" or "safe", is known as the "equivalent property damage only index (EPDO index)". This method is described in the PIARC Road Safety Manual (*PIARC Road Safety Manual*, 2003). The EPDO index attaches a greater importance to more serious trauma by ascribing to each crash a weight that is a function of the worst level of injury sustained by one of the accident victims. Based on the average EPDO value of all locations in the reference population, a critical value – usually two times of the mean EPDO in the reference population – was determined and a validation dataset was constructed with safe road sections (having an EPDO value below the critical value) and hotspots (having an EPDO value above the critical value). Some descriptive statistics for the validation dataset are given in Table 2.9.

Table 2.9 Summary Statistics for Validation Data; Separately for Hotspots and Safe Locations

Safe locations	**minimum**	**maximum**	**mean**	**standard deviation**
Number of crashes	0	14	1.57895	2.98177
AADT	15000	45000	30052.63	9835.11
Crash risk	0	0.00033	5.424E-05	9.101E-05
RP index	0.08791	0.34225	0.1694	0.03947
Hotspots	**minimum**	**maximum**	**mean**	**standard deviation**
Number of crashes	6	52	19.37778	11.03154
AADT	15000	45000	26600	9128.88
Crash risk	0.00018	0.00195	7.533 E-04	3.762 E-04
RP index	0.20572	0.75559	0.36004	0.11104

Subsequently, the RPI was calculated for each road section in the database based on the performance of each section on the decision criteria presented in this study. The measures for establishing this decision matrix were observed in the field and some of them, like traffic condition criteria, were derived from a present traffic database. Comparing the RPI values with known information about the safety of each section enables us to validate our approach and to identify a boundary of RPI values that discriminates safe from unsafe sections. Figure 2.1 shows the distribution of RPI values across safe and unsafe road sections in the database.

Figure 2.1 clearly shows that RPI values are much higher (on average a value of 0.36) for hotspots than for safe road sections (on average a value of 0.17). This means that the prioritization model presented in this chapter can be considered as a practical predictive identification model to identify hotspots when no or insufficient crash data are available. Furthermore, a Spearman's rank correlation analysis was carried out on the validation data to investigate the strength of the correlation between the ranks produced by the RPI index and crash risk (i.e. number of crashes divided by AADT per location). It turns out that this correlation equals 0.86 which indicates a strong correlation between the proposed RP Index and crash risk.

As it was mentioned before, there is unfortunately no reliable comprehensive crash database system available in Iran which makes the identification of hotspots using classical crash data driven methods impossible. As long as such valid and reliable crash database does not exist, road engineers in Iran or any other developing country can rely on the approach presented in this chapter to identify unsafe road sections.

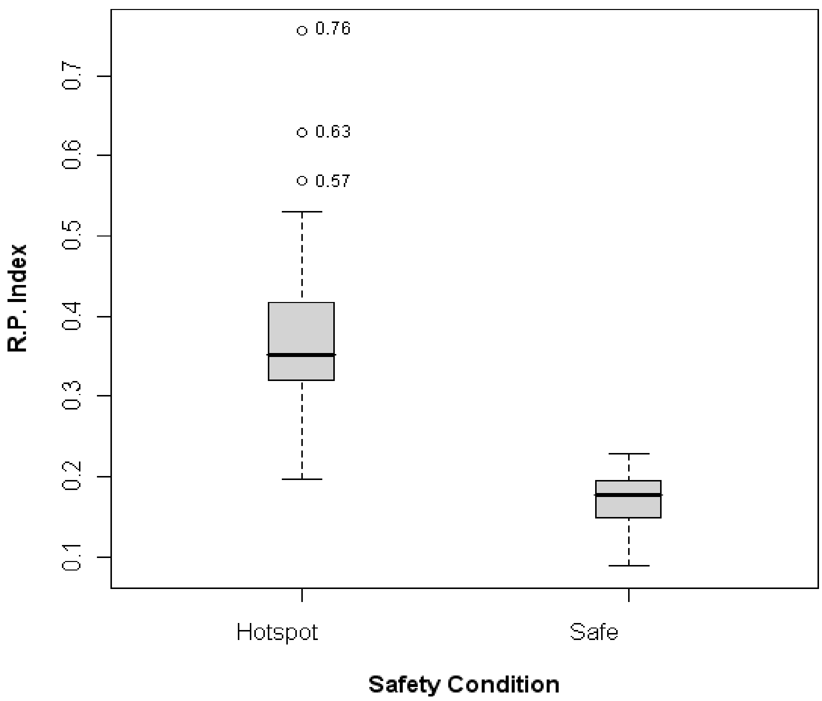

Figure 2.1 RPI values across safe and hazardous sections.

2.5.2. Sensitivity Analysis

In application of MCDM methods, the assessment of the data plays a crucial role. Indeed, the results obtained by application of a MCDM method are strongly related to the actual values assigned to these data. Since uncertainties may be present in the data, great care has to be taken when the results of such a method are interpreted. To facilitate this task, a number of methods have been proposed in literature, mainly focused on the assessment and influence of the weights on the ranking of the different alternatives (Wolters and Mareschal, 1995; Shi et al., 2007). In this chapter, sensitivity analysis was therefore performed in several ways.

The first type of sensitivity analysis was carried out in order to find the most critical criterion. Triantaphyllou and Sanchez (Triantaphyllou and Sánchez, 1997) proposed a method in which the notion of criticality is defined in terms of "smallest change" in any criterion's weight value which will change the ranking of alternatives. Based on the data used in this chapter, it turns out that criterion "B_3" (section located in steep slope) is the most critical one. In fact, a change of 14% to this criterion's weight will cause a change in the alternatives' ranking (this percentage of change is the smallest for all 14 criteria).

The second type of sensitivity analysis is carried out by removing each expert's opinion in order to find out how sensitive the calculated weights of each criterion are to each expert's opinion. In other words, all weights were recalculated after each of expert's opinions was removed. The results show a very small value of standard deviation for all criteria's recalculated weights, which means the weight values of all criteria are not very sensitive to one expert's opinion (Table 2.10).

Finally, a third type of sensitivity analysis was carried out to study the effects of changing each criterion's weight on the ranking of the different alternatives and their safety outcome (e.g. in Braglia et al., 2003; Opricovic and Tzeng, 2004; Ateş et al., 2006; Önüt and Soner, 2008). The greatest change would be achieved when a criterion is totally removed. Therefore, the computation of the RPI and prioritizing all alternatives have been carried out by increasing/decreasing each criterion's weight and at last, totally removing the criterion. It turns out that no changes appear in the safety condition of any of the alternatives after removing each criterion. In other words, the condition of each road section as being safe or hotspot does not change after removing a criterion from the procedure. Furthermore, the results showed that the ranking of the different alternatives is quite robust when removing any of the criteria.

Spearman's coefficients of rank correlation between the original ranking and the ranking after removing each criterion are presented in Table 2.10.

To conclude, the sensitivity analysis shows that the model is very strong to changes in the data used in this study.

Table 2.10 Standard Deviation for all Criteria's Recalculated Weight by Removing Each Expert's Opinion and Spearman's Coefficients of Rank Correlation

Criterion	Standard deviation	Spearman's coefficient
(A)	0.0213	0.980
(B)	0.0165	0.994
(C)	0.0193	0.986
(D)	0.0174	0.803
(E)	0.0155	0.586
(A_1)	0.0224	0.994
(A_2)	0.0218	0.992
(A_3)	0.0251	0.988
(A_4)	0.0216	0.997
(A_5)	0.0243	0.989
(B_1)	0.0230	0.997
(B_2)	0.0271	1.000
(B_3)	0.0181	0.991
(C_1)	0.0209	0.998
(C_2)	0.0213	1.000
(C_3)	0.0235	0.997
(C_4)	0.0247	0.986

2.6. Conclusions

The lack of reliable and valid traffic accident data puts a serious limitation on the use of classical crash prediction and hotspot identification methods in Iran, similar to many other developing countries. Yet, the road safety situation in Iran is so urgent that there is a need for alternative methods that can identify hazardous road locations based on other kinds of information. In this chapter, a Delphi method combined with a MCDM procedure was, therefore, utilized to identify and prioritize crash hotspots, based on several criteria that were considered by experts as relevant for the problem, such as geometric characteristics, traffic conditions, physical conditions and location characteristics.

More specifically, expert opinions about relevant hotspot identification criteria and their relative importance were obtained by means of a Delphi experiment to define a unique weight vector. A weight vector illustrates the relative importance of each criterion in comparison to others. This vector can exclusively be derived from any single decision matrix. This implies that for different decision matrices (i.e. different alternatives but similar criteria) this weight vector will be different. The advantage of having a unique weight vector is that it allows comparing different sets of alternatives (i.e. different locations) when new locations are surveyed and required to be identified and prioritized, since the weight vector is identical for all datasets and locations.

Subsequently, a MCDM procedure, namely TOPSIS, was adopted to prioritize locations based on their performance on each of the criteria. One of the advantages of the MCDM method is its compensatory nature, i.e. its possibility of trade-offs between several decision criteria. This means that the model will be more comprehensive with regard to using all pertinent criteria instead of using only a few of them. Moreover, the proposed framework is validated against a set of safe and unsafe road sections. The results confirm the capability of the presented approach in identifying hazardous locations, however, more extensive validation is probably required to corroborate the findings of this study.

Clearly, this rather qualitative selection approach does not make classic hotspot identification methods based on historical accident data obsolete. In fact, once more and detailed crash statistics for individual road locations become available, it is advisable to complement this qualitative method with statistical crash prediction and hotspot identification models since they rely on well-founded statistical, objective and data driven methodologies. However, until then, the procedure presented in this chapter gives road engineers a practical tool which can be used in systematic search for (potentially) dangerous accident location identification and prioritization. It also gives road authorities the opportunity of proactively identifying the most dangerous locations among a number of locations and treating them within road safety programs.

The results of the presented framework reveal the usefulness of adopting qualitative methodologies in road safety analysis, specifically in the context of absence or a lack of crash data. An important asset of the proposed methodology is its proactive attribute. This method can be adapted to new datasets, comprised of new locations, to predict the safety level of those locations. Furthermore, the proposed method is relatively simple and practical and can easily be adjusted to other regions/countries by collecting local expert's judgments.

Despite the advantages listed above, there are some possible limitations that might undermine the accuracy of identification/prioritization outcome. One concern might be due to the subjective approach that has been followed in defining relevant criteria and their relative importance. In other words, there is a chance that the final chosen criteria and their relative importance are to some extent biased, since they rely on human's judgments. To diminish this possible negative aspect, it is advised to carefully choose experts' panel (the more experts are involved, the better insight into the problem is expected), meticulously accomplish the Delphi technique through several rounds, calibrate the weight vector by coupling the results of this method with other qualitative/qualitative methods and finally further validate the results when new crash data becomes available.

References

Aittoniemi, E., 2008. Potential safety impacts of in-vehicle information services. VTT Technical Research Centre of Finland.

Asgharpoor, M.J., 2003. Group Decision Making and Game Theory in Operation Research. Publication of Tehran University, Tehran.

Ateş, N., Çevik, S., Kahraman, C., Gülbay, M., Erdoğan, S., 2006. Multi Attribute Performance Evaluation Using a Hierarchical Fuzzy TOPSIS Method, in:

Kahraman, C. (Ed.), Fuzzy Applications in Industrial Engineering, Studies in Fuzziness and Soft Computing. Springer Berlin / Heidelberg, pp. 537–572.

Braglia, M., Frosolini, M., Montanari, R., 2003. Fuzzy TOPSIS approach for failure mode, effects and criticality analysis. Quality and Reliability Engineering International 19, 425–443.

Brijs, T., Van den Bossche, F., Wets, G., Karlis, D., 2006. A model for identifying and ranking dangerous accident locations: a case study in Flanders. Statistica Neerlandica 60, 457–476.

Brijs, T., Karlis, D., Van den Bossche, F., Wets, G., 2007. A Bayesian model for ranking hazardous road sites. Journal of the Royal Statistical Society: Series A (Statistics in Society) 170, 1001–1017.

Charlton, S.G., O'Brien, T.G., 2002. Handbook of Human Factors Testing and Evaluation. Taylor & Francis.

Cheng, W., Washington, S.P., 2005. Experimental evaluation of hotspot identification methods. Accident Analysis & Prevention 37, 870–881.

Chu, A.T.W., Kalaba, R.E., Spingarn, K., 1979. A comparison of two methods for determining the weights of belonging to fuzzy sets. Journal of Optimization Theory and Applications 27, 531–538.

Cojazzi G., Fogli D., Grassini G., De Gelder P., Gryffroy D., Bolado R., Hofer E., Virolainen R., Coe I.M., Bassanelli A., Puga J., Papazoglou I., Zuchuat O., Cazzoli E., Eyink J., Guida G., Pinola L., Pulkkinen U., Simola K., von Winterfeldt D., Valeri A., 2001. Benchmark exercise on expert judgment techniques in PSA Level 2. Nuclear Engineering and Design 209, 211–221.

Donnell, E., Harwood, D., Bauer, K., Mason, J., Pietrucha, M., 2002. Cross-Median Collisions on Pennsylvania Interstates and Expressways. Transportation Research Record: Journal of the Transportation Research Board 1784, 91–99.

Elvik, R., 2008. Comparative Analysis of Techniques for Identifying Locations of Hazardous Roads. Transportation Research Record: Journal of the Transportation Research Board 2083, 72–75.

Fichtner, J., 1986. On deriving priority vectors from matrices of pairwise comparisons. Socio-Economic Planning Sciences 20, 341–345.

Fowles, J., 1978. Handbook of futures research. Greenwood Press.

Globalis - an Interactive World Map, 2008. UN common database, http://globalis.gvu.unu.edu/indicator_detail.cfm?Country=IR&IndicatorID=132, (accessed 23 April 2008).

Golshan, R., 2005. Designing a model to identify safety indicators for city entrances, Master thesis, Iran University of Science and Technology, Iran.

Helmer, O., 1977. Problems in futures research: Delphi and causal cross-impact analysis. Futures 9, 17–31.

Hwang, C.-L., Lin, M.-J., 1986. Group Decision Making under Multiple Criteria: Methods and Applications, 1st ed. Springer.

Hwang, C.-L., Yoon, K., 1981. Multiple attribute decision making: methods and applications : a state-of-the-art survey. Springer-Verlag.

Hydén, C., 1993. Theories on traffic safety evaluation connected to new road transport informatics (RTI), in: Proceedings of the 6th ICTCT Workshop. Presented at the 6th ICTCT workshop, Salzburg, Austria.

Iranian Police Statistics, 2008. Available at: http://www.bbc.co.uk/persian/iran/story/2008/04/080423_straffic-punishment.shtml (accessed 23 April 2008).

Jamson, S., Wardman, M., Batley, R., Carsten, O., 2008. Developing a driving Safety Index using a Delphi stated preference experiment. Accident Analysis & Prevention 40, 435–442.

Kim, T.-G., Donnell, E.T., Lee, D., 2008. Use of cultural consensus analysis to evaluate expert feedback of median safety. Accident Analysis & Prevention 40, 1458–1467.

Lin, Y.-H., Lee, P.-C., Chang, T.-P., Ting, H.-I., 2008. Multi-attribute group decision making model under the condition of uncertain information. Automation in Construction 17, 792–797.

Linstone, H.A., Turoff, M., 2002. Delphi Method : Techniques and Applications (digital version). Upper Saddle River, NJ: Addison-Wesley.

Montazeri, A., 2004. Road-traffic-related mortality in Iran: a descriptive study. Public Health 118, 110–113.

Önüt, S., Soner, S., 2008. Transshipment site selection using the AHP and TOPSIS approaches under fuzzy environment. Waste Management 28, 1552–1559.

Opricovic, S., Tzeng, G.-H., 2004. Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. European Journal of Operational Research 156, 445–455.

PIARC Road Safety Manual, 2003. PIARC Technical Committee on Road Safety (C13), France.

Pulkkinen, U., Holmberg, J., 1997. A method for using expert judgments in PSA, STUK-YTO-TR 129. Finnish Centre for Radiation and Nuclear Safety, Helsinki, Finland.

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15, 234–281.

Saaty, T.L., 1980. The Analytic Hierarchy Process. McGraw-Hill, New York.

Saaty, T.L., 2000. Fundamentals of Decision Making and Priority Theory With the Analytic Hierarchy Process, 1st ed. RWS Publications.

Sawalha, Z., Sayed, T., 2006. Transferability of accident prediction models. Safety Science 44, 209–219.

Shi, Y., Zelený, M., Olson, D.L., Stam, A., 2007. Advances in Multiple Criteria Decision Making and Human Systems Management: Knowledge and Wisdom : In Honor of Professor Milan Zeleny. IOS Press.

Shih, H.-S., 2008. Incremental analysis for MCDM with an application to group TOPSIS. European Journal of Operational Research 186, 720–734.

Simola, K., Virolainen, R., 2000. The STUK/VTT expert Judgments Methodology and its Use in the Phase 2: Benchmark Exercise in Expert Judgments Techniques in PSA Level 2 (BE-EJT). Finnish Centre for Radiation and Nuclear Safety, Helsinki, Finland.

Triantaphyllou, E., Sánchez, A., 1997. A Sensitivity Analysis Approach for Some Deterministic Multi-Criteria Decision-Making Methods*. Decision Sciences 28, 151–194.

Turner, S.A., Persaud, B.N., Chou, M., Lyon, C., Roozenburg, A., 2007. International Crash Experience Comparisons Using Prediction Models. Presented at the Transportation Research Board 86th Annual Meeting.

Venkata Rao, R., 2007. Decision Making in the Manufacturing Environment: Using Graph Theory and Fuzzy Multiple Attribute Decision Making Methods. Springer.

Wolters, W.T.M., Mareschal, B., 1995. Novel types of sensitivity analysis for additive MCDM methods. European Journal of Operational Research 81, 281–290.

Yoon, K., Hwang, C.-L., 1995. Multiple Attribute Decision Making: An Introduction. SAGE, Thousand Oaks, CA.

Zekavat, S., 2006. Road safety: government policy, in: Expert Group Meeting on the Development of the Asian Highway Network: Regional Experiences and Lessons in Financial Highway Infrastructure and Improving Road Safety. Bangkok, Thailand.

3. A Simulation-Based Traffic Safety Evaluation of a 4-leg Intersection Based on Traffic Volume and Speed Limit

Pirdavani, A., et al., (2010), Advances in Transportation Studies 22 (A), pp 43–52, DOI: 10.4399/97888548354814.

Pirdavani, A., et al., (2009), Electronic Proceedings of the 2nd Road Safety and Simulation Conference, Paris, France.

Pirdavani, A., et al., (2010), Proceedings of the 15th International Conference Road Safety on Four Continents. p. 1229-1239, Abu-Dhabi, UAE.

3.1. ABSTRACT

In many cases, crashes at intersections account for over 50% of all urban road accidents. The need to reduce these crashes has fostered considerable research on the development and evaluation of traffic safety at intersections. This chapter introduces a micro-level behavioral method for estimating crash potential at 4-leg signalized and unsignalized intersections at different traffic characteristics' categories, such as traffic volume and speed limit. Given that speeding is recognized as a major contributing factor in traffic crashes, in this study it is emphasized on the contribution of speed in the traffic safety situation of intersections. In this study, proximal safety indicators which represent the temporal and spatial proximity characteristics of unsafe interactions and near-accident situations are employed to evaluate the safety level of unsignalized and signalized intersections. Results of the analyses confirm that an increase in speed limit on roads results in a more dangerous situation. Moreover, merits and shortcomings in application of S-Paramics as a microsimulation software in road safety evaluation are addressed.

3.2. Introduction

Intersections present special safety concerns because of unsafe driving actions and maneuvers that result in traffic conflicts with a potential for real crashes. These include conflicts in vehicle trajectories for different intersection approaches, pedestrian conflicts, unexpected changes in vehicle speeds, unexpected lane changes, etc. A number of recent studies of crashes for North American urban roads report that over 50% of reported road crashes take place in proximity to intersections (Cunto and Saccomanno, 2007).

Speeding is recognized as a major contributing factor in traffic crashes, specifically for intersections. Numerous studies are conducted to explain the relationship between speed and road safety: detailed reviews of which are provided elsewhere (Skszek, 2004; Kweon and Kockelman, 2005). Kweon and Kockelman (Kweon and Kockelman, 2005) reported the three of the most important factors of their study to be:

(1) controls for speed conditions in models of crash counts,

(2) use of disaggregate roadway data permitting tight control of design factors, and

(3) specification and evaluation of various count models for panel data.

Several studies solely evaluated the effect of speed enforcement on speed (e.g. in Retting and Farmer, 2003; Champness et al., 2005) or on traffic safety (e.g. in Elvik, 1997; Hess, 2004), while others evaluated both speed and safety (e.g. in Chen et al., 2002; Ha et al., 2003; Hess, 2004, 2004; Cunningham et al., 2005; Goldenbeld and van Schagen, 2005). In an evaluation study (Cunningham et al., 2005), the associations of mean speed, the percentage of speed limit violators, the number of injury accidents, and the number of serious casualties were assessed by comparing two different situations; the development on the roads that were subject to targeted speed enforcement and the development on similar roads without targeted enforcement. Both the mean speed and the percentage of speed limit violators decreased during the targeted enforcement program (Cunningham et al., 2005). Another research (Goldenbeld and van Schagen, 2005) presents the results of a comprehensive analysis of the impact of the speed enforcement program on speeding behavior, crashes, and the economic impact of crashes. The impact on speeding behavior was estimated using generalized least square estimation, in which the observed speeds and the speeding frequencies during the program period were compared to those during other periods.

The relationship between traffic flow and road safety has been reported in several studies (e.g. in Vogel, 2003; Qin et al., 2004; Benedetto et al., 2007; Ye et al., 2009). Therefore, different traffic situations and categories are considered to evaluate their safety level.

Experience shows that microscopic traffic simulation is able to improve the knowledge of risks within a traffic flow (Archer, 2005). In fact, microscopic traffic simulation helps to evaluate and optimize different routing strategies without having to realize tests in the real field. While these tools are mainly used to assess the performance level of road networks in terms of flow, speed or travel time, the possibilities offered by these software tools in order to evaluate road safety remain however limited.

The potential benefits of adopting a micro-level simulation approach were initially recognized by Darzentas et al. (1980). Although the use of micro-level simulation in the safety field has resulted in some resistance due to its data intensive characteristic and the inherent problems of accurately representing a complex crash situation (Archer, 2000). Researchers have attempted to overcome these limitations by using surrogate safety measures within the context of a more aggregate micro-level approach. Arguably, it is a more effective safety assessment strategy which involves the use of proximal safety indicators (sometimes referred to as surrogate safety measures) that represent the temporal and spatial proximity characteristics of unsafe interactions and near-crash events. The main advantage of such measures is related to their resource-effectiveness, given that they occur more frequently than accidents and require relatively short periods of observation in order to establish statistically adequate and reliable results. Such surrogate measures include, time-to-collision (TTC), time extended TTC (TET), post-encroachment time (PET) and deceleration rate (DR), etc. (Minderhoud and Bovy, 2001; Gettman and Head, 2003; Huguenin et al., 2005; El Faouzi et al., 2007; Cunto and Saccomanno, 2008). Gettman and Head (Gettman and Head, 2003) described a project which identified surrogate measures that can be collected from commercial simulation models for evaluation of the relative safety of intersection design alternatives or existing facilities. The results of their analyses pointed out that TTC, PET, and DR are the best available measures. It was emphasized that TTC, PET, and DR can also be used to measure the severity of the conflict (Gettman and Head, 2003). El Faouzi et al. (2007) carried out a study to assess the risk associated to traffic situations using surrogate safety indicators. Their findings show that these safety indicators are able to expose safety problems before crashes emerge (El Faouzi et al., 2007).

In another study (Huguenin et al., 2005), a new safety indicator is proposed which is called "unsafety density" (UD). The concept of the unsafety parameter is based on the direct interaction between couple of vehicles, which seem to be in need of safety treatments. The UD parameter takes into account only potential for rear-end collision and is, therefore, particularly planned for highways network assessments. This indicator allows highlighting the difference in safety level between a fluid and a congested traffic flow situation, which cannot be shown by using traditional macroscopic outputs like speed, flow or occupancy. In another study (Libreros et al., 2007), a new microsimulator called "ValSim" was developed. This microsimulator allows researchers to relate the skewed angle at intersections (merging or crossing) to the driver's angle of visibility for both direct vision and indirect vision through rearview mirrors. ValSim aims to allow designers to evaluate the configuration and geometry of an intersection by dynamical analysis in the geometric design process, and to verify possible conflicts at merging as well as at skewed crossings due to lack of visibility. This software simulates the driver's behavior while carrying out the entry or crossing maneuver. For each timeframe, it calculates the blind spot zones and a possible visibility related conflict is highlighted (Libreros et al., 2007).

3.3. Application of Proximal Safety Indicators

While the use of statistical models based on historical accident data are most common in crash analysis, there are availability and quality problems associated with the data on which they are based. This approach is also considered "reactive" in nature rather than "proactive", where a significant number of accidents must occur before the problem is identified and suitable safety countermeasures are implemented. Understanding these problems, researchers have recently proposed frameworks for "proactive" safety planning; i.e. planning that is not entirely based on historical accident data, but uses other measures such as the use of safety indicators and predictive models (Archer, 2005).

An alternative and/or complementary approach to crash analysis is to measure the more frequent occurrence of near-accidents using proximal safety indicators. These indicators have been suggested as an alternative to the use of crash data since they are believed to have an established relationship with accident occurrence. These indicators are defined as measures of accident proximity, based on the temporal and/or spatial measures that reflect the "closeness" of road-users (or their vehicles), in relation to a projected point of collision (Archer, 2005).

A key advantage (and prerequisite) of proximal safety indicators is that they occur considerably more frequently than accidents. This suggests the need for a significantly shorter study period to establish statistically reliable results. Furthermore, the use of proximal safety indicators is also a more resource-efficiency and ethically appealing alternative for fast, reliable and effective road safety assessment.

A number of related criteria that can be used to identify the usefulness of proximal safety indicators have been identified, suggesting that they must (Archer, 2005):

1. Complement crash data and be more frequent than crash occurence
2. Have a statistical and causal relationship to crashes
3. Have the characteristics of "near-accidents" in a hierarchical continuum that describes all severity levels of road-user interactions with accidents at the highest level and very safe passages with a minimum of interaction at the lowest level

There are several number of proximal safety indicators like TTC, time integrated TTC (TIT), TET, PET, time-to-zebra (TTZ), DR, deceleration-to-safety time (DST), proportion of stopping distance (PSD), shock-wave frequency, time-to-line crossing (TLC) and standard deviation of lateral position (SDLP) which have been applied recently in different studies (Archer, 2005).

In this research, two different safety indicators are used to evaluate the safety situation of signalized and unsignalized intersections. For signalized intersections, time headway is chosen to serve as a safety indicator. Time headway is one of the indicators that are used to estimate the criticality of a certain traffic situation. It has been defined as the elapsed time between the front of the lead vehicle passing a point on the roadway and the front of the following vehicle passing the same point. In other words, time headway is measured by taking the time that passes between two vehicles' reaching the same location. Different countries have slightly different rules with regard to the legal or recommended safe time headway or distance. In the US, e.g. several driver training programs state that it is impossible to safely follow a vehicle with headway of less than 2 seconds (Michael et al., 2000). In Sweden the National Road Administration recommends time headway of 3 seconds in rural areas, and the police use time headway of 1 second as orientation for imposing fines. In Germany, the recommended minimum distance is "half the speedometer", which means a car traveling at 80 km/hr should keep a distance of at least 40 m. This rule translates to a recommended time headway of 1.8 seconds. Fines are imposed when the time headway is smaller than 0.9 second (Vogel, 2003). Besides time headway, there are several other safety indicators that have been used in safety evaluation studies. For instance, TTC is a very helpful indicator for such purposes but the difficulty in its computation process brought us to this conclusion that time headway is a useful, easy to obtain and meaningful indicator. In a comparative research, Vogel (2003) recommended that authorities use headway as criterion for tailgating, because it is easy to measure, it is easily understandable and interpretable, and most importantly, it is directed against potential danger, which effectively prevents dangerous TTC values from occurring at all. On the other hand, by defining a safety threshold for headway, it is possible to purify and filter the data in order to avoid too much work on the procedure of TTC calculation. In a car-following situation (e.g. vehicles are approaching a signalized intersection and a rear-end accident may occur) TTC can never be smaller than the time gap between the lead and the following vehicle. Thus, if the two values are to be compared, it seems reasonable to keep out the cases that are not safety critical with respect to any of the two measures. In other words, for large values of headway, having small TTCs is impossible and having a TTC value bigger than the headway is not unsafe. As a result, filtering the data according to the critical headway threshold will help in reducing computational works. Therefore, a short headway can be interpreted as potential danger, because only vehicles that travel with short headways have the possibility to produce small TTC values and then causing accidents.

In this research S-Paramics (SIAS, 2010) a microsimulation software is adopted to investigate whether changing speed limit under different traffic volume categories will affect traffic safety or not. Different simulation runs were carried out at different traffic volume and speed limit categories in order to make the survey as comprehensive as possible.

Headway is one of the outputs of the simulator. To achieve headways of all vehicles on different approaches, several loop detectors are defined on the network to collect the required data.

To evaluate the safety situation of unsignalized intersections, PET is chosen. This indicator is known to be one of the most commonly used proximal safety indicators in the literature. This measure is used to evaluate situations in which two road-users that are on a collision course, pass over a common spatial point or area with a temporal difference that is below a predetermined threshold (Archer, 2005). This measure represents the difference in time between the passages of the "*offended*" and "*conflicted*" road-users over a common conflict zone (i.e. area of potential collision). This makes PET not only a useful "*objective*" measure, but also one that is less resource-demanding than TTC with regard to data-extraction process and not requiring constant recalculations at each time-step during a safety critical event (Archer, 2005). An example representing the calculation of PET is illustrated in Figures 3.1(a) and 3.1(b).

This example indicates the position of the two vehicles involved in the safety critical event at the start and end of the PET calculation process.

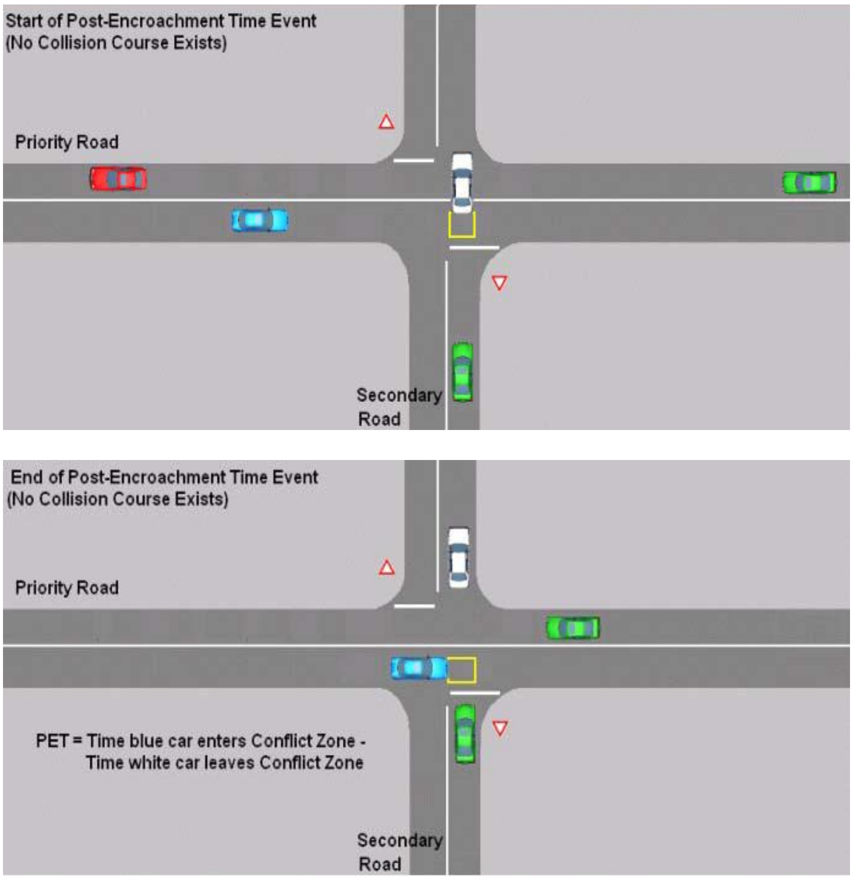

Figure 3.1 (a) and 3.1 (b): Example of the calculation of a Post-Encroachment Time event (Archer, 2005).

Outputs of the microsimulator don't provide any direct information about the safety situation. Therefore, a procedure should be adopted to derive desired safety measures out of the raw output information. Thus, four loop detectors are defined on outgoing links of the four approaches of the intersection. These detectors are located after the conflict zones, so PET values can be obtained easily. These detectors will collect required data such as speed and position of each vehicle. In the context of traffic safety evaluation, data should be as precisely as possible. This issue will become clearer if taking into account that all conflict events will usually take place in less than 2 or 3 seconds. Hence, simulation rate is set at 10 steps per second. It means that all required data is gathered and available for each tenth of a second.

To simplify the process of PET computation, four different conflict zones are assumed and defined for each intersection on which all of the possible conflicts will occur. These four different conflict zones at a 4-leg intersection are depicted in Figure 3.2. For this study, it is also assumed that each roadway on each direction contains 2 lanes.

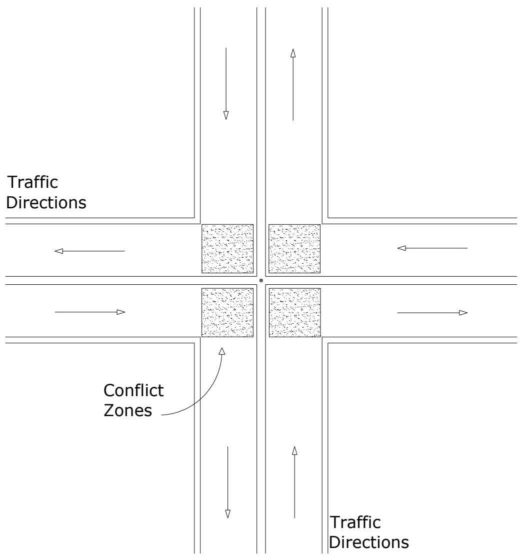

Figure 3.2 conflict zones at intersection and traffic directions.

3.4. Simulation Results

The main objective of this chapter is to evaluate the safety condition of signalized and unsignalized 4-leg intersections under different traffic conditions. One of the major concerns for traffic safety at intersections is speed limit on different roadways. Therefore, different scenarios based on different speed limits and traffic volumes are implemented to perform a comprehensive study. Traffic volume measures are defined in a way that no traffic congestion would occur. Arguably in a situation like traffic congestion, driver's behavior is not the same as in a normal situation. In other words, vehicles will not drive at their desired speed in traffic congestion; thus, evaluating the safety performance at different speed limits will be infeasible.

For signalized intersections traffic volume on major roads is assumed to vary from 750 vehicle per hour (VPH) to 2000 VPH and on minor roads it is supposed to be from 350 VPH to 1000 VPH. Also speed limits are assumed to vary from 45 kilometer per hour (km/hr) to 85 km/hr on major roads and on minor roads from 35 km/hr to 60 km/hr.

Analyzing the results shows a significant difference in headway distribution at different levels of speed limits. When drivers drive faster, distribution at low values of headway is denser. This fact indicates that a higher speed limit which actually ends to a higher level of operating speed will produce smaller values of headway which is more dangerous in terms of traffic safety. However, the mean value of headway for all speed limits will be approximately constant, if there is no change in traffic demand, because the mean time headway is just depended on traffic flow. Figure 3.3 depicted a "Density Estimate Distribution" of one scenario with the flow rate of 1000 VPH for the major road. In order to have a better view, the distribution graph is limited to 3 seconds. As it is shown, the distribution for the highest speed limit is the densest for small values of headway and vice versa. It means that at high speeds, simulated values of headway are smaller than in lower speed limit situations.

As can be seen from Figure 3.3, there is a kind of shift between higher and lower speed limits. This might be due to different car-following models being used in the microscopic simulation software for different speed limits.

The modeled unsignalized intersection is presumed to be a two way stop control intersection. Therefore, vehicles on major roads have the priority and vehicles on the minor road have to stop at the stop lines. As discussed before, traffic demand should be set in a way to avoid traffic congestion. Traffic volume on major roads is assumed to vary from 500 VPH to 650 VPH and on minor roads it is supposed to be from 150 VPH to 250 VPH. Moreover, speed limits are assumed to vary from 45 km/hr to 75 km/hr for major roads and on minor roads from 35 km/hr to 50 km/hr.

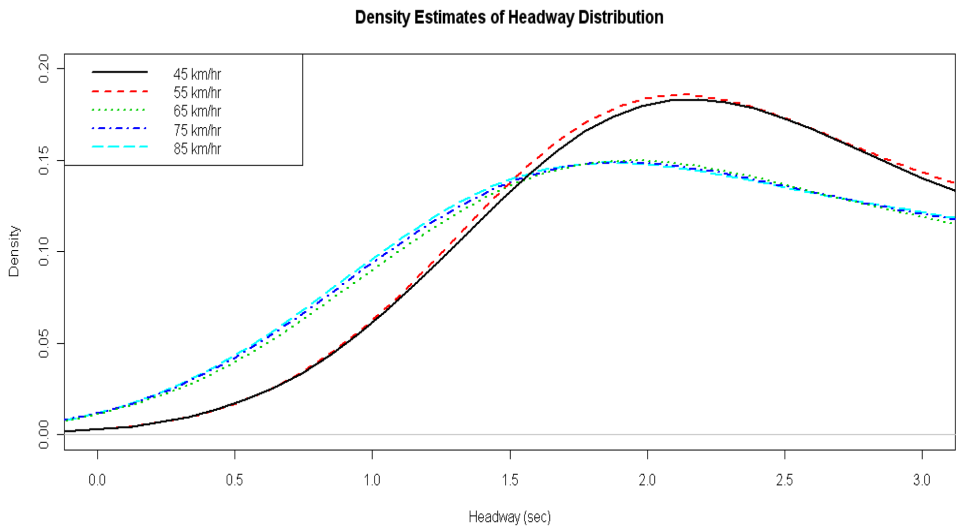

Figure 3.3 Density estimates of headway distribution for 1000 VPH on major road for a signalized intersection.

At these volume levels and speed limits no traffic congestion was seen and vehicles were driven at their desired speed; so the output data is reliable and not affected by other parameters. In order to assess the impact of variable speed limits, each scenario is compared with the others at the same level of traffic volume. Because the simulator assigns traffic demand stochastically, to account for uncertainty, simulation runs have been carried out 10 times for each scenario and the mean values of PET were calculated. Analyzing the results indicates that there is a significant change in traffic safety condition in terms of PET values. It turned out that by increasing speed limits on both roadways, mean values of PET will decrease. This means that in such situations that drivers are driving faster on the major roadway, drivers on the minor roadway will accept a smaller gap to cross over the intersection and lower PET measures specify a worse safety situation at higher levels of speed limit.

Furthermore, it was found that increasing the traffic volume on both major and minor roadways, will conduce to a decrease of mean PET. Arguably at the higher level of traffic volume, vehicles on major roads will face more conflicting vehicles from minor roads on conflict zones. In addition, vehicles on minor roads have to cross over the intersection, accepting shorter gap times because of more traffic on major roads. By increasing traffic volume, the probability of finding large and safe gap times becomes less and consequently PET values will become shorter.

Mean values of PET are shown in Table 3.1 regarding the speed limits and traffic volume on major and minor roads. Also for a better understanding, a "Density Estimate Distribution" of one scenario (Maj=500 VPH, Min=150 VPH) is depicted in Figure 3.4. As can be seen from Figure 3.4, the distribution for highest speed limit is the densest for small values of PET and vice versa. This means that at high speed limits, calculated values of PET are smaller than at lower speed limits.

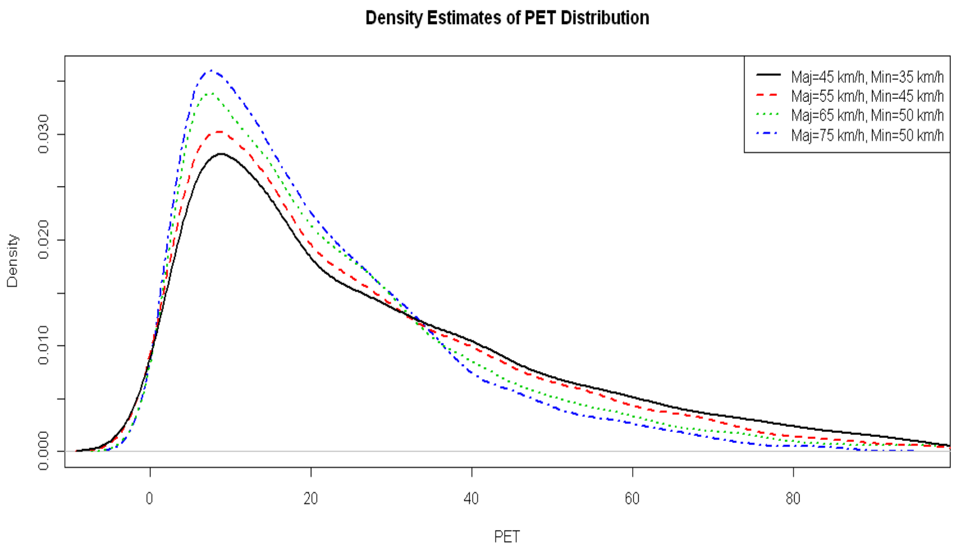

Figure 3.4 Density estimates of PET distribution for the traffic condition "Maj=500 VPH, Min=150 VPH" for an unsignalized intersection.

Table 3.1 Mean PET Values for Different Speed and Traffic Situations

Volume/Speed	Maj=45 km/hr Min=35 km/hr	Maj=55 km/hr Min=45 km/hr	Maj=65 km/hr Min=50 km/hr	Maj=75 km/hr Min=50 km/hr
Maj=650 VPH, Min=250 VPH	23.6629	21.1768	20.7834	19.7155
Maj=650 VPH, Min=200 VPH	24.3998	21.8657	21.5988	20.1679
Maj=600 VPH, Min=200 VPH	24.387	22.7361	21.4859	20.1545
Maj=600 VPH, Min=150 VPH	28.23	26.4424	24.7612	24.226
Maj=550 VPH, Min=200 VPH	24.2051	22.8496	22.6376	21.5428
Maj=550 VPH, Min=150 VPH	29.0857	26.4458	25.846	24.9584
Maj=500 VPH, Min=200 VPH	22.1836	21.8768	21.0508	20.1744
Maj=500 VPH, Min=150 VPH	25.8352	24.8912	23.9772	23.3489

3.5. Conclusions and Future Research

This chapter presents a traffic safety evaluation of signalized and unsignalized intersections using microsimulation and by means of time headway and the PET indicator. Applying time headway provides a safety evaluation of a signalized intersection, while adopting PET as a safety indicator provides useful comparisons for assessing the safety situation at an unsignalized intersection by different scenarios based on different traffic volume and speed limit categories.

In this research, the practical merits and drawbacks of microsimulation modeling have been demonstrated with varying traffic volumes and speed limits on major and minor roads. The application shows how changes in speed limits and also traffic volume will affect the safety level of intersections.

Results indicated that increasing the speed limits on both major and minor approaches will deteriorate safety level; its magnitude will be larger for higher ranges of traffic volumes. Besides, increasing traffic volume up to the point that doesn't cause any traffic congestion will worsen the safety situation. Mean values of PET will decrease by increasing speed limits and also traffic volumes. Also at higher speed levels, the number of observed short headways is more than at lower speed limit levels. It means that speeding will make the traffic safety situation worse.

This study also shows that drivers' behavior which have been defined in the microsimulator "S-Paramics" is sensitive to changes in speed limits as a policy measure. Nonetheless, it is still a black box how changing the speed limit or traffic demand will affect drivers' behavior and other inter-models like car following models used in the simulator. There are also some shortcomings of adopting commercial microsimulation software to perform safety analysis. As can be seen from Figure 3.3, the headway distribution of different speed limits are different. In other words, it seems that for different speed limits, different car-following models are applied. Not having control on what is being used in the software is a concern when using a software that, at first place, was not intended to be utilized for safety analysis. Moreover, uncontrolled behavior of cars in specific circumstances (e.g. when cars are approaching intersections) makes commercial microsimulation models less appropriate to perform safety analysis.

This study does not cover all of the most commonly used safety proximal indicators like TTC and its derived sub-indicators. TTC and its derived sub-indicators could be used to investigate many types of collisions like transverse, rear-end and converging collisions, while PET can just explore transverse collisions. Nevertheless, besides the simple procedure of calculating PET and its indication of the safety situation, the authors believe that this safety indicator yields promising results for a safety evaluation, specifically for unsignalized intersections where most of the collisions are transverse collisions.

For the future studies it would be ideal to further evaluate safety levels of intersections by means of other types of proximal safety indicators. A comparison study including implementation of other kinds of safety proximal indicators will also lead researchers to a better understanding of this proactive traffic safety evaluation approach.

References

Archer, J., 2000. Developing the potential of micro-simulation modeling for traffic safety assessment. Corfu, pp. 233–246.

Archer, J., 2005. Indicators for traffic safety assessment and prediction and their application in micro-simulation modelling: A study of urban and suburban intersections, PhD thesis, Royal Institute of Technology Stockholm, Sweden.

Benedetto, C., De Blasiis, M., Calvi, A., 2007. Risk of vehicle rear end collision in function of traffic flow. Presented at the International Conference Road Safety and Simulation, Rome, Italy.

Champness, P., Sheehan, M., Folkman, L., 2005. Time and distance halo effects of an overtly deployed mobile speed camera. Presented at the Road Safety Research, Policing and Education Conference, Wellington, New Zealand.

Chen, G., Meckle, W., Wilson, J., 2002. Speed and safety effect of photo radar enforcement on a highway corridor in British Columbia. Accident Analysis & Prevention 34, 129–138.

Cunningham, C.M., Hummer, J.E., Moon, J.-P., 2005. An Evaluation of the Safety Affects of Speed Enforcement Cameras in Charlotte, NC. Institute for Transportation Research and Education North Carolina State University, Raleigh, North Carolina.

Cunto, F., Saccomanno, F., 2007. Microlevel Traffic Simulation Method for Assessing Crash Potential at Intersections. Presented at the Transportation Research Board 86th Annual Meeting, Washington DC, United States.

Cunto, F., Saccomanno, F.F., 2008. Calibration and validation of simulated vehicle safety performance at signalized intersections. Accident Analysis & Prevention 40, 1171–1179.

Darzentas, J., Cooper, D.F., Storr, P.A., McDowell, M.R.C., 1980. Simulation of road traffic conflicts at T-junctions. SIMULATION 34, 155–164.

El Faouzi, N.-E., de Mouzon, O., Chung, E., Pham, M.-H., 2007. Road Safety Indicators: Results in Vaud Canton. Presented at the International Conference Road Safety and Simulation, Rome, Italy.

Elvik, R., 1997. Effects on Accidents of Automatic Speed Enforcement in Norway. Transportation Research Record: Journal of the Transportation Research Board 1595, 14–19.

Gettman, D., Head, L., 2003. Surrogate Safety Measures from Traffic Simulation Models. Transportation Research Record: Journal of the Transportation Research Board 1840, 104–115.

Goldenbeld, C., van Schagen, I., 2005. The effects of speed enforcement with mobile radar on speed and accidents: An evaluation study on rural roads in the Dutch province Friesland. Accident Analysis & Prevention 37, 1135–1144.

Ha, T.-J., Kang, J.-G., Park, J.-J., 2003. The effects of automated speed enforcement systems on traffic-flow characteristics and accidents in Korea. Institute of Transportation Engineers. ITE Journal.

Hess, S., 2004. Analysis of the Effects of Speed Limit Enforcement Cameras: Differentiation by Road Type and Catchment Area. Transportation Research Record: Journal of the Transportation Research Board 1865, 28–34.

Huguenin, F., Torday, A., Dumont, A.-G., 2005. Evaluation of traffic safety using microsimulation. Presented at the 5th Swiss Transport Research Conference, Monte Verità / Ascona, Swiss.

Kweon, Y.-J., Kockelman, K., 2005. Safety Effects of Speed Limit Changes: Use of Panel Models, Including Speed, Use, and Design Variables.

Transportation Research Record: Journal of the Transportation Research Board 1908, 148–158.

Libreros, L., Garcia, A., Contreras, J., 2007. A new Microsimulator to Evaluate Road Safety at Skewed Intersections. Presented at the International Conference Road Safety and Simulation, Rome, Italy.

Michael, P.G., Leeming, F.C., Dwyer, W.O., 2000. Headway on urban streets: observational data and an intervention to decrease tailgating. Transportation Research Part F: Traffic Psychology and Behaviour 3, 55–64.

Minderhoud, M.M., Bovy, P.H.L., 2001. Extended time-to-collision measures for road traffic safety assessment. Accident Analysis & Prevention 33, 89–97.

Qin, X., Ivan, J.N., Ravishanker, N., 2004. Selecting exposure measures in crash rate prediction for two-lane highway segments. Accident Analysis & Prevention 36, 183–191.

Retting, R., Farmer, C., 2003. Evaluation of Speed Camera Enforcement in the District of Columbia. Transportation Research Record: Journal of the Transportation Research Board 1830, 34–37.

SIAS, 2010. S-Paramics Microsimulation, SIAS Limited, Reference Manual S-Paramics, Edinburgh, UK.

Skszek, S.L., 2004. Actual Speeds on the Roads Compared to the Posted Limits (No. FHWA-AZ-04-551). Arizona Department of Transportation.

Vogel, K., 2003. A comparison of headway and time to collision as safety indicators. Accident Analysis & Prevention 35, 427–433.

Ye, X., Pendyala, R.M., Washington, S.P., Konduri, K., Oh, J., 2009. A simultaneous equations model of crash frequency by collision type for rural intersections. Safety Science 47, 443–452.

4. Developing Zonal Crash Prediction Models with a Focus on Application of Different Exposure Measures

Pirdavani, A., et al., (2012), Transportation Research Record: Journal of the Transportation Research Board, in press.

Pirdavani, A., et al., (2012), DVD Compendium of the 91st Transportation Research Board Conference, Washington D.C., USA.

4.1. ABSTRACT

Assessing the safety impacts of travel demand management (TDM) policies is essential to be carried out by means of a proactive approach. Since TDM policies are typically implemented at an aggregate level, crash prediction models (CPMs) should also be developed at a similar level of aggregation. These models should match better with the resolution at which TDM evaluations are performed. Therefore, zonal crash prediction models (ZCPMs) are considered to construct the association between observed crashes and a set of predictor variables in each zone. This is carried out by the generalized linear modeling (GLM) procedure with the assumption of Negative Binomial (NB) error distribution. Different exposure, network and socio-demographic variables of 2200 traffic analysis zones (TAZs) are considered as predictors of crashes in the study area, Flanders, Belgium. To this end, an activity-based transportation model framework is applied to produce exposure measurements. Crash data used in this study consist of recorded injury crashes between 2004 and 2007. In this chapter, different ZCPMs are developed to predict the number of injury crashes (NOICs); including fatal, severely and slightly injury crashes. To assess the impacts of different exposure measures on model performance, these models are classified into three different groups, i.e.:

(1) flow-based models,

(2) trip-based models and

(3) models based on a combination of the two.

The results show a considerable improvement of the model performance when both trip-based and flow-based exposure variables are used simultaneously in the model's formulation. The ultimate objective of this chapter is to provide a predictive tool at the planning-level which can be applied on different TDM policies to evaluate their traffic safety impacts.

4.2. Introduction

For many years, researchers have attempted to investigate the negative impacts of growing travel demand on traffic safety by predicting the number of crashes (NOCs) based on the patterns they have learnt from crashes that occurred in the past. Traditionally, this reactive approach consists of different phases such as; identification, diagnosis and improvement of unsafe locations, so called hotspots. From the ethical point of view, this reactive approach is not acceptable because it requires several years of crashes to occur in order to identify and treat safety problems. Thus, providing a more proactive approach, which is capable of evaluating road safety at the planning-level, is essential. This proactive approach is increasingly being paid attention to by researchers and practitioners in the last few years. Dealing with traffic safety at the planning-level requires the ability to integrate TDM policies into a crash predicting context. TDM policies are usually performed and evaluated at a more aggregate level than just on the level of individual intersections or road section. However, local TDM implementations like adding capacity to a segment of road may also be conducted. Typically the before/after analysis of such an infrastructure adjustment is carried out locally despite the fact that such an adjustment may have broader consequences. Thus, the impact of adopting a TDM strategy on transportation or traffic safety should be evaluated at a higher level rather than merely the local consequences. Therefore, application of CPMs at a zonal level like TAZ leads to ZCPM.

The main goal of this study is, therefore, to develop ZCPMs that can be used to evaluate the traffic safety effects of conducted TDM policies. Exposure is an important determinant of traffic safety. Therefore, it is needed to assess the exposure under different TDMs to be able to evaluate their traffic safety impacts.

To this end, the FEATHERS (Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS) activity-based transportation model is applied on the Flemish population. The FEATHERS framework (Janssens et al., 2007) was developed in order to facilitate the development of activity-based models for transportation demand in Flanders, Belgium. Currently, the framework is fully operational at the level of Flanders. The real-life representation of Flanders is embedded in an agent-based simulation model which consists of over six million agents, each agent representing one member of the Flemish population. A sequence of 26 decision trees is used in the scheduling process and decisions are based on a number of attributes of the individual (e.g. age, gender), of the household (e.g. number of cars) and of the geographical zone (e.g. population density, number of shops). For each individual with its specific attributes, the model simulates whether an activity (e.g. shopping, working, leisure activity …) is going to be carried out or not. Subsequently, amongst others, the location, transport mode (available modes in FEATHERS are "car driver", "car passenger", "public transportation" and "slow mode" including pedestrians and cyclists) and duration of the activity are determined, taking into account the attributes of the individual (Kochan et al., 2008). As such, the FEATHERS activity-based model can provide the exposure measure, number of trips (NOTs), by means of (time-of-day dependent) origin-destination (OD) matrices. Assigning these OD matrices of car trips to the Flemish road network provides other exposure variables like vehicle kilometers traveled (VKT) and vehicle hours traveled (VHT). These three different exposure variables together with network and socio-demographic variables are then used to construct the ZCPMs. Since the exposure which comes out of the activity-based model is sensitive to TDM policies, these ZCPMs are also TDM sensitive.

The structure of this chapter is as follows. Initially, the literature about crash prediction modeling at the zone level will be reviewed. In the next sections the data preparation and model development will be demonstrated. Finally, the results of ZCPMs will be shown followed by the final conclusions.

4.3. Background

CPMs can be categorized in two different levels: the local level (e.g. road and intersection) and the regional level (e.g. TAZ). Usually CPMs at the local level aim to predict the safety benefits/detriment of infrastructure improvements. These models are not typically designed to evaluate traffic safety impacts of TDM policies; thus, application of CPMs at a higher aggregation level will be more practical (Tarko et al., 2008). Recently, the application of ZCPMs became more popular amongst researchers because of their ability to estimate the effect of different TDM policies on traffic safety. This has been initially introduced by Levine et al. (1995). In their study a set of both socio-economic and network variables were chosen to predict the NOCs in different TAZs. They estimated a linear relationship between different explanatory variables and the NOCs. Several researchers have examined the association of a collection of network infrastructure, socio-demographic and socio-economic variables and weather conditions with the NOCs at the level of TAZs (e.g. in Amoros et al., 2003; Noland and Oh, 2004, 2004; Noland and Quddus, 2004; De Guevara et al., 2004; Aguero-Valverde and Jovanis, 2006; Wier et al., 2009; Quddus, 2008; Huang et al., 2010). De Guevara et al. (2004) developed planning-level ZCPMs for the city of Tucson, Arizona. They considered many socio-demographic and network variables in their model construction. They concluded that predictors such as population density, the number of persons younger than 17 years old as a percentage of the total population, the number of employees, the intersection density, the percentage of miles of principal arterials, the percentage of miles of minor arterials and the percentage of miles of urban collectors are significant predictors for the NOCs. In a study carried out by Wier et al. (2009) it was shown that traffic volume, population size, the proportion of arterial streets without public transit, the proportion of the population living in poverty and the number of people aged 65+ as a percentage of the total population, were significantly good predictors. Moreover, Noland and Quddus (2004) concluded that TAZs with high employment density had more traffic crashes while in urbanized more densely populated TAZs fewer crashes have been observed.

Hadayeghi at al. (Hadayeghi et al., 2003, 2006, 2007, 2010a, 2010b; Hadayeghi, 2009) have been working on ZCPMs for several years. In one of their first studies, it was shown that the number of accidents in a TAZ increases when the VKT, major and minor road length, total employed labor force, household population and intersection density increase and it decreases with a higher posted speed and a higher level of congestion in the TAZ (Hadayeghi et al., 2003). Hadayeghi at al. (2006) investigated the temporal transferability of the ZCPMs by applying models constructed on 1996 data to predict the NOCs for each TAZ in 2001 for the City of Toronto. In another research, twenty-three regression models were developed to examine the relationships between several types of transportation planning variables and collision frequency. The results showed the potential of planning-level safety models to serve as decision support tools for planners to consider safety in the planning phase (Hadayeghi et al., 2007). Hadayeghi at al. (2010b) conducted the same research but this time they applied geographically weighted Poisson regression (GWPR) instead of taking the GLM approach. The major difference between these two types of models is that GWPR models allow the model coefficient estimates to vary spatially for each TAZ. This very important additional attribute of these models provides some extra information as it takes the spatial location of a crash into consideration. Lovegrove and Sayed (2006) concluded that quantifying the relationship between the zonal characteristics such as exposure, network, socio-demographic and TDM variables (e.g. total commuters from each zone or commuter density) and crashes at a zonal level provides a predictive tool to predict the NOCs in a TAZ. They have used the GLM techniques to develop ZCPMs for both urban and rural areas across the Greater Vancouver Regional District (GVRD). The results of their study show that increasing signal density, intersection density per unit area and lane kilometers, arterial-local intersections in rural areas and total arterial road lane kilometers will lead to an increase in the NOCs. On the contrary, an increase in the number of three-leg intersections and local road lane kilometers will decrease the NOCs in a TAZ. Lovegrove and Sayed (2007) further developed a set of ZCPMs for a "black-spot" study in GVRD. These sets of ZCPMs consist of an exposure variable (i.e. VKT) and other network, socio-demographic and TDM variables. The results of this study also confirmed that ZCPMs have the ability to round out traditional reactive road safety improvement programs. An at al. (2011) found VHT, the number of intersections and the number of households with low income level to be correlated with the NOCs in TAZs.

To conclude, many variables such as traffic volume, VHT, VKT, population, employment, level of income, urbanization degree (i.e. the degree of urbanization is categorized into three different levels and, therefore, represented by two dummy variables; "Urban" and "Suburban". When "Urban" and "Suburban" metrics in a TAZ are both 0, then this TAZ is located in a rural area), traffic intensity, number of intersections and intersection density, speed and road length are confirmed in different studies to be significant predictors of crashes at the zonal level.

Recently, some researchers constructed ZCPMs by associating the NOCs in a TAZ with trip production/attraction and other network characteristics. Abdel-Aty et al. (2011a) identified and prioritized important variables which can be associated with crashes per TAZ by means of the classification and regression trees (CART) technique. It was shown that this methodology will be helpful in incorporating proactive safety measures for long range transportation planning. Abdel-Aty et al. (2011b) also developed different ZCPMs for different crash severity levels using the NOTs as the exposure variable. They concluded that different sets of predictors should be considered based on the type or severity of crashes (e.g. total trip productions and attractions provide better model fit for the total and peak hour crashes while severe crashes were better predicted by different trip motive related variables). Naderan and Shahi (2010) investigated the feasibility of associating travel demand in urban areas with crash frequencies in each TAZ. They developed a series of ZCPMs using the NOTs produced/attracted as predictors. They concluded that these models provide the basic tool for evaluating TDM policies in urban transportation planning in terms of traffic safety as the application of a specific TDM policy may reduce trip productions of a specific motive.

The drawback of considering only trips as an exposure variable is that the impact of trip time, trip length, route choice, intrazonal traffic and transit traffic on a TAZ will be neglected. The number of produced/attracted trips might be an acceptable indicator of how busy or active a TAZ is or how much people are exposed to unsafe situations, but it always leaves out the effects of through traffic which is just passing through a TAZ neither having their origin or destination in that TAZ. It is a well known relationship in literature that road crashes are tightly linked to traffic exposure (e.g. in Hadayeghi et al., 2003, 2006, 2007, 2010a, 2010b; Lovegrove and Sayed, 2006,2007; Hadayeghi, 2009; An et al., 2011; Abdel-Aty et al., 2011a, 2011b). Therefore, having a more informative measure of exposure, is expected to result in a better crash prediction. This study demonstrates the impact of applying different exposure variables on the performance of ZCPMs.

4.4. Data Preparation

The required information to construct ZCPMs consists of exposure, network and socio-demographic data accompanied with the crash data. These data should be collected for the whole study area and also be aggregated to the zonal level. The study area in this research is the Dutch speaking region in northern Belgium, Flanders. Flanders has over 6 million inhabitants, about 60% of the population of Belgium. As already mentioned, an activity-based model within the FEATHERS framework is applied on the Flemish population to derive the in depth information of Flemish people's travel behavior and travel demand for a null-scenario (current situation). The basic outputs of FEATHERS are activity-travel schedules/diaries. These can then be aggregated to OD matrices. These OD matrices include the NOTs for each traffic mode at different disaggregation levels (i.e. age, gender, day of the week, time of a day and motive). This traffic demand is then assigned to the network to obtain detailed exposure measures at the network level (i.e. VHT and VKT). VKT is calculated based on the distance traveled on each link (i.e. by multiplying the flow on each link by link's length), however, VHT is derived from the exact time that vehicles spent on each link (i.e. impact of congestion on travel time is accounted for). These network level exposure measures are then aggregated to TAZ level. This has been carried out at the zonal level comprising of 2200 TAZs. The average size of TAZs is 6.09 square kilometers with standard deviation of 4.78 square kilometers. In addition, for each TAZ a set of variables including socio-demographic and network variables were derived to construct the ZCPMs. The crash data used in this study consist of a geo-coded set of injury crashes that have occurred during the period 2004 to 2007. Table 4.1 shows a list of selected variables, together with their definition and descriptive statistics, which have been used in developing the ZCPMs presented in this chapter.

4.5. Model Development

Crash data consist of non-negative integers, so using ordinary least-squares regression which serves continuous dependent variables (e.g. time) is not an option (Lord and Mannering, 2010). For decades, researchers applied Poisson Regression models for crash prediction analysis. Because of the natural characteristics of crash data that variance does not necessarily equals to mean, application of Poisson Regression models becomes risky as it might bias the results by making parameter estimates inconsistent (Lord and Mannering, 2010). To overcome this problem, the Negative Binomial (NB) model, which allows the variance to differ from the mean, was applied as an extension of the Poisson model. The NB model is the most commonly used model in crash data modeling (Lord and Mannering, 2010). Application of CPMs at TAZ level has been initially introduced by Levine at al. based on a Linear Regression (Levine et al., 1995). As mentioned before, application of the NB model in crash prediction analysis became popular amongst many researchers (e.g. in Amoros et al., 2003; Hadayeghi et al., 2003, 2006, 2007; Noland and Oh, 2004, 2004; Noland and Quddus, 2004; Lovegrove, 2005; Aguero-Valverde and Jovanis, 2006; Lovegrove and Sayed, 2006, 2007; Lovegrove and Litman, 2008; Hadayeghi, 2009; Naderan and Shahi, 2010; Lord and Mannering, 2010; Abdel-Aty et al., 2011b; An et al., 2011). This is due to the fact that usually crash data have a greater variance compared to the mean, therefore, NB model can handle this over-dispersion better. In this study, the NB models were developed within the generalized linear modeling (GLM) framework.

Reviewing the literature for different model forms showed that the following model has been widely used by different researchers (e.g. in Lovegrove, 2005; Hadayeghi, 2009; Abdel-Aty et al., 2011b; An et al., 2011):

Table 4.1 List of Explanatory Variables for the ZCPMs with Their Definition and Descriptive Statistics

	Variable	Definition	Average	Min	Max	SD[a]
	Crash	total NOICs observed in a TAZ	36.03	0	326	41.58
Exposure variables	Number of Trips	average daily number of trips originating/destined from/to a TAZ	2765.8	0	18111.4	2869.8
	Total Flow	average annual daily traffic (AADT) in a TAZ (vehicle)	96414.5	70.9	4423325	181695
	VHT	total daily vehicle hours traveled in a TAZ	608.26	1.50	9998.6	930.29
	VKT	total daily vehicle kilometers traveled in a TAZ	52533.8	84.06	985192	90715.2
	Motorway Flow	AADT of motorways in a TAZ (vehicle)	37724.96	0	3881777	146757.5
	Motorway VHT	total daily vehicle hours traveled on motorways in a TAZ	260.52	0	9762.5	832.97
	Motorway VKT	total daily vehicle kilometers traveled on motorways in a TAZ	27471.82	0	946152.8	84669.53
	Other Roads Flow	AADT of other roads in a TAZ (vehicle)	58690.29	0	734152.5	73632.5
	Other Roads VHT	total daily vehicle hours traveled on other roads in a TAZ	348.51	0	3777.69	358.76
	Other Roads VKT	total daily vehicle kilometers traveled on other roads in a TAZ	26662.85	0	303237.6	28133.04
	V/C	average volume to capacity in a TAZ	0.0478	0	0.5697	0.0422

Network variables	Speed	average speed limit in a TAZ (km/hr)	69.4	31	120	10.91
	Capacity	hourly average capacity of links in a TAZ	1790.1	1200	7348.1	554.6
	Area	total area of a TAZ in square kilometers	6.09	0.09	45.22	4.78
	No. of Links	number of links in a TAZ	39.27	1	230	30.46
	Link Length	total length of the links in a TAZ (km)	15.86	0.39	87.95	10.79
	Link Density	link length per square kilometers in a TAZ	3.37	0.03	20.44	2.41
	Intersection	total number of intersections in a TAZ	5.8	0	40	5.9
	Intersection Density	number of intersection per square kilometers	1.76	0	50.63	3.39
	Motorway	presence of motorway in a TAZ describes as below: "No" represented by 0 "Yes" represented by 1	0	0	1	-[b]
	Urban	Is the TAZ in an urban area? "No" represented by 0 "Yes" represented by 1	0	0	1	-
	Suburban	Is the TAZ in a suburban area? "No" represented by 0 "Yes" represented by 1	0	0	1	-

Socio-demographic variables	Driving License	average driving license ownership in a TAZ describes as below: "No" represented by 0 "Yes" represented by 1	1	0	1	-
	Income Level	average income of residents in a TAZ describes as below: "Monthly salary less than 2249 Euro" represented by 0 "Monthly salary more than 2250 Euro" represented by 1	1	0	1	-
	Work Status	average work status of the residents in a TAZ describes as below: "Don't work" represented by 0 "Work" represented by 1	1	0	1	-
	Population	total number of inhabitants in a TAZ	2614.52	0	15803	2582.6
	Population Density	population per square kilometers	774.14	0	14567.4	1398.4
	Adults Population	total number of adult inhabitants in a TAZ	1796.06	0	12014	1823.5
	Adults Population Density	adults population per square kilometers	542.85	0	10444.8	1013.4

a: Standard deviation b: Data not applicable

$$E(C) = \beta_0 \times (Exposure)^{\beta_i} \times e^{\sum \beta_i x_i} \tag{1}$$

Where;

$E(C)$: expected crash frequency,

β_0 and β_i : model parameters,

$Exposure$: exposure variable (e.g. VHT, VKT or NOTs), and

x_i : other explanatory variables.

Logarithmic transformation of Equation (1) when considering only one exposure variable yields:

$$ln[E(C)] = ln(\beta_0) + \beta_1 ln(Exposure) + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n \tag{2}$$

Several models were constructed to associate the relationship between crash frequency and the explanatory variables while the main focus is on the application of different exposure measures. The models can be categorized into three different groups based on the type of exposure measure that was utilized, i.e.:

(1) flow-based models,
(2) trip-based models and
(3) models based on a combination of the two.

Flow-based models were constructed by regressing the NOICs in each TAZ on VHT or VKT, as the exposure variables, and a selection of network and socio-demographic variables listed in Table 4.1. Trip-based models use the same network and socio-demographic variables but use NOTs as the exposure variable. In the third type of models, both flow and trip based variables are included simultaneously as measures of exposure.

Initially in the flow-based models, VKT and VHT were computed based on all types of roads together. The Preliminary results showed that all models overestimated the NOICs in the TAZs in which a significant length of motorway was present. Based on this observation it was concluded that it is necessary to make the distinction between different road types (i.e. motorways and non-motorway roads) and consider each of their exposure values separately.

Coefficients were estimated by using a forward selection procedure by taking the intercept and one of the exposure variables for the starting point and then additional candidate variables were selected from the available data described in Table 4.1. When the exposure variable is included in the model, the next step is to include the second variable with the smallest p-value (Amato and Satake, 2009). This procedure continues until the remaining candidate variables have a p-value higher than 0.05. At this point the final model is obtained. For each model, the multicollinearity phenomenon was also checked for by means of the variance inflation factor (VIF) for all variables. As a common rule of thumb, 10 is defined (Kutner et al., 2004) as a cut off value meaning that if the VIF is higher than 10 then multicollinearity is high. VIF values for final chosen models are shown in Table (4.2). These results suggest that multicollinearity is not a problem in developing ZCPMs.

Table 4.2 VIF Values Among Explanatory Variable for Selected Models

	Model #5	**Model #6**
Coefficients	**VIF**	**VIF**
log(Number of Trips)	5.136244	5.150353
log(Motorways VKT)	-	2.142915
log(Other Roads VKT)	-	1.99657
log(Motorways VHT)	2.214077	-
log(Other Roads VHT)	2.10465	-
Capacity	2.048942	1.951593
Intersection	1.930527	1.941688
Income level	1.095113	1.095456
Urban	1.290003	1.303834
Suburban	1.249959	1.25879
Population	3.929244	3.959617

The data used in this study consists of the information from 2200 TAZs. This provides a sufficient number of cases with respect to sample size. Therefore, for model development, 70% of the TAZs were chosen randomly as training set and the rest of 30% were used as a test set. Applying the developed models to the test sets shows a significant correlation between the observed and the predicted NOICs for all developed models (see Figure 4.1). The Pearson Correlation Coefficients (PCC) (Amato and Satake, 2009) of all developed models shown in Table 4.3 indicate that all models are capable of predicting the NOICs quite well; however, more statistical tests are needed to be able to select the best fitted model (from the alternatives). This analysis will be carried out in the next section. The regression results of the developed models are presented in Table 4.3.

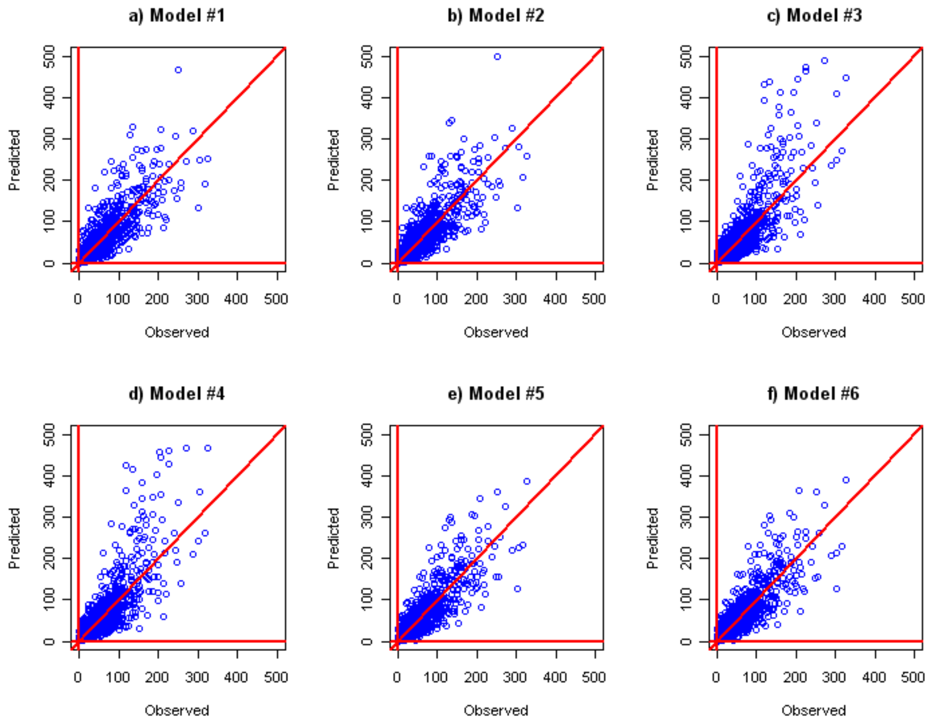

Figure 4.1 Correlations between the observed and the predicted NOICs.

4.6. Results

As already mentioned before, models were constructed by regressing the NOICs in each TAZ on the natural logarithmic transformation of NOTs, VHT or VKT, as the exposure variables, in addition to other network and socio-demographic variables. Different statistical tests were used to assess the goodness-of-fit of each developed model. The results of the analysis show that one or two exposure variables together with the variables (See Table 4.1 for a more detailed description of the variables) V/C, Capacity, Speed, Intersection, Income level, Population, Urban and Suburban were statistically significant at the 95% confidence level.

Table 4.3 Regression Results of the Developed ZCPMs

	Model #1	**Model #2**	**Model #3**
Coefficients	**Estimate**	**Estimate**	**Estimate**
(Intercept)	-3.713e+00	-1.885e+00	-3.744e-01
log(Number of Trips)	7.338e-01	6.659e-01	- [a]
log(Motorways VKT)	-	-	-
log(Other Roads VKT)	-	-	-
log(Motorways VHT)	-	-	1.025e-02
log(Other Roads VHT)	-	-	4.143e-01
log(V/C)	-	1.883e-01	-
Speed	9.806e-03	-	-
Capacity	3.157e-04	3.190e-04	4.317e-04
Intersection	4.184e-02	4.433e-02	3.220e-02
Income level	-1.160e-01	-1.308e-01	-5.401e-02
Urban	3.111e-01	-	3.182e-01
Suburban	3.487e-02	-	1.624e-01
Population	-	-	1.372e-04
Deviance/DF[b]	1.1386	1.1365	1.1331
AIC[c]	17353	17255	17196
MSPE[d]	581.04	583.59	1141.17
PCC[e]	0.8458	0.8448	0.8247
R^2	0.7155	0.7138	0.6802

a: Data not applicable

b: Degree of Freedom (DF)

c: Akaike Information Criterion (AIC)

d: Mean Squared Prediction Error (MSPE)

e: The Pearson Correlation Coefficient (PCC) between observed and predicted crash values

Continued Table 4.3 Regression Results of the Developed ZCPMs

	Model #4	**Model #5**	**Model #6**
Coefficients	**Estimate**	**Estimate**	**Estimate**
(Intercept)	-2.226e+00	-2.886e+00	4.141e+00
log(Number of Trips)	- [a]	4.676e-01	4.520e-01
log(Motorways VKT)	9.472e-03	-	7.744e-03
log(Other Roads VKT)	4.267e-01	-	3.132e-01
log(Motorways VHT)	-	7.717e-03	-
log(Other Roads VHT)	-	3.040e-01	-
log(V/C)	-	-	-
Speed	-	-	-
Capacity	3.913e-04	4.220e-04	3.894e-04
Intersection	3.271e-02	2.844e-02	2.888e-02
Income level	-5.879e-02	-1.056e-01	-1.071e-01
Urban	4.848e-01	2.287e-01	3.520e-01
Suburban	2.047e-01	5.712e-02	9.095e-02
Population	1.299e-04	2.340e-05	2.293e-05
Deviance/DF[b]	1.1302	1.1365	1.1355
AIC[c]	17166	16921	16918
MSPE[d]	1089.96	482.41	489.74
PCC[e]	0.8276	0.8709	0.8697
R^2	0.6848	0.7584	0.7564

a: Data not applicable

b: Degree of Freedom (DF)

c: Akaike Information Criterion (AIC)

d: Mean Squared Prediction Error (MSPE)

e: The Pearson Correlation Coefficient (PCC) between observed and predicted crash values

For all models, exposure variables were positively associated with the NOICs in each TAZ. As the NOTs, VHT or VKT increases, total NOICs also tends to increase. Many studies found a similar association between VKT (e.g. in Karlaftis and Tarko, 1998; Hadayeghi et al., 2003, 2006, 2007, 2010a, 2010b;

Lovegrove, 2005; Lovegrove and Sayed, 2006, 2007; Hadayeghi, 2009), VHT (e.g. in An et al., 2011), NOTs (e.g. in Naderan and Shahi, 2010; Abdel-Aty et al., 2011b, 2011b) and NOICs per TAZ. Among the road capacity related variables, V/C is found to be a significant predictor only for trip-based model (i.e. Model #2), whereas "capacity" is turned out to be a better predictor of the NOICs for other models (Table 4.3).

Positive correlation of average speed limit and number of intersections with NOICs per TAZ can be observed for all models. This positive relationship has also been reported in other studies (e.g. in Hadayeghi et al., 2003; Noland and Quddus, 2004; Lovegrove and Sayed, 2006; Abdel-Aty et al., 2011b; An et al., 2011). This can be explained as injury crashes are more likely to occur at higher speeds. In general, intersections have a higher risk of experiencing conflicts compared to road links because of their natural design, therefore, there are more crashes expected to occur in TAZs that have a higher number of intersections. Population is also found to have a positive association with the NOICs. It can be explained as in the TAZs with a higher number of inhabitants, there will be more people exposed to unsafe situations compared to TAZs with fewer inhabitants. The same association has been recognized by other researchers (e.g. in Levine et al., 1995; De Guevara et al., 2004; Huang et al., 2010).

As it can be observed in Table 4.3, all of the constructed models showed negative association with “Income Level” unlike other explanatory variables, which have positive signs. These results are in line with other studies’ findings. It has been shown in many studies that poverty has a positive relationship with the crashes that occurred in a TAZ (e.g. in Amoros et al., 2003; De Guevara et al., 2004; Noland and Oh, 2004, 2004; Aguero-Valverde and Jovanis, 2006; Wier et al., 2009; Huang et al., 2010; An et al., 2011). The negative sign for the “Income Level” variable indicates that TAZs with a higher income level are expected to have fewer crashes compared to less prosperous TAZs. In all of the models, coefficient estimate of the variable “Urban” is more than the coefficient estimate of “Suburban”. This means that the model correctly predicts more crashes for more urbanized TAZs. If a TAZ is located in a rural area, both variables relevant to urbanization degree will be zero and subsequently there will be less crashes predicted for rural areas. This is in line with the findings of Huang at al. (2010) that counties with a higher level of urbanization are associated with a higher crash risk. In general, it can be concluded that all of the parameters' signs are along the lines of theoretical expectations and findings of other previously published studies.

To select the best fitted model, different criteria were taken into consideration. The Akaike Information Criterion (AIC) is a measure of the relative goodness of fit. AIC is defined as:

$$AIC = 2k - 2ln(L) \tag{3}$$

Where;

k : the number of parameters in the model and

L : the maximized value of the likelihood function for the estimated model.

Models may be ranked according to their AIC values in a sense that the preferred model is the one with the minimum AIC value.

Another measure that has been used in comparative analysis between different models is Mean Squared Prediction Error (MSPE) (Hadayeghi, 2009). The MSPE is the sum of the squared differences between predicted and observed crashes divided by the sample size. The MSPE is defined as:

$$MSPE = \sum_{i=1}^{n}(\mathrm{P_i} - \mathrm{O_i})^2 / \mathrm{n} \tag{4}$$

Where;

$\mathrm{P_i}$: predicted NOICs for i_{th} TAZ,

$\mathrm{O_i}$: observed NOICs for i_{th} TAZ and

n: total number of TAZs

Comparing different models performance shows a significant improvement in models which are developed based on both types of exposure variables; NOTs and assigned traffic on the network (i.e. VHT or VKT). The maximum AIC value stands for the trip-based model (i.e. Model #1) (See Table 4.3). It indicates that this model achieves the poorest fit on the data compared to other models. The model which includes V/C as an exposure related model also performs badly (i.e. Model #2). Flow-based models (i.e. Models #3 and #4) provide a better fit compared to trip-based model according to their AIC values; however their MSPE values are the greatest among other models. Goodness-of-fit measures for combined exposure models (i.e. Models #5 and #6) signify the better performance of these models compared to the ones that include only one of the exposure variables. As can be seen in Table 4.3, Models #5 and #6 provide almost equally the minimum MSPE and AIC values, therefore, it can be concluded that these models are the best fitted models. In Figure 4.2, the observed and the predicted NOICs are displayed for each TAZ. Red TAZs have higher observed/predicted NOICs compared to blue TAZs. By comparing Figures 4.2(a), 4.2(b) and 4.2(c), a relatively similar pattern can be noticed. This is an indication that the final chosen models (i.e. Models #5 and #6) are capable of predicting the NOICs quite well.

4.7. Conclusions

In this study, a proactive approach was presented, which can be applied in transportation safety planning. Different ZCPMs were developed in order to associate the number of injury crashes (NOICs) with different exposure, network and socio-demographic variables at the zonal level comprising of 2200 TAZs in Flanders, Belgium. To this end, Negative Binomial (NB) models were developed within the generalized linear modeling (GLM) framework. This approach has been widely followed by other researchers when the crash data is over-dispersed. The results of the analysis showed the different performance of the models when different exposure variables are considered to be included in the model development. The contributing variables and their measure of effectiveness also provide some information that can be used by researchers and practitioners to evaluate the impact of each variable in predicting crashes. This can serve as a useful traffic safety input into transportation projects at the planning-level.

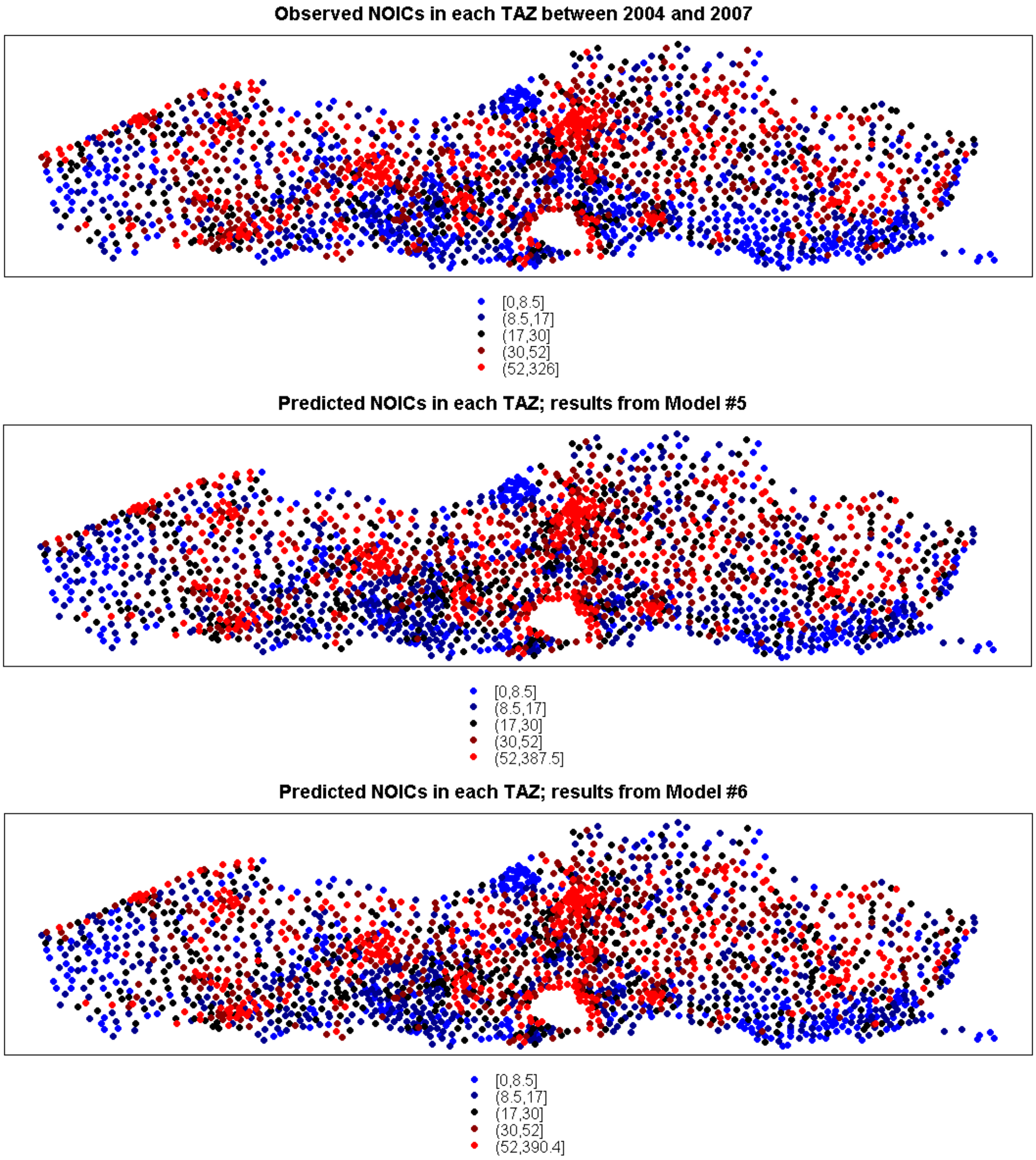

Figure 4.2 Graphical representation of observed and expected NOICs for each TAZ in Flanders.

In order to assess TDM's traffic safety implications, it is essential to have TDM sensitive exposure measures. Activity-based transportation models provide an adequate range of in-depth information about individuals' traveling behavior to realistically simulate and evaluate TDM strategies. The advantage of using activity-based transportation models is that these models can be adjusted to simulate different TDM scenario and, therefore, a wide range of traffic safety evaluation studies can be carried out based on their output. In this study, traffic demand was prepared by the activity-based transportation model's outputs, FEATHERS.

Based on the results presented in this paper the following conclusions can be drawn:

- Different exposure, network and socio-demographic variables have been considered for model development. Coefficients were estimated by using a forward selection procedure. Variable selection has been carried out until all remaining variables have a p-value higher than 0.05. For each developed model, multicollinearity was checked and the results didn't show any multicollinearity issues. Positive or negative association of all selected variables with the NOICs has been checked by comparing them with the results of other studies reported in the literature. The results found are in line with what can be found in literature.
- Sole use of NOTs originating/destining from/to a TAZ for crash prediction will result in missing some important information about the characteristics of travel demand; i.e. NOTs, as an exposure variable, does not contain information on trip time, trip length and route choice. Moreover, transit traffic which is just passing through a TAZ can have a significant share of the exposure observed in a TAZ. This part of the exposure is left out by only using the NOTs. Thus, other exposure variables which are sensitive to the impacts of trip assignment should be taken into account. This has been carried out by assigning the traffic demand to the network using an equilibrium assignment and by computing exposure variables that are sensitive to the assignment like VHT and VKT.
- Different models were developed based on the different measures of exposure that were generated. These models were categorized into three groups according to the exposure measures used as independent variables, i.e.:

  (1) flow-based models,

(2) trip-based models and

(3) models based on a combination of the two.

- The results of the model comparison showed that the models that contain the combination of two exposure variables outperform the models which only have one of the exposure variables (NOTs or VHT/VKT) in their formulation. Therefore, considering the application of both flow-based and trip-based exposure variables in ZCPM construction is recommended. The necessity of considering both exposure measures can also be implicitly perceived by looking at the VIF values for these two variables (see Table 4.2). Since these variable are not correlated with each other, it can be concluded that each of these variables carry different information and, therefore, both should be considered simultaneously in model construction.

Spatial and temporal transferability of the model and application of other model types would also be some directions for future studies.

References

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011a. Zonal Level Safety Evaluation Incorporating Trip Generation Effects. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011b. Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis & Prevention 38, 618–625.

Amato, P.P., and Satake, E., 2009. *Introductory Statistics: A First Course*. 1st edition, Kendall Hunt Publishing.

Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Accident Analysis & Prevention 35, 537–547.

An, M., Casper, C., Wu, W., 2011. Using Travel Demand Model and Zonal Safety Planning Model for Safety Benefit Estimation in Project Evaluation. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

De Guevara, F.L.D., Washington, S., Oh, J., 2004. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. Transportation Research Record: Journal of the Transportation Research Board 1897, 191–199.

Hadayeghi, A., 2009. Use of Advanced Techniques to Estimate Zonal Level Safety Planning Models and Examine Their Temporal Transferability, PhD thesis, Department of Civil Engineering, University of Toronto.

Hadayeghi, A., Shalaby, A., Persaud, B., 2003. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. Transportation Research Record: Journal of the Transportation Research Board 1840, 87–95.

Hadayeghi, A., Shalaby, A., Persaud, B., 2007. Safety Prediction Models: Proactive Tool for Safety Evaluation in Urban Transportation Planning Applications. Transportation Research Record: Journal of the Transportation Research Board 2019, 225–236.

Hadayeghi, A., Shalaby, A., Persaud, B., 2010a. Development of Planning-Level Transportation Safety Models using Full Bayesian Semiparametric Additive Techniques. Journal of Transportation Safety & Security 2, 45–68.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010b. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. Accident Analysis & Prevention 42, 676–688.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accident Analysis & Prevention 38, 579–589.

Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-Level Crash Risk Analysis in Florida. Transportation Research Record: Journal of the Transportation Research Board 2148, 27–37.

Janssens, D., Wets, G., Timmermans, H.J.P., Arentze, T.A., 2007. Modelling Short-Term Dynamics in Activity-Travel Patterns: Conceptual Framework of the Feathers Model. Presented at the 11th World Conference on Transport Research, Berkeley CA, USA.

Karlaftis, M.G., Tarko, A.P., 1998. Heterogeneity considerations in accident modeling. Accident Analysis & Prevention 30, 425–433.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., 2008. Assessing the Impact of Fuel Cost on Traffic Demand in Flanders Using Activity-Based Models. Presented at the Travel Demand Management TDM, Vienna, Austria.

Kutner, M. H., Nachtsheim, C.J., Neter, J., 2004. *Applied Linear Regression Models*. 4th ed. McGraw-Hill.

Levine, N., Kim, K.E., Nitz, L.H., 1995. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. Accident Analysis & Prevention 27, 675–685.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44, 291–305.

Lovegrove, G., Sayed, T., 2007. Macrolevel Collision Prediction Models to Enhance Traditional Reactive Road Safety Improvement Programs. Transportation Research Record: Journal of the Transportation Research Board 2019, 65–73.

Lovegrove, G.R., 2005. Community-Based, Macro-Level Collision Prediction Models, PhD thesis, University of British Columbia.

Lovegrove, G.R., Litman, T., 2008. Using Macro-Level Collision Prediction Models to Evaluate the Road Safety Effects of Mobility Management Strategies: New Empirical Tools to Promote Sustainable Development. Presented at the Transportation Research Board (TRB) 87th Annual Meeting, Washington D.C. USA.

Lovegrove, G.R., Sayed, T., 2006. Macro-level collision prediction models for evaluating neighbourhood traffic safety. Canadian Journal of Civil Engineering 33, 609–621.

Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: Introducing crash generation concept. Accident Analysis & Prevention 42, 339–346.

Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention 36, 525–532.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis & Prevention 36, 973–984.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. Accident Analysis & Prevention 40, 1486–1497.

Tarko, A., Inerowicz, M., Ramos, J., Li, W., 2008. Tool with Road-Level Crash Prediction for Transportation Safety Planning. Transportation Research Record: Journal of the Transportation Research Board 2083, 16–25.

Wier, M., Weintraub, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis & Prevention 41, 137–145.

5. Spatial Analysis of Injury Crashes in Flanders, Belgium; Application of Geographically Weighted Regression Technique

Pirdavani, A., et al., (2013), DVD Compendium of the 92nd Transportation Research Board Conference, Washington D.C., USA.

5.1. ABSTRACT

Generalized linear models (GLMs) are the most widely used models utilized in crash prediction studies. These models illustrate the relationships between the dependent and explanatory variables by estimating fixed global estimates. Since the crash occurrences are often spatially heterogeneous and are affected by many spatial variables, the existence of spatial correlation in the data is examined by means of calculating Moran's I measures (Moran, 1950) for dependent and explanatory variables. The results indicate the necessity of considering the spatial correlation when developing crash prediction models. The main objective of this research is to develop different zonal crash prediction models (ZCPMs) within the geographically weighted generalized linear models (GWGLM) framework in order to explore the spatial variations in association between number of injury crashes (NOICs) (including fatal, severely and slightly injury crashes) and other explanatory variables. Different exposure, network and socio-demographic variables of 2200 traffic analysis zones (TAZs) are considered as predictors of crashes in the study area, Flanders, Belgium. To this end, an activity-based transportation model framework is applied to produce exposure measurements. Crash data used in this study consist of recorded crashes between 2004 and 2007. GWGLMs are developed using a Poisson error distribution and are often referred to as geographically weighted Poisson regression (GWPR) models. Moreover, the performances of developed GWPR models are compared with their corresponding GLMs. The results show that GWPR models outperform the GLM models; this is due to the capability of GWPR models in capturing the spatial heterogeneity of crashes.

5.2. Introduction

For many years, researchers have attempted to investigate the negative impacts of growing travel demand on traffic safety by predicting the number of crashes based on the patterns they have learnt from crashes that occurred in the past. This should lead to providing a predictive tool which is capable of evaluating road safety at the planning-level. Dealing with traffic safety at the planning-level requires the ability to integrate travel demand management (TDM) policies into a crash predicting context. TDM policies are usually performed at a more aggregate level than just on the level of an individual intersection or road section. Thus, the impact of adopting a TDM strategy on transportation or traffic safety should be evaluated at a higher level rather than merely the local consequences. Application of crash prediction models (CPMs) at a macro level like TAZ leads to a type of prediction models commonly referred to as ZCPMs. ZCPMs have been initially introduced by Levine et al. based on linear regression models (Levine et al., 1995a). However, the most common modeling framework for ZCPMs is the GLM framework (e.g. in Amoros et al., 2003; Hadayeghi et al., 2003, 2006, 2007, 2010; Noland and Oh, 2004; Noland and Quddus, 2004; De Guevara et al., 2004; Lovegrove, 2005; Aguero-Valverde and Jovanis, 2006; Lovegrove and Sayed, 2006, 2007; Lovegrove and Litman, 2008; Hadayeghi, 2009; Lord and Mannering, 2010; Naderan and Shahi, 2010; Abdel-Aty et al., 2011; An et al., 2011). Within a GLM framework, fixed coefficient estimates explain the association between the explanatory variables and the dependent variable. In other words, a single model tries to fit the observed data for all locations (i.e. TAZs) similarly. Expectedly, different spatial variation may be observed for different explanatory variables especially where the study area is relatively large. Neglecting this spatial variation may deteriorate the predictive power of ZCPMs.

Spatial variation is known to be an important aspect of traffic safety analysis and in particularly crash prediction modeling. Inclusion of spatial variation in traffic safety studies has been reported by many researchers. In one of the earliest studies, the spatial relationship between activities which generate trips and motor vehicle accidents was studied and applied to the City and County of Honolulu (Levine et al., 1995b). Different spatial patterns for different variables such as population, employment and road characteristics were identified. LaScala et al. (2000) found that significant spatial relationships exist between specific environmental and demographic characteristics of the City and County of San Francisco and pedestrian injury crashes. Flahaut et al. (2003) presented different methods for identifying and delimiting accidents black-zones. This was an application of spatial correlation in defining accident black-zones which share similar characteristics. A similar study was carried out by Moons et al. (2009) where the structure of the underlying road network is taken into account by applying Moran's I to identify crash hot zones. In another study by Flahaut (2004), it was indicated that spatial autocorrelation should be integrated in the modeling process if spatial data are being studied. He concluded that spatial models in comparison to non-spatial models, do not overestimate the significance of explanatory variables; thus, spatial variation should be considered to analyze spatial data. Geurts et al. (2005) investigated the clustering phenomenon in road accidents. This was an application of spatial analysis in traffic safety that aims to analyze the characteristics of specific zones on which more accident occur. Spatial correlation was found to be significant in injury crashes in a study conducted for the State of Pennsylvania at the county level (Aguero-Valverde and Jovanis, 2006). Aguero-Valverde and Jovanis (2008) further explored the effect of spatial correlation in models of road crash frequency at the segment level. The results of their study highlighted the importance of including spatial correlation in road crash modeling studies. The models with spatial correlation show significantly better fit compared to the Poisson lognormal models. The existence of clusters in the spatial arrangement of pedestrian crashes was reported by (Cottrill and Thakuriah, 2010). They supported their conclusions by computing Moran's I value and presenting the local indicators of spatial association (LISA) significance map of crashes. Huang et al. (2010) performed a county-level road safety analysis for the state of Florida. They reported that significant spatial correlations in crash occurrence were identified across adjacent counties.

This spatial variation which is often referred to as "spatial non-stationarity" is overlooked by the GLMs. Following such a modeling approach ends up with a set of fixed global variable estimates which are the same for different TAZs; however, it is possible that an explanatory variable which is found to be a significant predictor of crashes in some TAZs might not be a powerful predictor in other TAZs. There are different spatial modeling techniques that have been applied by many researchers in the crash prediction field. Auto-logistic models, conditional auto-regression (CAR) models, simultaneous auto-regression (SAR) models, spatial error models (SEM), generalized estimating equation (GEE) models, full-Bayesian spatial models, Bayesian Poisson-lognormal models are some of the most employed techniques to conduct spatial modeling in traffic safety (e.g. in Levine et al., 1995b; Miaou et al., 2003; Flahaut, 2004; Aguero-Valverde and Jovanis, 2006, 2008; Wang and Abdel-Aty, 2006; Quddus, 2008; Wang et al., 2009; Guo et al., 2010; Huang et al., 2010; Siddiqui et al., 2012). The output of these models are still fixed coefficient estimates for all locations, however the spatial variation is taken into account.

Another solution for taking the spatial variation into account is developing a set of local models, so called geographically weighted regression (GWR) models (Fotheringham et al., 2002). These models rely on the calibration of multiple regression models for different geographical entities. The GWR approach has mainly been followed in health, economic and urban studies. Also a few studies have been carried out in the transportation field using this technique (e.g. in Zhao and Park, 2004; Chow et al., 2006; Du and Mulley, 2006; Páez, 2006; Clark, 2007; Blainey, 2010).

In traffic safety, Hadayeghi et al. (2003) developed GWR models to investigate spatial variations in the model relationships. The results of the GWR models indicated an improvement in model predictability by means of an increased R^2. In another study (Delmelle and Thill, 2008), bicycle crashes were studied in Buffalo, New York. Density of development, physical road characteristics, socioeconomic and demographic variables were the selected explanatory variables. Given the spatial nature of these variables, a GWR model was developed and showed a better performance compared with the conventional model. An inter-province difference in traffic accidents in Turkey was studied by Erdogan (2009). Different spatial autocorrelation analyses were performed to see whether the accidents are clustered or not. Since the results of these analyses indicated non-stationarity in the data, a GWR model was developed. They also showed that the GWR model performs better than the ordinary least square (OLS) model.

The GWR technique can be adapted to the GLM models and form Geographically Weighted Generalized Linear Models (GWGLMs) (Fotheringham et al., 2002). GWGLMs are able to serve the count data (such as number of crashes) while simultaneously accounting for the spatial non-stationarity. Hadayeghi et al. (2010) used the GWR technique in conjunction with the GLM framework using the Poisson error distribution. They developed different geographically weighted Poisson regression (GWPR) models to associate the relationship between crashes and a number of predictors. The results of the comparisons between GLMs and GWPR models revealed that the GWPR models outperform the GLMs since they are capable of capturing spatially dependent relationships.

The first objective of this chapter is to examine the existence of spatial correlation in the dependent and other explanatory variables available in the data. The main objective of this study is then to develop different ZCPMs within the GWGLM framework in order to explore the spatial variations in association between crashes and other explanatory variables. Moreover, the performance of GWGLMs will be compared with the GLMs developed in an earlier study (Pirdavani et al., 2012). In this study GWGLMs are developed using a Poisson error distribution; henceforth, we refer to these models as GWPR models.

5.3. Methodology

5.3.1. Data Preparation

The required information to construct the prediction models consists of exposure, network and socio-demographic data accompanied with the crash data. These data should be collected for the whole study area and also be aggregated to the zonal level. The study area in this research is the Dutch speaking region in northern Belgium, Flanders. Flanders has over 6 million inhabitants, about 60% of the population of Belgium.

Exposure is an important determinant of traffic safety. Therefore, it is needed to assess the exposure under different scenarios to be able to evaluate their traffic safety impacts. To this end, the FEATHERS (Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS) activity-based transportation model is applied on the Flemish population. The FEATHERS framework (Janssens et al., 2007) was developed in order to facilitate the development of activity-based models for transportation demand in Flanders, Belgium. The real-life representation of Flanders is embedded in an agent-based simulation model which consists of over 6 million agents, each agent representing one member of the Flemish population. A sequence of 26 decision trees is used in the scheduling process and decisions are based on a number of attributes of the individual (e.g. age, gender), of the household (e.g. number of cars) and of the geographical zone (e.g. population density, number of shops). For each individual with its specific attributes, the model simulates whether an activity (e.g. shopping, working, leisure activity …) is going to be carried out or not. Subsequently, amongst others, the location, transport mode (available modes in FEATHERS are "car driver", "car passenger", "public transportation" and "slow mode" including pedestrians and cyclists) and duration of the activity are determined, taking into account the attributes of the individual (Kochan et al., 2008). As such, the FEATHERS activity-based model can provide the exposure measure, number of trips (NOTs), by means of time-of-day dependent origin-destination (OD) matrices for all traffic modes. Assigning the OD matrices of car trips to the Flemish road network provides other exposure variables like vehicle kilometers traveled (VKT) and vehicle hours traveled (VHT). VKT is calculated based on the distance traveled on each link (i.e. by multiplying the flow on each link by link's length), however, VHT is derived from the exact time that vehicles spent on each link (i.e. impact of congestion on travel time is accounted for). These network level exposure measures are then aggregated to the zonal level comprising of 2200 TAZs. The average size of TAZs is 6.09 square kilometers with a standard deviation of 4.78 square kilometers. In addition, for each TAZ a set of variables including socio-demographic and network variables were derived. The crash data used in this study consist of a geo-coded set of injury crashes (including fatal, severely injured and slightly injured crashes) that have occurred during the period 2004 to 2007 and were provided by the Flemish Ministry of Mobility and Public Works. Table 5.1 shows a list of variables, together with their definition and descriptive statistics, which have been used in developing the models presented in this chapter.

5.3.2. Essentiality of Conducting Spatial Analysis

Previous research has indicated that there might be significant spatial correlations in crash occurrence across different locations (e.g. in Erdogan, 2009; Cottrill and Thakuriah, 2010; Hadayeghi et al., 2010; Huang et al., 2010; Siddiqui et al., 2012). Therefore, it is essential to check for the existence of spatial correlation of dependent and explanatory variables. This can be carried out by means of different statistical tests such as Moran's autocorrelation coefficient commonly referred to as Moran's I (Moran, 1950; Fotheringham et al., 2002). Moran's I is an extension of the Pearson product-moment correlation coefficient to a univariate series. It may be expected that in the existence of spatial patterns, close observations are more likely to be similar than those far apart. This is often referred to as Tobler's first law of geography (Tobler, 1970). Moran's I can be formulated as follows:

$$Moran\acute{}s\ I = \frac{n}{SumW}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad (1)$$

Where n is the number of cases (number of TAZs in our study), $\overline{x}$ is the mean of x_i's, x_i is i^{th} observation of a variable, w_{ij} is the weight between cases i and j, and SumW is the sum of all w_{ij}'s:

$$SumW = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \qquad (2)$$

Table 5.1 List of Explanatory Variables for the ZCPMs with Their Definition and Descriptive Statistics

	Variable	Definition	Average	Min	Max	SD[a]
	Crash	total NOICs observed in a TAZ	36.03	0	326	41.58
Exposure variables	Number of Trips	average daily number of trips originating/destined from/to a TAZ	2765.8	0	18111.4	2869.8
	Total Flow	average annual daily traffic (AADT) in a TAZ (vehicle)	96414.5	70.9	4423325	181695
	VHT	total daily vehicle hours traveled in a TAZ	608.26	1.50	9998.6	930.29
	VKT	total daily vehicle kilometers traveled in a TAZ	52533.8	84.06	985192	90715.2
	Motorway Flow	AADT of motorways in a TAZ (vehicle)	37724.96	0	3881777	146757.5
	Motorway VHT	total daily vehicle hours traveled on motorways in a TAZ	260.52	0	9762.5	832.97
	Motorway VKT	total daily vehicle kilometers traveled on motorways in a TAZ	27471.82	0	946152.8	84669.53
	Other Roads Flow	AADT of other roads in a TAZ (vehicle)	58690.29	0	734152.5	73632.5
	Other Roads VHT	total daily vehicle hours traveled on other roads in a TAZ	348.51	0	3777.69	358.76
	Other Roads VKT	total daily vehicle kilometers traveled on other roads in a TAZ	26662.85	0	303237.6	28133.04
	V/C	average volume to capacity in a TAZ	0.0478	0	0.5697	0.0422

Network variables	Speed	average speed limit in a TAZ (km/hr)	69.4	31	120	10.91
	Capacity	hourly average capacity of links in a TAZ	1790.1	1200	7348.1	554.6
	Area	total area of a TAZ in square kilometers	6.09	0.09	45.22	4.78
	No. of Links	number of links in a TAZ	39.27	1	230	30.46
	Link Length	total length of the links in a TAZ (km)	15.86	0.39	87.95	10.79
	Link Density	link length per square kilometers in a TAZ	3.37	0.03	20.44	2.41
	Intersection	total number of intersections in a TAZ	5.8	0	40	5.9
	Intersection Density	number of intersections per square kilometers	1.76	0	50.63	3.39
	Motorway	presence of motorway in a TAZ describes as below: “No” represented by 0 “Yes” represented by 1	0	0	1	-[b]
	Urban	Is the TAZ in an urban area? “No” represented by 0 “Yes” represented by 1	0	0	1	-
	Suburban	Is the TAZ in a suburban area? “No” represented by 0 “Yes” represented by 1	0	0	1	-

Socio-demographic variables	Driving License	average driving license ownership in a TAZ describes as below: “No” represented by 0 “Yes” represented by 1	1	0	1	-
	Income Level	average income of residents in a TAZ describes as below: “Monthly salary less than 2249 Euro” represented by 0 “Monthly salary more than 2250 Euro” represented by 1	1	0	1	-
	Work Status	average work status of the residents in a TAZ describes as below: “Don’t work” represented by 0 “Work” represented by 1	1	0	1	-
	Population	total number of inhabitants in a TAZ	2614.52	0	15803	2582.6
	Population Density	population per square kilometers	774.14	0	14567.4	1398.4
	Adults Population	total number of adult inhabitants in a TAZ	1796.06	0	12014	1823.5
	Adults Population Density	adults population per square kilometers	542.85	0	10444.8	1013.4

a: Standard deviation b: Data not applicable

There are different ways to define the weight matrix (Lee and Wong, 2001). The simplest solution is using a binary matrix. Each cell of the binary matrix has a weight of 1 if the corresponding geographical units are neighbors and if they are not adjacent to each other, the corresponding cell has a value of 0. This method does not seem to be very efficient to be used in spatial analysis. The values of the matrix heavily depend on the size and the number of neighboring zones and in most of the cases these values are 0. Besides adjacency which can describe the spatial relationship among neighboring geographical entities, distance is a powerful measure which can explain this spatial relationship quite well. There are different ways to define the distance between two geographical entities. In this study the geographical entities are TAZs. TAZs are represented by their centroid; therefore, the distance can then be calculated by using the TAZs' centroid information. Adjacent TAZs have short distances from each other while distant TAZs have larger distance values. Thus, the weights are defined as the inverse of distances between each pair of TAZs. In other words, the weight is inversely proportional to the distance between two TAZs. According to (Lee and Wong, 2001), the strength of many spatial relationships has been found to diminish more than proportionally to the distance between different geographical features. Therefore, the squared distance is sometimes used to represent the weights. In this study the weight matrix is defined as:

$$w_{ij} = \frac{1}{{d_{ij}}^2} \tag{3}$$

Where d_{ij} is the distance between the centroid of the i^{th} and the j^{th} TAZ.

The value of Moran's I varies from -1 representing complete spatial dispersion to 1 indicating full spatial clustering. Table 5.2 presents the Moran's I values for the selected variables used in the model construction. It can be seen that all of the selected variables show a significant spatial clustering. Table 5.2 also includes the significance level of Moran's I values by means of p-values and Z-scores. Z-scores can be derived as follows:

$$Z(MI_i) = \frac{O(MI_i) - E(MI_i)}{SD(MI_i)} \tag{4}$$

Where $Z(MI_i)$ is the Z-score of Moran's I of variable i, $O(MI_i)$ is the Observed Moran's I of variable i, $E(MI_i)$ is the expected Moran's I of variable i and $SD(MI_i)$ is the standard deviation of Moran's I of variable i. As can be seen from Table 5.2, "Urban" is the most significant spatial clustered variable among all variables. The results presented in Table 5.2 indicate the necessity of considering this spatial correlation when developing crash prediction models.

Table 5.2 Moran's I Statistics for Dependent and Explanatory Variables

Variable	**Observed Moran's I**	**Z-score**	**Spatial status**
Crash	0.211	37.648	Non-stationary
log(Number of Trips)	0.263	46.834	Non-stationary
log(Motorway VHT)	0.149	26.547	Non-stationary
log(Other Roads VHT)	0.179	31.956	Non-stationary
log(Motorway VKT)	0.156	27.771	Non-stationary
log(Other Roads VKT)	0.166	29.593	Non-stationary
log(V/C)	0.301	53.471	Non-stationary
Capacity	0.121	21.541	Non-stationary
Intersection	0.199	35.556	Non-stationary
Urban	0.437	78.088	Non-stationary
Suburban	0.239	42.539	Non-stationary
Income Level	0.187	33.318	Non-stationary
Population	0.154	27.423	Non-stationary

5.4. Model Construction

5.4.1. Generalized Linear Model

Reviewing the literature for different model forms showed that the following model has been widely used in different studies (e.g. in Lovegrove, 2005; Hadayeghi, 2009; Abdel-Aty et al., 2011; An et al., 2011):

$$E(C) = \beta_0 \times (Exposure)^{\beta_1} \times e^{\sum_{i=2}^{n} \beta_i x_i} \tag{5}$$

Where;

$E(C)$ is the expected crash frequency, β_0 and β_i are model parameters, $Exposure$ is the exposure variable (e.g. VHT, VKT or NOTs) and $x_i's$ are the other explanatory variables.

Logarithmic transformation of Equation (5) when considering only one exposure variable yields:

$$ln[E(C)] = ln(\beta_0) + \beta_1 ln(Exposure) + \beta_2 x_2 + \beta_3 x_3 + \ldots + \beta_n x_n \tag{6}$$

Several models were constructed to associate the relationship between crash frequency and the explanatory variables (Pirdavani et al., 2012). Coefficients were estimated by using a forward selection procedure by taking the intercept and one of the exposure variables for the starting point and then additional candidate variables were selected from the available data described in Table 5.1. The parameter estimations in the GLM models are often called global estimates as the spatial correlation is not taken into account.

5.4.2. Geographically Weighted Generalized Linear models

The Geographically Weighted form of Equation (6) would be:

$$ln[E(C)(\boldsymbol{l}_i)] = ln(\beta_0(\boldsymbol{l}_i)) + \beta_1(\boldsymbol{l}_i) ln(Exposure) + \beta_2(\boldsymbol{l}_i) x_2 + \cdots + \beta_n(\boldsymbol{l}_i) x_n \tag{7}$$

The output of these models will be different location-specific estimates for each case (here each TAZ). All variable estimates are functions of each location (here the centroid of each TAZ), $\boldsymbol{l}_\mathrm{i} = (x_\mathrm{i}, y_\mathrm{i})$ representing the x and y coordinates of the i^{th} TAZ. The main purpose of developing geographically weighted models is that these models allow the estimates to vary where different spatial correlation among the explanatory variables exists. If the aim is estimating parameters for a model at a specific location, expectedly the locations nearby this location have a greater impact on this estimation compared with the locations which are far from it. This impact can be expressed by a weighting function. This weighting function is conditioned on the location $\boldsymbol{l}_i$ and is therefore different for each location (Fotheringham et al., 2002). The weights are derived from a weighting scheme which is commonly referred to as a kernel. There are two kernels which are frequently used to generate the weighting scheme; the Gaussian and the bi-square functions which can be formulated as follows:

Gaussian function: $\mathrm{W_{ij}} = e^{-0.5(\frac{d_{\mathrm{ij}}}{\mathrm{b}})^2}$ (8)

bi-square function: $\mathrm{W_{ij}} = \begin{cases} (1-\left(\frac{d_{\mathrm{ij}}}{\mathrm{b}}\right)^2)^2 & \text{if} \quad d_{\mathrm{ij}} < b \\ 0 & \text{otherwise} \end{cases}$ (9)

Where $\mathrm{W_{ij}}$ represents the measure of contribution of location j when calibrating the model for location i, d_{ij} is the distance between locations i and j and b is the bandwidth (Fotheringham et al., 2002). It is reported in the literature (e.g. in Guo et al., 2008; Hadayeghi et al., 2010) that selection of the kernel function and accordingly the bandwidth is very critical as the model might be very sensitive to this selection. However, Fotheringham et al. (2002) indicated that regarding the fit of the model, the choice of a bandwidth is more important than the shape of the kernel. As a rule of thumb, when the sample locations are commonly positioned across the study area, then a kernel with a fixed bandwidth is a suitable choice for modeling. On the contrary, when the sample locations are clustered in the study area, it is generally better to apply an adaptive kernel; i.e., having larger bandwidth where sample locations are sparser and applying smaller bandwidth for denser sample locations. Adaptive bandwidth will be displayed as a quantile of the number of adjacent locations (TAZs) which will influence the weighting function (e.g. in Table 5.3 and for Model #6, the bandwidth value is 0.03369; this means that 3.369% of the adjacent TAZs, 74 TAZs out of 2200 TAZs, should be selected to calculate the weighting function for each TAZ).

Despite the fact that parameters estimation depends on the weighting function, selecting an appropriate bandwidth is a more crucial task. There are different approaches that can be used in bandwidth selection. Cross-validation (CV) is a technique in which the optimal bandwidth size is determined by minimizing the CV score which is formulated as follows:

$$CV = \sum_{i=0}^{n} (\mathrm{y_i} - \hat{\mathrm{y}}_{\neq \mathrm{i}})^2 \quad (10)$$

Where n is the number of TAZs and $\hat{y}_{\neq i}$ is the fitted value of y_i when the i^{th} case is left out during the calibration process.

Another method to derive the bandwidth which provides a trade-off between Goodness-of-fit and degrees of freedom is minimizing the Akaike Information Criterion (AIC) (Fotheringham et al., 2002). It is reported by (Nakaya et al., 2005) that in the case of local regression, given the fact that the degrees-of-freedom are likely to be small, including a small sample bias adjustment in the AIC definition is recommended. This will lead to a corrected AIC often referred to as AICc. The formulations of AIC and AICc are as follows:

$$AIC = D(b) + 2K(b) \tag{11}$$

and

$$AICc = AIC + 2\frac{K(b)(K(b)+1)}{n - K(b) - 1} \tag{12}$$

Where D and K are respectively the deviance and the effective number of parameters in the model with bandwidth b and n denotes the number of TAZs.

In this study, both the CV and AICc methods were applied to determine the most appropriate bandwidth. The results reveal that in case of applying the AICc method, the optimum derived bandwidths are very close to each other no matter which kernel function is used. The computed bandwidths following the CV approach are slightly different from the ones derived by the AICc approach. Since the model selection is based on the minimum AICc values, only the bandwidths derived by the AICc approach will be used in model development. At this stage, it is time to find the most appropriate bandwidth/kernel combination which leads to the best model fit. This is carried out by constructing different models in which different combinations of kernel functions and fixed/adaptive bandwidths are applied. The variables used to fit these models are the same variables found to be significant variables in our earlier research (Pirdavani et al., 2012). The results of these models are shown in Table 5.3 represented by the minimum, maximum, 1^{st} quartile, median and 3^{rd} quartile of the parameter estimates. These results will be discussed in the next section of this chapter.

Table 5.3 GWPR Models Based on Different Bandwidths and Kernel Functions

	Model #1	**Model #2**	**Model #3**	**Model #4**
Coefficients	**Estimates**	**Estimates**	**Estimates**	**Estimates**
(Intercept)	-6.13, -1.34 (-4.41,-3.77,-2.96)[a]	-5.039, -0.329 (-3.35,-2.84,-2.01)	-6.276, -1.47 (-4.405,-3.75,-2.97)	-5.281, -0.495 (-3.32,-2.87,-2.01)
log(Number of Trips)	0.153, 0.991 (0.37,0.49,0.59)	0.146, 0.998 (0.365,0.473,0.581)	0.177, 0.922 (0.38,0.504,0.59)	0.171, 0.9416 (0.37,0.485,0.571)
log(Motorways VKT)	-0.0208, 0,0307 (-0.007,0.0006,0.011)	-	-0.02, 0.0306 (-0.008,2e-4,0.011)	-
log(Other Roads VKT)	0.155, 0.452 (0.21,0.25,0.295)	-	0.154, 0.457 (0.213,0.252,0.296)	-
log(Motorways VHT)	-	-0.0377, 0.0524 (-0.016,-0.002,0.013)	-	-0.0424, 0.0523 (-0.02,-0.002,0.014)
log(Other Roads VHT)	-	0.1595, 0.4596 (0.228,0.269,0.334)	-	0.1735, 0.4651 (0.23,0.266,0.334)
Capacity	-1.1e-4, 6.76e-4 (1.7e-4,3.2e-4,4.3e-4)	-5.1e-5, 6.4e-4 (2.1e-4,3.4e-4,4.5e-4)	-1.4e-4, 6.47e-4 (1.7e-4,3.3e-4,4.3e-4)	1.14e-5, 6.24e-4 (2.1e-4,3.4e-4,4.4e-4)
Intersection	-0.006, 0.046 (0.021, 0.027,0.031)	-0.0046, 0.046 (0.02,0.027,0.029)	-0.0035, 0.0428 (0.021,0.028,0.029)	-0.0021, 0.0437 (0.02,0.026,0.029)
Income level	-0.582, 0.592 (-0.14,-0.075,0.01)	-0.6316, 0.589 (-0.14,-0.067,0.021)	-0.592, 0.644 (-0.13,-0.066,-2e-4)	-0.6385, 0.6723 (-0.13,-0.06,0.01)
Urban	-0.172,0.778 (0.282,0.398,0.528)	-0.2363, 0.6839 (0.196,0.327,0.458)	-0.1748, 0.7566 (0.278,0.39,0.508)	-0.25, 0.66 (0.19,0.319,0.435)
Suburban	-0.12, 0.473 (0.064,0.127,0.216)	-0.1451, 0.449 (0.055,0.12,0.199)	-0.093, 0.3856 (0.063,0.11,0.207)	-0.1297, 0.3679 (0.052,0.11,0.198)
Population	-1.1e-4, 9.18e-5 (-1e-6,2.3e-5,3.3e-5)	-1.13e-4, 9.19e-5 (-8e-7,2.2e-5,3.1e-5)	-9.7e-5, 8.44e-5 (1e-6,2.3e-5,3.2e-5)	-1.03e-4, 8.57e-5 (2e-6,2.2e-5,3.1e-5)
Bandwidth	144870.4	145241.1	457314.8	451464.4
AICc	11097.74	11032.39	11363.54	11236.86
MSPE	239.99	236.81	267.33	262.29
PCC	0.928	0.929	0.919	0.921

a: minimum, maximum, (1st quartile, median, 3rd quartile) of the parameter estimates.

Continued Table 5.3 GWPR Models Based on Different Bandwidths and Kernel Functions

	Model #5	**Model #6**	**Model #7**	**Model #8**
Coefficients	**Estimates**	**Estimates**	**Estimates**	**Estimates**
(Intercept)	-6.32, -1.156 (-4.39,-3.64,-2.84)	-5.215, -0.1736 (-3.26,-2.75,-1.92)	-5.659, -1.627 (-4.492,-3.65,-2.75)	-4.348, -0.728 (-3.35,-2.75,-1.72)
log(Number of Trips)	0.1375, 0.7652 (0.37,0.48,0.59)	0.1471, 0.7665 (0.36,0.462,0.57)	0.201, 0.714 (0.37,0.48,0.597)	0.227, 0.7085 (0.36,0.459,0.578)
log(Motorways VKT)	-0.027, 0,0217 (-0.006,0.001,0.0098)	-	-0.022, 0.0245 (-0.0066,7e-4,0.012)	-
log(Other Roads VKT)	0.1298, 0.4188 (0.21,0.25,0.303)	-	0.1517, 0.3799 (0.222,0.248,0.303)	-
log(Motorways VHT)	-	-0.0385, 0.0366 (-0.013,-0.002,0.011)	-	-0.0423, 0.0395 (-0.01,-0.003,0.017)
log(Other Roads VHT)	-	0.1269, 0.4684 (0.229,0.27,0.343)	-	0.1862, 0.4058 (0.23,0.266,0.343)
Capacity	-1.5e-4, 7.61e-4 (1.7e-4,3.1e-4,4.3e-4)	-8.8e-5, 7.1e-4 (2.1e-4,3.3e-4,4.4e-4)	-3.15e-6, 6.69e-4 (1.5e-4,3.1e-4,4.3e-4)	8e-5, 6.29e-4 (2e-4,3.3e-4,4.5e-4)
Intersection	0.005, 0.052 (0.02, 0.026,0.031)	-0.0042, 0.053 (0.02,0.026,0.03)	0.0062, 0.0445 (0.019,0.025,0.045)	0.007, 0.0428 (0.019,0.024,0.028)
Income level	-0.526, 0.498 (-0.195,-0.072,0.01)	-0.5875, 0.5099 (-0.19,-0.064,0.023)	-0.397, 0.53 (-0.16,-0.074,0.007)	-0.4328, 0.5135 (-0.15,-0.07,0.01)
Urban	-0.137,0.783 (0.291,0.394,0.56)	-0.2487, 0.6552 (0.204,0.33,0.48)	-0.221, 0.733 (0.301,0.4,0.558)	-0.165, 0.614 (0.22,0.33,0.468)
Suburban	-0.102, 0.384 (0.07,0.145,0.233)	-0.1299, 0.376 (0.063,0.139,0.215)	-0.081, 0.3466 (0.076,0.14,0.223)	-0.1237, 0.3137 (0.067,0.13,0.204)
Population	-5.4e-5, 9.23e-5 (9.6e-7,2.2e-5,3.7e-5)	-5.6e-5, 9.26e-5 (-2e-6,2.1e-5,3.5e-5)	-3.2e-5, 8.58-5 (1e-6,2.7e-5,3.2e-5)	-2.5e-5, 7.68e-5 (9e-7,2.6e-5,3.6e-5)
Bandwidth	0.03371	0.03369	0.21565	0.23607
AICc	10713.24	10605.15	11267.41	11301.82
MSPE	238.27	234.82	255.36	258.1
PCC	0.929	0.93	0.923	0.923

Models presented in Table 5.3 are defined as follows:

- Model #1: GWPR model with fixed bandwidth, Gaussian kernel function and VKT as the exposure variable
- Model #2: GWPR model with fixed bandwidth, Gaussian kernel function and VHT as the exposure variable
- Model #3: GWPR model with fixed bandwidth, bi-square kernel function and VKT as the exposure variable
- Model #4: GWPR model with fixed bandwidth, bi-square kernel function and VHT as the exposure variable
- Model #5: GWPR model with adaptive bandwidth, Gaussian kernel function and VKT as the exposure variable
- Model #6: GWPR model with adaptive bandwidth, Gaussian kernel function and VHT as the exposure variable
- Model #7: GWPR model with adaptive bandwidth, bi-square kernel function and VKT as the exposure variable
- Model #8: GWPR model with adaptive bandwidth, bi-square kernel function and VHT as the exposure variable

5.4.3. Measuring the Goodness-of-fit

There are a number of Goodness-of-fit measures which are used to evaluate the model performance.

The Pearson's product moment correlation coefficient which is often referred to as PCC is a measure of linear relationship between two variables. This measure is calculated to quantify the correlation between the observed and predicted NOICs. A higher PCC value for a model indicates that this model predicts the observed data better. The PCC is defined as follows:

$$PCC = \frac{\sum_{i=1}^{n}(\mathrm{P_i} - \overline{\mathrm{P}})(\mathrm{O_i} - \overline{\mathrm{O}})}{\sqrt{\sum_{i=1}^{n}(\mathrm{P_i} - \overline{\mathrm{P}})^2}\sqrt{\sum_{i=1}^{n}(\mathrm{O_i} - \overline{\mathrm{O}})^2}} \tag{13}$$

Where $\mathrm{P_i}$ is the predicted NOICs for i^{th} TAZ, $\overline{\mathrm{P}}$ is the average predicted NOICs, $\mathrm{O_i}$ is the observed NOICs for i^{th} TAZ, $\overline{\mathrm{O}}$ is the average observed NOICs and n denotes the total number of TAZs.

Another measure that has been used in comparative analysis between different models is the Mean Squared Prediction Error (MSPE) (Hadayeghi, 2009). The MSPE is the sum of the squared differences between predicted and observed crashes divided by the sample size. A lower value of MSPE indicates a better model fit. The MSPE is defined as follows:

$$MSPE = \sum_{i=1}^{n} (\mathrm{P_i} - \mathrm{O_i})^2 \Big/ \mathrm{n} \qquad (14)$$

The measures of goodness-of-fit are calculated and the results for different models are shown in Tables 5.3 and 5.4.

5.5. Discussion on Model Results

5.5.1. Finding the Best Fitted Model

A common rule-of-thumb in the use of AICc is that if the difference in AICc values between two models is more than 2, there is a substantial difference in the performance of the two models (Nakaya et al., 2005). As can be seen in Table 5.3, model #6 outperforms all other models by means of having the minimum AICc value which is far lower than the AICc values of all other models. Model #6 is fitted using an adaptive bandwidth and a Gaussian kernel function for the weighing function. It can be concluded that for the given data, utilizing adaptive bandwidth and the Gaussian kernel function will result in the best model fit. Therefore, this combination will be used to fit different models by which we aim to compare the performance of GWPR models against GLM models.

Comparable with our previous research (Pirdavani et al., 2012) in which different GLM models were developed, similar GWPR models are constructed to evaluate the benefits of accounting for the spatial autocorrelation which exists in the data. Here, the GWPR models and their corresponding GLM models are summarized and their performances together with their goodness-of-fit measures are presented in Table 5.4. Table 5.4 shows the models in which both trip and flow related exposure measures are taken into account. As mentioned earlier, there are a number of measures that have been used in comparative analysis between different models; e.g. PCC and MSPE. Comparing AICc, PCC and MSPE measures in Table 5.4 shows that GWPR models outperform the GLM models. Between these models, Model #6 outperforms Model #5 by having lower AICc and MSPE and higher PCC measures.

Table 5.4 GLM and GWPR Models Based on Both Trip- and Flow-Based Exposure Variables

	Model #9	**Model #10**	**Model #5**	**Model #6**
Coefficients	**Estimates**	**Estimates**	**Estimates**	**Estimates**
(Intercept)	-4.141e+00	-2.886e+00	-6.32, -1.156 (-4.39,-3.64,-2.84) [a]	-5.215, -0.1736 (-3.26,-2.75,-1.92)
log(Number of Trips)	4.520e-01	4.676e-01	0.1375, 0.7652 (0.37,0.48,0.59)	0.1471, 0.7665 (0.36,0.462,0.57)
log(Motorways VKT)	7.744e-03	-	-0.027, 0,0217 (-0.006,0.001,0.0098)	-
log(Other Roads VKT)	3.132e-01	-	0.1298, 0.4188 (0.21,0.25,0.303)	-
log(Motorways VHT)	-	7.717e-03	-	-0.0385, 0.0366 (-0.013,-0.002,0.011)
log(Other Roads VHT)	-	3.040e-01	-	0.1269, 0.4684 (0.229,0.27,0.343)
Capacity	3.894e-04	4.220e-04	-1.5e-4, 7.61e-4 (1.7e-4,3.1e-4,4.3e-4)	-8.8e-5, 7.1e-4 (2.1e-4,3.3e-4,4.4e-4)
Intersection	2.888e-02	2.844e-02	0.005, 0.052 (0.02, 0.026,0.031)	-0.0042, 0.053 (0.02,0.026,0.03)
Income level	-1.071e-01	-1.056e-01	-0.526, 0.498 (-0.195,-0.072,0.01)	-0.5875, 0.5099 (-0.19,-0.064,0.023)
Urban	3.520e-01	2.287e-01	-0.137,0.783 (0.291,0.394,0.56)	-0.2487, 0.6552 (0.204,0.33,0.48)
Suburban	9.095e-02	5.712e-02	-0.102, 0.384 (0.07,0.145,0.233)	-0.1299, 0.376 (0.063,0.139,0.215)
Population	2.293e-05	2.340e-05	-5.4e-5, 9.23e-5 (9.6e-7,2.2e-5,3.7e-5)	-5.6e-5, 9.26e-5 (-2e-6,2.1e-5,3.5e-5)
Bandwidth	-	-	0.03371	0.03369
AIC	16918	16921	10641	10533
MSPE	489.74	482.41	238.27	234.82
PCC	0.869	0.871	0.929	0.931

a: minimum, maximum, (1^{st} quartile, median, 3^{rd} quartile) of the parameter estimates.

Model #9: GLM Negative Binomial model using VKT and NOT

Model #10: GLM Negative Binomial model using VHT and NOT

Model #5: GWPR model with adaptive bandwidth and Gaussian kernel using VKT and NOT

Model #6: GWPR model with adaptive bandwidth and Gaussian kernel using VHT and NOT

5.5.2. Further Investigation on the Selected Model

As stated earlier, the results of the GWPR models are presented as sets of locally estimated coefficients often referred to as 5-number summaries (minimum, 1st quartile, median, 3rd quartile and maximum). Unlike spatially stationary models (e.g. GLM models) which have a unique estimate for each variable, variable estimates for GWPR models vary across the space and sometimes have different and unexpected signs.

As can be seen from the results of different GWPR models, some variable estimates have different signs. Unlike some other studies (e.g. Hadayeghi et al., 2010) which report on this trend to happen for their most significant variables, in our study all of the most significant variables have similar signs in line with our expectations. "log(Number of Trips)" and "log(Other Roads VHT)" as the most significant variables always have positive signs for all local estimates. However, the signs of other coefficients are not always the same. To have a better view on these differences, local variable estimates are depicted in Figure 5.1. This issue which is often referred to as "the problem with counterintuitive signs" has already been reported in many studies (e.g. in Wheeler and Tiefelsdorf, 2005; Chow et al., 2006; Wheeler and Calder, 2007; Hadayeghi et al., 2010).

One explanation for this problem would be the existence of multicollinearity among some variables for some locations. It is quite possible that some variables at some locations are locally correlated while no global multicollinearity is observed among the explanatory variables. The other reason could be due to the basis of calibrating GWPR models. GWPR models are somehow developed locally based on the defined bandwidth and the chosen weighing function described earlier. Presumably for some locations, some variables might not be significant variables; therefore, it is possible that the local models produce some unexpected variable signs for those insignificant variables. The latter reason can be easily investigated by calculating t-Statistics.

Figure 5.2 depicts the results of p-values for all explanatory variables of Model #6. In this figure, significant variables at any location are colored in green while insignificant variables are depicted in red. By comparing Figures 5.1 and

5.2 it can be concluded that the p-values for all of the locations with unexpected coefficient signs are insignificant at the 95% confidence level. For instance, the variable "Urban" is expected to have a positive association with the crash frequency (Huang et al., 2010). As can be seen from Figure 5.1, only a few TAZs show negative association with the NOICs (TAZs colored in light blue). When comparing this figure with the corresponding map in Figure 5.2, it is evident that in these TAZs, "Urban" is not a significant predictor. The same phenomenon can be perceived for all other explanatory variables that the TAZs with unexpected variable signs are always the TAZs on which those variables are insignificant predictors.

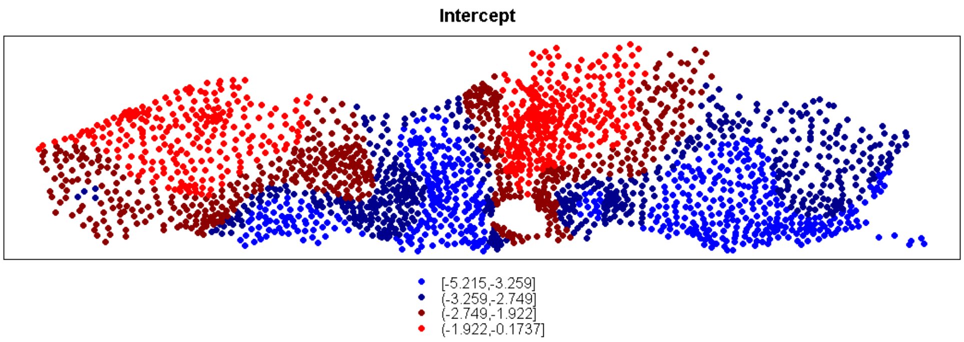

Figure 5.1.a Graphical representations of local variable estimates for "Intercept" from Model #6.

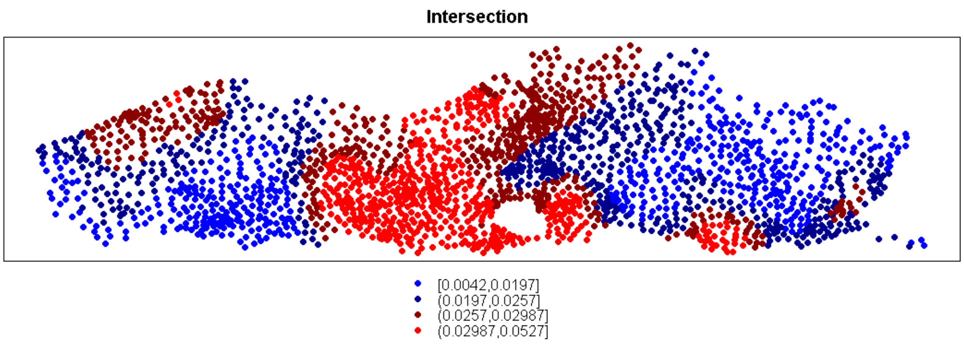

Figure 5.1.b Graphical representations of local variable estimates for "Intersection" from Model #6.

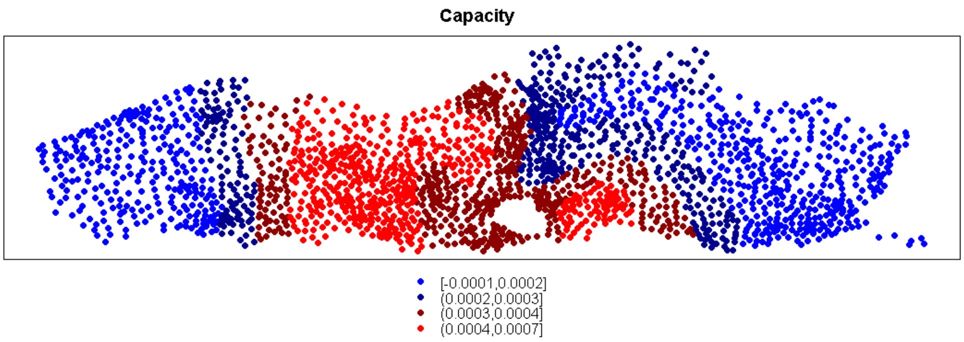

Figure 5.1.c Graphical representations of local variable estimates for "Capacity" from Model #6.

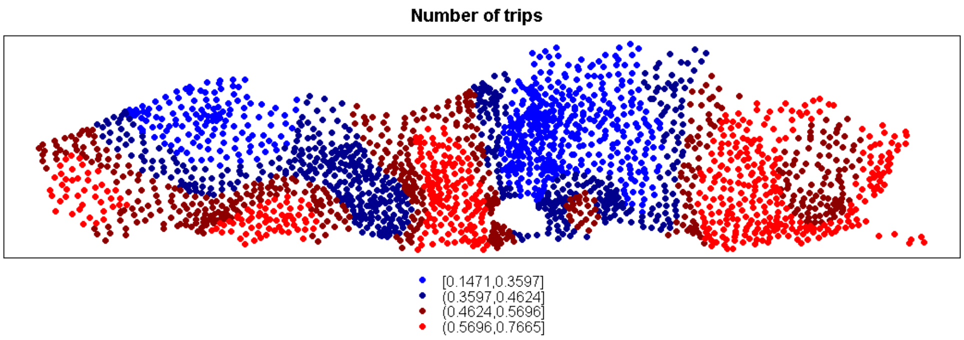

Figure 5.1.d Graphical representations of local variable estimates for "Number of car trips" from Model #6.

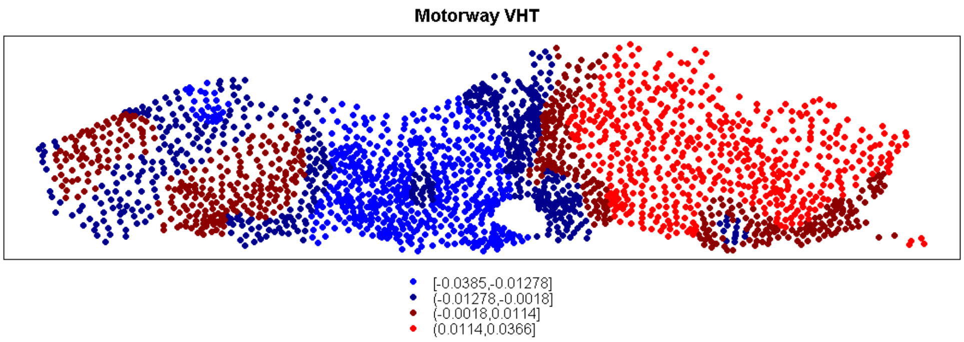

Figure 5.1.e Graphical representations of local variable estimates for "Motorway VHT" from Model #6.

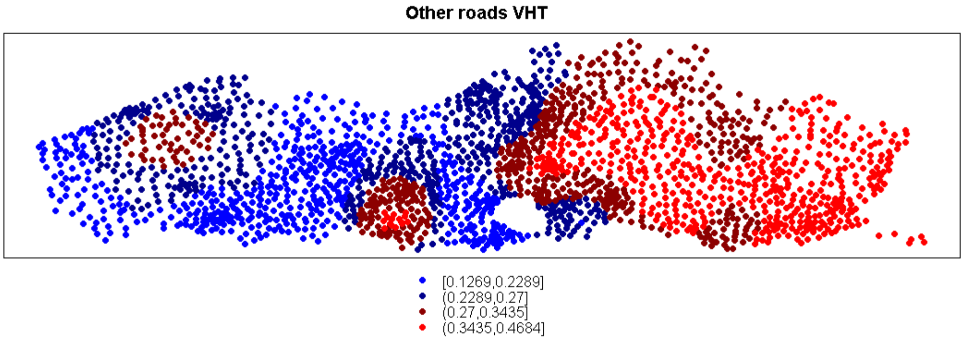

Figure 5.1.f Graphical representations of local variable estimates for "Other roads VHT" from Model #6.

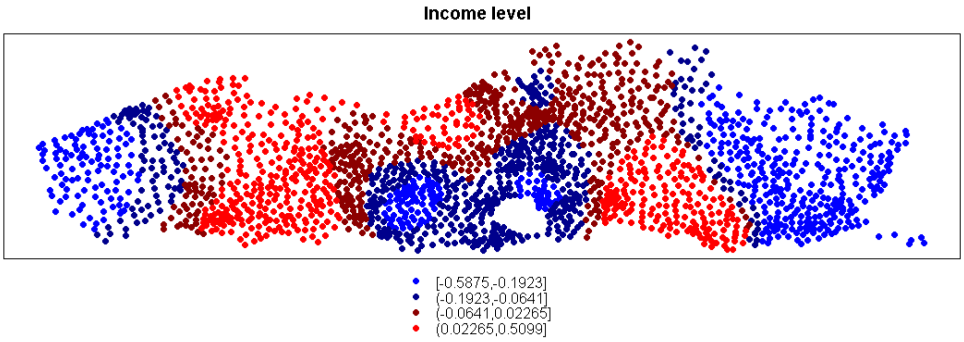

Figure 5.1.g Graphical representations of local variable estimates for "Income level" from Model #6.

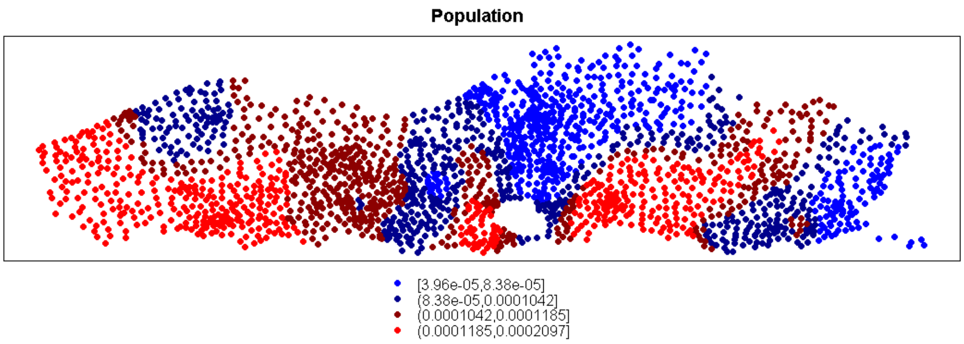

Figure 5.1.h Graphical representations of local variable estimates for "Population" from Model #6.

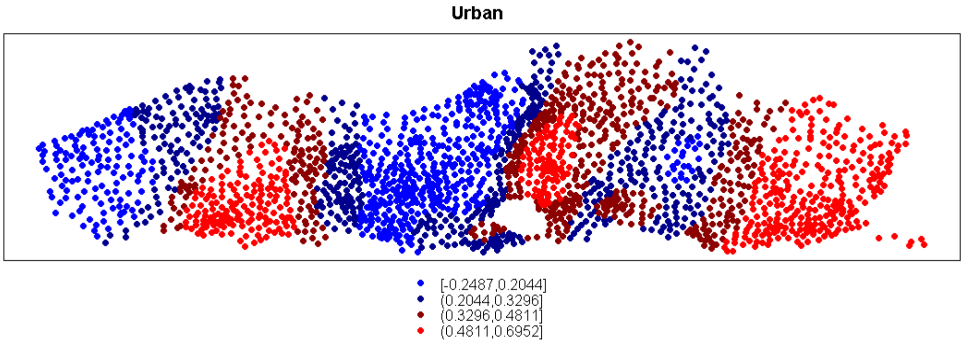

Figure 5.1.i Graphical representations of local variable estimates for "Urban" from Model #6.

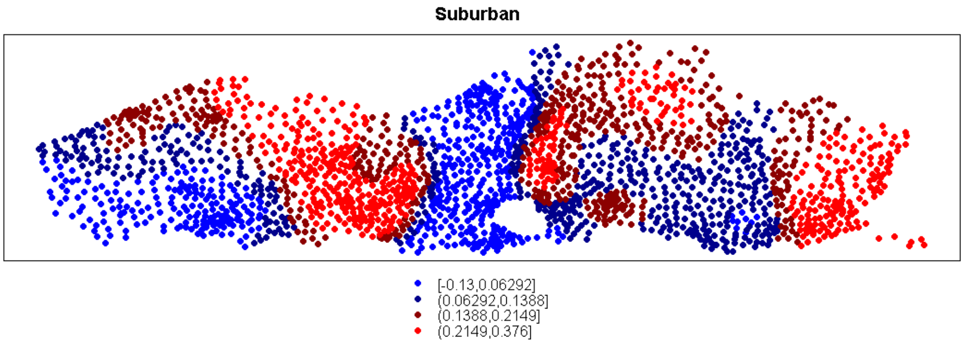

Figure 5.1.j Graphical representations of local variable estimates for "Suburban" from Model #6.

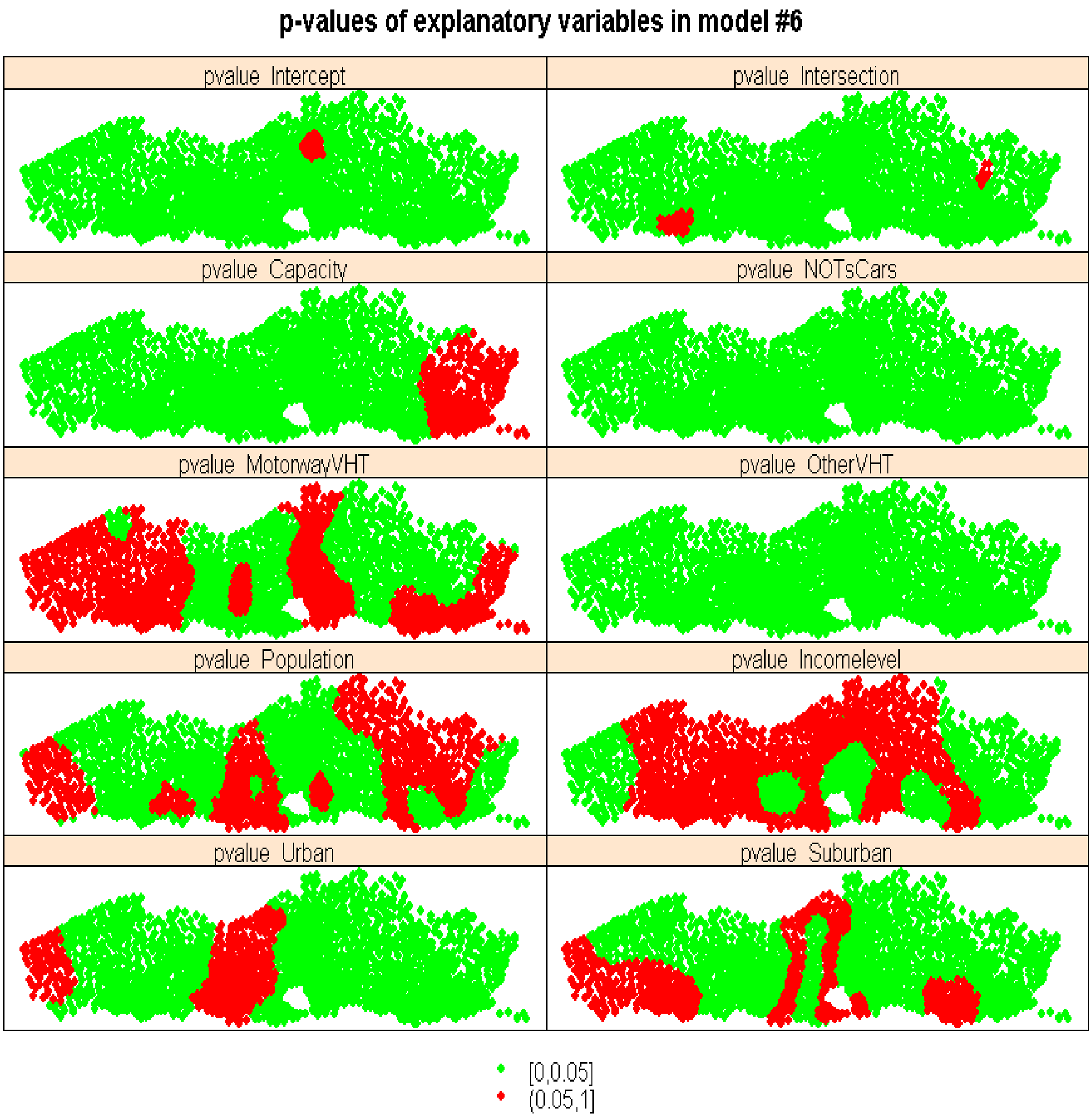

Figure 5.2 Graphical representations of p-values of all explanatory variables for Model #6.

Generally, the GWPR models outperform the GLM models because of their capability in capturing spatial heterogeneity. As can be seen from Figure 5.3, observed and predicted NOICs are having almost the same pattern. This is an indication of how well these models are able to fit the observed data.

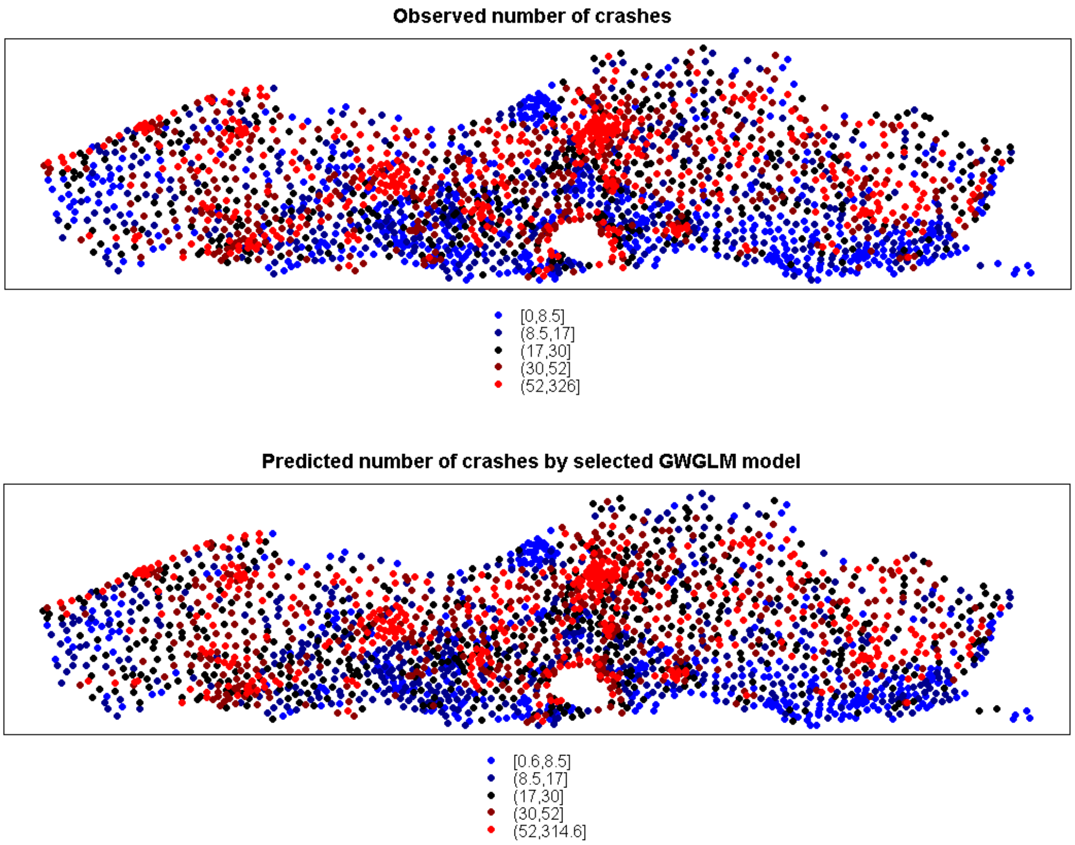

Figure 5.3 Observed and predicted (results from Model #6) NOICs.

5.6. Conclusions

The results of GLM models are a set of fixed coefficient estimates which represent the average relationship between the dependent variable and other explanatory variables for all locations. These relationships are assumed to be constant across space. However, these explanatory variables are often found to be spatially heterogeneous especially when the study area is large enough to cover different traffic volume, urbanization and socio-demographic patterns. In this chapter, we first aim to investigate the presence of spatial variation of dependent and different explanatory variables which are being used in developing crash prediction models. This was carried out by computing Moran's I statistics for dependent and selected explanatory variables. The results revealed the necessity of considering spatial correlation when developing crash prediction models. Therefore, different Geographically Weighted Poisson Regression (GWPR) models were developed, using different exposure, network and socio-demographic variables. GWPR models allow the estimations to vary where different spatial correlation among the variables exists. Hence, the association between NOICs and other explanatory variables are formed by means of different local models for each TAZ. Comparing all developed models show that GWPR models always perform better than GLM models. This is due to the fact that GWPR models are capable of capturing the spatial heterogeneity of crashes.

In construction of GWPR models different actions need to be taken. An important task is computing the most proper bandwidth and selecting the most suitable kernel function. For the current data, adaptive bandwidth with Gaussian kernel function resulted in the best model fit. Since AICc value is the measure to select the best fitted model, the AICc method is followed to compute bandwidth. This method relies on producing a minimum AICc measure and has advantages compared to the cross-validation method.

The selected GWPR model (Model #6) predicts the number of crashes properly. The Pearson's product moment correlation coefficient for this model is 0.93 which is quite high. Comparing the graphical representation of observed and predicted NOICs shows a very similar pattern; this indicates that the NOICs are very well predicted by the explanatory variables included in the model for most of the TAZs across the study area.

Despite the promising results of GWPR models and their potential in capturing spatial variation, these models are not spatially transferable. This is due to the fact that GWPR models produce local parameter estimates (i.e. local models) for each TAZ which are influenced by their adjacent TAZs. Therefore, different models need to be developed for different study areas.

References

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011. Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis & Prevention 38, 618–625.

Aguero-Valverde, J., Jovanis, P.P., 2008. Analysis of Road Crash Frequency with Spatial Models. Transportation Research Record: Journal of the Transportation Research Board 2061, 55–63.

Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Accident Analysis & Prevention 35, 537–547.

An, M., Casper, C., Wu, W., 2011. Using Travel Demand Model and Zonal Safety Planning Model for Safety Benefit Estimation in Project Evaluation. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Blainey, S., 2010. Trip end models of local rail demand in England and Wales. Journal of Transport Geography 18, 153–165.

Chow, L.-F., Zhao, F., Liu, X., Li, M.-T., Ubaka, I., 2006. Transit Ridership Model Based on Geographically Weighted Regression. Transportation Research Record: Journal of the Transportation Research Board 1972, 105–114.

Clark, S.D., 2007. Estimating local car ownership models. Journal of Transport Geography 15, 184–197.

Cottrill, C.D., Thakuriah, P. (Vonu), 2010. Evaluating pedestrian crashes in areas with high low-income or minority populations. Accident Analysis & Prevention 42, 1718–1728.

De Guevara, F.L.D., Washington, S., Oh, J., 2004. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in

Tucson, Arizona. Transportation Research Record: Journal of the Transportation Research Board 1897, 191–199.

Delmelle, E.C., Thill, J.-C., 2008. Urban Bicyclists: Spatial Analysis of Adult and Youth Traffic Hazard Intensity. Transportation Research Record: Journal of the Transportation Research Board 2074, 31–39.

Du, H., Mulley, C., 2006. Relationship Between Transport Accessibility and Land Value: Local Model Approach with Geographically Weighted Regression. Transportation Research Record: Journal of the Transportation Research Board 1977, 197–205.

Erdogan, S., 2009. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. Journal of Safety Research 40, 341–351.

Flahaut, B., 2004. Impact of infrastructure and local environment on road unsafety: Logistic modeling with spatial autocorrelation. Accident Analysis & Prevention 36, 1055–1066.

Flahaut, B., Mouchart, M., Martin, E.S., Thomas, I., 2003. The local spatial autocorrelation and the kernel method for identifying black zones: A comparative approach. Accident Analysis & Prevention 35, 991–1004.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression the analysis of spatially varying relationships. John Wiley & Sons Ltd, West Sussex, England.

Geurts, K., Thomas, I., Wets, G., 2005. Understanding spatial concentrations of road accidents using frequent item sets. Accident Analysis & Prevention 37, 787–799.

Guo, F., Wang, X., Abdel-Aty, M., 2010. Modeling signalized intersection safety with corridor-level spatial correlations. Accident Analysis & Prevention 42, 84–92.

Guo, L., Ma, Z., Zhang, L., 2008. Comparison of bandwidth selection in application of geographically weighted regression: a case study. Canadian Journal of Forest Research 38, 2526–2534.

Hadayeghi, A., 2009. Use of Advanced Techniques to Estimate Zonal Level Safety Planning Models and Examine Their Temporal Transferability, PhD thesis, Department of Civil Engineering, University of Toronto.

Hadayeghi, A., Shalaby, A., Persaud, B., 2003. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. Transportation Research Record: Journal of the Transportation Research Board 1840, 87–95.

Hadayeghi, A., Shalaby, A., Persaud, B., 2007. Safety Prediction Models: Proactive Tool for Safety Evaluation in Urban Transportation Planning Applications. Transportation Research Record: Journal of the Transportation Research Board 2019, 225–236.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. Accident Analysis & Prevention 42, 676–688.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accident Analysis & Prevention 38, 579–589.

Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-Level Crash Risk Analysis in Florida. Transportation Research Record: Journal of the Transportation Research Board 2148, 27–37.

Janssens, D., Wets, G., Timmermans, H.J.P., Arentze, T.A., 2007. Modelling Short-Term Dynamics in Activity-Travel Patterns: Conceptual Framework of the Feathers Model. Presented at the 11th World Conference on Transport Research, Berkeley CA, USA.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., 2008. Assessing the Impact of Fuel Cost on Traffic Demand in Flanders Using Activity-Based Models. Presented at the Travel Demand Management TDM, Vienna, Austria.

LaScala, E.A., Gerber, D., Gruenewald, P.J., 2000. Demographic and environmental correlates of pedestrian injury collisions: a spatial analysis. Accident Analysis & Prevention 32, 651–658.

Lee, J., Wong, D.W.S., 2001. Statistical analysis with ArcView GIS. John Wiley & Sons, Inc., United States of America.

Levine, N., Kim, K.E., Nitz, L.H., 1995a. Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. Accident Analysis & Prevention 27, 663–674.

Levine, N., Kim, K.E., Nitz, L.H., 1995b. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. Accident Analysis & Prevention 27, 675–685.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44, 291–305.

Lovegrove, G., Sayed, T., 2007. Macrolevel Collision Prediction Models to Enhance Traditional Reactive Road Safety Improvement Programs. Transportation Research Record: Journal of the Transportation Research Board 2019, 65–73.

Lovegrove, G.R., 2005. Community-Based, Macro-Level Collision Prediction Models, PhD thesis, University of British Columbia.

Lovegrove, G.R., Litman, T., 2008. Using Macro-Level Collision Prediction Models to Evaluate the Road Safety Effects of Mobility Management Strategies: New Empirical Tools to Promote Sustainable Development. Presented at the Transportation Research Board (TRB) 87th Annual Meeting, Washington D.C. USA.

Lovegrove, G.R., Sayed, T., 2006. Macro-level collision prediction models for evaluating neighbourhood traffic safety. Canadian Journal of Civil Engineering 33, 609–621.

Miaou, S.-P., Song, J.J., Mallick, B.K., 2003. Roadway Traffic Crash Mapping: A Space-Time Modeling Approach. Journal of Transportation and Statistics 6, 33–57.

Moons, E., Brijs, T., Wets, G., 2009. Identifying Hazardous Road Locations: Hot Spots versus Hot Zones. Presented at the International Conference on Computational Science and Its Applications (ICCSA), Perugia, Italy.

Moran, P. A. P., 1950. Notes on Continuous Stochastic Phenomena. *Biometrika* 37 (1-2), 17–23.

Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: Introducing crash generation concept. Accident Analysis & Prevention 42, 339–346.

Nakaya, T., Fotheringham, A.S., Brunsdon, C., Charlton, M., 2005. Geographically weighted Poisson regression for disease association mapping. Statistics in Medicine 24, 2695–2717.

Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention 36, 525–532.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis & Prevention 36, 973–984.

Páez, A., 2006. Exploring contextual variations in land use and transport analysis using a probit model with geographical weights. Journal of Transport Geography 14, 167–176.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., 2012. Developing Zonal Crash Prediction Models with a Focus on Application of Different Exposure Measures. Transportation Research Record: Journal of the Transportation Research Board, in press.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. Accident Analysis & Prevention 40, 1486–1497.

Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis & Prevention 45, 382–391.

Tobler W., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46(2), 234-240.

Wang, C., Quddus, M.A., Ison, S.G., 2009. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. Accident Analysis & Prevention 41, 798–808.

Wang, X., Abdel-Aty, M., 2006. Temporal and spatial analyses of rear-end crashes at signalized intersections. Accident Analysis & Prevention 38, 1137–1150.

Wheeler, D., Calder, C., 2007. An assessment of coefficient accuracy in linear regression models with spatially varying coefficients. Journal of Geographical Systems 9, 145–166.

Wheeler, D., Tiefelsdorf, M., 2005. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. Journal of Geographical Systems 7, 161–187.

Zhao, F., Park, N., 2004. Using Geographically Weighted Regression Models to Estimate Annual Average Daily Traffic. Transportation Research Record: Journal of the Transportation Research Board 1879, 99–107.

6. Evaluating the Road Safety Effects of a Fuel Cost Increase Measure by means of Zonal Crash Prediction Modeling

Pirdavani, A., et al., (2012), Accident Analysis & Prevention, DOI: 10.1016/j.aap.2012.04.008.

Pirdavani, A., et al., (2012), DVD Compendium of the 91st Transportation Research Board Conference, Washington D.C., USA.

6.1. ABSTRACT

Travel demand management (TDM) consists of a variety of policy measures that affect the transportation system's effectiveness by changing travel behavior. The primary objective to implement such TDM strategies is not to improve traffic safety, although their impact on traffic safety should not be neglected. The main purpose of this study is to evaluate the traffic safety impact of conducting a fuel-cost increase scenario (i.e. increasing the fuel price by 20%) in Flanders, Belgium. Since TDM strategies are usually conducted at an aggregate level, crash prediction models (CPMs) that are used to evaluate such strategies should also be developed at a geographically aggregated level. Therefore, zonal crash prediction models (ZCPMs) are considered to present the association between observed crashes in each zone and a set of predictor variables. To this end, an activity-based transportation model framework is applied to produce exposure metrics which will be used in the prediction models. This allows us to conduct a more detailed and reliable assessment while TDM strategies are inherently modeled in the activity-based models unlike traditional models in which the impact of TDM strategies are assumed. The crash data used in this study consist of fatal and injury crashes observed between 2004 and 2007. In this study, different ZCPMs are developed to predict the number of injury crashes (NOICs) (disaggregated by different severity levels and crash types) for both the null and the fuel-cost increase scenario. The results show a considerable traffic safety benefit of conducting the fuel-cost increase scenario apart from its impact on the reduction of the total vehicle kilometers travelled (VKT). A 20% increase in fuel

price is predicted to reduce the annual VKT by cars by almost 5 billion (11% of the total annual VKT in Flanders), which causes the total NOCs to decline by 2.8%.

6.2. Introduction

It is beneficial to know the consequences of TDM strategies e.g. on traffic safety, which is considered to be an external side effect. Road crashes are known as one of the negative impacts of growing travel demand. For many years, researchers have attempted to investigate this impact by predicting the NOCs based on patterns they learned from crashes that occurred in the past. From an ethical point of view, this reactive approach is not acceptable because it requires several years of crashes to occur in order to identify and treat safety problems. Providing a more proactive approach, capable of evaluating road safety at the planning-level is therefore essential. In the last few years, researchers and practitioners have increasingly applied this proactive approach. Dealing with traffic safety at the planning level requires the ability to integrate a crash predicting context into TDM strategies. TDM consists of several policies and strategies which aim to overcome transportation problems in different ways, e.g. changing travel behavior, making transportation systems more efficient or reducing travel demand. In general, TDM strategies are implemented to improve transportation systems' efficiency. However, their potential traffic safety impact should not be ignored. TDM strategies improve the transport system efficiency by means of mode shift (e.g. using public transportation instead of cars, biking for short distance trips or carpooling), travel time shift (e.g. avoiding traffic peak-hours by leaving home/the work place earlier or later), or travel demand reduction (e.g. teleworking) (VTPI, 2011). TDM strategies are usually performed and evaluated at geographically aggregated levels rather than merely the level of individual intersections or road sections. Therefore, the impact of adopting a TDM strategy on transportation or traffic safety should also be evaluated at a level higher than merely the local consequences.

The application of crash prediction models (CPMs) at a geographically aggregated level like traffic analysis zone (TAZ) leads to a zonal crash prediction model (ZCPM). Until now, ZCPMs have hardly ever been incorporated in TDM strategies. The main goal of this study is, therefore, to integrate ZCPMs with a fuel-cost increase scenario to evaluate the traffic safety effects of conducting such a TDM strategy by means of a simulation-based analysis of the impact of fuel price on the travel demand in Flanders, Belgium. This way, the behavioral impact of the TDM scenario in terms of traffic demand is incorporated in the analysis. By assigning traffic demand to the road network and using this information at zonal level, the impact of responses to TDM, such as changes in trip planning, route choice and modal choice are incorporated in the analysis. This study is an assessment exercise which illustrates the impact of a 20% increase in fuel-related costs on traffic safety. It is essential to note that this 20% increase in fuel-related costs (expressed in € per kilometer) is not an optimized value and is the result of a change in fuel price (expressed in € per liter) on the one hand and fuel economy (liter per kilometer) on the other hand.

To account for severity of crashes, different ZCPMs are developed at different severity levels; i.e. "fatal + severe injury" and "slight injury" crashes. Additionally, in order to represent the mode shift effects, crashes are disaggregated into two different types namely "Car-Car" and "Car-Slowmode" crashes ("Slowmode" comprises pedestrians and cyclists). Accordingly, four different ZCPMs are developed to explain the traffic safety impacts of the fuel-related cost increase at different crash type/severity levels.

It is necessary to indicate that the FEATHERS model (Bellemans et al., 2010) models the transportation demand of a population who is aware of the state of the transportation network. Hence, the assumed travel times during the activity-travel planning phase are in correspondence with the travel times obtained after assigning the total traffic demand to the road network (this is achieved through iteration). This means that the model is a steady state model and that no transients are modeled. Moreover, the model is a short term model in the sense that it does not assume a shift in the composition of the vehicle fleet as a result of the change in fuel cost. This assumption is justified by the slower time scale of vehicle fleet composition (in the order of several months to even years). Also changes in the location of businesses and/or the location choice for living (i.e. land use characteristics) occur at a far slower time scale than the adaptation of travel behavior to changing fuel cost triggers.

The structure of this chapter is as follows. Initially, we will review the literature. Then the activity-based model which is used in this study will be briefly introduced. In the next sections, the data preparation and the fuel-cost increase scenario evaluation process will be demonstrated. Finally, the results of this evaluation will be shown followed by the final conclusions and discussions.

6.3. Background

TDM strategies have hardly ever been implemented to improve traffic safety. Their main objectives are usually the reduction of congestion and emission, as well as travel cost and energy by means of reducing travel demand and consequently vehicle distance travelled. Nevertheless, apart from improving the efficiency of the transportation system and economic/environmental benefits, identifying the traffic safety impacts of a TDM strategy, can strengthen the implementation of such a strategy. It is a well-known fact in literature that road crashes are tightly linked to traffic exposure. Therefore, strategies that reduce travel demand or distance travelled, or cause a modal shift to a safer mode (e.g. from car to public transportation) tend to reduce the NOCs (Litman, 2006; 2011).

Lovegrove and Litman (2008) applied community-based, macro-level CPMs to calculate the road safety impacts of three mobility management strategies: smart growth, congestion pricing and improving transport options. They assumed the effect of implementing these strategies on different explanatory variables of the CPMs (e.g. they found that the smart growth strategy will decrease VKT by 15%). Based on these assumptions the expected NOCs have been calculated for each TDM strategy. The results indicated that mobility management strategies can significantly improve traffic safety.

Fuel-related costs are a major component of each motor vehicle's operating expenses. By increasing the fuel price as a TDM strategy, people tend to travel less by car, and instead use public transportation, carpool, or shift towards slow modes (biking and walking), etc. Thus, traffic crashes are expected to decrease as a result of a reduction in the number of car kilometers traveled. Fuel-related costs have an impact on traffic safety through changes in travel demand. Grabowski and Morrisey (2004) reported a relatively stable number of fatal motor vehicle crashes despite new traffic safety laws and vehicle innovations over a period of time. Their explanation was that the price of gasoline declined, which resulted in more vehicle miles traveled and potentially more fatalities. Chi et al. (2010) also investigated the impact of gasoline price changes on different types of crashes at a more disaggregated level for different ages and genders. In their reactive approach, they developed models to predict traffic crashes based on explanatory variables like exposure, gasoline price, alcohol consumption, seat belt usage, etc. It was concluded that an increase in gasoline price both has a short-term and intermediate-term effect on reducing total traffic crashes. One of the longer-term effects of a fuel cost increase is the change of the fleet composition to more fuel-economic vehicles, which can partially compensate the increased fuel price by an increased fuel economy. In literature it is described (Goodwin et al., 2004; Litman, 2010) that the fuel price elasticity of fuel consumption ranges from -0.25 to -0.6, the elasticity of fuel efficiency ranges from 0.3 to 0.4 and the vehicle mileage elasticity ranges from –0.1 to -0.3.

CPMs can be categorized in two different levels: the local level (road and intersection) and the regional level (e.g. TAZ). Usually, local level CPMs aim to predict the safety benefits/detriment of infrastructural improvements. These models are not typically designed to evaluate traffic safety impacts of TDM strategies; thus, the application of CPMs at a higher aggregation level will be more practical (Tarko et al., 2008). The application of CPMs at TAZ level has been initially introduced by Levine at al. (1995). In their study, a set of both socio-economic and road network variables were chosen to predict the NOCs in TAZs. They developed a linear relationship between different explanatory variables and the NOCs.

Recently, the application of ZCPMs became more popular amongst researchers. Several researchers examined the association of a collection of network infrastructure variables, socio-demographic and socio-economic variables and weather conditions with the NOCs in TAZs (e.g. in Amoros et al., 2003; Noland and Oh, 2004; Noland and Quddus, 2004; Aguero-Valverde and Jovanis, 2006; Quddus, 2008; Wier et al., 2009; Huang et al., 2010). It was found that the number of lanes, road length and road density were significantly correlated with the NOCs. As for the demographic variables, it was found that TAZs with a higher percentage of population under the poverty level and a higher percentage of population in the young and also elderly age groups have the potential of increasing crash risk. It was also found that the traffic safety situation is worse for TAZs with lower income and education levels and a higher unemployment rate compared to relatively affluent TAZs. In another study by Wier et al. (2009) it was shown that traffic volume, population size, the proportion of arterial streets without public transit, the proportion of population living in poverty, and the number of people aged 65+ as percentage of the total population, were significantly good predictors of crashes. Moreover, Noland and Quddus (2004) concluded that TAZs with high employment density had more traffic crashes, while in urbanized more densely populated TAZs fewer crashes were observed. De Guevara et al. (2004) developed planning-level ZCPMs for the city of Tucson, Arizona. They considered many socio-demographic and road network variables in their model construction. They concluded that predictors such as population density, the number of persons younger than 17 years old as a percentage of the total population, the number of employees, the intersection density, the percentage of miles of principal arterials, the percentage of miles of minor arterials and the percentage of miles of urban collectors are significant predictors for the NOCs.

Hadayeghi et al. (2003; 2006; 2007; 2010a; 2010b) have been working on ZCPMs for several years. In one of their first studies, it was shown that the number of accidents in a TAZ increases when the VKT, major and minor road length, total employed labor force, household population, and intersection density increase whereas it decreases with a higher posted speed and a higher level of congestion in the TAZ (Hadayeghi et al., 2003). Hadayeghi et al. (2006) investigated the temporal transferability of the ZCPMs by applying models constructed on 1996 data to predict the NOCs for each TAZ in 2001 for the City of Toronto. They concluded that the models are not transferable statistically but that VKT, socioeconomic and demographic parameters are significantly stable over time. In another research, twenty-three regression models were developed to examine the relationships between several types of transportation planning variables and collision frequency. Models were developed for each planning category individually and in combination with other categories. A comparison of the models' performance showed that the comprehensive models are performing statistically better than the individual models. The results showed the potential of planning-level safety models to provide decision support tools for planners to consider safety in the planning phase (Hadayeghi et al., 2007). Hadayeghi at al. (2010a) conducted the same research but this time they applied geographically weighted Poisson regression (GWPR) instead of taking a generalized linear modeling (GLM) approach. The major difference between these two types of models is that GWPR models allow the model coefficient estimates to vary spatially for each TAZ. This very important additional attribute of these models provides some extra information as it takes the spatial location of a crash into consideration.

Lovegrove and Sayed (2006) concluded that quantifying the relationship between the zonal characteristics such as exposure, network, socio-demographic and TDM variables and crashes at a zonal level provides a predictive tool to predict the NOCs in a TAZ. They used GLM techniques to develop ZCPMs for both urban and rural areas across the Greater Vancouver Regional District (GVRD). Their results show that increasing signal density, intersection density per unit area and per lane kilometers, arterial-local intersections in rural areas and total arterial road lane kilometers will lead to an increase in the NOCs. On the contrary, an increase in the number of three-leg intersections and local road lane kilometers will decrease the NOCs in a TAZ. Lovegrove and Sayed (2007) further developed a set of ZCPMs for a "black-spot" study in the GVRD. These sets of ZCPMs consist of an exposure variable (e.g. VKT) and other network, socio-demographic and TDM variables. The results of this study also confirmed that ZCPMs have the potential to complement traditional reactive road safety improvement programs.

Recently, some researchers constructed ZCPMs by associating the NOCs in a TAZ with trip production/attraction and other network characteristics. Abdel-

Aty et al. (2011a) identified and prioritized important variables which can be associated with crashes per TAZ by means of the Classification and Regression Trees technique. It was shown that this methodology will be helpful in incorporating proactive safety measures for long-range transportation planning. They (Abdel-Aty et al., 2011b) also developed different ZCPMs for different crash severity levels and concluded that different sets of predictors should be considered based on the type or severity of crashes (e.g. total trip productions and attractions provide a better model fit for the total and peak hour crashes while severe crashes were best predicted by different trip-related variables). Naderan and Shahi (2010) investigated the possibility of associating travel demand in urban areas with crash frequencies in each TAZ. They developed a series of ZCPMs using the number of trips (NOTs) produced/attracted as predictors. They concluded that these models provide the basic tool for evaluating TDM scenarios in urban transportation planning in terms of traffic safety as the application of a specific TDM scenario may reduce trip productions of a specific motive. The drawback of considering only trips as an exposure variable is that the impact of trip time, trip length, route choice, intrazonal traffic and transit traffic on a TAZ will be neglected. The number of produced/attracted trips might be an acceptable indicator of how busy or active a TAZ is or how much people are exposed to dangerous situations, but it always leaves out the effects of through traffic which is just passing through a TAZ neither having their origin or destination in that TAZ.

Although most of the above-mentioned studies were trying to demonstrate their potential as a predictive tool at the planning level, so far not much attention has been paid to the application of these models to evaluate the effect of TDM's on traffic safety. There are very few attempts at estimating the road safety benefits of applying a specific TDM strategy. In a study conducted by Lovegrove and Litman (2008), they assumed the effect of implementing these strategies on different explanatory variables of the CPMs. Based on these assumptions the expected NOCs were calculated for each TDM strategy. For instance, it was concluded that a smart growth strategy of more compact and multi-modal land use development patterns may increase traffic safety by means of reducing crash frequency per capita by 20% and 29% for total and severe crashes respectively. An et al. (2011) found vehicle hours travelled (VHT), the number of intersections and the number of households with low income levels to be correlated with the NOCs in TAZs. After running two add-capacity projects in the Pikes Peak region and applying the results in their developed ZCPMs for the do-nothing scenario and both project scenarios, total crashes for both projects were estimated to decrease respectively by 0.1% and 0.06% when compared with the do-nothing scenario.

According to literature, exposure is the most important predictor of crashes. Therefore, having a more informative measure of exposure is expected to result in a better crash prediction. When a TDM scenario is performed, it basically changes the exposure compared with the null scenario. Thus, it is essential to predict the exposure metrics as accurately as possible. Activity-based models help with this as they are able to simulate the scenarios and in this case, they model the decision process of individuals with respect to the changes in fuel price. This is the key advantage of applying activity-based models rather than making educated guesses about the impact of fuel-related cost changes on travel demand in order to obtain exposure. In the next section, the activity-based model is briefly introduced and its contribution to the fuel-cost increase scenario evaluation process is described.

6.4. Impact of Fuel-Related Cost on Traffic Demand

Traditionally, travel was assumed to be the result of four subsequent decisions which were modeled separately, also referred to as four-step models. More recently, several researchers claimed that travel has an isolated role in these models and the reason why people undertake trips is neglected completely. This is why activity-based models have been taken into consideration. The main difference between four-step models and activity-based transportation models is that the latter try to predict interdependencies between several facets of activity profiles (Davidson et al., 2007). Hence, activity-based models are designed to keep the linkages between the travel decisions of individual members of a single household. Interactions among family members such as the use of household vehicles, sharing household responsibilities or performing joint activities, often affect and in many cases largely determine people’s travel. Four-step models that ignore such linkages, misstate people’s responses to TDM strategies. It is shown that activity-based models are capable of treating TDM strategies and policy issues whereas four-step models become ineffective (Vovsha and Bradley, 2006).

6.4.1. FEATHERS Framework

The FEATHERS (Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS) framework (Janssens et al., 2007) was developed in order to facilitate the development of activity-based models for transportation demand in Flanders, Belgium. The scheduling engine that is currently implemented in the FEATHERS framework is based on the scheduling engine that is present in the Albatross system (Arentze and Timmermans, 2004). Currently, the framework is fully operational at the level of Flanders. The real-life representation of Flanders is embedded in an agent-based simulation model which consists of over six million agents, each agent representing one member of the Flemish population. A sequence of 26 decision trees, derived by means of the CHi-squared Automatic Interaction Detector (CHAID) algorithm, is used in the scheduling process and decisions are based on a number of attributes of the individual (e.g. age, gender), of the household (e.g. number of cars) and of the geographical zone (e.g. population density, number of shops). For each agent with its specific attributes, the model simulates whether an activity (e.g. shopping, working, leisure activity …) is going to be carried out or not. Subsequently, amongst others, the location, transport mode (available modes in FEATHERS are "car driver", "car passenger", "public transportation" and "slow mode" including pedestrians and cyclists) and duration of the activity are determined, taking into account the attributes of the individual (Kochan et al., 2008). Traffic demand is subsequently assigned to the road network in such a way that an equilibrium is established between transportation demand and supply (Bellemans et al., 2010), which results in a time-dependent traffic state on the road network. In order to run, calibrate and validate the activity-based model, three major types of data are required (Kochan et al., 2011); data describing the environment (e.g. population density, level of service of the transportation networks), a synthetic population which is simulated and activity-travel data originating from a representative sample of the population from which the human behavior is derived.

6.4.2. Implementation of Fuel-Cost Increase Scenario in FEATHERS

An important asset of activity-based models in this context is their integrated approach towards activities and travel. Due to this approach, it can be taken into account that certain trips, which are linked to activities that are not so flexible (e.g. work activities) are less likely to be altered under changing traffic system conditions than others (e.g. leisure activities). In addition, activity-based models are not only able to predict a change in the demand for travel, but they also predict shifts between different modes of transport and the reallocation of activities due to the imposed measures. Providing a structured approach to agent-based modeling of activities and travel for individuals, the FEATHERS framework is able to account for TDM strategies. For instance, when applying a fuel cost increase scenario, FEATHERS can predict the impact on the NOTs, modal shift and changes in trip time and length. Price changes can have an impact on different facets of travel, affecting the NOTs people undertake, their destination, route, mode, travel time, type of vehicle (including size, fuel efficiency and fuel type) and parking location and duration. Therefore, in order to predict the impact of price changes like fuel price, the scheduling engine has to be structured to account for those changes. In this scheduling engine, price and cost parameters are incorporated in the decision trees related to activity selection, timing, trip-chaining, location and mode choices. The extended decision trees or Parametric Action Decision Trees combine conventional decision trees and parametric action assignment rules yielding a model that is sensitive for travel-costs scenarios (Arentze and Timmermans, 2005).

In this study, fuel-related cost is assumed to increase by 20% as a result of an increase in fuel price. We consider the short term effect and can as such neglect the rebound effect caused by a changing fuel economy of the fleet (i.e. the fuel economy is considered to remain constant, which results in the reduction of fuel consumption to be equal to the reduction of the VKT). One might question how the impact of this global fuel price increase will be sensible at zonal level. As mentioned earlier, each zone has its own characteristics: the level of income, availability of public transportation, major activity types, etc.

are different from zone to zone. These differences result in different travel behavior and more specifically different mode choice by the inhabitants of each zone. Therefore, despite the fact that the fuel price increase is applied globally, its impact is dissimilar from zone to zone.

6.5. Data Preparation

The study area in this research is the Dutch-speaking region in northern Belgium, Flanders. Flanders has over 6 million inhabitants, or about 60% of the population of Belgium. As already mentioned before, an activity-based model within the FEATHERS framework is applied on the Flemish population to derive the in-depth information of Flemish people's travel behavior and travel demand for a null-scenario (current situation) and some TDM scenarios like increasing fuel price, teleworking, etc. FEATHERS produces traffic demand by means of origin-destination (OD) matrices. These OD matrices include the number of trips for each traffic mode at different disaggregation levels (i.e. age, gender, day of the week, time of day and motive). This traffic demand is then assigned to the road network to obtain detailed exposure metrics at network level. To carry out the assignment of car trips to the road network, the user equilibrium method was selected. The fundamental nature of equilibrium assignment is that travelers will strive to find the shortest path (e.g. minimum travel time) from origin to destination, and network equilibrium occurs when no traveler can decrease travel effort by shifting to a new path. This is an optimal condition, in which no user will gain from changing travel paths once the system is in equilibrium. Exposure metrics are then geographically aggregated to the TAZ level. This has been carried out at zonal level, comprising 2,200 TAZs in Flanders. The average size of TAZs is 6.09 square kilometers with a standard deviation of 4.78 square kilometers. In addition, for each TAZ a set of variables including socio-demographic and road network variables were derived to construct ZCPMs. The crash data used in this study consist of a geo-coded set of injury crashes based on different severity levels that occurred during the period 2004 to 2007. Table 6.1 shows a list of selected variables, together with their definition and descriptive statistics, which have been used in developing the ZCPMs presented in this chapter.

Table 6.1 Selected Variables to Develop ZCPMs

	Variable	Definition	Average	Min	Max	SD
Dependent variables	CCFS	total Car-Car/Fatal and Severe injury crashes observed in a TAZ	2.82	0	21	3.06
	CCSL	total Car-Car/Slight injury crashes observed in a TAZ	19.22	0	199	20.77
	CSFS	total Car-Slow mode/Fatal and Severe injury crashes observed in a TAZ	1.36	0	16	2.08
	CSSL	total Car-Slow mode/Slight injury crashes observed in a TAZ	10.07	0	202	17.81
Exposure variables	NOTs Car	average daily number of car trips originating/arriving from/at a TAZ	2765.8	0	18111.4	2869.8
	NOTs Slow	average daily number of slow-mode trips originating/arriving from/at a TAZ	1082.2	0	9134	1352.2
	Motorway VKT	average daily vehicle kilometers traveled on motorways in a TAZ	27471.82	0	946152.8	84669.53
	Other Roads VKT	average daily vehicle kilometers traveled on other roads in a TAZ	26662.85	0	303237.6	28133.04
Network variables	Capacity	hourly average capacity of links in a TAZ	1790.1	1200	7348.1	554.6
	Intersection	total number of intersections in a TAZ	5.8	0	40	5.9
	Urban	Is the TAZ in an urban area? "No" represented by 0 "Yes" represented by 1	0	0	1	-[a]
	Suburban	Is the TAZ in a suburban area? "No" represented by 0 "Yes" represented by 1	0	0	1	-
Socio-demographic variables	Income Level	average income of residents in a TAZ described as below: "Monthly salary less than 2249 Euro" represented by 0 "Monthly salary more than 2250 Euro" represented by 1	1	0	1	-

a: Data not applicable

6.6. Fuel-Cost Increase Scenario Evaluation

6.6.1. Model Development

Crash data consist of non-negative integers, so using ordinary least-squares regression which serves continuous dependent variables (e.g. time) is not an option (Lord and Mannering, 2010). Given the observed overdispersion in the crash data in this study, it was chosen to model the data using the Negative Binomial (NB) model which enables the modeling of overdispersed data. The NB model is the most commonly used model in crash data modeling (Lord and Mannering, 2010).

Reviewing the literature for different model forms showed that the following GLM model has been widely used in different studies (e.g. in Lovegrove, 2005; Hadayeghi, 2009; Abdel-Aty et al., 2011b; An et al., 2011; Pirdavani et al., 2012):

$$E(C) = \beta_0 \times (Exposure)^{\beta_1} \times e^{\sum_{i=2}^{n} \beta_i x_i} \qquad (1)$$

Where;

$E(C)$ is the expected crash frequency, β_0 and β_i are model parameters, $Exposure$ is the exposure variable (e.g. VKT or NOTs) and x_i's are the other explanatory variables.

In this study, different ZCPMs were constructed within the GLM framework, using the explanatory variables listed in Table 6.1. The models can be categorized into three different groups based on the type of exposure metric that was utilized, i.e.:

(1) flow-based models,

(2) trip-based models and

(3) models based on a combination of the two.

Flow-based models were constructed by associating the NOCs in each TAZ with VHT or VKT, as the exposure variables, and the road network and socio-demographic variables listed in Table 6.1. Trip-based models use the same road network and socio-demographic variables, but NOTs as the exposure variable. In the third type of model, both flow and trip-based variables are included simultaneously as metrics of exposure.

Coefficients were estimated using a forward selection procedure where one of the exposure variables is taken as the starting point, subsequently, selecting additional candidate variables. The analysis results revealed that the combination of exposure variables provides a better model fit; i.e. the models which simultaneously have both NOTs and VHT/VKT as the exposure variables overperform the flow-based or trip-based models (Pirdavani et al., 2012).

For each model, the multicollinearity phenomenon was also checked for by means of the variance inflation factor (VIF) for all variables. As a common rule of thumb, 10 is defined (Kutner et al., 2004) as a cut off value meaning that if the VIF is higher than 10 then multicollinearity is high. VIF values for final chosen models are shown in Table (6.2). These results suggest that multicollinearity is not a problem in developing ZCPMs.

Table 6.2 VIF Values Among Explanatory Variable

	Model #1 (CCFS)	**Model #2 (CCSL)**	**Model #3 (CSFS)**	**Model #4 (CSSL)**
Coefficients	**VIF**	**VIF**	**VIF**	**VIF**
ln(NOTs Car)	2.348508	1.86523	-	-
ln(NOTs Slow)	-	-	1.766268	2.363723
ln(Motorway VKT)	1.035405	1.025774	1.019962	1.036635
ln(Other Roads VKT)	1.911223	1.797954	1.727161	1.911578
Income Level	-	1.095849	1.092445	1.096245
Capacity	1.059042	1.079604	1.075026	1.084524
Intersection	1.006336	1.067071	1.065409	1.071994
Urban	1.161919	-	-	1.161927
Suburban	1.246187	-	-	1.246572

Table 6.3 provides the final models' coefficient estimates. These are the models by which the fuel-cost increase scenario is being evaluated. For all models, most of the exposure variables were positively associated with the NOCs in each TAZ. As the NOTs and VKT increase, the total NOCs also tends to increase. Many studies found similar associations between VKT (e.g. in Lovegrove, 2005; Hadayeghi et al. 2010a; 2010b) or NOTs (e.g. in Naderan and Shahi, 2010; Abdel-Aty et al., 2011a; 2011b) and NOCs per TAZ. The only exception was observed where VKT on motorways was negatively associated with "Car-Slowmode" crashes. This is in line with the fact that the more VKT carried out on motorways (where there is no "Slowmode" traffic) instead of other roads, less vulnerable road users are exposed to unsafe situations.

Table 6.3 Model Estimates for the Final Chosen ZCPMs

	Model #1 (CCFS)	**Model #2 (CCSL)**	**Model #3 (CSFS)**	**Model #4 (CSSL)**
Coefficients	**Estimates**	**Estimates**	**Estimates**	**Estimates**
(Intercept)	-4.356e+00	-4.539e+00	-7.357e+00	-6.802e+00
ln(NOTs Car)	1.162e-01	3.990e-01	-	-
ln(NOTs Slow)	-[a]	-	7.445e-01	9.005e-01
ln(Motorway VKT)	1.464e-02	1.784e-02	-2.528e-02	-1.267e-02
ln(Other Roads VKT)	3.693e-01	3.379e-01	2.144e-01	2.280e-01
Income Level	-	-1.116e-01	-1.643e-01	-1.268e-01
Capacity	3.342e-04	4.140e-04	2.018e-04	1.716e-04
Intersection	3.128e-02	3.195e-02	1.959e-02	1.397e-02
Urban	-4.455e-01	-	-	5.678e-01
Suburban	-2.194e-02	-	-	2.673e-01
PCC[b]	0.615	0.831	0.721	0.855

a: Data not applicable

b: The Pearson Correlation Coefficient (PCC) between observed and predicted crash values

A positive correlation of the number of intersections with the NOCs per TAZ is observed for all models. This positive relationship has also been reported in other studies (e.g. Hadayeghi et al., 2003; de Guevara et al., 2004; An et al., 2011). In general, intersections have a higher risk of experiencing conflicts compared to road links because of their natural design, therefore, there are more crashes expected to occur in TAZs that have a higher number of intersections. As can be observed in Table 6.3, all of the constructed models showed a negative association with "Income Level". This is similar to other studies' findings that poverty has a positive relationship with the number of crashes that occurred in a TAZ (e.g. Quddus, 2008; Wier et al., 2009; Huang et al., 2010).

The degree of urbanization is categorized into three different levels and, therefore, represented by two dummy variables; "Urban" and "Suburban". When "Urban" and "Suburban" metrics in a TAZ are both 0, then this TAZ is located in a rural area. In Model #1, the coefficient estimate for the variables "Urban" and "Suburban" have negative signs. This is in line with our expectations that CCFS crashes are expected to occur more frequently in rural areas where cars are driven at higher speeds. On the contrary, the positive association between "Urban" and "Suburban" variables and CSSL crashes reveals that the model correctly predicts more crashes of this type in more urbanized TAZs.

6.6.2. Traffic Safety Evaluation Process

To compute the changes in exposure, OD matrices for the fuel cost increase scenario will be derived for scenario evaluation by running the activity-based transportation model. After assigning this demand to the road network, all required variables become available to set up the evaluation task. Now, the final ZCPMs are applied and crashes are predicted for each TAZ. The traffic safety evaluation can then be conducted by comparing the NOCs predicted by the final ZCPMs for the null and the fuel-cost increase scenario. Figure 6.1 depicts the conceptual framework of the traffic safety evaluation process.

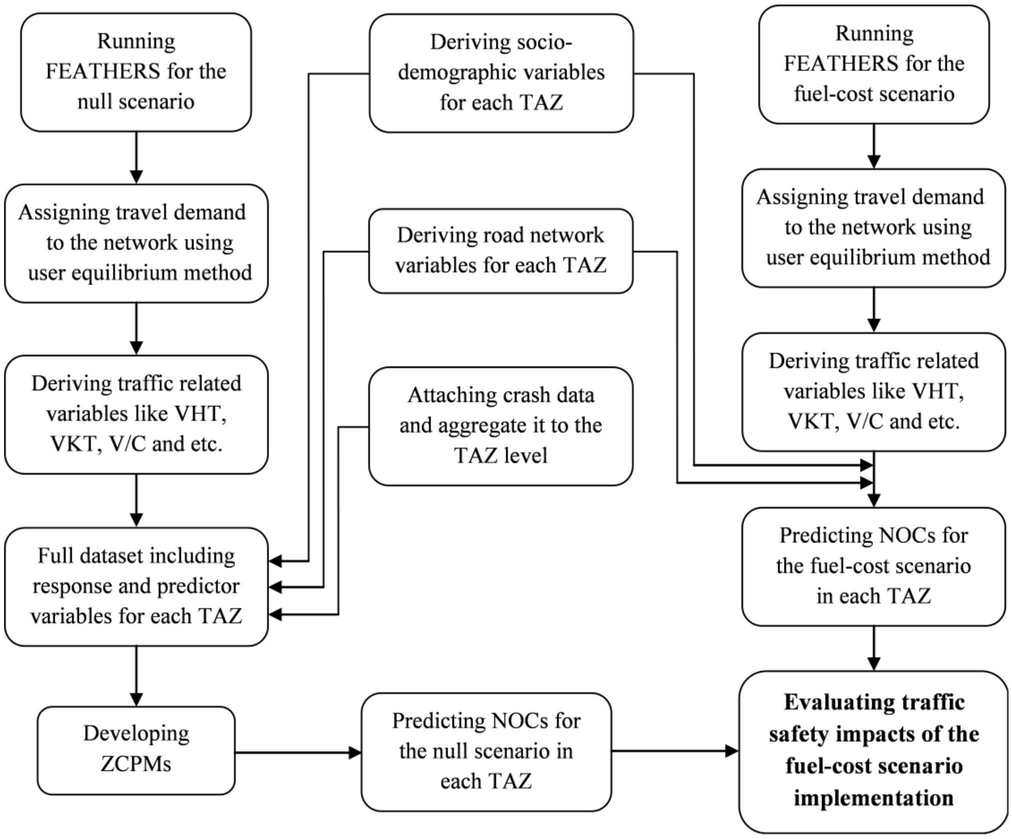

Figure 6.1 Conceptual framework of the traffic safety evaluation process.

6.7. Validation

There are different types of validation methodologies that can be performed to check the validity of predictive models output. The simples method which is often called "the apparent validity" which verifies model performance on the same sample used to develop predictive model. This method gives an optimistic estimate of model performance that cannot be reliable. An alternative to this method is to validate the performance on population underlying the sample. A key characteristic of this method is that data for model development and evaluation are both random samples from the same underlying population. This method which is often called "internal validation" enables us to obtain reliable estimate of performance, at least for a population similar to the development sample by which the reproducibility and generalizability of predictive models can be ensured.

Different internal validation techniques can be executed to accomplish the validation task (Steyerberg et al., 2001). A simple technique is split-sample meaning that a random split is made, resulting in e.g. a 50% training and a 50% validation dataset (sometimes called test dataset). A more complex technique is cross-validation that uses the same principle, but alternates the training and test sets (e.g. 50:50 split means 2 training and test rounds (2-fold cross-validation) while a 90:10 split signifies 10 training and test rounds (10-fold cross-validation)). The most extreme variant of cross-validation is the "jack-knife" or "Leave-one-out" technique, which uses a single observation from the original sample as the validation data, and the remaining observations as the training data.

In this study we carried out a 10-fold cross-validation is utilized to estimate how accurately the developed ZCPMs will perform when they are coupled with the unseen datasets. The analyses results show a significant correlation between the observed and the predicted NOICs for all developed models and for all cross-validation sets. This indicates that all models are capable of predicting the NOICs quite well; however, more statistical tests are needed to assure robustness and generalizability of the models. With this respect, a robustness index (RI) is defined as follows:

$$RI = \frac{RMSE-Validation}{RMSE-Calibration} \quad (2)$$

Where (RMSE-Validation) is the root mean square error of a model against a validation dataset,

$$RMSE - Validation = \sqrt{\sum_{i=1}^{n} \left(\mathrm{Y_i} - \widehat{\mathrm{Y}}_\mathrm{i}\right)^2 \Big/ \mathrm{n}} \quad (3)$$

(RMSE-Calibration) is the root mean square error of a model against dataset used to develop the model:

$$RMSE - Calibration = \sqrt{\sum_{j=1}^{m} \left(\mathrm{Y_j} - \hat{\mathrm{Y}}_\mathrm{j}\right)^2 \Big/ \mathrm{m}} \tag{4}$$

Where;

$\hat{\mathrm{Y}}_\mathrm{i}$, $\hat{\mathrm{Y}}_\mathrm{j}$: predicted NOICs for i_{th} and j_{th} TAZ,

$\mathrm{Y_i}$: observed NOICs for i_{th} TAZ from validation dataset,

$\mathrm{Y_j}$: observed NOICs for j_{th} TAZ from calibration dataset,

n: total number of TAZs used in validation and

m: total number of TAZs used in calibration.

The key feature of the RI is that it quantifies the predictive accuracy of the model (RMSE-Validation) relative to the expected accuracy (RMSE-Calibration). Lower RI values (less than 1) indicate a robust model, with very good accuracy against independent data (upon which the RMSE-Validation is calculated), compared with the expected accuracy based on the data used in constructing the model (upon which the RMSE-Calibration is calculated). It is unlikely to have models that function much better on unseen data compared with training data which results in extreme low RI values. Therefore, it can be concluded that having RI values close to 1 represent robustness of a model. RI is calculated for each cross-validation dataset and the results are shown in Table 6.4. As can be seen in Table 6.4, average value of RI for all model types are very close to 1. This is indicating that all models are robust and generalizable. Another important conclusion that can be derived by looking at the RI values is that these models do not over fit the training data. Over-fitting occurs when RMSE-Calibration is much smaller than RMSE-Validation (i.e. a model very well fits the data on which it is developed but it cannot properly predict an unseen data sample from the same population) and, therefore, RI values would be way larger than 1.

Table 6.4 Robustness Index (RI) of ZCPMs Based On Different Cross-validation Datasets

Model	**RI value**	**Model**	**RI value**	**Model**	**RI value**	**Model**	**RI value**
CCFS1	1.040396	CCSL1	0.846416	CSFS1	1.011576	CSSL1	0.981209
CCFS2	0.952812	CCSL2	0.927468	CSFS2	0.925979	CSSL2	1.05221
CCFS3	0.958401	CCSL3	0.784589	CSFS3	0.811194	CSSL3	0.872586
CCFS4	0.95928	CCSL4	0.937306	CSFS4	0.863302	CSSL4	0.737996
CCFS5	1.063179	CCSL5	1.052041	CSFS5	1.104713	CSSL5	1.09632
CCFS6	1.130811	CCSL6	1.874544	CSFS6	1.043589	CSSL6	1.309239
CCFS7	1.091476	CCSL7	1.120564	CSFS7	1.093095	CSSL7	1.048313
CCFS8	0.926128	CCSL8	0.671674	CSFS8	0.953564	CSSL8	0.924771
CCFS9	1.028027	CCSL9	1.161244	CSFS9	1.231513	CSSL9	1.006333
CCFS10	0.943546	CCSL10	0.718683	CSFS10	1.019129	CSSL10	1.044715
Average CCFS	1.012223	Average CCSL	0.994631	Average CSFS	1.006294	Average CSSL	1.004991

CCFS: Car-Car/Fatal and Severe injury crashes

CCSL: Car-Car/Slight injury crashes

CSFS: Car-Slow mode/ Fatal and Severe injury crashes

CSSL: Car-Slow mode/Slight injury crashes

6.8. Results

Before describing the traffic safety impact of the fuel-cost increase scenario, it would be beneficial to have a look at the changes made to the more traffic-related attributes playing a role in the whole chain. As described before, increasing fuel-related costs will affect and increase the total travel expenses of motor vehicle trips. As a result, people will start comparing the relative costs of travelling and may consider a shift to other available transportation modes. For instance, short-distance trips can be substituted by public transportation (e.g. bus or tram) or slow mode (i.e. biking or walking) or long-distance trips may shift towards public transportation (e.g. train) or be substituted by carpooling.

Comparing OD matrices derived from the activity-based model for both the null and the fuel-cost increase scenario, will enable us to perceive any changes in NOTs for different modes and will also allow us to figure out if any mode shift will occur. The results of these comparisons revealed that the fuel-cost increase scenario reduces the average daily NOTs carried out by car (-4.1%) and in contrast, leads to an increase in the average daily NOTs by other modes (Car Passenger by +2.9%, Public transportation by +4.4% and Slow mode by +2.9%). In addition to these changes observed at the global level, it is also interesting to describe these changes at the TAZ level. It was observed that more urbanized areas experience a higher reduction of car trips as well as a higher increase of other mode trips when compared to less urbanized areas. On the contrary and in some particular TAZs like the ones nearby the Flemish borders, fewer mode shifts occurred. The reason might be that the cross-border public transportation offer is not a convenient option or no cross-border public transportation service is available in these TAZs. Thus, many travelers who are travelling across borders prefer to take their car since they cannot easily find a substitution mode.

Although the NOTs may represent an acceptable indication of exposure, it ignores the impact of transit traffic which is just passing through a TAZ. As already mentioned before, NOTs do not contain any information about trip time, trip length, route choice, intrazonal traffic and transit traffic. Therefore, investigating the impact of a fuel-cost increase scenario cannot be practically carried out by merely considering the changes in the NOTs starting or arriving in a TAZ. Thus, other exposure variables which can account for the impact of trip assignment should be taken into account. As a result, inclusion of the flow-related variables (e.g. VKT) in the prediction models is essential.

The analyses show that the average values of the VKT and VHT decrease after implementing the fuel-cost increase scenario. Not surprisingly, the reductions for motorways are higher than for other roads (i.e. reduction in motorway VHT by 16.9%, other roads VHT by 10.3%, motorway VKT by 13.3% and other roads VKT by 9.8%). This can be explained by the fact that the majority of reduced long-distance trips are carried out on motorways. It also indicates that this scenario has a somewhat higher effect on long-distance trips compared to short-distance trips. It can also be noticed that the reduction in VHT is slightly higher than the reduction in VKT. A stronger decrease of travel time compared to travel distance can be explained by a decreased level of congestion on the roads. As the rebound effect of the vehicle fleet economy was not taken into account in this short-term analysis, we find that the reductions in VKT under the scenario are in line with the reductions in fuel consumption reported in literature (Goodwin et al., 2004; Litman 2010).

Predictably, in the fuel-cost increase scenario the total number of predicted "Car-Car" crashes decreases compared to the null scenario. This is due to reduced exposure as the main predictor of crashes. The results show that CCFS and CCSL crashes are predicted to decrease by 250 and 2080 respectively for a period of 4 years (-4% and -4.8%). On the contrary, for "Car-Slowmode" crashes a slight increase is observed for both fatal-severe and slight injury crashes. Taken as a whole, CSFS and CSSL crashes are predicted to increase by 15 and 185 respectively for the same period of 4 years (+0.5% and +0.8%). Figure 6.2 represents the violin plots of changes in NOCs after the fuel-cost increase scenario implementation. The violin plot is a synergistic combination of the box plot and the density trace (Hintze and Nelson, 1998). These plots retain much of the information of box plots (except for the individual outliers), besides providing information about the distributional characteristics of the data. In these plots, the wider the violin, the more data points are associated to that value. Moreover, the white dots indicate the median; black boxes show the upper and lower quartile and the vertical black lines denote the upper and lower whiskers.

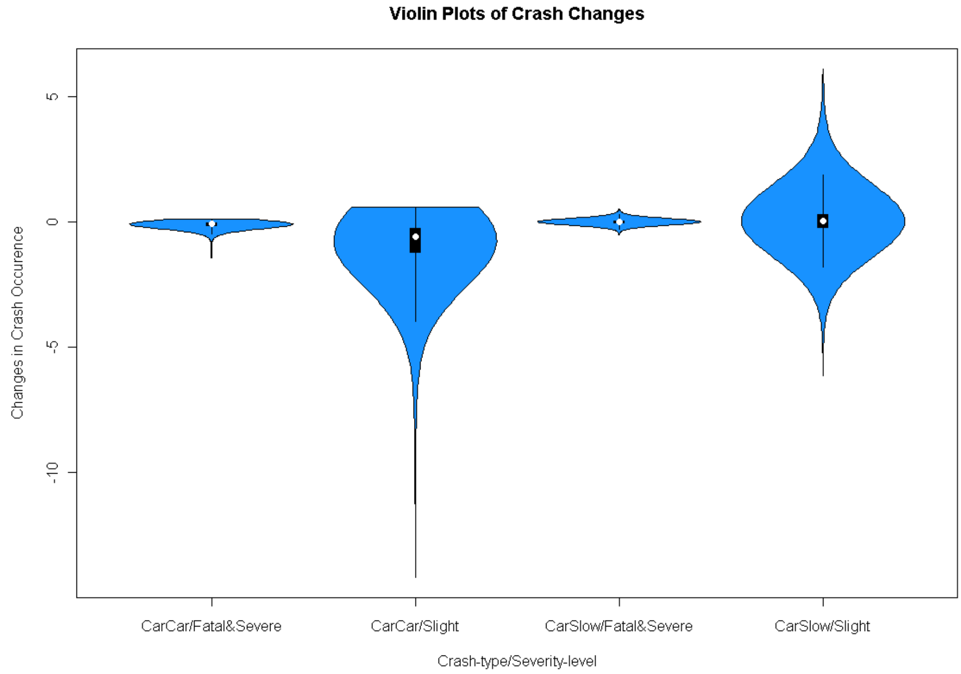

Figure 6.2 Violin plots of changes in crash occurrence after the fuel-cost increase scenario implementation

In the development of "Car-Slowmode" models, both car and Slowmode-related exposure variables were used. Following the implementation of the fuel-cost increase scenario and as a result of mode shift, the number of car trips decreased whereas the number of Slowmode trips increased. However, these mode shifts are not always similar in all TAZs; i.e. more urbanized areas have a higher number of mode shifts; consequently more Slowmode-related crashes are predicted for these areas. An illustration of changes in the NOCs for all TAZs may present a better pattern on how different TAZs are affected by the scenario. In Figure 6.3, the changes in the predicted NOCs are displayed for each TAZ. Figure 6.3 reveals that the reductions in CCFS and CCSL crashes are greater for urban areas and generally smaller for TAZs close to the Flemish borders. As explained earlier, CSFS and CSSL crashes are predicted to increase in more urbanized areas; this is evident from the corresponding maps in Figure 6.3 where concentrations of red dots stand for the major cities in Flanders.

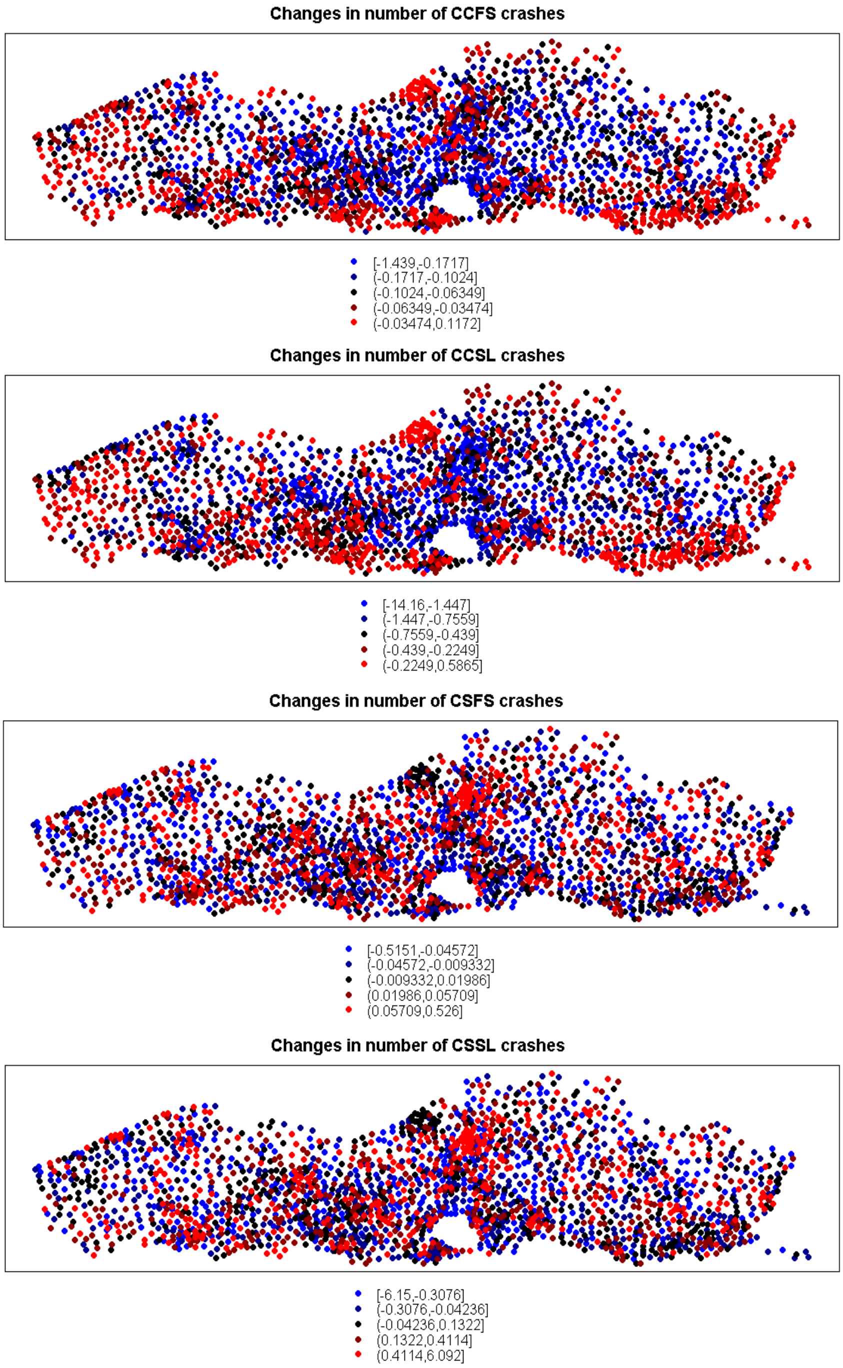

Figure 6.3 Changes in NOCs in each TAZ after the fuel-cost increase scenario implementation.

6.9. Conclusions

In this study, a zonal crash prediction modeling approach has been integrated into a fuel-cost increase scenario to assess this TDM strategy's impact on traffic safety. This assessment deals with the relatively short-term impact of fuel price changes and it has been carried out by applying an activity-based travel demand model to derive the exposure metrics. Based on the results of the analyses, the following conclusions can be drawn:

- Activity-based transportation models provide an adequate range of in-depth information about individuals' travel behavior to realistically simulate and evaluate TDM strategies. The advantage of these models is that the impact of applying a TDM strategy will be accounted for, for each individual, throughout a decision making process. Activity-based models provide more reliable information since, unlike traditional models, TDM strategies are inherently accounted for in these models.
- Unlike some prior studies where a reactive approach was generally used to evaluate the traffic safety impact of a TDM strategy, in this study a proactive methodology was applied. This was carried out in an assessment exercise by assuming a 20% increase in fuel price.
- Using only NOTs originating/destinating from/in a TAZ for crash prediction and consequently evaluating the safety impact of a TDM strategy will lead to a lack of some important information about the characteristics of reduced trips; i.e. NOTs, as an exposure variable, does not contain information on trip time, trip length and route choice. For this reason, other exposure variables which are sensitive to the impact of trip assignment should also be taken into account. This was done by assigning traffic demand to the road network adopting the user equilibrium assignment method.
- The results of the comparison analysis revealed that the fuel-cost increase scenario has many impacts such as mode shift and reductions of total travel demand, total crash occurrence, VKT and VHT. On the whole, there is a reduction of on average 105,000 daily trips (all types of modes) as a result of the fuel-cost increase scenario. This scenario also causes a reduction of

nearly 5 billion VKT by cars per year, almost 11% of the total annual VKT by cars in Flanders.

- The total NOCs is predicted to decrease by 2,130 for a period of 4 years. However, changes in the NOCs for different crash-type/severity-levels are not identical. As a result of an increase in the NOTs for the "Slowmode" category, crashes which involve vulnerable road users are predicted to increase; i.e. CSFS and CSSL crashes increased by 0.5% and 0.8% respectively. On the other hand, CCFS and CCSL crashes are predicted to decrease by 4% and 4.8% respectively. This reveals that the fuel-cost increase scenario affects different road users differently. The traffic safety situation slightly deteriorates for vulnerable road users; nevertheless, there are noticeable safety benefits for "Car-Car" crashes and for the overall traffic safety situation.
- When considering the changes in the NOCs at the TAZ level, it was found that the maximum reduction of "Car-Car" crashes and the maximum increase of "Car-Slowmode" crashes were both observed in urban areas (i.e. cities). It can be concluded that in cities, in contrast to other areas, there is a higher likelihood of finding an alternative mode for cars. In contrast, the TAZs in less urbanized regions and the TAZs nearby the borders usually lack good public transportation services. Therefore, it is expected that we will not see many trips shift from cars to other modes in less urbanized areas and consequently there is a more stable traffic safety situation in these TAZs despite conducting the fuel-cost increase scenario.

This chapter presents a new extension to the application of ZCPMs incorporated into TDM strategies. The results show the ability of ZCPMs as a reliable predictive tool which can be used during the planning level of transportation projects.

The fuel-cost increase scenario studied in this research investigated the relatively short-term effects of an increased fuel price. In other words, the model is a short-term model in the sense that neither a shift in the composition of the vehicle fleet, nor changes in the location of businesses and/or the location choice for living as a result of the change in fuel cost are assumed.

Crashes are known to be a function of two components; exposure and risk. It is therefore likely that a fuel price increase will impact people's driving behavior and their speed choice; i.e. drivers might try to reduce their fuel consumption by driving more slowly. As a result, it can be assumed that the risk component will also decrease after the fuel-cost increase scenario implementation. In this study however, only the changes in the exposure component were taken into account, whereas the risk component was assumed to be constant. This might be a limitation of this study. If we were to include the risk component in this study as well, however, the traffic safety benefits might be expected to be even larger than predicted in this study.

Although some clear benefits (e.g. global traffic safety improvement or VKT reduction) are noticeable from the fuel-cost increase scenario, it would be beneficial to extend this study by including other TDM strategies in order to present a comprehensive traffic safety evaluation package.

In this study, as a first attempt, the methodology relied on the aggregate daily traffic information. Activity-based models are capable of providing disaggregate travel characteristics. Hence, different types of disaggregation based on time of day, day of the week, age, gender, motive, etc are on the list of potential future research in order to take full advantage of the output of activity-based models.

References

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011a. Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones. In Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011b. Zonal Level Safety Evaluation Incorporating Trip Generation Effects. In Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis & Prevention, 38(3), 618-625.

Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Accident Analysis & Prevention, 35(4), 537-547.

An, M., Casper, C., Wu, W., 2011. Using Travel Demand Model and Zonal Safety Planning Model for Safety Benefit Estimation in Project Evaluation. In Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Arentze, T.A., Timmermans, H.J.P., 2004. ALBATROSS – Version 2.0 – A learning based transportationoriented simulation system. EIRASS (European Institute of Retailing and Services Studies).

Arentze, T.A., Timmermans, H.J.P., 2005. Incorporating Parametric Action Decision Trees in Computational Process Models of Activity-Travel Behavior: Theory and Illustration. Advanced OR and AI Methods in Transportation, 567-572.

Bellemans, T., Kochan, B., Janssens, D., Wets, G., Arentze, T.A., Timmermans, H.J.P., 2010. Implementation Framework and Development Trajectory of the FEATHERS Activity-Based Simulation Platform. Transportation Research Record, 2175, 111-119.

Chi, Guangqing, Cosby, A.G., Quddus, M.A., Levinson, D., Gilbert, P.A., 2010. Gasoline Prices and Traffic Safety in Mississippi. Journal of Safety Research, 41(6), 493-500.

Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J., Picado, R., 2007. Synthesis of first practices and operational research approaches in activity-based travel demand modeling. Transportation Research Part A: Policy and Practice, 41(5), 464-488.

Grabowski, D.C., Morrisey, M.A., 2004. Gasoline prices and motor vehicle fatalities. Journal of Policy Analysis and Management, 23(3), 575-593.

De Guevara, F.L., Washington, S.P., Oh, J., 2004. Forecasting crashes at the planning level. A simultaneous negative binomial crash model applied in Tucson, Arizona. Transportation Research Record, 1897, 191-199.

Goodwin, P., Dargay, J., Hanly, M., 2004. Elasticities of road traffic and fuel consumption with respect to price and income: a review. Transport Reviews, 24(3), 275–292.

Hadayeghi, A., 2009. Use of Advanced Techniques to Estimate Zonal Level Safety Planning Models and Examine Their Temporal Transferability. PhD thesis, Department of Civil Engineering, University of Toronto.

Hadayeghi, A., Shalaby, A., Persaud, B., 2010. Development of Planning-Level Transportation Safety Models using Full Bayesian Semiparametric Additive Techniques. Journal of Transportation Safety & Security, 2(1), 45-68.

Hadayeghi, A., Shalaby, A., Persaud, B., 2003. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. Transportation Research Record, 1840(1), 87-95.

Hadayeghi, A., Shalaby, A., Persaud, B., 2007. Safety Prediction Models: Proactive Tool for Safety Evaluation in Urban Transportation Planning Applications. Transportation Research Record, 2019(-1), 225-236.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. Accident; Analysis and Prevention, 42(2), 676-688.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accident Analysis & Prevention, 38(3), 579-589.

Hintze, J. L., Nelson, R. D., 1998. Violin plots: A box plot-density trace synergism. American Statistician, 52 (2), 181-184.

Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-Level Crash Risk Analysis in Florida. Transportation Research Record, 2148(-1), 27-37.

Janssens, D., Wets, G., Timmermans, H.J.P., Arentze, T.A., 2007. Modelling Short-Term Dynamics in Activity-Travel Patterns: Conceptual Framework of the Feathers Model. In 11th World Conference on Transport Research, Berkeley CA, USA.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., 2008. Assessing the Impact of Fuel Cost on Traffic Demand in Flanders Using Activity-Based Models. In Travel Demand Management TDM, Vienna, Austria.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., 2012. Validation of an Activity-Based Traffic Demand Model for Flanders Implemented in the Feathers Simulation Platform. Book chapter in "Computational Intelligence for Traffic and Mobility", Atlantic Press, Forthcoming.

Levine, N., Kim, K.E., Nitz, L.H., 1995. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. Accident Analysis & Prevention, 27(5), 675-685.

Litman, T., 2006. Mobility Management Traffic Safety Impacts. In Transportation Research Board (TRB) 85th Annual Meeting, Washington D.C. USA.

Litman, T., Fitzroy, S., 2011. Safe Travels: Evaluating Mobility Management Traffic Safety Impacts. Victoria Transport Policy Institute, www.vtpi.org (last accessed 17.03.2012).

Litman, T., 2010. Changing Vehicle Travel Price Sensitivities: The Rebounding Rebound Effect. Victoria Transport Policy Institute, www.vtpi.org (last accessed 17.03.2012).

Litman, T., 2011. Pricing For Traffic Safety: How Efficient Transport Pricing Can Reduce Roadway Crash Risk. Victoria Transport Policy Institute, www.vtpi.org (last accessed 17.03.2012).

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice, 44(5), 291-305.

Lovegrove, G.R., Litman, T., 2008. Using Macro-Level Collision Prediction Models to Evaluate the Road Safety Effects of Mobility Management Strategies:

New Empirical Tools to Promote Sustainable Development. In Transportation Research Board (TRB) 87th Annual Meeting, Washington D.C. USA.

Lovegrove, G.R., Sayed, T., 2006. Macro-Level Collision Prediction Models for Evaluating Neighbourhood Traffic Safety. Canadian Journal of Civil Engineering, 33, 609-621.

Lovegrove, G.R., Sayed, T., 2007. Macrolevel Collision Prediction Models to Enhance Traditional Reactive Road Safety Improvement Programs. Transportation Research Record, 2019(-1), 65-73.

Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: Introducing crash generation concept. Accident Analysis & Prevention, 42(1), 339-346.

Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention, 36(4), 525-532.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis & Prevention, 36(6), 973-984.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., (2012). Developing Zonal Crash Prediction Models with a Focus on Application of Different Exposure Measures. In: Transportation Research Board (TRB) 91st Annual Meeting, Washington D.C. USA.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. Accident Analysis & Prevention, 40(4), 1486-1497.

Siddiqui, C., Abdel-Aty, M., Choi, K., Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis & Prevention, 45, 382-391.

Steyerberg, E.W., Harrell, F.E., Borsboom, G.J.J.M., Eijkemans, M.J.C., Vergouwe, Y., Habbema, J.F., 2001. Internal Validation of Predictive Models: Efficiency of Some Procedures for Logistic Regression Analysis. Journal of Clinical Epidemiology 54 (8), 774–781.

Tarko, A.P., Inerowicz, M., Ramos, J., Li, W., 2008. Tool with Road-Level Crash Prediction for Transportation Safety Planning. Transportation Research Record, 2083(-1), 16-25.

Vovsha, P., Bradley, M., 2006. Advanced Activity-Based Models in Context of Planning Decisions. Transportation Research Record: Journal of the Transportation Research Board, 1981(-1), 34-41.

VTPI, 2011. Online TDM Encyclopedia, Victoria Transport Policy Institute.

http://www.vtpi.org/tdm/index.php (last accessed 17.03.2012).

Wier, M., Weintraub, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis & Prevention, 41(1), 137-145.

7. Assessing the Road Safety Impacts of a Teleworking Policy by means of a Macro-level Crash Prediction Approach

Pirdavani, A., et al., (2013), DVD Compendium of the 92nd Transportation Research Board Conference, Washington D.C., USA.

7.1. ABSTRACT

Travel demand management (TDM) consists of a variety of policy measures that affect the effectiveness of transportation systems by changing travel behavior. The primary objective of such TDM strategies is not to improve traffic safety, although their impact on traffic safety should not be neglected. The main purpose of this study is to simulate the traffic safety impact of conducting a teleworking scenario (i.e. 5% of the working population engages in teleworking) in the study area, Flanders, Belgium. Since TDM strategies are usually conducted at a geographically aggregated level, crash prediction models (CPMs) that are used to evaluate such strategies should also be developed at an aggregate level. Given that crash occurrences are often spatially heterogeneous and are affected by many spatial variables, the existence of spatial correlation in the data is also examined. The results indicate the necessity of accounting for the spatial correlation when developing crash prediction models. Therefore, zonal crash prediction models (ZCPMs) within the geographically weighted generalized linear modeling (GWGLM) framework are developed to incorporate the spatial variations in association between the number of crashes (NOCs) (including fatal, severe and slight injury crashes recorded between 2004 and 2007) and other explanatory variables. Different exposure, network and socio-demographic variables of 2200 traffic analysis zones (TAZs) are considered as predictors of crashes. An activity-based transportation model framework is adopted to produce detailed exposure metrics. This enables to conduct a more detailed and reliable assessment while TDM strategies are inherently modeled in the activity-based models. In this study, several ZCPMs with different severity levels and crash types are developed to predict the NOCs for both the null and

the teleworking scenario. The models show a considerable traffic safety benefit of conducting the teleworking scenario due to its impact on the reduction of total vehicle kilometers traveled (VKT) by cars by 3.2%. Implementing the teleworking scenario is predicted to reduce the annual VKT by cars by 1.4 billion and total NOCs to decline by 2.6%.

7.2. Introduction

Urbanization and population growth together with employment and motor vehicle growth largely affect road transportation systems' performance. To diminish these negative impacts, different policy measures and strategies have been applied by authorities. These programs and strategies that promote more efficient use of transportation systems are generally called TDM strategies (Litman, 2003). TDM strategies consist of several policies and strategies which aim to overcome transportation problems by means of mode shift (e.g. using public transportation instead of cars, biking for short distance trips or carpooling), travel time shift (e.g. avoiding traffic peak-hours by leaving home/the work place earlier or later) or travel demand reduction (e.g. teleworking). In general, TDM strategies are implemented to improve transportation systems' efficiency. However, their potential secondary impacts such as traffic safety or environmental effects should not be overlooked.

"Teleworking" is a general term used when application of telecommunication systems substitutes for actual travel to the work place. Teleworking is one of the most popular and effective components of commute trip reduction programs (Litman and Fitzroy, 2012). Teleworking can significantly reduce participating employees' commute travel and consequently the total distance traveled. As mentioned earlier, TDM strategies usually have consequential impacts (e.g. impacts of reduced travel demand after applying a teleworking strategy) such as traffic safety, which is interesting to be investigated. To the best of our knowledge, traffic safety impacts of teleworking as a TDM strategy have not been investigated before in a proactive manner. The main goal of this study is, therefore, to evaluate the road safety impacts of a teleworking scenario by coupling ZCPMs with an activity-based model for

Flanders, Belgium. This way, the behavioral impact of the TDM scenario, in terms of traffic demand, is incorporated in the safety analysis. By assigning traffic demand to the road network, the impacts of responses to TDM, such as changes in trip planning, route choice and modal choice are incorporated into the analysis.

The most immediate and direct impacts of teleworking are travel demand and consequently a reduction of total distance traveled. Previous research has evaluated these impacts from individual and global points of views; i.e. some studies focused on the changes of only telecommuter's behavior and their travel pattern (individual) whereas other studies investigated the effects of a telecommuting strategy on a more global level.

Henderson and Mokhtarian (1996) compared the differences in non-telecommuting days and telecommuting days for a telecommuting group. They showed that vehicle miles traveled (VMT) and the number of daily trips reduced by 66.5% and 31.9%, respectively. Koenig et al. (1996) compared participants' telecommuting day travel behavior with their non-telecommuting behavior. They concluded that the number of person vehicle trips reduced by 27% while VMT decreased by 77%. Moreover, Mokhtarian and Varma (1998) compared several travel indicators between telecommuting days and non-telecommuting days for a sample of 72 center-based telecommuters in California. An average reduction of 11.9% in person miles traveled and 11.5% in VMT was found over a five-day work week.

In a study conducted in the USA by Nilles (1996), it was estimated that if 10% of the workforce telecommutes on any given day, total vehicle travel would decline by 4%. Results of another study (Choo et al., 2005) indicated that estimated VMT without telecommuting would have been 1.78% to 3.31% higher compared to the observed VMT, with a mean impact of 2.12%. In another study, Choo and Mokhtarian (2007) found that teleworking appears to reduce VMT as little as 0.34%. In contrast to the above mentioned studies which report a relatively modest impact of teleworking on distance traveled, other studies report quite higher numbers. For instance, Vu and Vandebona (2007) estimated a reduction of 10.8% to 15.46% in VKT after evaluating different teleworking scenarios in Australia. Dissanayake and Morikawa (2008) investigated the reductions of VKT for car and motorcycle travel after a telecommuting policy implementation in Bangkok. The results revealed that the telecommuting policy proposed in their study significantly reduces congestion and vehicle usage reduces by 18–20%.

Based on the literature, it can be concluded that although teleworking seems to decrease significantly the amount of VKT, individual estimations by different studies tend to vary strongly. This uncertainty was also reported by Choo et al. (2005) who claimed that a wide range of answers to the question of "what impact on travel?" can be obtained. They concluded that although teleworking has a statistically significant impact on reducing travel demand, the magnitude of this impact would not be very extraordinary. The main focus of this chapter is not to assess the magnitude of the impact of teleworking on distance traveled, however, it is important to assure that the estimates of our study are reasonable and in line with the findings of other studies. This is important in view of the fact that these estimates are the main input of the CPMs and, therefore, we would like to avoid any possible bias in the results of our policy assessment.

Kochan et al. (2011) studied the effects of teleworking on total distance traveled in Flanders, Belgium. It was reported that in 2002, in Flanders, the total distance traveled decreased by 1.6% where the proportion of teleworkers that telework on a working day was 3.8% (Kochan et al., 2011). These results are in line with the findings of literature. Therefore, our study will be based on the framework presented in Kochan et al. (2011), although we simulate a 5% of the working population engages in teleworking instead of 3.8% (detailed information about implementation of this teleworking scenario is provided in the next section of the chapter).

It can be concluded that the cause-effect relationship between teleworking and a reduction in VKT is well-established. Moreover, the relation between different types of exposure metrics (e.g. number of trips or VKT) and crashes has also been reported and well documented in literature (e.g. in Lovegrove, 2005; Hadayeghi et al., 2010a; Naderan and Shahi, 2010; Abdel-Aty et al., 2011a, 2011a; Pirdavani et al., 2012, in press) and although exposure might not be the direct cause of crash occurrence, it is a major predictive variable to estimate the number of crashes. Therefore, it is plausible to utilize the association between the teleworking scenario and the number of crashes so as to evaluating the traffic safety impacts of such TDM strategy.

The structure of this chapter is as follows. Initially, the activity-based model and the procedure of implementing the teleworking scenario will be briefly introduced. In the next sections the data preparation, model construction and the teleworking scenario evaluation process will be demonstrated. Finally, the results of this evaluation will be shown followed by the final conclusions.

7.3. Impact of Teleworking on Travel Demand

Traditionally, travel was assumed to be the result of four subsequent decisions which were modeled separately, also referred to as four-step models. More recently, several studies claim that travel plays a rather isolated role in these models and the reason why people undertake trips is neglected completely (Arentze and Timmermans, 2004). This gave rise to a new framework of models, called activity-based transportation models. The main difference between four-step models and activity-based transportation models is that the latter try to predict interdependencies between several facets of activity profiles (Davidson et al., 2007). The major advantages of activity-based models are that they deal with participation of various types of activities during a day. Moreover, a microsimulation approach which considers a high behavioral realism of individual agents is often adopted in these types of models (Kochan et al., 2011). Interactions between family members like using the household vehicles, sharing household responsibilities or performing joint activities affect people’s travel behavior. Four-step models that ignore such links are expected to misstate people’s responses to TDM strategies in some circumstances. As a result, activity-based models are capable of treating TDM strategies and policy issues more effectively compared to four-step models (Vovsha and Bradley, 2006).

7.3.1. FEATHERS Framework

The FEATHERS (Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS) framework (Janssens et al., 2007) was developed to facilitate the development of activity-based models for transportation demand in Flanders, Belgium. The scheduling engine that is currently implemented in the FEATHERS framework is based on the scheduling engine that is present in the Albatross system (Arentze and Timmermans, 2004). The real-life representation of Flanders is embedded in an agent-based simulation model which consists of over six million agents, each agent representing one member of the Flemish population. A sequence of 26 decision trees are used in the scheduling process and decisions are based on a number of attributes of the individuals (e.g. age, gender), the households (e.g. number of cars) and the geographical zones (e.g. population density, number of shops). For each agent with its specific attributes, the model simulates whether an activity (e.g. shopping, working, leisure activity, etc.) is going to be carried out or not. Subsequently, amongst others, the location, transport mode (available modes in FEATHERS are "car driver", "car passenger", "public transportation" and "slow mode" including pedestrians and cyclists) and duration of the activity are determined, taking into account the attributes of the individual (Kochan et al., 2008). Traffic demand is subsequently assigned to the road network in such a way that an equilibrium is established between transportation demand and supply (Bellemans et al., 2010), which results in a time-dependent traffic state on the road network. In order to run, calibrate and validate the activity-based model, three major types of data are required (Kochan et al., Forthcoming); data describing the environment (e.g. population density, level of service of the transportation networks), a synthetic population which is simulated and finally activity-travel data originating from a representative sample of the population from which human behavior is derived.

7.3.2. Implementation of Teleworking Scenario in FEATHERS

It is known from literature that one of the major advantages of the activity-based modeling approach is its sensitivity to scenarios that are generally important in transport planning and policy making (Arentze and Timmermans, 2005). In contrast to trip-based and tour-based models, activity-based models are sensitive to institutional changes in society in addition to land-use and transportation-system related factors. Such changes are related to work times and work durations of individuals and opening hours of stores or other facilities for "out-of-home" activities.

The scheduler in FEATHERS first starts with an empty schedule or diary and evaluates whether work activities will be included or not. If there will be any work activity, then the number of work activities (1 or 2 work activities), their starting times, durations and also the time in-between the work activities (in case 2 work activities are performed) will be estimated. In a second step the locations of the work activities are determined. The system sequentially assigns locations to the work activities in order of schedule position. This is carried out by systematically consulting a fixed list of specific decision trees (Kochan et al., 2011).

After the locations of the work activities are determined, the teleworking scenario gets involved. A dedicated procedure samples employees as teleworkers according to a preset distribution. The proportion of working population that is selected to telework on a working day is 5%. After selecting the teleworkers their work location(s) will be replaced by their home address and their schedules will be updated with this new information. This way teleworkers stay at home during their working episode. Now that the teleworking scenario is enforced, the scheduler returns back to normal scheduling and proceeds with the next decision steps, which are selection of work related transport modes (only for individuals not teleworking on that day), inclusion and time profiling of non-work fixed and flexible activities, determination of fixed and flexible activity locations and finally determination of fixed and flexible activity transport modes. More information about this procedure can be found in (Kochan et al., 2011).

7.4. Macro-Level Crash Prediction Approach

CPMs can be developed at different levels of aggregation, for instance, at the local level (road section or intersection) or at the regional level (e.g. TAZ). Recently, crash analyses at a regional level receive more and more attention. Several studies examined the association of a collection of zone-level factors such as traffic patterns, socio-demographic and socio-economic variables, land use patterns and weather conditions with crashes, aggregated by a specific spatial scale (e.g. in Amoros et al., 2003; Hadayeghi et al., 2003, 2007, 2010a, 2010b; Noland and Oh, 2004; Noland and Quddus, 2005; Aguero-Valverde and Jovanis, 2006; Lovegrove and Sayed, 2006; Quddus, 2008; Lovegrove and Litman, 2008; Wier et al., 2009; Huang et al., 2010; Pirdavani et al., 2012, in press). Macro-level crash analyses can provide important information enabling for instance in cross-sectional comparisons between different zones, or to identify safety problems in specific zones and, therefore, safety interventions could be implemented to improve the traffic safety situation (Huang et al., 2010). Furthermore, it is indispensable to take traffic safety into account already during the planning stage of transportation projects. To do so, traffic safety impacts of different transportation project alternatives should be compared and assessed by a number of factors which have zone-level characteristics (Huang et al., 2010).

Moreover, TDM strategies are usually performed and evaluated at geographically aggregated levels rather than merely at the level of individual intersections or road sections. Therefore, the impact of adopting a TDM strategy on transportation or traffic safety should also be evaluated at a level higher than the local consequences. Indeed, local level CPMs mostly aim to predict the safety effects of infrastructural improvements. However, these models are not typically designed to evaluate traffic safety impacts of TDM strategies; thus, the application of CPMs at a higher aggregation level will be more practical (Tarko et al., 2008).

7.5. Methodology

7.5.1. Data Preparation

The study area in this research is the Dutch-speaking region in northern Belgium, Flanders. Flanders has over 6 million inhabitants, or about 60% of the population of Belgium. As already mentioned before, an activity-based model within the FEATHERS framework is applied on the Flemish population to derive the in-depth information of Flemish people's travel behavior and travel demand for a null-scenario (current situation) and some TDM scenarios like teleworking, increasing fuel price, etc. FEATHERS produces the traffic demand by means of origin-destination (OD) matrices. These OD matrices include the number of trips for each traffic mode at different disaggregation levels (i.e. age, gender, day of the week, time of day and motive). This traffic demand is then assigned to the Flemish road network to obtain detailed exposure metrics at the network level. To carry out the assignment of vehicle trips to the road network, the user equilibrium method was selected. The fundamental nature of equilibrium assignment is that travelers will strive to find the shortest path (e.g. minimum travel time) from origin to destination, and network equilibrium occurs when no traveler can decrease his travel effort by shifting to a new path. This is an optimal condition, in which no user will gain from changing travel paths once the system is in equilibrium. Exposure metrics are then geographically aggregated to the TAZ level. This has been carried out at the zonal level, comprising 2,200 TAZs in Flanders. The average size of TAZs is 6.09 square kilometers with a standard deviation of 4.78 square kilometers. In addition, a set of socio-demographic and road network variables were collected for each TAZ (see Table 7.1).

According to the literature, "exposure" (i.e. number of trips (NOTs) and VKT) (e.g. in Lovegrove, 2005; Hadayeghi et al., 2010a, 2010b; Naderan and Shahi, 2010; Abdel-Aty et al., 2011a, 2011b; Pirdavani et al., 2012), "number of intersections" (e.g. in Hadayeghi et al., 2003; De Guevara et al., 2004; An et al., 2011; Pirdavani et al., 2012), "income level" (e.g. in Aguero-Valverde and Jovanis, 2006; Quddus, 2008; Wier et al., 2009; Huang et al., 2010; Pirdavani et al., 2012), "degree of urbanization" (e.g. in De Guevara et al., 2004; Pirdavani et al., 2012), "speed" (e.g. in Hadayeghi et al., 2003), "number of inhabitants" (e.g. in De Guevara et al., 2004; Wier et al., 2009), etc., are found to be important predictors of crashes. The crash data used in this study consist of a geo-coded set of fatal and injury crashes that occurred during the period 2004 to 2007. Table 7.1 shows a list of selected variables, together with their definition and descriptive statistics, which have been used in developing the ZCPMs presented in this chapter.

Table 7.1 Selected Variables to Develop ZCPMs

	Variable	Definition	Average	Min	Max	SD
Dependent variables	CCFS	total Car-Car/Fatal and Severe injury crashes observed in a TAZ	2.82	0	21	3.05
	CCSL	total Car-Car/Slight injury crashes observed in a TAZ	19.22	0	199	20.77
	CSFS	total Car-Slow mode/Fatal and Severe injury crashes observed in a TAZ	1.36	0	16	2.08
	CSSL	total Car-Slow mode/Slight injury crashes observed in a TAZ	10.07	0	202	17.81
Exposure variables	NOTs Car	average daily number of car trips originating/arriving from/at a TAZ	2944.5	0	18974.8	3031.6
	NOTs Slow	average daily number of slow-mode trips originating/arriving from/at a TAZ	1064.3	0	9289.8	1337.8
	Motorway VKT	average daily vehicle kilometers traveled on motorways in a TAZ	28449.6	0	1007437	87429.6
	Other Roads VKT	average daily vehicle kilometers traveled on other roads in a TAZ	27951.1	0	323688	29672.1
Network variables	Capacity	hourly average capacity of links in a TAZ	1790.1	1200	7348.1	554.6
	Intersection	total number of intersections in a TAZ	5.8	0	40	5.9
	Urban	Is the TAZ in an urban area? "No" represented by 0 "Yes" represented by 1	0	0	1	-[a]
	Suburban	Is the TAZ in a suburban area? "No" represented by 0 "Yes" represented by 1	0	0	1	-
Socio-demographic variables	Income Level	average income of residents in a TAZ described as below: "Monthly salary less than 2249 Euro" represented by 0 "Monthly salary more than 2250 Euro" represented by 1	1	0	1	-

a: Data not applicable

7.5.2. Motivation for Conducting Spatial Analysis

The most common modeling framework for ZCPMs is the GLM framework (e.g. in Amoros et al., 2003; Hadayeghi et al., 2003, 2006, 2007; Noland and Oh, 2004; Noland and Quddus, 2004; De Guevara et al., 2004; Lovegrove, 2005; Aguero-Valverde and Jovanis, 2006; Lovegrove and Sayed, 2006, 2007; Hadayeghi, 2009; Naderan and Shahi, 2010; Lord and Mannering, 2010; Abdel-Aty et al., 2011b; An et al., 2011; Pirdavani et al., 2012, in press). Within a GLM framework, fixed coefficient estimates explain the association between the dependent variable and a set of explanatory variables. In other words, a single model is fitted on the observed data for all locations (e.g. TAZs). However, not surprisingly different spatial variation, which is often referred to as "spatial non-stationarity", may be observed for different explanatory variables especially where the study area is relatively large. Neglecting this spatial variation may deteriorate the predictive power of ZCPMs and also has impacts on the significance of explanatory variables.

Checking for the existence of spatial correlation of dependent and explanatory variables can be carried out by means of different statistical tests such as "Moran's autocorrelation coefficient" commonly referred to as Moran's I (Lee and Wong, 2001).

Moran's I is an extension of the Pearson product-moment correlation coefficient to a univariate series. It may be expected that in the existence of spatial patterns, close observations are more likely to be similar than those far apart. Moran's I can be formulated as follows:

$$Moran's\ I = \frac{n}{SumW} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad (1)$$

Where n is the number of cases (number of TAZs in our study), $\overline{x}$ is the mean of x_i's, w_{ij} is the weight between cases i and j, and SumW is the sum of all w_{ij}'s:

$$SumW = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \quad (2)$$

The value of Moran's I varies from -1 representing complete spatial dispersion to 1 indicating full spatial clustering. Table 7.2 shows the Moran's I values for the selected variables used in the model construction. It is evident that all of the selected variables show significant spatial clustering. Table 7.2 also includes the significance level of Moran's I values by means of p-values and Z-scores. Z-scores can be derived as follows:

$$Z(MI_i) = \frac{O(MI_i) - E(MI_i)}{SD(MI_i)} \quad (3)$$

Where $Z(MI_i)$ is the Z-score of Moran's I of variable i, $O(MI_i)$ is the Observed Moran's I of variable i, $E(MI_i)$ is the expected Moran's I of variable i and $SD(MI_i)$ is the Standard-deviation of Moran's I of variable i. The results presented in Table 7.2 indicate the necessity of considering spatial autocorrelation when developing crash prediction models.

7.5.3. Model Construction

Inclusion of spatial variation in traffic safety studies has been considered by several researchers. However, there are different spatial modeling techniques that can be applied. Auto-logistic models, conditional auto-regression (CAR) models, simultaneous auto-regression (SAR) models, spatial error models (SEM), generalized estimating equation (GEE) models, Full-Bayesian spatial models, Bayesian Poisson-lognormal models are some of the most employed techniques to conduct spatial modeling in traffic safety (e.g. in Levine et al., 1995; Miaou et al., 2003; Flahaut, 2004; Aguero-Valverde and Jovanis, 2006, 2008; Wang and Abdel-Aty, 2006; Quddus, 2008; Wang et al., 2009; Guo et al., 2010; Huang et al., 2010; Siddiqui et al., 2012). The output of these models are still fixed variable estimates for all locations, however spatial variation is taken into account.

Another solution for taking spatial variation into account is developing a set of local models, so called geographically weighted regression (GWR) models (Fotheringham et al., 2002). These models rely on the calibration of multiple regression models for different geographical entities.

Table 7.2 Moran's *I* Statistics for Dependent and Selected Explanatory Variables

Variable	**Observed Moran's *I***	**Z-score**	**Spatial status**
CCFS	0.091	16.17	Non-stationary
CCSL	0.173	30.928	Non-stationary
CSFS	0.140	25.074	Non-stationary
CSSL	0.219	39.242	Non-stationary
log(NOTs Car)	0.177	31.606	Non-stationary
log(NOTs Slow)	0.222	39.648	Non-stationary
log(Motorway VKT)	0.067	12.016	Non-stationary
log(Other Roads VKT)	0.101	18.042	Non-stationary
Capacity	0.121	21.54	Non-stationary
Intersection	0.199	35.556	Non-stationary
Urban	0.437	78.088	Non-stationary
Suburban	0.239	42.539	Non-stationary
Income Level	0.187	33.318	Non-stationary

CCFS: Car-Car/Fatal and Severe injury crashes

CCSL: Car-Car/Slight injury crashes

CSFS: Car-Slow mode/ Fatal and Severe injury crashes

CSSL: Car-Slow mode/Slight injury crashes

The GWR technique can be adapted to GLM models (i.e. extend GLM models) and form geographically weighted generalized linear models (GWGLMs) (Fotheringham et al., 2002). GWGLMs are able to model count data (such as the number of crashes) while simultaneously accounting for spatial non-stationarity. Hadayeghi et al. (2010b) used the GWR technique in conjunction with the GLM framework using the Poisson error distribution. They developed different geographically weighted Poisson regression (GWPR) models to associate the relationship between crashes and a set of predictors. The comparison between GLMs and GWPR models revealed that the GWPR models clearly outperform the GLMs since they are capable of capturing spatially dependent relationships.

Reviewing the literature for different model forms showed that the following GLM model has been widely used in different studies (e.g. in

Lovegrove, 2005; Hadayeghi, 2009; Abdel-Aty et al., 2011b; An et al., 2011; Pirdavani et al., 2012):

$$E(C) = \beta_0 \times (Exposure)^{\beta_1} \times e^{\sum_{i=2}^{n} \beta_i x_i} \quad (4)$$

Where;

$E(C)$ is the expected crash frequency, β_0 and β_i are model parameters, $Exposure$ is the exposure variable (e.g. VKT or NOTs) and x_i's are the other explanatory variables. Logarithmic transformation of Equation (4) when considering only one exposure variable yields:

$$ln[E(C)] = ln(\beta_0) + \beta_1 ln(Exposure) + \beta_2 x_2 + \beta_3 x_3 + \ \dots + \beta_n x_n \quad (5)$$

The Geographically Weighted form of Equation (5) would be:

$$ln[E(C)(\boldsymbol{l}_i)] = ln(\beta_0(\boldsymbol{l}_i)) + \beta_1(\boldsymbol{l}_i) ln(Exposure) + \beta_2(\boldsymbol{l}_i) x_2 + \cdots + \beta_n(\boldsymbol{l}_i) x_n \quad (6)$$

The output of these models will be different location-specific estimates for each case (here each TAZ). All variable estimates are functions of each location (here the centroid of each TAZ), $\boldsymbol{l}_\mathrm{i} = (x_\mathrm{i}, y_\mathrm{i})$ representing the x and y coordinates of the i^{th} TAZ.

To account for severity of crashes, different models are developed at different severity levels; i.e. "fatal + severe injury" and "slight injury" crashes. Moreover, TDM scenarios have different safety impacts on different road users.

For instance, if implementing a TDM scenario results in transferring individuals out of private vehicles to non-motorized modes, the safety level of car users might be improved, but injury risk for pedestrians or cyclists is increased. Therefore, to address this issue, crashes are further disaggregated into two types namely "Car-Car" and "Car-Slowmode" crashes ("Slowmode" comprises pedestrians and cyclists) and different models are fitted for these different crash types. Hence, four GWPR models are developed to associate the relationship between crash frequency and the explanatory variables. These models are constructed using a SAS macro program (Chen and Yang, 2012). The selected models are shown in Table 7.3 represented by the minimum, maximum, 1^{st} quartile, median and 3^{rd} quartile of the parameter estimates.

Table 7.3 Model Estimates for the Final Chosen ZCPMs

	Model #1 (CCFS)	Model #2 (CCSL)	Model #3 (CSFS)	Model #4 (CSSL)
Coefficients	**Estimates**	**Estimates**	**Estimates**	**Estimates**
(Intercept)	-9.763, -2.692 (-6.517, -5.569, -4.445)[a]	-7.356, -3.077 (-5.611,-4.944,-4.196)	-11.797, -5.453 (-7.889,-7.317,-6.833)	-10.897, -3.994 (-6.574,-6.075,-5.63)
log(NOTs Car)	-0.035, 0.632 (0.093, 0.184,0.268)	0.194, 0.622 (0.352, 0.424,0.479)	-	-
log(NOTs Slow)	-	-	0.484, 1.222 (0.616, 0.745,0.838)	0.621, 1.165 (0.794, 0.917,1.008)
log(Motorways VKT)	-0.036, 0.047 (-0.002, 0.013,0.022)	-0.022, 0.041 (0.001, 0.011,0.018)	-0.073, 0.023 (-0.04, -0.02,-0.007)	-0.054, 0.044 (-0.019,-0.008,0.004)
log(Other Roads VKT)	0.169, 0.669 (0.348, 0.42,0.465)	0.171, 0.632 (0.296, 0.342,0.395)	-0.05, 0.511 (0.163, 0.239,0.311)	0.0243, 0.361 (0.133,0.178,0.229)
Capacity	2.8 e-5, 1.003e-3 (3.3e-4,4.5e-4,6.3e-4)	6.5 e-6, 9.8e-4 (3.5e-4,4.8e-4,6.3e-4)	-4.2e-4, 8.2e-4 (3.3e-5,1.6e-4,3.5e-4)	-7.02e-4, 6.06e-4 (-8.4e-5,4.2e-5,1.9e-4)
Intersection	-0.0296, 0.0611 (0.007,0.019,0.029)	-0.0096, 0.0484 (0.017,0.022,0.026)	-0.063, 0.086 (0.003,0.012,0.023)	-0.0523, 0.056 (0.005,0.015,0.027)
Income level	-	-0.467, 0.637 (-0.185,-0.109,0.053)	-0.562, 1.97 (-0.25,-0.129,0.089)	-0.658, 2.525 (-0.209,-0.078,0.062)
Urban	-1.829, -0.017 (-0.89,-0.68,-0.37)	-	-	-0.193, 1.216 (0.359,0.619,0.86)
Suburban	-0.85, 0.138 (-0.4,-0.29,-0.147)	-	-	-0.219, 0.841 (0.165,0.325,0.409)
PCC[b]	0.735	0.907	0.789	0.952

a: minimum, maximum, (1st quartile, median, 3rd quartile) of the parameter estimates.
b: The Pearson Correlation Coefficient (PCC) between observed and predicted crash values.

7.5.4. Traffic Safety Evaluation Process

Road crashes are known to be a function of two components; exposure and risk. By implementing the teleworking scenario, the risk component is kept constant and only the exposure factor will be changed. To compute the changes in exposure, OD matrices for both the null scenario and the teleworking scenario were derived from FEATHERS for scenario evaluation. After assigning the travel demand to the road network, all required variables become available to set up the evaluation task. Now, the final ZCPMs (see Table 7.3) are applied and crashes are predicted for each TAZ. The traffic safety evaluation can then be conducted by comparing the NOCs predicted by the final ZCPMs for the null and the teleworking scenario. Figure 7.1 depicts the conceptual framework of the traffic safety evaluation process in more detail.

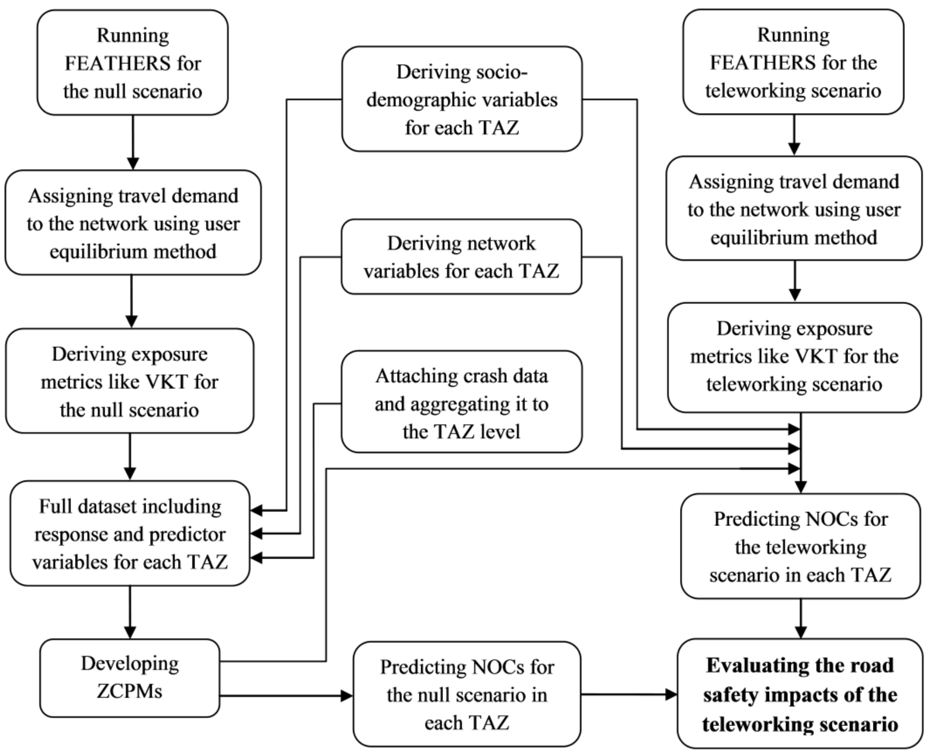

Figure 7.1 Conceptual framework of the traffic safety evaluation process.

7.6. Validation and Sensitivity Analysis

GWR models aim at identifying spatial heterogeneities in regression models of geo-referenced data. The spatial variability of the estimated local regression coefficients is usually examined to verify whether the underlying data shows signs of local deviations from a global regression model. In this regard, mapping the spatial GWR coefficient patterns associated with each variable may reveal some information. This approach, however, ignores possible dependencies among the local coefficient estimates linked with different variables. These dependencies can be expressed as the correlation between several sets of local coefficient estimates associated with different variables at all locations. Strong dependencies among the local coefficient estimates imply the fact that coefficients are not uniquely defined and as such, any convincing interpretation cannot be derived (Wheeler and Tiefelsdorf, 2005).

Due to the greater complexities of the GWR estimation procedure that conceivably causes interrelationships among the local estimates, it is essential to check for multicollinearity among local coefficient estimates. There are frequently used exploratory tools available to discover possible multicollinearity, such as bivariate scatter plots (See Appendix I) or bivariate correlation coefficients, however, a more statistical oriented measure that adopts a simultaneous view to identify multicollinearity is variance inflation factor (VIF). The VIF quantifies the severity of multicollinearity. It provides an index that measures how much the variance of an estimated regression coefficient is increased because of collinearity. Analyzing the magnitude of multicollinearity is carried out by considering the size of the VIF. As a common rule of thumb, 10 is defined (Kutner et al., 2004) as a cut off value meaning that if the VIF is higher than 10 then multicollinearity is high. VIF values among local coefficient estimates of models are shown in Table 7.4. These results suggest that multicollinearity among local coefficient estimates is not a problem in any of the developed models.

Table 7.4 VIF Among Local Coefficient Estimates of GWR Models

	Model #1 (CCFS)	Model #2 (CCSL)	Model #3 (CSFS)	Model #4 (CSSL)
Coefficients	**VIF value**	**VIF value**	**VIF value**	**VIF value**
log(NOTs Car)	4.737612	2.867101	-	-
log(NOTs Slow)	-	-	2.04558	1.899709
log(Motorways VKT)	3.587583	1.966296	2.712529	2.514306
log(Other Roads VKT)	1.681959	1.298513	1.627484	1.557763
Capacity	5.073428	2.356621	2.158348	2.541915
Intersection	2.849059	2.057096	1.9835	1.949019
Income level	-	1.205415	1.27295	1.37345
Urban	2.932739	-	-	1.853614
Suburban	2.781285	-	-	1.981406

CCFS: Car-Car/Fatal and Severe injury crashes

CCSL: Car-Car/Slight injury crashes

CSFS: Car-Slow mode/ Fatal and Severe injury crashes

CSSL: Car-Slow mode/Slight injury crashes

Due to the nature of GWR models which are location specific models, validation cannot be accomplished by conventional methods described in Chapter 6 (see section 6.7). Unlike traditional regression modeling in which a general model is fitted on training dataset and validated on a test dataset, GWR models are a series of local models, therefore, the concept of training and testing cannot be applied in the context of GWR models. However, a new framework is proposed in this research by which sensitivity of the predictability power of fitted models is checked. To this end, the whole dataset is randomly divided into 10 segments. In each round of model fitting one segment is left out, therefore, there will be 9 different models fitted for each single data point (here TAZ). Each of these models are developed by using the derived information from the neighboring TAZs. In this case, neighboring TAZs are changed in each round of model fitting for each TAZ. Robustness of the prediction models can be confirmed by checking the variability of predictions derived from 9 different models that are fitted for each TAZ. In case of having an acceptable low variation in predictions, it could be concluded that models are not sensitive to presence/absence of specific vicinity TAZs. Moreover, a low variation in predictions further confirms presence of spatial correlation and the right choice of bandwidth, meaning that missing information of left out TAZs are properly substituted by presence of other TAZs that have similar characteristics to the excluded TAZs. Comparing predictions of different local fitted models revealed a high predictive accuracy, substantiating the robustness of models. Since displaying the complete results is too much space-consuming, a few examples of this exercise are illustrated in Appendix II.

7.7. Results

In order to better understand the traffic safety impacts of the teleworking scenario, it is interesting to have a look at the changes in the traffic-related attributes playing a role in the whole chain due to the teleworking scenario. It turns out that the teleworking scenario reduces the average number of daily car trips by 1.5%, car passenger trips by 0.2%, public transportation trips by 1.9% and slow mode trips by 1%. Moreover, the analyses show that the total VKT by cars decreases by 3.2% after implementing the teleworking scenario.

Furthermore, in the teleworking scenario the total predicted number of crashes decreases compared to the null scenario as a result of this reduced exposure. The results show that the total number of CCFS and CCSL crashes is predicted to decrease by 175 and 1200 units respectively per 4 year period (-2.8% and -2.8%). Likewise, the total number of CSFS and CSSL crashes is predicted to decrease by 75 and 470 units respectively for the same period of 4 years (-2.5% and -2.1%). Additionally, Figure 7.2 shows violin plots of changes in the NOCs after the teleworking scenario implementation. A violin plot is a synergistic combination of a box plot and a density trace (Hintze and Nelson, 1998). These plots retain much of the information of box plots (except for the individual outliers), while also providing information about the distributional characteristics of the data. In these plots, the wider the violin, the more data points are associated to that value. The white dots indicate the median, the black boxes show the upper and lower quartile and the vertical black lines denote the upper and lower whiskers.

As can be seen in Figure 7.2, NOCs have increased in some TAZs as a result of an increase in travel demand and exposure in those specific TAZs while they are decreased in most TAZs. The reason for the increase of NOCs in a limited number of TAZs – specifically for CSFS and CSSL crashes – is due to an increase of exposure in those TAZs. This increase in the exposure can be explained by a secondary effect of the teleworking scenario where the remaining activities in teleworkers daily trip schedules - other than work trips - (e.g. shopping activity in the chain of home-work-shopping-home) are still carried out and substituted by other modes (e.g. Slowmode) and avoided work trips by teleworkers are partially substituted by extra generated traffic (e.g. generated traffic for shopping, bringing children to school, etc.).

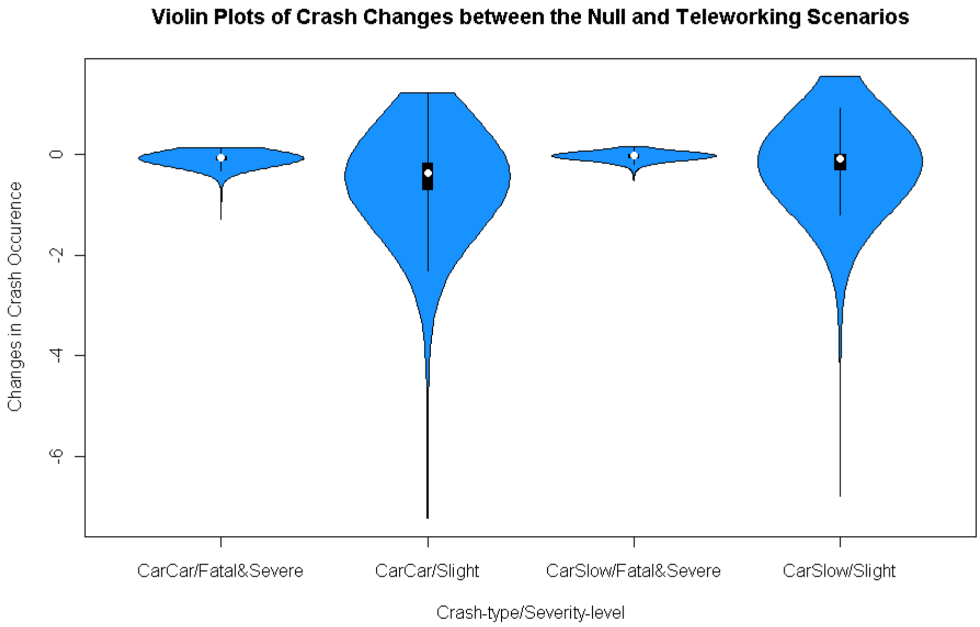

Figure 7.2 Violin plots of changes in crash occurrence after the teleworking scenario implementation.

In the development of CSFS and CSSL models, both car and Slowmode-related exposure variables were used. Following the implementation of the teleworking scenario, the total number of car and Slowmode trips decreased. However, these changes are not always similar in all TAZs. In fact, in more urbanized areas, the NOTs reduces more heavily and, therefore, also the NOCs reduces more rapidly in these areas. This can be explained by the fact that most of the commuters commute to urbanized areas. An illustration of changes in the NOCs for all TAZs may present a better indication on how different TAZs are affected by the scenario. In Figure 7.3, the changes in the predicted NOCs are displayed for each TAZ. Figure 7.3 reveals that the reductions in CCFS and CCSL crashes are greater for urban areas. As explained earlier, CSFS and CSSL crashes are also predicted to decrease more substantially in more urbanized areas; this is evident from the corresponding maps in Figure 7.3 where concentrations of blue dots stand for the major cities in Flanders.

7.8. Conclusions

In this study, the traffic safety impacts of a teleworking scenario are evaluated. To this end, ZCPMs are coupled with the activity-based model, FEATHERS. Based on the results of the analyses, the following conclusions can be drawn:

Activity-based transportation models provide an adequate range of in-depth information about individuals' travel behavior to realistically simulate and evaluate TDM strategies. The main advantage of these models is that the impact of applying a TDM strategy will be accounted for, for each individual, throughout a decision making process instead of applying the scenario on a general population level. Activity-based models, therefore, provide more reliable travel information since, unlike traditional models, TDM strategies are inherently accounted for in these models. Activity-based models follow a disaggregate modeling approach and as such, allow for a more detailed analysis of the reduction of travel demand due to the implementation of the teleworking scenario.

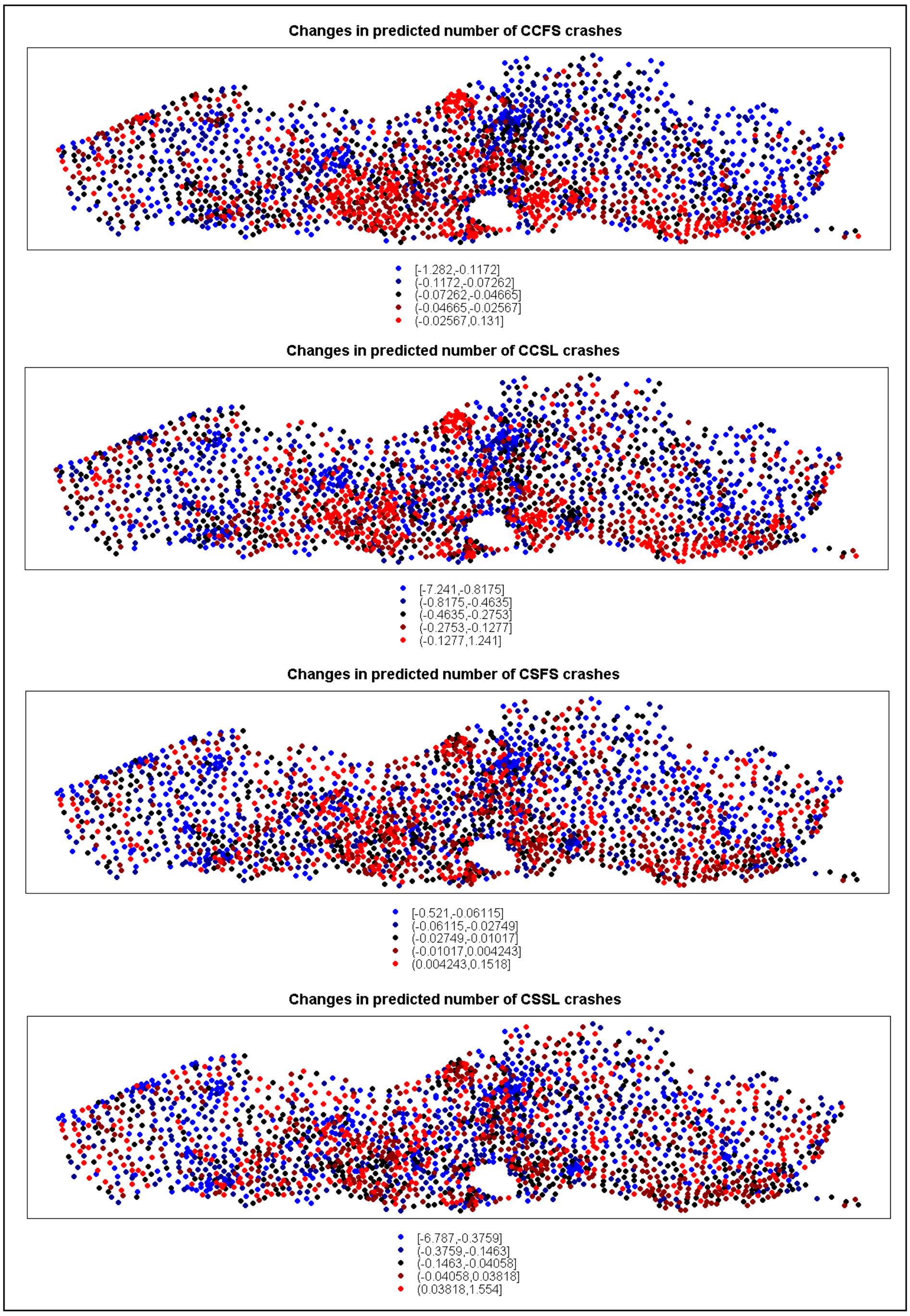

Figure 7.3 Changes in NOCs in each TAZ after the teleworking scenario implementation.

Analyzing crashes at a zonal level provides important information that enables us to compare traffic safety of different zones. This information is used to identify safety problems in specific zones and consequently, implementing safety interventions to improve the traffic safety situation. Furthermore, traffic safety should be taken into account during the planning stage of transportation projects. This can be carried out by associating the NOCs with a number of factors which have macro-level characteristics, such as socio-demographic, network level exposure, etc. Moreover, TDM strategies are usually performed at geographically aggregated levels. Therefore, it seems more appropriate to also evaluate the traffic safety impacts of TDM strategies at a zonal level.

In crash analysis, predictor variables are often found to be spatially heterogeneous especially when the study area is large enough to cover different traffic volume, urbanization and socio-demographic patterns. The results of the analysis confirm the presence of spatial variation of dependent and different explanatory variables which are used in developing crash prediction models. This was examined by computing Moran's *I* statistics for the dependent and selected explanatory variables. The results reveal the necessity of considering spatial correlation when developing crash prediction models. Therefore, different zonal GWPR models were developed, using different exposure, network and socio-demographic variables.

The results of the comparison analysis confirm that the teleworking scenario has many impacts such as the reduction of total travel demand, VKT and total crash occurrence. On the whole, there is an average reduction of 167,000 daily trips (all types of modes) as a result of the teleworking scenario. This scenario also causes a reduction of 1.4 billion VKT by cars per year, almost 3.2% of the total annual VKT by cars in Flanders.

The total NOCs is predicted to decrease by 1915 per 4 year period. As a result of the teleworking scenario and the average reduction in travel demand, CCFS, CCSL, CSFS and CSSL crashes are predicted to decrease by 2.8%, 2.8%, 2.5% and 2.1% respectively. This illustrates that teleworking can positively affect traffic safety of different road users and that noticeable safety benefits can be achieved. However, these positive impacts are slightly lower for "Car-Slowmode" crashes.

When considering the changes in the NOCs at the TAZ level, it turns out that especially urbanized areas (cities) benefit most from a general reduction of "Car-Car" and "Car-Slowmode" crashes.

Finally, this chapter presents an extension to the application of ZCPMs incorporated into TDM strategies. The results show the ability of ZCPMs as a predictive tool which can be used during the planning stage of transportation projects. Nevertheless, also some limitations of this study should be mentioned.

A constraint in application of GWPR models is that these models are not spatially transferable. This is due to the fact that GWPR models produce local parameter estimates (local models) for each TAZ which are influenced by their adjacent TAZs. Therefore, different models need to be developed for different study areas.

The teleworking scenario studied in this research investigated the relatively short-term effects of simulating 5% of the working population as teleworkers. In other words, the model is a short-term model in the sense that neither a shift in the composition of the vehicle fleet or car ownership, nor changes in the location of businesses and/or the location choice for living as a result of the teleworking scenario are assumed. Indeed in the longer run, it can be expected that teleworkers tend to change their living location and live closer to their working place and, therefore, the magnitude of trip reduction can be diminished.

Moreover, the real power of activity-based models has not yet been fully incorporated. In this study, the methodology relied on the aggregate daily traffic information. Activity-based models are however capable of providing disaggregate travel characteristics by differentiating between many household and person characteristics like gender, age, number of cars, etc. Hence, different types of disaggregation based on time of day, age, gender and motive are on the list of potential future research in order to take full advantage of the output of activity-based models.

References

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011a. Zonal Level Safety Evaluation Incorporating Trip Generation Effects. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Abdel-Aty, M., Siddiqui, C., Huang, H., 2011b. Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis & Prevention 38, 618–625.

Aguero-Valverde, J., Jovanis, P.P., 2008. Analysis of Road Crash Frequency with Spatial Models. Transportation Research Record: Journal of the Transportation Research Board 2061, 55–63.

Amoros, E., Martin, J.L., Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. Accident Analysis & Prevention 35, 537–547.

An, M., Casper, C., Wu, W., 2011. Using Travel Demand Model and Zonal Safety Planning Model for Safety Benefit Estimation in Project Evaluation. Presented at the Transportation Research Board (TRB) 90th Annual Meeting, Washington D.C. USA.

Arentze, T.A., Timmermans, H.J.P., 2004. ALBATROSS – Version 2.0 – A learning based transportation oriented simulation system. EIRASS (European Institute of Retailing and Services Studies).

Arentze, T.A., Timmermans, H.J.P., 2005. Incorporating Parametric Action Decision Trees in Computational Process Models of Activity-Travel Behavior: Theory and Illustration 567–572.

Bellemans, T., Kochan, B., Janssens, D., Wets, G., Arentze, T., Timmermans, H., 2010. Implementation Framework and Development Trajectory of FEATHERS Activity-Based Simulation Platform. Transportation Research Record: Journal of the Transportation Research Board 2175, 111–119.

Chen, V.Y.-J., Yang, T.-C., 2012. SAS macro programs for geographically weighted generalized linear modeling with spatial point data: Applications to health research. Computer Methods and Programs in Biomedicine 107, 262–273.

Choo, S., Mokhtarian, P.L., 2007. Telecommunications and travel demand and supply: Aggregate structural Equation models for the US. Transportation Research Part A: Policy and Practice 41, 4–18.

Choo, S., Mokhtarian, P.L., Salomon, I., 2005. Does telecommuting reduce vehicle-miles traveled? An aggregate time series analysis for the U.S. Transportation 32, 37–64.

Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J., Picado, R., 2007. Synthesis of first practices and operational research approaches in activity-based travel demand modeling. Transportation Research Part A: Policy and Practice 41, 464–488.

De Guevara, F.L.D., Washington, S., Oh, J., 2004. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. Transportation Research Record: Journal of the Transportation Research Board 1897, 191–199.

Dissanayake, D., Morikawa, T., 2008. Impact assessment of satellite centre-based telecommuting on travel and air quality in developing countries by exploring the link between travel behaviour and urban form. Transportation Research Part A: Policy and Practice 42, 883–894.

Flahaut, B., 2004. Impact of infrastructure and local environment on road unsafety: Logistic modeling with spatial autocorrelation. Accident Analysis & Prevention 36, 1055–1066.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression the analysis of spatially varying relationships. John Wiley & Sons Ltd, West Sussex, England.

Guo, F., Wang, X., Abdel-Aty, M., 2010. Modeling signalized intersection safety with corridor-level spatial correlations. Accident Analysis & Prevention 42, 84–92.

Hadayeghi, A., 2009. Use of Advanced Techniques to Estimate Zonal Level Safety Planning Models and Examine Their Temporal Transferability, PhD thesis, Department of Civil Engineering, University of Toronto.

Hadayeghi, A., Shalaby, A., Persaud, B., 2003. Macrolevel Accident Prediction Models for Evaluating Safety of Urban Transportation Systems. Transportation Research Record: Journal of the Transportation Research Board 1840, 87–95.

Hadayeghi, A., Shalaby, A., Persaud, B., 2007. Safety Prediction Models: Proactive Tool for Safety Evaluation in Urban Transportation Planning Applications. Transportation Research Record: Journal of the Transportation Research Board 2019, 225–236.

Hadayeghi, A., Shalaby, A., Persaud, B., 2010a. Development of Planning-Level Transportation Safety Models using Full Bayesian Semiparametric Additive Techniques. Journal of Transportation Safety & Security 2, 45–68.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010b. Development of planning level transportation safety tools using Geographically Weighted Poisson Regression. Accident Analysis & Prevention 42, 676–688.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accident Analysis & Prevention 38, 579–589.

Henderson, D.K., Mokhtarian, P.L., 1996. Impacts of center-based telecommuting on travel and emissions: Analysis of the Puget Sound Demonstration Project. Transportation Research Part D: Transport and Environment 1, 29–45.

Hintze, J.L., Nelson, R.D., 1998. Violin Plots: A Box Plot-Density Trace Synergism. The American Statistician 52, 181–184.

Huang, H., Abdel-Aty, M., Darwiche, A., 2010. County-Level Crash Risk Analysis in Florida. Transportation Research Record: Journal of the Transportation Research Board 2148, 27–37.

Janssens, D., Wets, G., Timmermans, H.J.P., Arentze, T.A., 2007. Modelling Short-Term Dynamics in Activity-Travel Patterns: Conceptual Framework of the Feathers Model. Presented at the 11th World Conference on Transport Research, Berkeley CA, USA.

Kochan, B., Bellemans, T., Cools, M., Janssens, D., Wets, G., 2011. An estimation of total vehicle travel reduction in the case of telecommuting: Detailed analyses using an activity-based modeling approach. Presented at the European Transportation Conference, Glasgow, Scotland, UK.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., Forthcoming. Validation of an Activity-Based Traffic Demand Model for Flanders Implemented in the Feathers Simulation Platform, in: Computational Intelligence for Traffic and Mobility. Atlantic Press.

Kochan, B., Bellemans, T., Janssens, D., Wets, G., 2008. Assessing the Impact of Fuel Cost on Traffic Demand in Flanders Using Activity-Based Models. Presented at the Travel Demand Management TDM, Vienna, Austria.

Koenig, B.E., Henderson, D.K., Mokhtarian, P.L., 1996. The travel and emissions impacts of telecommuting for the state of California Telecommuting Pilot Project. Transportation Research Part C: Emerging Technologies 4, 13–32.

Kutner, M.H., Nachtsheim, C.J., Neter, J., 2004. *Applied Linear Regression Models*. 4th ed, McGraw-Hill.

Lee, J., Wong, D.W.S., 2001. Statistical analysis with ArcView GIS. John Wiley & Sons, Inc., United States of America.

Levine, N., Kim, K.E., Nitz, L.H., 1995. Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators. Accident Analysis & Prevention 27, 675–685.

Litman, T., 2003. The Online TDM Encyclopedia: mobility management information gateway. Transport Policy 10, 245–249.

Litman, T., Fitzroy, S., 2012. Safe Travels: Evaluating Mobility Management Traffic Safety Impacts. Victoria Transport Policy Institute.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44, 291–305.

Lovegrove, G., Sayed, T., 2007. Macrolevel Collision Prediction Models to Enhance Traditional Reactive Road Safety Improvement Programs. Transportation Research Record: Journal of the Transportation Research Board 2019, 65–73.

Lovegrove, G.R., 2005. Community-Based, Macro-Level Collision Prediction Models, PhD thesis, University of British Columbia.

Lovegrove, G.R., Litman, T., 2008. Using Macro-Level Collision Prediction Models to Evaluate the Road Safety Effects of Mobility Management Strategies: New Empirical Tools to Promote Sustainable Development. Presented at the Transportation Research Board (TRB) 87th Annual Meeting, Washington D.C. USA.

Lovegrove, G.R., Sayed, T., 2006. Macro-level collision prediction models for evaluating neighbourhood traffic safety. Canadian Journal of Civil Engineering 33, 609–621.

Miaou, S.-P., Song, J.J., Mallick, B.K., 2003. Roadway Traffic Crash Mapping: A Space-Time Modeling Approach. Journal of Transportation and Statistics 6, 33–57.

Mokhtarian, P.L., Varma, K.V., 1998. The trade-off between trips and distance traveled in analyzing the emissions impacts of center-based telecommuting. Transportation Research Part D: Transport and Environment 3, 419–428.

Naderan, A., Shahi, J., 2010. Aggregate crash prediction models: Introducing crash generation concept. Accident Analysis & Prevention 42, 339–346.

Nilles, J.M., 1996. What does telework really do to us? World Transport Policy & Practice 2, 15–23.

Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention 36, 525–532.

Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis & Prevention 36, 973–984.

Noland, R.B., Quddus, M.A., 2005. Congestion and safety: A spatial analysis of London. Transportation Research Part A: Policy and Practice 39, 737–754.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., 2012. Developing Zonal Crash Prediction Models with a Focus on Application of Different Exposure Measures. Transportation Research Record: Journal of the Transportation Research Board, in press.

Pirdavani, A., Brijs, T., Bellemans, T., Kochan, B., Wets, G., in press Evaluating the road safety effects of a fuel cost increase measure by means of zonal crash prediction modeling. Accident Analysis & Prevention, in press, doi:10.1016/j.aap.2012.04.008.

Quddus, M.A., 2008. Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. Accident Analysis & Prevention 40, 1486–1497.

Siddiqui, C., Abdel-Aty, M., Choi, K., 2012. Macroscopic spatial analysis of pedestrian and bicycle crashes. Accident Analysis & Prevention 45, 382–391.

Tarko, A., Inerowicz, M., Ramos, J., Li, W., 2008. Tool with Road-Level Crash Prediction for Transportation Safety Planning. Transportation Research Record: Journal of the Transportation Research Board 2083, 16–25.

Vovsha, P., Bradley, M., 2006. Advanced Activity-Based Models in Context of Planning Decisions. Transportation Research Record: Journal of the Transportation Research Board 1981, 34–41.

Vu, S.T., Vandebona, U., 2007. Telecommuting and Its Impacts on Vehicle-km Travelled. Presented at the International Congress on Modelling and Simulation, University of Canterbury, Christchurch, New Zealand.

Wang, C., Quddus, M.A., Ison, S.G., 2009. Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England. Accident Analysis & Prevention 41, 798–808.

Wang, X., Abdel-Aty, M., 2006. Temporal and spatial analyses of rear-end crashes at signalized intersections. Accident Analysis & Prevention 38, 1137–1150.

Wheeler, D., Tiefelsdorf M., 2005. Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression. Journal of Geographical Systems 7 (2), 161–187.

Wier, M., Weintraub, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis & Prevention 41, 137–145.

8. Conclusions and Discussion

One of the most essential ingredients of a successful road safety analysis is having good quality data. Certainly, several types of information at different levels of detail are required, whether to comprehend the causality of crash occurrence, to identify crash hotspots, to develop crash prediction models, or to implement safety countermeasures and evaluate policy scenarios. In this context, historical crash data are the most important source of information that has been employed to carry out different types of road crash analysis. Although the presence of reliable crash data is essential to carry out crash analysis, there are numerous situations where reliable crash data are not available. Under these difficult circumstances, road safety analysis is challenging, but can be assessed indirectly by different means, such as gathering expert knowledge, conflict observation or road safety audit. These methodologies are different from the conventional road safety analyses since they are independent of historical crash data. On the contrary and in the case of most developed countries, several data collection systems have been set up to better understand transportation systems and their road safety level.

Road crash analysis can be performed by means of two main approaches; reactive and proactive. When crash data are available and safety interventions are immediately required, a reactive approach is often adopted. This approach is comprised of a set of safety interventions to improve the safety situation of existing crash hotspots. The reactive approaches are usually more costly as they aren't implemented until crashes have occurred. Traditional reactive road safety processes comprise activities such as data collection and management, identification of crash hotspots and furthermore analysis, development and realization of safety countermeasures. On the other hand, in the absence of crash data or at the planning level of transportation projects when no crash has occurred yet, a proactive approach seems to be more appropriate. A proactive approach has a preventive attribute and tries to confront safety problems before crashes have occurred. Road safety audits, conflict observations and policy impact assessment are examples of proactive approaches.

Microscopic crash analysis usually concerns a single infrastructural element such as a road segment or an intersection. On the contrary, macroscopic crash analysis concerns a relatively large area (e.g. TAZ level) and associates the probability of crash occurrence with a set of variables that are generally macroscopically characterized. In general, micro-level crash analysis is carried out in a reactive way and by utilizing micro-level CPMs.

TDM strategies, traffic policy measures or traffic forecasts for transportation projects are often adopted at a macro-level rather than on an individual intersection or road segment. Moreover, TDM measures try to change travel behavior of road users individually, however the impact of changed behavior on the transportation system or road safety should be addressed collectively and at an aggregate level. Therefore, micro-level crash analysis becomes less appropriate to evaluate safety levels of transportation projects or TDM strategies. When new transportation projects are proposed, macro-level safety impact evaluation is required. This is to ensure that new projects do not have an opposing safety impact on the whole transportation network. As such, to successfully perform a planning level road safety evaluation, it is necessary to utilize macro-level CPMs.

8.1. Overview of the Results

In this section, the results of the six chapters are discussed in relation to the research objectives that were stated in Chapter 1.

Chapter 2 discussed a proactive methodology, applied to identify and prioritize crash hotspots in the absence of crash data. In this chapter, a Delphi procedure combined with a multiple criteria decision making (MCDM) procedure were employed to prioritize crash hotspots. More specifically, expert's opinions about relevant identification criteria and their relative importance were obtained from a Delphi experiment. Subsequently, a MCDM procedure was adopted to prioritize road sections based on their performance on each of the selected criteria. One of the advantages of the MCDM method is its compensatory nature, i.e. its possibility of trade-offs between several decisions criteria. This means that the prioritization model will be more comprehensive with regard to using all pertinent criteria instead of using only a few of them. Furthermore, the proposed methodology is fairly simple and practical. Moreover, validation against a set of safe and unsafe road sections shows the capability of the presented methodology in identifying crash hotspots, although more extensive validation is required to corroborate the findings of this study. In summary, it can be concluded that the presented methodology is a reliable tool to analyze road crashes in the absence of crash data. Moreover, the results show the usefulness of adopting qualitative methodologies in the context of road safety. Qualitative approaches provide detailed insights into behaviors and values, to a depth of understanding which is often not possible using quantitative methods. However, this rather qualitative approach could be complemented by classic hotspot identification methods based on historical crash data.

In Chapter 3, a microscopic proactive methodology was presented to assess the traffic safety of signalized and unsignalized intersections. This was carried out by using a microsimulation software and by means of time headway and a proximal safety indicator, namely post-encroachment time (PET). To this end, several sets of scenarios based on different traffic volume and speed limit categories were defined and safety indicators were measured for each of these scenarios. The analyses in Chapter 3 revealed that increasing the speed limit on both major and minor approaches will deteriorate the safety level whilst its magnitude will be larger for higher ranges of traffic volume. Moreover, increasing traffic volume up to the point that doesn’t cause any traffic congestion will worsen the safety situation. Also at higher speed levels, there are more observed number of short headways compared with lower speed levels. Based on the results of Chapter 3, it can be concluded that driving at higher speed limits is a potential threat for traffic safety and will worsen the road safety situation.

Regarding the possible changes in behavioral aspects, it was found that drivers’ behavior in the microsimulator “S-Paramics” is sensitive to changes in speed limit. This might be due to different car-following models that are applied in the microsimulator. Not having control on what is being used in the software is a shortcoming when using a commercial software that at the first place was not intended to be utilized for safety analysis. Moreover, uncontrolled behavior of cars in specific circumstances (e.g. when cars are approaching intersections) makes commercial microsimulation models less appropriate to perform safety analysis. In other words, S-Paramics as a commercial microsimulation software - which was initially intended to assess traffic flow performance and not traffic safety - does not realistically simulate interactions between vehicles. This is due to the fact that in this type of commercial microsimulation software, average time headway (important to do traffic flow performance analysis) is correctly set whereas distributional (i.e. microscopic or individual) time headway (important to do safety analysis) is not necessarily accurate. There is an ongoing progress in calibrating and validating microsimulation models, mainly by providing realistic driver behavior related models (e.g. car-following model) that is pertinent to the real nature of human driving and not normative behaviors. Therefore, it is recommended to use properly calibrated microsimulation models to calculate proximal safety indicators due to the shortcomings stated above.

Chapter 4 presented a proactive approach which can be applied in the safety planning stage of transportation projects. Different zonal crash prediction models (ZCPMs) were developed to associate different exposure, network and socio-demographic variables with the number of injury crashes at the zonal level comprising of 2200 traffic analysis zones (TAZs) in Flanders, Belgium. The analyses in Chapter 4 revealed that sole use of number of trips (NOTs) - representing trip production/attraction of a TAZ - for crash prediction results in missing important information about the characteristics of travel demand. In other words, NOTs - as an exposure measure - does not contain information on trip time, trip length and route choice. Moreover, transit traffic which just passes through a TAZ can have a significant share of exposure observed in a TAZ. This part of exposure is left out by only using the NOTs as the exposure variable. Thus, other exposure variables which are sensitive to the impacts of trip assignment should be taken into account. This has been carried out by assigning the traffic demand by cars to the road network and computing exposure variables that are sensitive to the assignment like vehicle hours traveled (VHT) and vehicle kilometers traveled (VKT). Reviewing the results of the model comparison (i.e. comparing model's PCC and MSPE) showed that the models that contain the combination of two exposure variables outperform the models which only have one of the exposure variables (i.e. NOTs, VKT or VHT) in their formulation. Therefore, considering application of both flow-based (e.g. VHT or VKT) and trip-based (e.g. NOTs) exposure variables in ZCPM construction is strongly recommended.

In Chapter 5 of this dissertation, existence of spatial correlation in the data is examined and further included when developing ZCPMs. Application of generalized linear models (GLM) with the assumption of Negative Binomial error distribution is the most popular technique in crash prediction analysis. The results of GLM models are sets of fixed coefficient estimates which represent the average relationship between the dependent variable and other explanatory variables for all locations. These relationships are assumed to be constant across space; however, these explanatory variables are often found to be spatially heterogeneous especially where the study area is large enough to cover different traffic volume, urbanization and socio-demographic patterns. Presence of spatial variation was investigated by computing the Moran's *I* statistic for dependent and selected explanatory variables. The results revealed the necessity of considering spatial correlation when developing crash prediction models. Therefore, different geographically weighted Poisson regression (GWPR) models were developed, using different exposure, network and socio-demographic variables. GWPR models allow the estimations to vary where different spatial correlation among the variables exists. Comparing all developed models show that GWPR models always perform better than GLM models. This is due to the fact that GWPR models are capable of capturing the spatial heterogeneity of crashes. The selected GWPR model predicts the number of crashes properly; i.e. the Pearson's product moment correlation coefficient for this model is 0.93 which is quite a high correlation measure. Furthermore, comparing the graphical representation of the observed and predicted number of injury crashes (NOICs) shows a very similar pattern for crash occurrence. This indicates that the NOICs are very well predicted by the explanatory variables (e.g. "Other roads VKT", "NOTs Cars", "Intersection" are among the most significant variables) included in the model for most of the TAZs across the study area. In summary, it can be concluded that spatial correlation should always be taken into account regardless of the methodology being applied to account for this phenomenon.

In Chapter 6, ZCPMs derived from Chapter 4 have been integrated into a fuel-cost increase scenario to assess this TDM strategy’s impact on traffic safety. This assessment has been carried out by applying an activity-based travel demand model to derive the exposure metrics. In order to assess TDM’s traffic safety implications, it is essential to have TDM sensitive exposure measures. Activity-based transportation models provide an adequate range of in-depth information about individuals’ travel behavior to realistically simulate and evaluate TDM strategies. The advantage of using activity-based transportation models compared with traditional transportation models is that the activity-based models can be adjusted to simulate different TDM scenarios.

Unlike some prior studies where a reactive approach was generally used to evaluate the traffic safety impact of a fuel cost increase, in this study a proactive methodology was applied. This was carried out in an assessment exercise by assuming a 20% increase in fuel price. The results of the comparison analysis revealed that this fuel-cost increase scenario has an impact on total travel demand, total crash occurrence, VKT and VHT and mode shift. Overall, there is a reduction of on average 105,000 daily trips (all types of modes) as a result of the fuel-cost increase scenario. This scenario also causes a reduction of 5 billion VKT by cars per year, almost 11% of the total annual VKT by cars in Flanders.

The total number of crashes (NOCs) is predicted to decrease by 2,130 crashes for a period of 4 years. However, changes in the NOCs for different crash-type/severity-level are not identical. As a result of an increase in the NOTs for the “Slowmode” category, crashes which involve cyclists and pedestrians (i.e. vulnerable road users) are predicted to increase; i.e. CSFS and CSSL crashes increased by 0.5% and 0.8% respectively. On the other hand, CCFS and CCSL crashes are predicted to decrease by 4% and 4.8% respectively. This reveals that the fuel-cost increase scenario affects different road users differently. The traffic safety situation slightly deteriorates for vulnerable road users; nevertheless, there are noticeable safety benefits for “Car-Car” crashes and for the overall traffic safety situation.

In addition, when considering the changes in the NOCs at the TAZ level, it was found that the maximum reduction of “Car-Car” crashes and the maximum increase of "Car-Slowmode" crashes were both observed in urban areas (i.e. cities). It can be concluded that in cities, in contrast to other areas, there is a higher likelihood of finding an alternative mode for cars. In contrast, the TAZs in less urbanized regions and the TAZs nearby the Flemish borders usually lack good public transportation services. Therefore, we will not see many trips shifting from cars to other modes in less urbanized areas and consequently there is a more stable traffic safety situation in these TAZs despite conducting the fuel-cost increase scenario.

The fuel-cost increase scenario studied in this research investigated the relatively short-term effects of an increased fuel price. In other words, the model is a short-term model in the sense that neither a shift in the composition of the vehicle fleet, nor changes in the location of businesses and/or the location choice for living as a result of the change in fuel cost are assumed. Furthermore, crashes are known to be a function of two components; exposure and risk. It is therefore likely that a fuel price increase will impact people's driving behavior and their speed choice; i.e. drivers might try to reduce their fuel consumption by driving more slowly. As a result, it can be assumed that the risk component will also decrease after the fuel-cost increase scenario implementation. In Chapter 6 however, only the changes in the exposure component were taken into account, whereas the risk component was assumed to be constant. If we would include the risk component as well, the traffic safety benefits might be expected to be even larger than predicted in this chapter.

Chapter 7 analyzed the traffic safety impacts of a teleworking scenario. To this end, ZCPMs derived from Chapter 5 are coupled with the teleworking scenario simulated by the activity-based model, FEATHERS. The results of the comparison analysis confirmed that the teleworking scenario has an impact on total travel demand, VKT and total crash occurrence. On the whole, there is an average reduction of 167,000 daily trips (all types of modes) as a result of the teleworking scenario. This scenario also causes a reduction of 1.4 billion VKT by cars per year, almost 3.2% of the total annual VKT by cars in Flanders.

The total NOCs is predicted to decrease by 1915 crashes per 4 year period. As a result of the teleworking scenario and the average reduction in travel demand, CCFS, CCSL, CSFS and CSSL crashes are predicted to decrease by 2.8%, 2.8%, 2.5% and 2.1% respectively. This illustrates that teleworking can positively affect traffic safety of different road users and that noticeable safety benefits can be achieved. However, these positive impacts are slightly lower for "Car-Slowmode" crashes, since the NOTs by Slowmode reduced slightly less than the NOTs by cars.

When considering the changes in the NOCs at the TAZ level, it turns out that especially urbanized areas benefit most from a general reduction of "Car-Car" and "Car-Slowmode" crashes. It can be concluded that in cities, in contrast to other areas, there is a higher likelihood of finding people who telework.

Despite the average reduction of crash occurrence after the teleworking scenario implementation, NOCs have increased in a limited number of TAZs. The reason for the increase of NOCs in those TAZs – specifically for CSFS and CSSL crashes – is due to an increase of exposure in those TAZs. This increase in the exposure can be explained by a secondary effect of the teleworking scenario where the remaining activities in teleworkers daily trip schedules - other than work trips - (e.g. shopping activity in the chain of home-work-shopping-home) are still carried out and substituted by other modes (e.g. Slowmode) and avoided work trips by teleworkers are partially substituted by extra generated traffic (e.g. generated traffic for shopping, bringing children to school, etc.).

8.2. Reflections

This section offers a discussion of several issues and recommendations for further research that are related to the present research and which brings the results into a broader perspective.

First, the results strongly confirm the literature review findings that in different circumstances, crash analysis should be performed at different levels of disaggregation. In case of a lack of reliable crash data, methodologies such as gathering expert's knowledge (see Chapter 2) or conflict observation (see Chapter 3) are found to be capable tools in analyzing road crashes. These two methodologies were adopted at the level of individual road elements (i.e. intersection or road segment). As a general recommendation, less data intensive methodologies carried out at a micro-level are more suitable approaches for crash analysis in the absence of crash data. On the other hand and when reliable data are available, it is possible to perform a more complex crash analysis.

Since considering traffic safety during the planning stage of transportation projects is recognized as an essential task among the decision makers and road engineers, application of these models (e.g. crash prediction models) can better express the road safety level of those transportation projects. This can be carried out by associating the NOCs with a number of factors such as socio-demographic, network level exposure and etc., which generally have macro-level characteristics. Moreover, as it was mentioned in Chapters 6 and 7, TDM strategies are usually performed at geographically aggregated levels. Therefore, it seems more appropriate to also evaluate the traffic safety impacts of TDM strategies at an aggregated level. In summary, it can be concluded that in case of availability of reliable data, traffic safety assessment of transportation projects or TDM strategies can be better performed at a macro-level (e.g. TAZ level) by means of zonal crash prediction models (ZCPMs).

Second, comparing the performance of several prediction models, it can be confirmed that selecting appropriate exposure measures is a very crucial task in developing CPMs. As confirmed by the results of Chapter 4, the sole use of NOTs originating/destining from/to a TAZ, as the exposure metric, will result in missing some important information about the characteristics of the existing travel demand. In other words, NOTs, as an exposure variable, does not contain information on trip time, trip length and route choice. Moreover, transit traffic which just passes through a TAZ can have a significant share of the exposure observed in that TAZ. This part of the exposure is left out when only using the NOTs as the exposure metric. Thus, other exposure variables which are sensitive to the impacts of trip assignment should be taken into account. This has been carried out by assigning the traffic demand to the road network and by computing exposure variables such as VHT and VKT. In summary, it can be concluded that both types of exposure (i.e. trip-based exposure and flow-based exposure) should be considered in the development of CPMs since each of them have different exposure related characteristics.

Third, the results of the analyses demonstrate the usefulness of activity-based models for the assessment of TDM scenarios. Indeed, an important asset of activity-based models is their integrated approach towards activities and travel. Due to this approach, it can be taken into account that certain trips, which are linked to the activities that are not so flexible (e.g. work activities) are less likely to be altered under changing traffic system conditions than other activities (e.g. leisure activities). In addition, activity-based models are not only able to predict changes in the demand for travel, but they also predict shifts between different modes of transport and the reallocation of activities due to the imposed measures. Providing a structured approach to agent-based modeling of activities and travel for individuals, the FEATHERS framework is able to account for TDM strategies. For instance, when applying a fuel cost increase scenario, FEATHERS can predict the impact on the NOTs, modal shift and changes in trip time and length. Moreover, activity-based transportation models provide an adequate range of in-depth information about individuals' travel behavior to realistically simulate and evaluate TDM strategies. The advantage of these models is that the impact of applying a TDM strategy will be accounted for, for each individual, throughout a decision making process. Activity-based models provide more reliable information since, unlike traditional models, TDM strategies are inherently accounted for in these models. Until now, the methodology implemented in this dissertation relied on the aggregate daily traffic information. Activity-based models are capable of providing disaggregate travel characteristics. Hence, different types of disaggregation based on time of day, day of the week, age, gender, motive, and etc., are on the list of potential future research in order to take full advantage of activity-based models. These disaggregation are expected to have constructive policy implications. In other words, having a better insight into different safety consequences of a potentially implementable policy scenario enables policy makers to enhance their road safety programs. Furthermore, unveiling more detailed traffic safety impacts of a traffic policy measure and highlighting its benefits can also strengthen the implementation of that policy among policy makers.

Forth, another issue that needs further exploration is the update of crash data and exploration of new explanatory variables to be included in the development process of CPMs. Currently, the crash data used in this study, consist of a geo-coded set of injury crashes that occurred during the period 2004 to 2007. Quality of crash data can be improved by including more recent crash data and, moreover, by including other types of crashes, like property damage only crashes. Examining new explanatory variables, besides disaggregating current variables based on different categories such as time of day, day of the week, gender or age, might help in a better prediction of crashes. Moreover, improving the prediction model structure by applying new modeling frameworks (e.g. multivariate models based on copulas) is currently being undertaken and further investigation is required within our future research plans.

Finally, the review of applied statistical models in the crash analysis literature revealed that the most commonly used model in crash data modeling is the GLM framework by assuming the NB distribution error. The same modeling framework was developed in Chapter 4 and further employed in Chapter 6 of this dissertation. The results of the analysis confirm that this modeling framework is an appropriate approach to model crash data. However, the crash occurrences are often spatially heterogeneous and are affected by many spatial variables, especially where the study area is large enough to cover different traffic volume, urbanization and socio-demographic patterns. Since the results revealed the necessity of considering spatial correlation, geographically weighted regression (GWR) models within a GLM framework were utilized to develop ZCPMs. These models allow the estimations to vary where different spatial correlation among the variables exists. The results of the model performance comparisons revealed that indeed the GWPR models always outperform the GLM models. This is due to the capability of GWPR models in capturing spatial heterogeneity of crashes. There are several approaches studied in the literature to account for this spatial correlation, however, it was not within the scope of this dissertation to compare different methodologies. In summary, it can be concluded that in the case of having spatially heterogeneous data, accounting for this phenomenon will improve the predictability power of CPMs.

8.3. Policy Recommendations

The previous section offers several recommendations and directions for future research while its emphasis was mainly on methodological recommendations rather than policy recommendations. Nonetheless, a few recommendations with respect to the studied policies can be formulated as follows.

First, the results of the crash analysis at intersections demonstrated the close relationship between driving at higher speed limits and near-crash events. It is concluded that increasing the speed limit on both major and minor approaches of an intersection will deteriorate the safety level, and its magnitude will be greater for higher ranges of traffic volumes. Besides, increasing traffic volume up to the point that no traffic congestion occurs, will worsen the safety situation. Mean values of PET decreased by increasing speed limit and traffic volume. Also at higher speed levels, there are more short headways observed compared with lower speed levels. In summary, it can be concluded that increasing the speed limit will make the traffic safety situation worse. Therefore, relevant policies by means of speed reduction programs or where applicable, restricting speed limit are strongly recommended to improve safety level of roads and particularly intersections.

Second, the results of the comparison analyses suggest that the fuel-cost increase scenario has several effects such as a reduction of total travel demand, total crash occurrence, VKT and VHT and mode shift. This scenario also causes a reduction of nearly 5 billion VKT by cars per year, almost 11% of the total annual VKT by cars in Flanders. Moreover, the total NOCs is predicted to decrease by 2,130 for a 4 year period. However, changes in the NOCs for different crash types are not identical. CSFS and CSSL crashes (crashes which involve cyclists and pedestrians; i.e. vulnerable road users) are predicted to increase, while CCFS and CCSL crashes are predicted to decrease. This shows that the fuel-cost increase scenario has a kind of mixed effect that affects different road users differently. The traffic safety situation slightly deteriorates for vulnerable road users; nevertheless, there are noticeable safety benefits for "Car-Car" crashes and for the overall traffic safety situation. In addition, the greatest magnitude of changes in the NOCs were observed in urbanized areas.

This is due to the higher likelihood of finding an alternative mode for cars in more urbanized areas where avoided car trips are easily shifted towards other modes (e.g. biking or walking). Therefore, car trips decrease, Slowmode trips increase and consequently, the maximum reduction of "Car-Car" crashes and the maximum increase of "Car-Slowmode" crashes are both observed in urban areas. In summary, adopting a fuel-cost increase policy can generally be recommended from the road safety point of view and due to its positive impacts on crash frequency reduction. However, the slight negative effect on the traffic safety level of vulnerable road users requires special attention. These negative impacts can be diminished by improving cycle paths infrastructure, improving public transportation efficiency, etc. Moreover, it should be noticed that these positive impacts are fully realizable, only in the short term. In the long run and due to the shift in the composition of the vehicle fleet or changes in the location of businesses and/or the location choice for living, these positive impacts might erode. Hence, fuel-cost increase strategies should be considered as medium term effective TDM policies, with respect to their road safety impacts.

The third policy recommendation is related to the safety impacts of the teleworking scenario, studied in Chapter 7. The results of the comparison analyses confirm that all types of crashes are predicted to decrease as a result of conducting the teleworking scenario. This illustrates that teleworking can positively affect traffic safety of different road users and that noticeable safety benefits can be achieved by adopting this TDM policy. However, despite the average reduction of NOCs (for all types of studied crashes) in the study area, crashes in a limited number of TAZs – specifically for CSFS and CSSL crashes – increased. This can be explained by a secondary effect of the teleworking scenario where the remaining trips in teleworkers daily trip schedule (i.e. non-work-related trips such as shopping, bringing children to school, and etc.) are replaced by other modes. Non-work related trips are typically short distance trips and are covered by walking, biking or public transportation. Therefore, the number of non-car trips are expected to increase in some TAZs and as a result of this, CSFS and CSSL crashes are predicted to increase in those TAZs. In summary, it can be recommended that implementing a teleworking scenario should be accompanied by improving level-of-service of public transportation, providing safer cycle paths, and etc. Assessing the impacts of simultaneously implementation of these measures could be another extension for future research.

Finally, due to the observed mixed effects of TDM scenarios on safety levels of different road users, decision makers and road engineers are strongly recommended to make a distinction between different road users when carrying our any safety assessment. Moreover, combined policies might complement each other and accordingly, desired safety benefits might be realized with more confidence. Another policy related issue that needs further exploration is the safety assessment of other TDM policies; i.e. separately assessment of an aging population, a public transportation level-of-service improvement and their combination with the studied policies in this dissertation are on the list of the future research agenda.

Appendices

Appendix I: Bivariate Scatter Plots of Local Coefficient Estimates

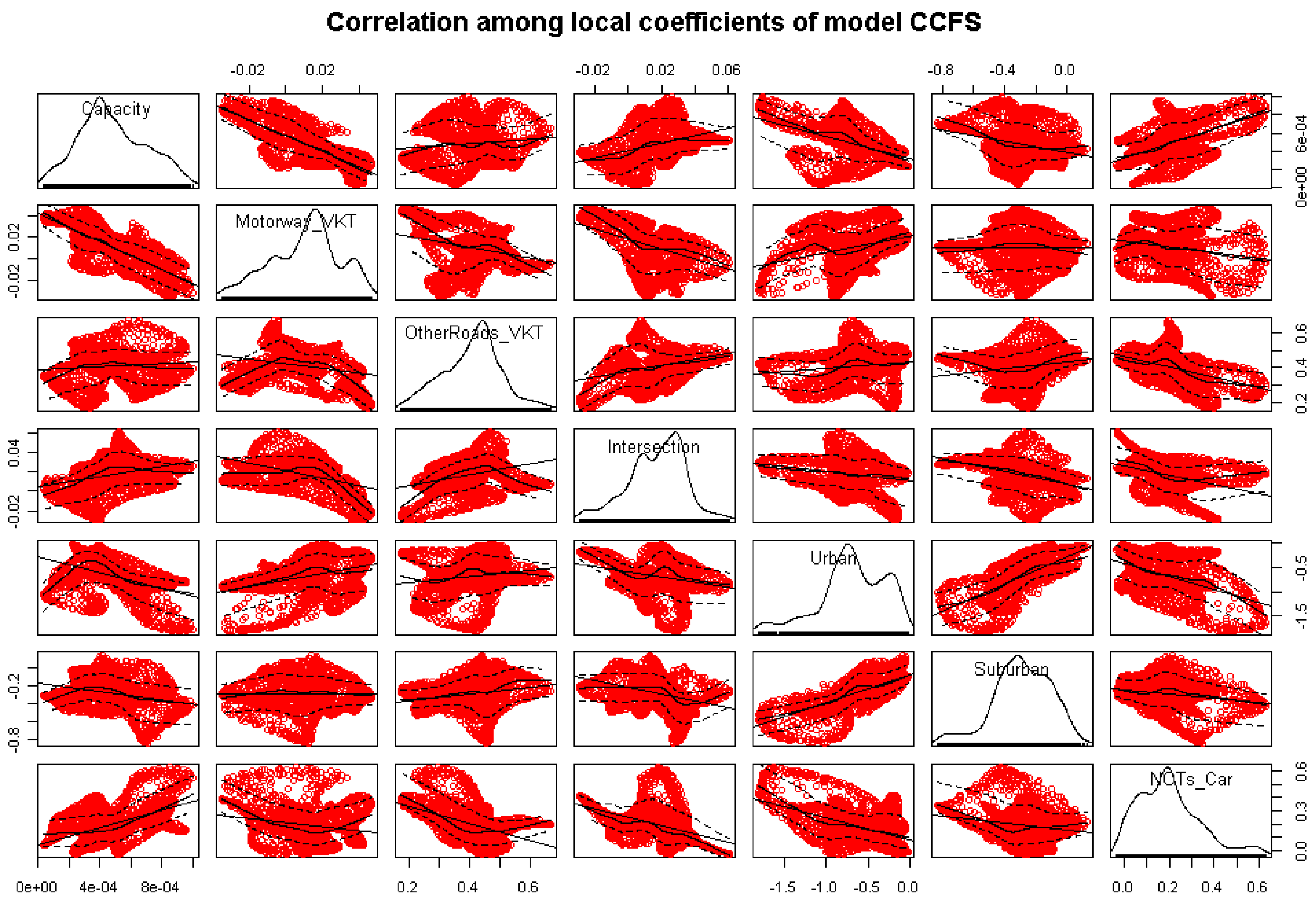

Figure AI-1 Bivariate scatter plots of local coefficient estimates for model CCFS.

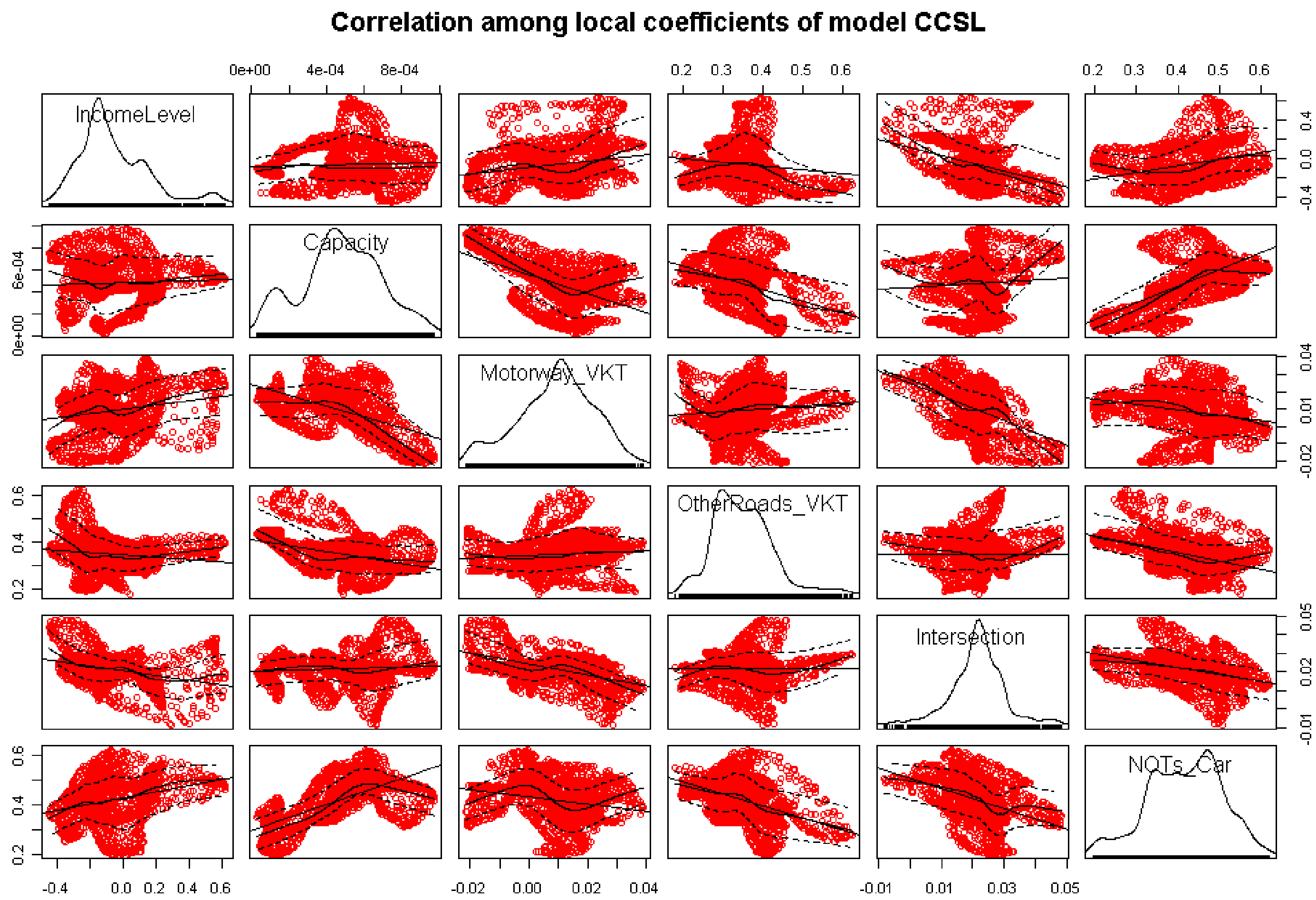

Figure AI-2 Bivariate scatter plots of local coefficient estimates for model CCSL.

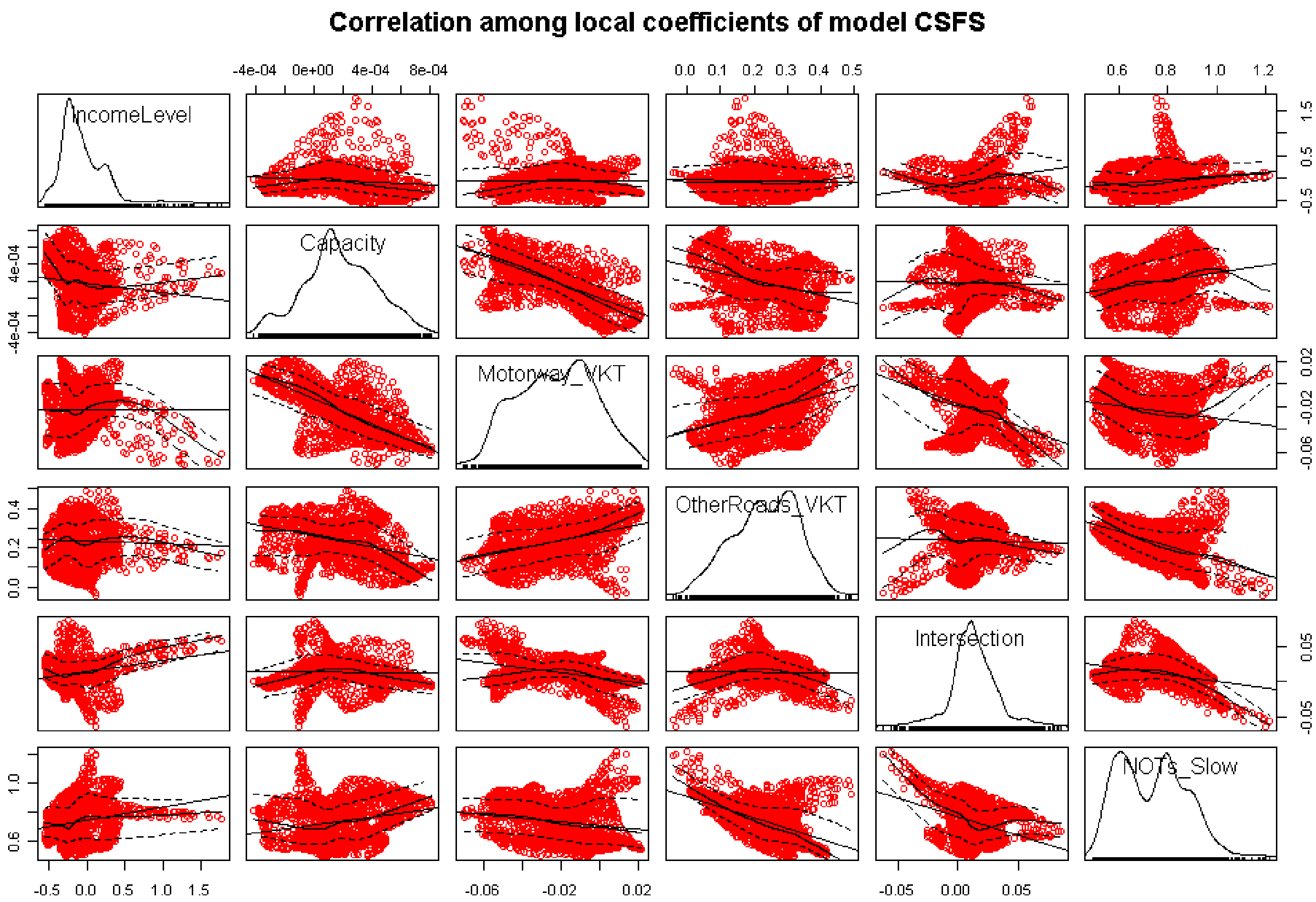

Figure AI-3 Bivariate scatter plots of local coefficient estimates for model CSFS.

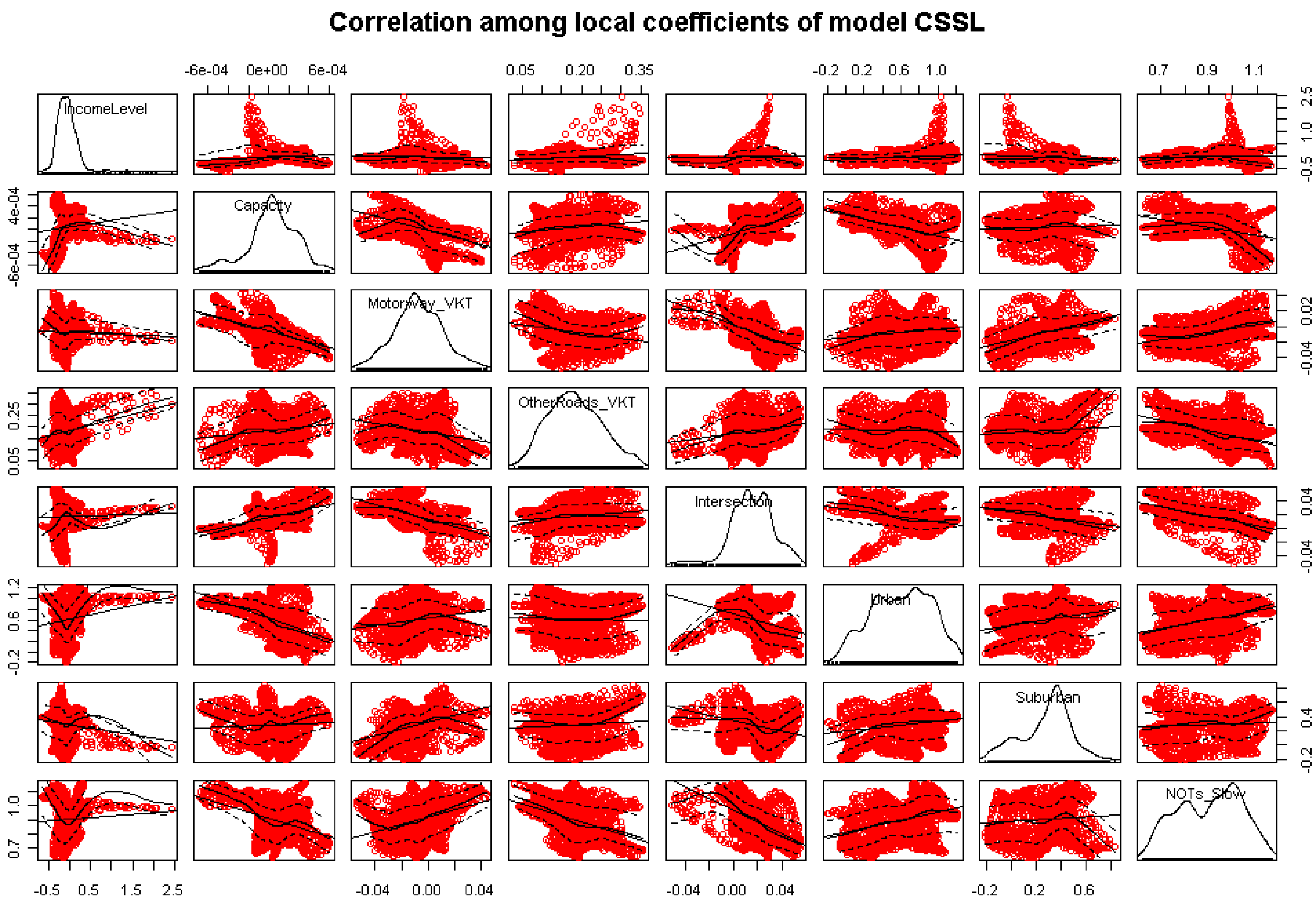

Figure AI-4 Bivariate scatter plots of local coefficient estimates for model CSSL.

Appendix II: Descriptive Statistics of an Example of Different Local Fitted Models at 99% Confidence Interval

Table AII-1 Results of Different Local Fitted Models for TAZ #2

Fitted model #	Observed number of CCFS crashes	Predicted number of CCFS crashes		
1	12	10.69954789	Mean	11.06542468
2	12	10.86517278	STDEV[a]	0.458422145
3	12	10.47775955	CI[b]	0.512668765
4	12	11.22889212	LB[c]	10.55275592
5	12	11.80425896	UB[d]	11.57809345
6	12	11.2696978		
7	12	11.2448744		
8	12	11.50260848		
9	12	10.49601019		
Fitted model #	**Observed number of CCSL crashes**	**Predicted number of CCSL crashes**		
1	91	78.78810318	Mean	79.81138598
2	91	82.07545584	STDEV	1.831160644
3	91	78.76930636	CI	2.047847987
4	91	78.99614116	LB	77.76353799
5	91	79.68390943	UB	81.85923396
6	91	79.12014815		
7	91	79.1409907		
8	91	78.05875445		
9	91	83.66966451		

[a] STDEV: standard deviation

[b] CI: confidence interval

[c] LB: lower bound

[d] UB: upper bound

Continued Table AII-1 Results of Different Local Fitted Models for TAZ #2

Fitted model #	Observed number of CSFS crashes	Predicted number of CSFS crashes		
1	2	2.727956328	Mean	2.944180737
2	2	2.676639623	STDEV[a]	0.241216253
3	2	2.961407873	CI[b]	0.269760176
4	2	2.973940562	LB[c]	2.674420562
5	2	3.342271122	UB[d]	3.213940913
6	2	2.969661942		
7	2	2.904076808		
8	2	2.667674532		
9	2	3.273997848		
Fitted model #	Observed number of CSSL crashes	Predicted number of CSSL crashes		
1	23	21.19860886	Mean	21.79908682
2	23	20.7303711	STDEV	0.779124382
3	23	21.75530003	CI	0.871320767
4	23	23.31339789	LB	20.92776606
5	23	21.31917755	UB	22.67040759
6	23	21.27039001		
7	23	22.00122234		
8	23	22.33806334		
9	23	22.26525028		

[a] STDEV: standard deviation

[b] CI: confidence interval

[c] LB: lower bound

[d] UB: upper bound

List of Publications (2009-2013)

2013

Proceedings Paper

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Kochan, Bruno & Wets, Geert, (2013). Assessing the Impacts of a Teleworking Policy on Crash Occurrence: The Case of Flanders, Belgium.
In: DVD Compendium of the 92nd Transportation Research Board Conference, Washington D.C., USA.

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Kochan, Bruno & Wets, Geert, (2013). Spatial Analysis of Fatal and Injury Crashes in Flanders, Belgium; Application of Geographically Weighted Regression Technique.
In: DVD Compendium of the 92nd Transportation Research Board Conference, Washington D.C., USA.

2012

Journal Contribution

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Kochan, Bruno & Wets, Geert, (2012). Evaluating the Road Safety Effects of a Fuel Cost Increase Measure by means of Zonal Crash Prediction Modeling.
In: Accident Analysis and Prevention, doi:10.1016/j.aap.2012.04.008.

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Kochan, Bruno & Wets, Geert, (2012). Application of Different Exposure Measures in Developing Planning-level Zonal Crash Prediction Models.
In: Transportation Research Record: Journal of the Transportation Research Board, in press.

Dhondt, Stijn; Pirdavani, Ali; Macharis, Cathy; Bellemans, Tom & Putman, Koen, (2012). Translating Road Safety into Health Outcomes Using a Quantitative Impact Assessment Model.
In: Injury Prevention, doi:10.1136/injuryprev-2011-040286.

Dhondt, Stijn; Kochan, Bruno; Beckx, Carolien; Lefebvre, Wouter; Pirdavani, Ali; Degraeuwe, Bart; Bellemans, Tom; Int Panis, Luc; Macharis, Cathy & Putman, Koen, (2012). Integrated Health Impact Assessment of Travel Behavior: Model Exploration and Application to a Fuel Price Increase.

In: Environment International, in press.

Proceedings Paper

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Kochan, Bruno & Wets, Geert, (2012). Developing Zonal Crash Prediction Models with a Focus on Application of Different Exposure Measures.
In: DVD Compendium of the 91st Transportation Research Board Conference, Washington D.C., USA.

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Kochan, Bruno & Wets, Geert, (2012). Application of Zonal Crash Prediction Models in Traffic Safety Evaluation of a Fuel-Cost Increase Scenario Using an Activity-Based Transportation Model.
In: DVD Compendium of the 91st Transportation Research Board Conference, Washington D.C., USA.

2011

Journal Contribution

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom & Wets, Geert, (2011). Travel Time Evaluation of a U-Turn Facility Comparison with a Conventional Signalized Intersection.
In: Transportation Research Record: Journal of the Transportation Research Board, 2223. p. 26-33, doi: 10.3141/2223-04.

Proceedings Paper

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom & Wets, Geert, (2011). Travel Time Evaluation of a U-Turn Facility and Its Comparison with a

Conventional Signalized Intersection.
In: DVD Compendium of the 90th Transportation Research Board Conference, Washington D.C., USA.

2010

Journal Contribution

Pirdavani, Ali; Brijs, Tom & Wets, Geert, (2010). A Multiple Criteria Decision-Making Approach for Prioritizing Accident Hotspots in the Absence of Crash Data.
In: Transport Reviews, 30(1). p. 97-113,
doi:10.1080/01441647.2010.534941.

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom & Wets, Geert, (2010). Evaluation of Traffic Safety at Un-signalized Intersections Using Microsimulation:A Utilization of Proximal Safety Indicators.
In: Advances in Transportation Studies, 22(B). p. 43-52,
doi: 10.4399/97888548354814.

Proceedings Paper

Pirdavani, Ali; Brijs, Tom & Wets, Geert, (2010). Multiple Criteria Decision-Making Approach for Prioritizing Accident Hot Spots in Developing Countries in the Absence of Crash Data.
In: DVD Compendium of the 89th Transportation Research Board Conference, Washington D.C., USA.

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom & Wets, Geert, (2010). A Simulation-Based Traffic Safety Evaluation of Signalized and Un-Signalized Intersections.
In: Proceedings of the 15th International Conference Road Safety on Four Continents. p. 1229-1239, Abu-Dhabi, UAE.

Fallah Zavareh, Mohsen; Pirdavani, Ali & Saffarzadeh, Mahmoud, (2010). A safety audit approach in ranking urban road crash hotspots using analytical hierarchy process (AHP). Case study: region 20 of Tehran.

In: Proceedings of the 15th International Conference Road Safety on Four Continents. p. 1142-1153, Abu-Dhabi, UAE.

2009

Proceedings Paper

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom & Wets, Geert, (2009). Evaluation of traffic safety at un-signalized intersections using microsimulation.

In: Electronic Proceedings of the 2nd Road Safety and Simulation Conference, Paris, France.

Pirdavani, Ali; Brijs, Tom; Bellemans, Tom; Wets, Geert & Vanhoof, Koen, (2009). Travel Time Evaluation of an innovative U-turn facility on urban arterial roadways.

In: Proceedings of International Conference on Urban, Regional Planning and Transportation, Paris, France.

Samenvatting

Verkeersslachtoffers veroorzaken een enorme sociale en economische kost die al decennialang behoren tot de grootste problemen in de volksgezondheid. Men verwacht een dramatische stijging van het huidige sterftecijfer in de komende twee decennia, tenzij verkeersongevallen grondig geanalyseerd worden en effectieve veiligheidsinterventies op de juiste wijze genomen worden zodat deze ongewenste toename verlaagd kan worden. Om verkeersongevallen te kunnen analyseren, moet men in eerste instantie relevante gegevens verzamelen. Deze gegevens worden in verschillende toepassingen gebruikt om de ongevalsoorzaken te begrijpen, ongevalshotspots te identificeren, ongevalspredictiemodellen te ontwikkelen, verschillende verkeersbeleidsmaatregelen te evalueren, etc. In dit opzicht vormen ongevalsgegevens de belangrijkste bron van informatie om ongevalsanalyses uit te voeren. Voor een succesvolle ongevalsanalyse moeten deze gegevens nauwkeurig, tijdig, volledig en betrouwbaar zijn. Hoewel de aanwezigheid van betrouwbare ongevalsgegevens een sleutelelement is bij ongevalsanalyses, zijn er situaties waar degelijke ongevalsgegevens niet beschikbaar zijn. In deze omstandigheden kan men de verkeersveiligheid indirect beoordelen door gebruik te maken van methodologieën die afwijken van de conventionele verkeersveiligheidsanalyse omdat ze niet afhankelijk zijn van historische ongevalsgegevens. In de meerderheid van de ontwikkelde landen, daarentegen, zijn er talrijke systemen voor gegevensverzameling beschikbaar die een meer complexe ongevalsanalyse vergemakkelijken.

Het eerste belangrijke doel van deze dissertatie is daarom het aanleveren van een verzameling van verschillende methodologieën voor de predictie en analyse van verkeersveiligheid, afhankelijk van de hoeveelheid en het type beschikbare gegevens. Zo zijn twee hoofdstukken (meer bepaald hoofdstuk 2 en 3) gewijd aan de beschrijving van twee methodologieën (het verzamelen van expertkennis en conflictobservatie door middel van microsimulatiemodellering) die worden toegepast voor het uitvoeren van ongevalsanalyse bij een gebrek aan ongevalsgegevens. Het tweede deel van deze dissertatie (hoofdstukken 4, 5, 6 en 7) is gewijd aan het ontwikkelen van ongevalspredictiemodellen op macro-niveau om de verkeersveiligheidsimpact van 2

mobiliteitsmanagementmaatregelen te evalueren (nl. een verhoging van de brandstofkosten met 20% en een telewerkenscenario waarin verondersteld wordt dat 5% van de werkende bevolking thuis werkt). Dit is de tweede belangrijkste bijdrage van deze dissertatie.

Een verkeersveiligheidsanalyse kan op verschillende manieren benaderd worden. Een van de meest gebruikte benaderingen is het ontwikkelen van voorspellingsmodellen. Deze aanpak wil verkeersveiligheidsproblemen verklaren en voorspellen door middel van analytische modellering. Voorspellingsmodellen kunnen ontwikkeld worden op verschillende aggregatieniveaus. Op microscopisch niveau hebben ongevalspredictiemodellen meestal betrekking op een infrastructureel element zoals een wegvak of een kruispunt. Macroscopische ongevalsanalyses, daarentegen bestrijken een relatief groot gebied en koppelen de ongevalskans aan een aantal variabelen met kenmerken op macro-niveau. Beperkingen bij het gebruik van micro-niveau ongevalspredictiemodellen voor een proactieve verkeersveiligheidsevaluatie, leiden naar het gebruik van ongevalspredictiemodellen op macro-niveau. In deze dissertatie werden inspanningen geleverd voor de ontwikkeling van ongevalspredictiemodellen op macro-niveau door ongevallen te koppelen aan een aantal voorspellende macro-niveau variabelen, zoals blootstelling, netwerk en socio-demografische variabelen. Deze modellen worden vervolgens gebruikt voor een proactieve evaluatie van de verkeersveiligheidsimpact van de twee mobiliteitsmanagementmaatregelen.

De beheersing van de blootstelling is een voor de hand liggende en uitvoerbare aanpak in een ongevalsverlagende strategie. Het reduceren van de vervoersvraag wordt beschouwd als een doeltreffende maatregel bij de maatregelen om de blootstelling te beheersen. De vervoersvraag kan verlaagd worden door het toepassen van bepaalde mobiliteitsmanagementstrategieën zoals prijsverhoging (bijvoorbeeld verhoging van de brandstofprijs), telecommunicatie (telewerken), enzovoort. Over het algemeen worden mobiliteitsmanagementstrategieën toegepast om de efficiëntie van vervoerssystemen te verbeteren; echter hun potentiële impact op de verkeersveiligheid mag niet genegeerd worden. Bovendien, door het identificeren van de positieve impact van mobiliteitsmanagementstrategieën op de verkeersveiligheid en hun voordelen aan te halen, kan de bereidwilligheid onder beleidsmakers om deze strategieën te implementeren versterkt worden. Zo draagt de studie, uitgevoerd in het tweede deel van dit eindwerk (d.w.z. hoofdstukken 6 en 7) bij tot de noodzaak van mobiliteitsmanagementstrategieën bij een verkeersveiligheidsevaluatie.

Om de impact van mobiliteitsmanagement op de verkeersveiligheid te beoordelen, moet men beschikken over mobiliteitsmanagement-gevoelige blootstellingsmaatregelen. Voor dit doel wordt er hier gebruik gemaakt van blootstellingsstatistieken geproduceerd door een activiteitengebaseerd vervoersmodel, genaamd "FEATHERS". Het gebruik van een activiteitengebaseerd vervoersmodel is gebaseerd op het feit dat een dergelijk model het verplaatsingsgedrag van personen op een realistische manier simuleert, en hierbij mobiliteitsmanagement-gevoelige blootstellingsmaatregelen produceert, waardoor zowel beleidsmakers als onderzoekers in staat zijn om mobiliteitsmanagementstrategieën zo nauwkeurig mogelijk te evalueren. Daarbij worden mobiliteitsmanagementstrategieën vaak geïmplementeerd en geanalyseerd op macro-niveau in plaats van op het niveau van een individueel kruispunt of wegvak. Bovendien beogen mobiliteitsmanagementstrategieën een wijziging van het verplaatsingsgedrag van individuele weggebruikers. Echter, de impact van een gewijzigd gedrag op het transportsysteem of op de verkeersveiligheid moet collectief en op een geaggregeerd niveau aangepakt worden. Daarom werd er een macroscopische benadering op het niveau van geografische zones uitgevoerd om voorspellingsmodellen te ontwikkelen. Deze modellen worden zonale ongevalspredictiemodellen genoemd omdat ze gebaseerd zijn op zonaal gebaseerde informatie.

Tot nu toe hebben we het gehad over de algemene onderzoeksmotivatie en het kader van dit eindwerk. Omdat deze dissertatie samengesteld is op basis van verschillende wetenschappelijke artikels willen wij in het volgende deel van deze samenvatting de gebruikte technieken en afgeleidde resultaten van elk hoofdstuk individueel kort samenvatten.

Hoofdstuk 2. In dit hoofdstuk wordt een kader voorgesteld om ongevalshotspots te identificeren en prioriteren bij afwezigheid van ongevalsgegevens. Om dit te bekomen worden twee belangrijke taken uitgevoerd. Bij de eerste taak worden relevante hotspotcriteria en hun relatieve belang bepaald door middel van een techniek gericht op het vergaren van expertinformatie, namelijk de Delphi-techniek. Het doel van de meeste Delphi-toepassingen is een betrouwbare en creatieve verkenning van ideeën of het genereren van geschikte informatie voor besluitvorming. Zodra de definitieve criteriagewichten bepaald zijn, worden ze toegepast in een multi-criteria beslissingsproces om een ongevalsprioriteringsmodel te ontwikkelen. Voor dit onderzoek hebben wij de “Technique for Order Preference Similarity to Ideal Solution” (TOPSIS) techniek toegepast om ongevalshotspots te prioriteren. De resultaten van het voorgestelde kader wijzen op het nut van kwalitatieve methodologieën voor verkeersveiligheidsanalyse, vooral in situaties waarbij degelijke ongevalsdata ontbreken of onvoldoende beschikbaar zijn. Kwalitatieve benaderingen bieden zodanig gedetailleerde inzichten in het gedrag en waarden die vaak niet mogelijk zijn met behulp van kwantitatieve methoden.

Hoofdstuk 3. In dit hoofdstuk wordt een methodologie ontwikkeld om de relatie tussen snelheid, verkeersvolume en veiligheidsniveaus op kruispunten te onderzoeken zodat deze ingezet kunnen worden wanneer ongevalsgegevens ontbreken. Hiervoor wordt gebruik gemaakt van een microsimulatiemodel, namelijk S-Paramics, om interacties tussen auto's te simuleren om op basis hiervan de veiligheidsniveaus van kruispunten te bepalen door middel van een berekening van proximale veiligheidsindicatoren. Deze indicatoren hebben een bijna-ongevalskenmerk en worden voorgesteld als alternatief voor ongevalsgegevens. Om een verkeersveiligheidsanalyse uit te voeren worden verschillende scenario’s gedefinieerd die gebaseerd zijn op verschillende verkeersvolumes en maximumsnelheidscategorieën. Verkeersveiligheidsindicatoren worden gemeten voor elk van deze scenario's. De analyseresultaten tonen aan dat het verhogen van de maximumsnelheid zowel op primaire als secundaire toegangswegen van een kruispunt de veiligheid zal doen dalen. Deze daling is het grootste bij hogere verkeersvolumes.

Hoofdstuk 4. In dit hoofdstuk worden verschillende statistische modellen ontwikkeld om het verband te leggen tussen waargenomen ongevallen en een reeks voorspellingsvariabelen. Gezien de waargenomen overdispersie in ongevalsgegevens, wordt het Negative Binomial (NB) model binnen het Generalized Linear Modeling (GLM) kader toegepast om ons in staat te stellen deze data te modelleren. De analyses in hoofdstuk 4 tonen aan dat gebruik van enkel het aantal verplaatsingen – dat de verplaatsingsproductie/attractie van een geografische zone vertegenwoordigt – bij het maken van ongevalspredictiemodellen, ertoe leidt dat belangrijke informatie ontbreekt over de kenmerken van de vervoersvraag, aangezien het aantal verplaatsingen geen informatie bevat over de reistijd, reisduur en routekeuze. Bovendien wordt doorgaand verkeer dat gewoon passeert door een geografische zone genegeerd wanneer men enkel het aantal verplaatsingen gebruikt als blootstellingsvariabele. Dit deel van de blootstelling kan een aanzienlijk deel vormen van de totale blootstelling die waargenomen wordt in een geografische zone. Daarom moet er ook rekening gehouden worden met andere blootstellingsvariabelen die gevoelig zijn voor de impact van verkeerstoedeling (bijv. afgelegde voertuigkilometers of afgelegde reistijd). Bij het vergelijken van de resultaten van de modellen zien we dat modellen die zowel verplaatsingsgebaseerde als verkeersstroomgebaseerde blootstellingsvariabelen bevatten beter presteren dan modellen die slechts één van de blootstellingsvariabelen bevatten.

Hoofdstuk 5. De analyseresultaten tonen op de aanwezigheid van ruimtelijke niet-stationariteit in de gegevens die gebruikt worden bij de ontwikkeling van zonale ongevalspredictiemodellen. Een van de oplossingen om rekening te houden met ruimtelijke variatie is de ontwikkeling van een reeks lokale modellen, de zogenaamde Geographically Weighted Regression (GWR) modellen. Deze modellen zijn afhankelijk van de kalibratie van meerdere regressiemodellen voor verschillende geografische entiteiten. De GWR-techniek wordt daarom aangepast aan GLM-modellen en vormen Geographically Weighted Generalized Linear Models (GWGLMs). Deze modellen zijn uitbreidingen van modellen die ontwikkeld werden in hoofdstuk 4 en zijn in staat om gehele getallen te modelleren (zoals het aantal ongevallen) terwijl ze tegelijkertijd de ruimtelijke niet-stationariteit verklaren. Daarom zijn ze bijzonder nuttig in de context van deze studie. Wanneer we alle ontwikkelde modellen vergelijken zien we dat de GWGLM-modellen altijd beter presteren dan de conventionele GLM-modellen.

Hoofdstuk 6. Zoals eerder vermeld, is het belangrijkste doel van deze dissertatie de evaluatie van de impact van twee mobiliteitsmanagement-scenario's op de verkeersveiligheid. De belangrijkste kenmerken van deze evaluatie zijn de koppeling van de verkeersveiligheidsevaluatie aan de output van een activiteitengebaseerd transportmodel en dat dit binnen een proactief kader wordt uitgevoerd. In dit hoofdstuk worden de modellen die ontwikkeld werden in hoofdstuk 4 gebruikt voor het uitvoeren van een eerste beoordelingsoefening waarbij de brandstofprijs verondersteld wordt te stijgen met 20%. De resultaten van de vergelijkende analyse tonen aan dat dit scenario van verhoogde brandstofkosten een impact heeft op de totale vervoersvraag, totale ongevalsaantallen, voertuigkilometers en afgelegde afstand en de modale shift. Ondanks een verbetering van de algemene verkeersveiligheid, zijn de veranderingen in het aantal ongevallen voor verschillende ongevalstypes/ernstmaat niet identiek. Na de implementatie van het brandstofkostenscenario verschuiven veel verplaatsingen van personenwagens naar andere transportmodi, zoals fietsen en wandelen (d.w.z. de zogenaamde 'trage modi'). Als gevolg van een toename van het aantal verplaatsingen in de categorie 'trage modi' wordt er over het algemeen een toename verwacht van het aantal ongevallen waarbij fietsers en voetgangers (d.w.z. zwakke weggebruikers) betrokken zijn, in tegenstelling tot ongevallen waarbij enkel auto's betrokken zijn. Hiervan wordt een daling verwacht. Dit toont aan dat een verhoging van de brandstofkosten verschillende weggebruikers op een andere manier beïnvloedt.

Hoofdstuk 7. In dit hoofdstuk worden de modellen die ontwikkeld werden in hoofdstuk 5 gebruikt om een tweede beoordelingsoefening uit te voeren waarbij wordt verondersteld dat 5% van de werkende bevolking thuis werkt op eender welke weekdag. De resultaten van de vergelijkende analyse bevestigen dat het telewerkenscenario een impact heeft op de totale vervoersvraag, afgelegde afstand en reistijd en de totale ongevalsaantallen. Het telewerkenscenario heeft een positieve invloed op de verkeersveiligheid van verschillende weggebruikers. Vanwege de aard van dit scenario en een waargenomen "reboundeffect", wordt er een iets kleinere vermindering van "trage modi" verplaatsingen verwacht dan de vermindering van autoritten. Daarom is de positieve impact op de verkeersveiligheid iets lager voor "Auto-trage modi"-ongevallen, dan voor "Auto-Auto"-ongevallen.