PEDESTRIAN SAFETY MODELS FOR URBAN ENVIRONMENTS WITH HIGH ROADSIDE ACTIVITIES

by

RAID A. KRAIDI

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Civil Engineering
School of Engineering
College of Engineering and Physical Science
The University of Birmingham
January 2018

ABSTRACT

Pedestrian safety is of paramount importance in road safety management. Yet to date there have been a limited number of models capable of capturing satisfactorily all aspects of pedestrian movement in road environments and subsequently suggesting measures to make roads safer for them. This study aimed at investigating such a model with emphasis on road side activities in urban environments. The work followed a methodology which analysed the problem at hand, suggested appropriate mathematical models, and tested the models.

Based on data gathered from appropriately chosen road links and other records, the study showed that the number of bus stoppings, parking activity and pedestrian crossing violations, the traffic speed variation, the number of intersecting side roads, main road traffic volume and intersecting traffic volume were the most significant risk factors related to pedestrian crash risk. In addition, the traffic operating speed was found to be the predominant factor determining the crash severity. Two models were produced to link these variables with: (a) the risk of a pedestrian crash using generalised regression models (GLM); and (b) with the crash severity using logistic regression. Model selection tests were conducted and it was found that the models can capture the pedestrian crash risk satisfactorily. In addition, these models were tested and validated using three evaluation measures including the goodness of fit assessment, a graphical exploration of the observed versus the predicted crash numbers and a field validation.

Furthermore, the study proposed the integration of the new models within the iRAP, an established road safety management system; it demonstrated how this may be achieved and examined their applicability if used in practice. The integration was systematically tested and

validated and it was found that the new model provided better results compared to the original iRAP.

The developed model, albeit estimated using data from urban roads in Birmingham in the UK, it can be transferred and ultimately used in developing countries subject to appropriate adaptation to local conditions and associated data.

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere thanks and gratitude to my supervisor Dr Harry Evdorides, for his guidance, advice and continued support throughout my research studies. He showed me different ways to approach a research question and the need to be persistent to accomplish any objective.

I also would like to thank my second supervisor Dr Michael Burrow and Professor Miles Tight for their advice and comments during the research and annual progress reviews.

I would like to offer my special thanks to James Bradford – from the International Road Assessment Programme – for his advice and support of this study. Thanks also to Mrs Janet Hingley for proofreading.

I would like to express my appreciation to my family: my parents and to my brothers and sisters for the support and encouragement they offered during my studies in the UK. Special thanks to my wonderful wife Rana and my kids Meena and Mustafa, without their love, support and encouragement my PhD studies would not have been possible.

Special thanks and acknowledgement to the Higher Committee for Education Development in Iraq (HCED) for awarding me a scholarship and other support to carry out this research.

Great appreciation goes to Birmingham City Council and the Department for Transport for providing the necessary data.

Finally, I offer my regards and thanks to all fellow PhD students and colleagues for their support and the friendly atmosphere during the completion of this research. I am also thankful for their prompt help, whenever I needed it.

TABLE OF CONTENTS

Chapter 10

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AADT	Average Annual Daily Traffic
AASHTO	American Association of State Highway and Transport Official
BS	Number of Bus Stoppings
CMF	Crash Modification Factors
CV	Coefficient of Speed Variation
DFT	Department for Transport
DMRB	Design Manual of Roads And Bridges
EuroRAP	European Road Assessment Programme
FHWA	Federal Highway Administration
GLM	Generalized Linear Modelling
GOF	Goodness of Fit
HSM	Highway Safety Manual
HV	Percentage of Heavy Vehicle Traffic
INT	Intersecting Traffic Volume
iRAP	International Road Assessment Programme
ITE	Institute of Transportation Engineers
JSTS	Journal of Society For Transportation and Traffic Studies
LMICs	Low And Middle-Income Countries
MSE	The Mean Squared Error
MSPE	The Mean Squared Prediction Error
NB	Negative Binomial Model
NCHRP	National Cooperative Highway Research Program

P	Roadside Parking
PAL	Pedestrian Walking Along Volume
PBA	Parking and Un-Parking, Boarding-Alight Activities
PC	Pedestrian Crossing Volume
PCV	Pedestrian Crossing Violations
SI	Serious Injuries’ Crashes
SR	Number of Intersecting Side Roads
SRS	Star Rating Score
TOI	Institute of Transport Economics
TRL	Transport Research Laboratory
usRAP	U.S. Road Assessment Program
V	Traffic Mean Speed
VIF	The Variance Inflation Factor
WHO	World Health Organization
ZIGP	Zero-Inflated Generalized Poisson Model

Chapter 1

INTRODUCTION

1.1 Background

The World Health Organization (2015) reported that the annual casualties due to road traffic accidents have reached 1.25 million fatalities, while non-fatal injuries are up to 50 million. However, the statistics of injuries worldwide cannot be considered as reliable due to reporting inconsistencies (Elvik et al., 2009). Although traffic injuries are among the leading causes of death globally, most traffic accidents are predictable and largely preventable.

In addition to the immeasurable impact of the massive toll of casualties on associated communities and families, road traffic accidents may cost a country's economy approximately 3-5 per cent per annum of the estimated gross national product (GNP) (Jacobs, Thomas, and Astrop 1999; World Health Organization, 2015; Ansari et al., 2000).

Pedestrians represent a significant percentage of crash fatalities. The statistics in Figure 1.1 show that pedestrian fatalities accounting for about 22% of total traffic deaths globally and for about 26% in Europe (WHO, 2015).

The road traffic safety problem is more severe in developing countries, where pedestrians and other vulnerable road users, albeit being dominant road users in the urban areas, are the least physically protected group (Downing et al., 2000). In addition, pedestrians constitute a significantly higher share in road traffic fatalities and injuries in low and middle-income countries (LMICs) than in high-income countries (World Health Organization, 2013b; Ogendi et al., 2013). As a representative example, the share of pedestrian deaths ranged between 40-60% of total road traffic fatalities between 1990 and 1999 in Delhi (Tiwari, 2001).

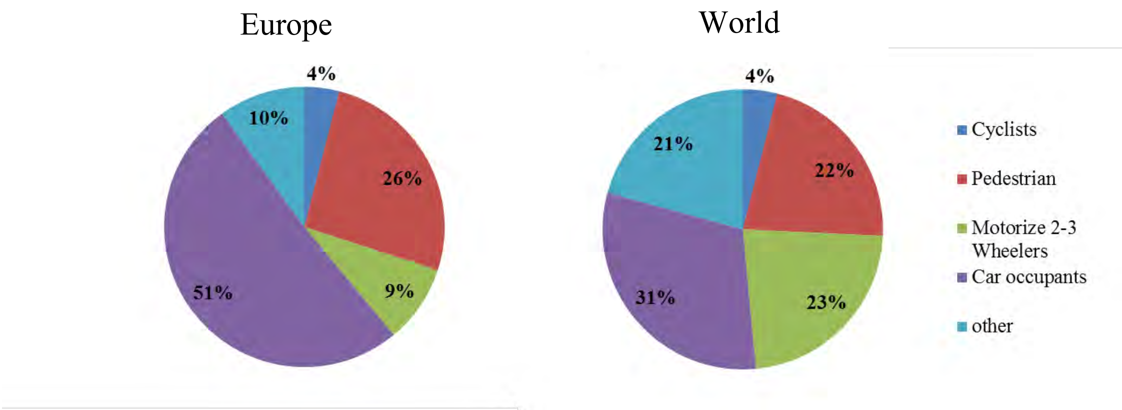


Figure 1.1 Road crash fatalities' statistics by types of road user (WHO, 2015)

1.2 Problem Definition

Pedestrians are considered as the most vulnerable because of their fragility and slow movement. Subsequently, they have a higher risk of road crash potential than motorised vehicle occupants (Zhang et al., 2014). Every year, a large number of road traffic crashes involve pedestrians. In the UK, 25% of road fatalities were pedestrians in 2016. The total pedestrian-vehicle crashes exceeded 23000, from which there were 448 fatalities; while more than 5000 suffered serious injuries (DfT, 2016). Pedestrians tend to travel more in urban areas, where the traffic accident hazards are higher than in inter-urban areas (Elvik et al., 2009). For example, around 76% of pedestrian deaths in the United States and 70% in the European Union were in urban areas (WHO, 2013b). Therefore, there is an association between pedestrian safety and the surrounding environment. For instance, children living in urban areas face up to a five-times higher risk of having a pedestrian crash than for those living in rural areas (Petch and Henson, 2000).

Different types of land use and related activities have been associated with vehicle-pedestrian crashes, with a potential increase of risk linked with increased uses for mixed, retail, commercial and community purposes (Geyer et al., 2005; Wedagama et al., 2006; Kim et al., 2006; Loukaitou-Sideris et al., 2007). From an empirical prospect, the increasing traffic volume in urban areas increases the potential of pedestrian-vehicle conflicts (Lee and Abdel-Aty, 2005). Thus the probability of the pedestrian-vehicle crash risk is related to the number of vehicle-pedestrian interactions. Accordingly, pedestrian crashes are over-represented in urban areas, especially areas of higher road and roadside activities, where pedestrian activities, traffic volume and traffic speed variations are greater compared to other area settings (Zegger et al., 2002; Clifton and Fults, 2007; Zheng et al., 2010).

It has been stated that pedestrian casualties are mostly preventable and proven intervention measures are available (WHO, 2013b). Yet, successful interventions to make walking safer and protect pedestrians need an understanding of the causes and nature of the risk factors related to pedestrian-vehicle crashes.

For that reason, there appears to be a need to develop models suitable for application in urban areas with high road and roadside activities. This study seeks to develop a model to assess pedestrian safety and capture the effect of pedestrian and roadside activities' intensity, such as bus stoppings, parking, pedestrian crossings and violations, loading and un-parking events, together with road traffic speed characteristics.

Such a model could be a particularly useful tool for traffic planners and others with the ability to rank roadways with regard to pedestrians' crash risk; and also in the process of evaluating existing roads and prioritizing the roads which are in need of safety interventions. Furthermore, this approach may provide a better understanding of the causes and nature of risk factors related to pedestrian crashes while moving in road environments with high roadside activities in urban areas.

Ultimately, the findings of this study may be further tested and possibly calibrated in low and middle-income countries where some road links accommodate high pedestrian and roadside activities; which are to some extent, comparable with the selected area of emphasis in this study.

1.3 Aims and Objectives

1.3.1 Aim

The primary aim of this study is to enhance the safety of pedestrians moving in urban road environments with high roadside activities through the application of a modelling approach which considers the infrastructure and traffic attributes of the road environment and their associated risk to make walking safer and protect pedestrians.

1.3.2 Study objectives

To achieve the above aim, the study objectives were defined as follows:

1. To provide a theoretical review of available models used to assess pedestrian road safety in urban areas to determine their applicability in environments with high roadside and pedestrian activities.

2. To develop a new model estimating the safety of pedestrians moving in environments with high roadside activities in urban areas.
3. To identify the most important factors that affect pedestrian road safety in urban environments of high pedestrian and vehicle traffic.
4. To estimate crash risk factors using the developed models and to examine their integration in the International Road Assessment Programme (iRAP) star rating system.
5. To validate the developed models and to compare their outputs with those of the iRAP model using data collected from appropriately chosen sites.
6. To offer recommendations to improve pedestrians' safety based on the finding of the analysis conducted.

1.4 Thesis Layout

The thesis structured into ten further chapters as shown in Figure 1.2. Following this chapter, it is organised as follows:

Chapter 2 summarises literature reviews of previous studies about road safety and its global status; the contributing factors and their influence on pedestrian safety; the modelling approaches used to assess pedestrian safety; and the concept of a star rating score to assess road safety under the International Road Assessment program (iRAP).

Chapter 3 proposes the methodology of this study. It outlines the research approach to conceptualize the pedestrian safety model; theoretical background; modelling approach to develop models for the safety of pedestrians moving in environments with high roadside activities in urban areas; and accounts for the related risk factors and their effect on crash risks.

Chapter 4 gives the concept of a star rating score to assess road safety for pedestrians under the International Road Assessment Program (iRAP); and the need for enhancement, structure and the factors included in the iRAP pedestrians' model; furthermore, it proposes a methodology for enhancement.

Chapter 5 presents the data requirements for the model estimation; proposes the introduction of new aspects in pedestrian road safety modelling; describes the data requirements for the modelling process and giving the methods used for data collection.

Chapter 6 covers the data collection methodology, site selection and the statistical characteristics' summary of the collected data.

Chapter 7 presents the methods for modelling the pedestrian crash risk and the severity, checking the correlation between independent variables and describing the measures used in comparing and selecting the modelling method.

Chapter 8 discusses the process of model fitting; estimates the model's coefficients using the collected data; describes the selection measures and presents the measures used in the testing and validation of the developed model. This chapter describes the testing and validation of the developed model by assessing the goodness-of-fit (GOF) measures to examine how well the observed data set fit the developed model. It compares the observed and predicted crash numbers, to estimate the accuracy of the estimation percentage and conducts field validation to examine the performance of the suggested models in a different field data set.

Chapter 9 proposes an application of the developed models to enhance the current (iRAP) star rating system. This chapter consists of performing a sensitivity analysis using the developed models to examine the effect of various levels of input variables on the safety outputs; assessing the effect of independent variables on crash frequency potentials; and suggesting an approach to enhance the current iRAP star rating tool for pedestrians, by incorporating the developed risk factors.

Chapter 10 provides a discussion of this research with regard to: method selection and contributing factors; the evaluation of the success of the model by testing the correlation between contributing factors and validation of the developed models; the applicability of the developed models and the potential application and integration to enhance the current iRAP system; data issues with regard to the collection process and sample size; and also the limitations of this study.

Chapter 11 presents the conclusions and areas for further studies.

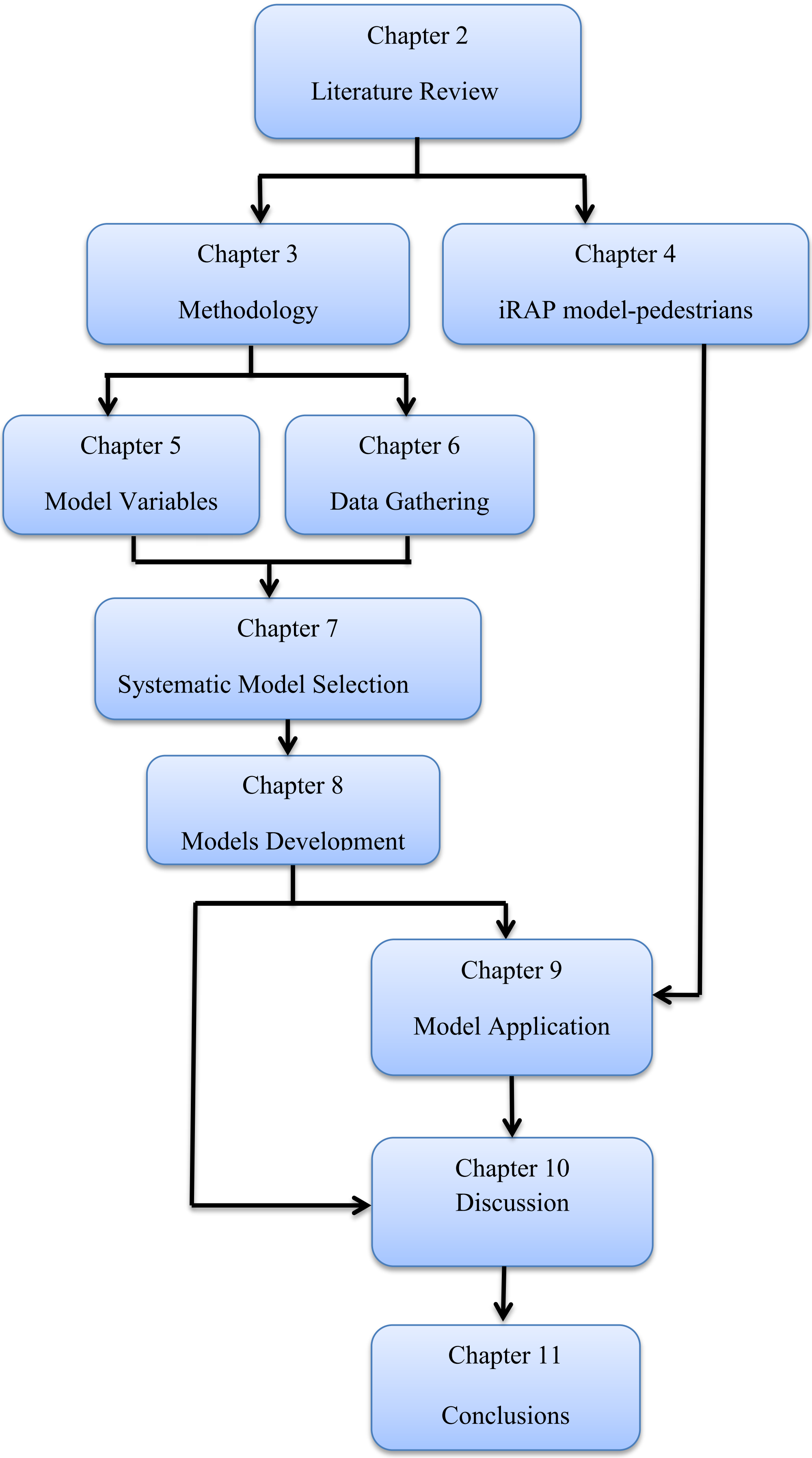


Figure 1.2 Thesis structure

Chapter 2

LITERATURE REVIEW

2.1 Introduction

As mentioned in Chapter 1, the primary aim of this study is to propose a modelling approach to enhance the safety of pedestrian and to examine its integration in the International Road Assessment Programme (iRAP) star rating system. Therefore, this chapter reviews previous work in this area and presents it in further five sections. The following section reviews road safety and its global status. The third and fourth sections discuss the contributing factors and their influence on pedestrian safety. The fifth section presents modelling approaches used to assess pedestrian safety; while the last section introduces the concept of the star rating score of the International Road Assessment Program (iRAP).

2.2 Road Safety

The high toll of traffic accidents fatalities and injuries is considered as one of the major health problems. It has been reported that the annual casualties due to road traffic accidents are about 1.24 million fatalities; while non-fatal injuries number 20 to 50 million (WHO, 2015). Moreover, traffic accidents are ranked as the ninth leading cause of fatalities globally and are projected to rise to become the seventh by 2030 (see Table 2.1) (WHO, 2015). Therefore, policies and engineering interventions need to be implemented to militate against this problem at both local and global scales.

In addition to the immeasurable impact of the huge toll of casualties on associated communities and families, road traffic accidents may cost a country's economy between 1 and 5 per cent per annum of the estimated gross national product (GNP) (Jacobs, Thomas, and Astrop 1999; WHO, 2013a; Ansari et al., 2000).

Table 2.1 WHO leading causes of death 2015- 2030.

	2015			2030		
Cause	Rank	% Deaths	Deaths per 100,000 population	Rank	% Deaths	Deaths per 100,000 population
Ischaemic heart disease	1	13.2	105	1	13.2	112
Stroke	2	11.7	92	2	12.2	104
Lower respiratory infections	3	5.6	44	4	5.0	43
Chronic obstructive pulmonary disease	4	5.6	44	3	6.5	55
Diarrhoeal diseases	5	3.2	25	9	2.3	20
HIV/AIDS	6	2.9	23	8	2.6	22
Trachea, bronchus, lung cancers	7	2.9	23	7	2.6	22
Diabetes mellitus	8	2.7	21	5	3.5	30
Road injury	9	2.5	20	7	2.6	22
Hypertensive heart disease	10	2.0	16	10	2.1	18

Pedestrians represent a significant percentage of traffic crash fatalities. The statistics in Figure 2.1 show the pedestrian deaths as a proportion of total traffic fatalities for selected regions (WHO, 2015).

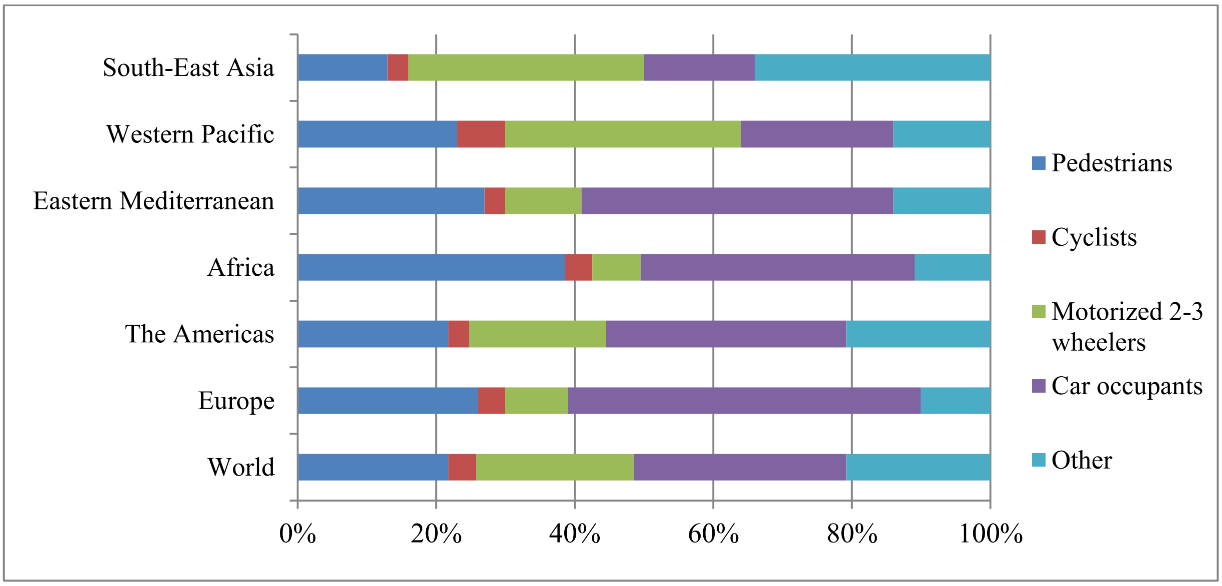


Figure 2.1 Road crash fatalities statistics by types of road users (WHO, 2015)

In less developed countries, vulnerable road users, such as pedestrians, cyclists and rickshaw drivers may be dominant in urban areas and are the least physically protected group (Downing et al., 2000; Zhang et al., 2014). Consequently, pedestrians constitute a significantly higher share in road traffic fatalities and injuries in low and middle-income countries (LMICs) than in high-income countries (WHO, 2013b; Ogendi et al., 2013).

The road infrastructure in transitional economies carries traffic volumes and compositions far beyond its designed capacities. In addition, rapid urbanisation and motorisation lead to traffic delays, congestion and accidents in urban centres. Furthermore, urban areas experience a high rate of conflicts between pedestrians and motorised traffic because of incompatible land uses and unplanned growth (Ross et al., 1994). In turn, this leads to uneven competition between different road user groups, high risk exposure, significant traffic speed variation and increased deterioration of the road network.

The road crashes occur as a result of the interaction of many risk factors found within traffic system. One conceptual view of traffic safety is to summarise these factors by using the Haddon matrix shown in Table 2.2 (Haddon, 1980). The matrix identifies the factors associated with humans, environment and the vehicles and their potential effect before, during and after the crash occurrence. In the pre-crash stage, the factors influence the risk of a crash; while during the crash event these factors influence its severity. For the post-crash stage, measures are related to emergency and recovery actions.

Table 2.2 Haddon matrix (Haddon, 1980)

Phase		Human	Vehicles and equipment	Environment
Pre-crash	Factors contributing to increased risk of crash	Training Education Behaviour (e.g. alcohol and drug), Attitudes Impairment Police-enforcement	Roadworthiness Lighting Braking Handling Speed management	Delineation Road geometry Speed limits Visibility Traffic conditions Pedestrian-facilities
Crash	Factors contributing to crash severity	Use of restraints Impairment	Occupant restraints Other safety devices Crash protective design	Safety barriers Roadside safety
Post-crash	Factors contributing to crash outcome	Emergency Medical treatment	Ease of access Fire risk	Restoration of road and traffic devices Congestion

Another view to conceptualise the traffic safety situation is to describe it through the multidimensional factors of *exposure*, *risk* and c*onsequence* (Nilsson 2004), where:

- Exposure: concerns all factors that describe the level of exposure of road users to risk
- Risk: covers all factors that influence the probability of being involved in a crash per unit of exposure
- Consequence: address all factors that influence the severity of a crash given that a crash event has occurred.

Chapman (1973) describes exposure as the degree to which a road user is exposed to traffic and roadway characteristics and the potential of being involved in an accident.

To improve road safety, one or more of these factors have to be targeted. Many variables may be associated with each of the factors and facilitate their understanding and ultimately their impact on the number of crashes and the severity

Several safety models investigate the effect of traditional traffic measures, such as traffic speed, traffic volume and geometric design on road safety (Harwood et al., 2008; Elvik et al., 2009; Ukkusuri et al., 2012). However, some others may affect pedestrian safety of developing areas, such as roadside activities (Lynam, 2012; Hosseinpour et al., 2014).

2.3 Variables Influencing Pedestrian Safety

The number of pedestrian fatalities is large compared with other road users groups (Peden, 2004). A similar study by Elvik (2009) showed that pedestrians are exposed to a significantly higher risk of injury crashes than car drivers in the current transport system. This section reviews therefore the contributing variables and their influence on the safety of pedestrians at street level. Previous studies have suggested there are different variables associated with pedestrians` accidents and can be varied according to road types and surrounding environment. However, four significant variables influencing pedestrian safety in general; intersection and roadway characteristics, pedestrian volume, traffic volume and speed variables, as presented below.

2.3.1 The effect of intersection and roadway characteristics on pedestrian safety

The effect of roadway and intersection geometric factors has been investigated in numerous crash modelling studies.

For intersections, the risk of a crash per driven unit of length is significantly higher at intersections' approaches compared to other parts of the road (Kulmala, 1995). Crash models for urban roads normally include minor junctions with local streets and exits. These are conflict points prone to a high risk of crash and leading to a more complex and interactive traffic environment. The intersection also brings a higher risk for vulnerable road users travelling alongside, especially when crossing side roads (Jonsson, 2005).

Abdel-Aty and Wang (2006) studied signalised road intersections in Florida and reported that the intersections' design has an impact on traffic accidents. For example, 3-leg intersections with protected left turn phase and exclusive lanes for right-turn on both directions in residential areas or open counties had lower accident incidents. While, shorter signal spacing, high speed limits, a larger number of phasing and the total number of lanes at intersections are associated with high accident frequencies.

Elvik et al. (2009) reviewed international studies investigating the impact of roundabouts on accident frequencies for multiple road users. The analysis reported a 30-40% reduction in injury crashes where roundabouts were built. The same reduction was indicated for pedestrians.

In terms of lane width at junctions, a reduction in accidents figures at junctions with wider lanes was reported (Elvik et al., 2009).

Shankar et al. (1996) found that road links that include curves with higher design speeds result in more crashes compared with segments with lower design speeds.

Zegeer et al. (1985) investigated pedestrian crashes in 15 U.S. cities at 1297 urban signalised intersections over a six-year period. The study performed an analysis of pedestrian crashes together with data for signals, pedestrian and traffic volumes and geometrics for signalised intersections. The analysis found that in addition to pedestrian and traffic volume, other features including residential area type, street width, bus stops and pedestrian signal type, may contribute significantly to accident frequencies.

Box (2002) reviewed studies investigating the impact of parallel parking and angle parking and the study concluded that road segments with angle parking showed crash rates of 1.5 to 3 times more than those in segments with a parallel parking configuration.

Hauer (2004) performed an analysis to estimate crashes frequencies on undivided four-lane roads in urban areas. The study showed that the variables correlated with crash frequencies included the number of commercial driveways, the posted speed limits, lane width and the presence of parking.

Garber and White (1996), conducted a study of 30 road sections in Virginia and found that traffic volume, average speed, left turn availability and number of side roads influence the crash rate for urban arterials.

Zegeer et al. conducted another study (2005) in 30 cities across the United States to determine the effects of crosswalks' marking at signalised intersections on pedestrian safety. This study found that crosswalk marking without other treatments was linked with higher pedestrian crash frequencies in multilane roads carrying more than 12000 vehicles per day at controlled sites. In addition, raised median presence compared to roads with no raised median correlated with significantly lower pedestrian crash frequencies on multilane roads.

Based on the reviewed studies, it may be seen that intersection and roadway design features could have significant effects on safety; and by considering these factors in the models, a better explanation of crash frequency and severity can be provided. Thus, improving infrastructure, geometry design and signage may help in improving road pedestrian safety.

2.3.2 The effect of pedestrian volume on pedestrian safety

Several studies showed that the volume of pedestrians is one of the most influential factors on pedestrian safety and associated crash frequencies.

Zegeer et al. (1985), in their study found that the pedestrian volume was the most influential factor on pedestrian safety and the increase of pedestrian volume may be associated with higher crash frequencies. The study also calculated an average rate of pedestrian accidents per intersection per year as a function of pedestrian volume. The pedestrian accidents' rate for intersections with a volume of less than 1200 pedestrians per day, intersections with 1200 pedestrians or more and intersections of 3500 pedestrians or more were experiencing an average of 0.178, 0.553, and 1.002 pedestrian accidents per site per year respectively.

Brude and Larson (1993) studied the influence of pedestrian volume on pedestrian safety in Swedish municipalities. The study selected 285 intersections with a daily crossing volume of 100 or more pedestrian, including 9 roundabouts, 155 unsignalized and 121 signalized intersections. The relationship between pedestrian volume and crashes was found to be significant and positive, where an increase of the pedestrians' volume leads to increase in crashes.

Lyon and Persaud (2002) studied 684 four-legged and 263 three-legged signalized intersections, and 122 three-legged stop controlled intersections in urban areas averaging 7.72, 4.05 and 1.3 pedestrian crashes over 11 years to develop a pedestrian crash prediction models for urban intersections in Toronto. The relationships between pedestrian volume and crashes in all studied intersections types were also found to be significant and positive. Moreover, models which included pedestrian volume were found to be more reliable. Also Zeeger et al. (2005), in their study of the safety outcomes of marked and unmarked crossing facilities, reported that pedestrian volume has a positive and significant relationship with pedestrian road accidents.

In contrast, other studies suggest that a higher number of vulnerable road users including pedestrians and cyclists leads to less risk for the individual user. This is explained by the fact that a higher flow of pedestrians makes the drivers more aware and alert to the possibility of pedestrian presence. This in turn might lead to more careful driving and a decrease in the crash risk (e.g. Jacobsen, 2003; Elvik, 2009; Agarwal, 2011; Jonsson, 2013).

Based on the above, this study will consider the volume of pedestrians as a candidate variable for the modelling process.

2.3.3 The effect of traffic flow on pedestrian safety

Among the most common influential factors in pedestrian crash risk and severity are traffic volumes and their compositions. Increasing traffic flow means more interactions between road users and higher exposure of pedestrians to risk. Previous studies suggested that the traffic flow volume appears to be the leading cause of crash frequency with an approximately negligible influence on crash severity (Wang and Abdel-Aty, 2008).

In addition to pedestrian volume, the traffic volume was also found by several studies to be a major contributing element to pedestrian accidents.

Zegeer et al. (1985) considered that the traffic volume is the second most influential factor on the size of pedestrian accidents' frequency. The study found that the pedestrian crash figures increased when the traffic volume increased to a certain pedestrian volume level.

Studies by Brude and Larson (1993) found that at intersections traffic volume has a positive relationship to pedestrian road accidents.

Lee and Abdel-Aty (2005) investigated the relationship between intersections' average traffic volume and pedestrian road crashes. The analysis revealed that the pedestrian accidents in intersections increased as the average traffic volume increased, due to higher pedestrian-vehicles conflicts. However, the increasing figures of pedestrian crashes are steeper at lower average traffic volumes values.

The above studies relied on average traffic volumes. Other studies used turning vehicle volumes to detect the association with pedestrian crashes. Lyon and Persaud (2002) developed a pedestrian crashes' predictive model to include left turning traffic volume. The study stated that the weight of the left turning proportion to the total incoming vehicles at signalised intersections had an important and positive effect on pedestrian accidents. Also, the left turning volume was found to have more effect than the total incoming traffic volume at stop-controlled intersections; and therefore it was included as an isolated factor in the developed model rather than use its proportion to overall traffic volume. Leden (2002) compared the pedestrian accidents involving right turning traffic with pedestrian road accidents with left turning vehicles at semi protected left turn intersections. The study reported that left turning traffic volume is associated with more pedestrian crashes than those associated with similar right turning traffic.

2.3.4 The effect of traffic speed on pedestrian safety

Understanding the vehicle- pedestrian crash is essential to develop effective and feasible countermeasures to decrease crashes and in turn save lives. To achieve this, the relationship between traffic speed and pedestrian accidents should be established.

Traffic speed is important factor and has been reported to influence road crashes both in terms of likelihood and severity (Elvik, 2009; Pulugurtha and Sambhara ,2011; Al-omari and Obaidat, 2013; Derry et al., 2010; Koushki and Ali, 2003; Hunter et al., 1996;Taylor et al.,2000).

In terms of accident likelihood, according to the laws of physics and road traffic theories, a higher stopping distance is needed in the case of higher driving speed, which could lead to an increase in the likelihood of accident occurrence (Wang, Quddus, and Ison, 2013). Pedestrians and cyclists could be mixed relatively safely with the traffic stream at speeds below 30 km/h (Lynam, 2012).

Figure 2.2 shows the relationship between speed and stopping distance for emergency braking. Vehicles travelling at higher speed require longer distances and times to stop to avoid a crash occurrence (Anderson et al., 1997).

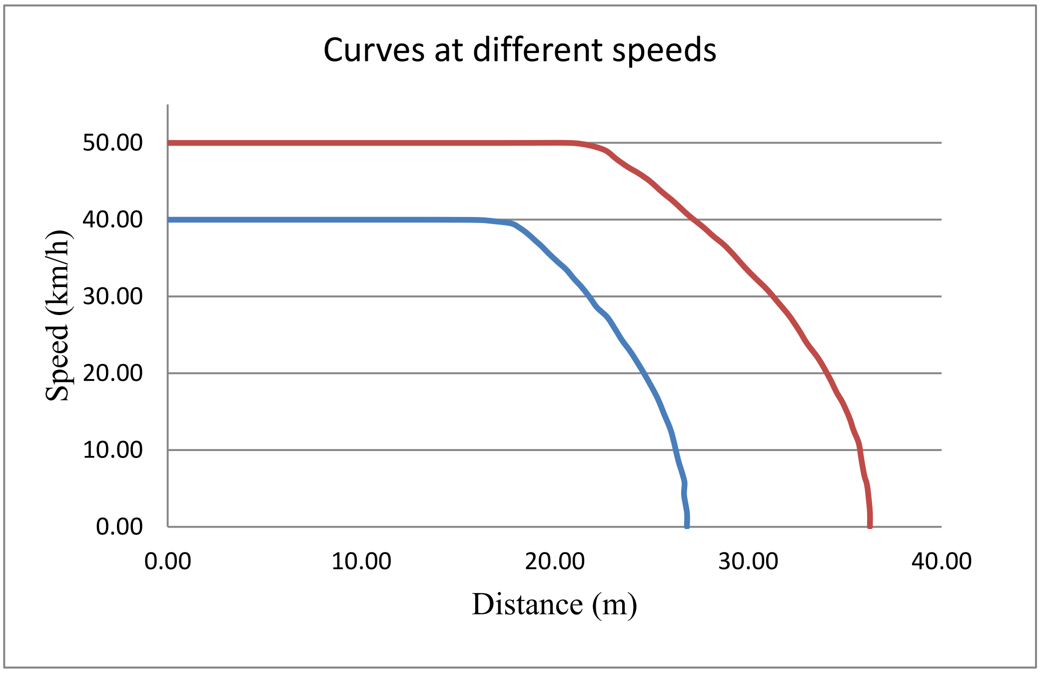


Figure 2.2 Speed and stopping distance for emergency braking (McLean et al., 1994)

In terms of severity, it has been reported that an increase of traffic speed may lead to a more severe outcome of an accident. In the event of a crash, higher the impact speed leads to more severe the crash outcome (Lin et al., 2003; Wier et al., 2009; Lin and Kraus, 2009).

For example, in the event of a collision between pedestrians and vehicles, the possibility of pedestrian survival decreases significantly above the speed of 30 km/h. The critical speed levels for survival in case of other forms of collisions are 50 km/h and 70 km/h for events of cars' side collisions and head on crashes respectively. These are illustrated in Figure 2.3 (Turner and Smith, 2013).

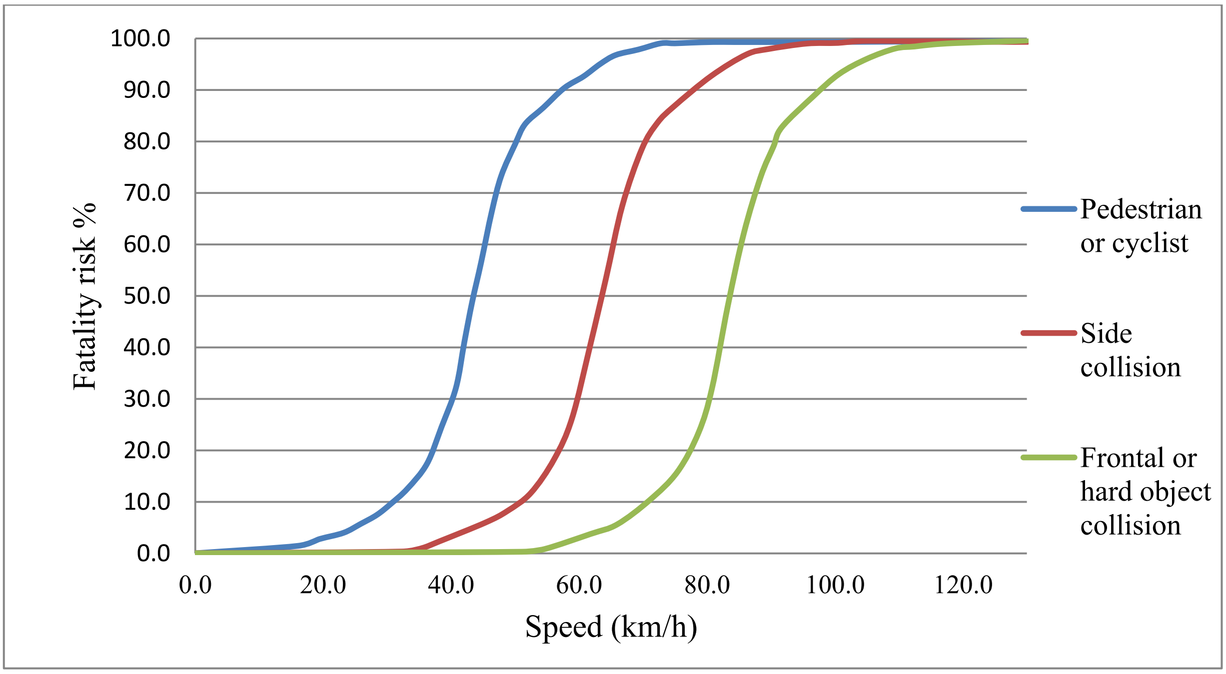


Figure 2.3 Fatality risk (Wramborg P., 2005)

Ashton et al. (1977) investigated the importance of impact speed on pedestrian accident severity. This study reported that pedestrian accidents occurring in rural areas had higher rates of sever and fatal accidents than those in urban links reflecting the higher traffic speeds in these areas. Also the study suggested that the severity of pedestrian injuries was associated with the impact speed and all of the severe and fatal injuries occurred at higher speeds levels. For adults and children, the study reported that severe and fatal accidents occurred at impact speeds higher than 50 km/h.

Accordingly, two speed levels affect a pedestrian crash: the initial speed linked to the possibility of crash avoidance and the impact speed which determines the severity of the crash. Also, the speed levels are correlated with other factors including road design and the presence of other road users because higher speed roads accommodate fewer intersections per length unit, less vulnerable road users' crossings and higher design standards (Jonsson, 2005).

Taylor et al. (2000) found that the possibility of accidents occurrence on a road segment change positively with the square of the traffic speed. Peden et al. (2004) reported that there is an exponential relationship between pedestrian fatalities' figures and traffic speed. Nilsson (2004) estimated that the exponential relationship is close to the fourth power of speed change for all road users' accidents fatalities.

2.4 The Effect of Factors Associated with Urban Environment

The interaction between pedestrians and drivers of motorised vehicles in urban roads is higher leading to increase in the incidence of conflicts between pedestrians and the drivers of motorised vehicles. Subsequently, the pedestrian-vehicle crashes are higher in urban environment , especially settings of higher road and roadside activities and commercial tracts, where pedestrians' activities, traffic volume and traffic speed variations are higher compared to other area settings such as residential areas (Clifton and Fults, 2007; Zheng et al., 2010). Regarding the contributing variables, in addition to the general factors of traffic speed, volume characteristics and geometric design, some other factors may also associated with pedestrian safety specifically in urban roads. Based on the literature review, it was felt that more factors need to be examined when considering pedestrian safety in urban environments with high roadside activities in urban areas including traffic speed variation, land uses, frequent road side and pedestrian activities. The effect of these factors on pedestrian safety on urban roads is not yet well understood. Inclusion of these factors may improve the developed models to assess pedestrian safety; therefore, there is a need to develop models taking these factors into consideration.

2.4.1 Speed variation

Speed variability may be attributed to drivers' speed-choice behaviour. For example, some drivers simply may wish to drive faster than others; while other reasons could be associated with different traffic conditions (e.g. varying levels of traffic flow), or a combination of the two (Taylor et al., 2000).

Several studies found that speed variation has an important effect on accident risk and attempted to quantify its influence on accident frequencies. Moreover, it has been also urged that speed variation is a problem that has a more significant effect on accident frequencies rather than speed itself.

Solomon (1964), one of the earlier researchers in this context, quantified the relationship between accident frequencies and speed variation after investigating a wide range of accidents' records and compared it with speed measurements and interview results.

The study introduced a U-shaped figure to explaine the relationship between accident frequencies and traffic speed. Accident rates were found to be increasing with higher deviation from the speed mean and reached the lowest value where the speed was closest to the mean. There were some debates regarding these results. Frith and Patterson (2001) suggested that the findings were most probably due to the way that the accidents and speeds were measured. Another critical view for the effect of speed variability is related to the fact that the strong correlation between speed's mean and variance possibly make it difficult to separate the effects of speed variance and mean speed on crash events (Elvik, Christensen and Amundsen, 2004).

Nevertheless, Lave (1985) found that the accident fatality figures were significantly associated with speed variance rather than speed itself.

There are other studies suggesting the influence of speed variation on accident risk (Garber and Ehrhart, 2000; Garber and Gadiraju, 1989).

Quddus (2013) investigated the relationship between speed variation, average speeds and accident rates using a range of motorways and A-class road sections around London. The study results showed that speed variation was associated with crash frequency where a 1% increase in speed variability leads to a 0.3% increase in crash rates.

Recently, the use of loop detector data has improved the ability to study the effect of speed variability in more accurate and well-controlled manner. The collected data from loop detectors can be used to construct traffic characteristics at a given site. It then becomes possible to determine if the period immediately before an accident occurred was characterized by a larger speed variance than other periods (Elvik, 2014). Using this method and data resource, Zheng et al. (2010) has shown that speed variation influences accident risk.

Elvik, Christensen and Amundsen (2004), showed that the number of conflicts between vehicles travelling at different speeds is higher than that between vehicles travelling at identical speeds in opposite directions. As the number of conflicts were considered as an indicator of a crash occurrence probability, therefore, the speed variance could be associated with an increase in this probability; and it would seem that this is indeed the case.

These relationships between speed variation and safety are not as well established as the relationship with speed limits and average speed. The theoretical bases of the relationship between traffic safety and speed variation can be noticed in the interaction between the expectations of drivers and vehicles. The higher the traffic speed variation, the more overtaking manoeuvres and consequently, the more situations with risk (Finch et al., 1994). Also, for the road user, larger speed variations may create more unexpected situations due to increased differences between the expected traffic speed and the actual speed, which may lead to collisions (Elvik et al., 1997, as cited in Jonsson, 2005, p.11).

2.4.2 Land uses

Many studies have investigated the relationship between road safety and land use types.

Ashton et al. (1977) examined the influence of vehicle design on pedestrian accident severity and reported that the majority of pedestrian accidents occurred in urban roads.

Chapman (1978) performed a detailed investigation to examine the road and environment factors associated with accidents rates across 4 towns in Southern England. Accidents rates were compared along the road sections passing through different land uses of study area on each side. The land uses on each side of the investigated roads were mainly shops, industrial, residential and other uses categorised as open. The study showed that a notably higher accident rates associated with sections of road passing through shopping areas. The accident rate was found to be over twice the average along road sections adjacent to at least one side of shopping development compared with other land uses. Although the accident rate was relatively higher for all road users, pedestrian and cyclist were particularly vulnerable as the number of pedestrian accidents was significantly more along these sections.

McGuigan (1982) conducted study in Lothian region to determine the relationship between accident rates, roadside development and traffic. The study reported that there are significant relationships between accidents and adjacent roadside development. Using regression equations, the analysis showed that accident rates vary according to the types of the adjacent land uses, and higher accident rates were associated with roads passing through shopping development.

Another work (Lawson, 1985, 1986) on Birmingham radial routes indicated the effect of land uses and roadside development on the differences of accidents at different locations. the results showed that there were high numbers of accidents along the roads adjacent to shops. The results of this study confirmed the findings of earlier studies of Champman (1978) and McGuigan (1982). Also, there were significantly more pedestrian accidents on roads passing through shopping areas than on roads passing through other forms of land use.

The Birmingham radial route data (Lawson, 1985, 1986) have been compared with those provided by Chapman (1978) as shown in Table 2.3.This shows the accident rates per kilometre per year and illustrates the high accidents associated with shopping areas. The accident rates show similarity when ranked, but those for Birmingham radial routes are lower. This may be attributed to the differences in the traffic volume.

Table 2.3 Accidents rates by land use types (Chapman, 1978 and Lawson, 1985, 1986).

Land use type (both sides)	Accident rate (Accident/Km/Year)	
	Chapman (1978)	Lawson (1985, 1986)
Shops/Shops	19.62	13.98
Residential/Residential	7.32	7.98
Industrial/Industrial	8.57	6.53
Shops//Residential	15.95	12.24
Residential/Industrial	11.47	8.51
Other	9.32	6.29

Similarly, for selected land use categories, Table 2.4 Shows close agreement when the accident rates provided by Chapman, Birmingham radial routes and Lothian region ranked according to adjacent land uses.

Table 2.4 Rank order of accident rates according to adjacent land use types (Chapman, 1978; McGuigan, 1982 and Lawson, 1985, 1986

Land use type (both sides)	Ranking accident rates according to land use type		
	Chapman (1978)	McGuigan (1982)	Lawson (1985, 1986)
Shops/Shops	1	1	1
Residential/Residential	5	3	3
Industrial/Industrial	4	4	4
Shops//Residential	2	2	2
Other	3	5	5

Table 2.3 and Table 2.4 Demonstrated that the accident rates on roads passing through shopping areas are consistently higher than on those passing through other types of land use.

More recently, Jonsson (2005) stated that land use comes in the second order of importance after traffic flow as an explanatory factor for most traffic accident types.

Harwood et al. (2008) found that the presence of schools and commercial structures have a significant impact on vehicle-pedestrian crashes.

Sawalha and Sayed (2001) reported that the land use nature had an important effect on accident occurrence as well as traffic volume, number of traffic lanes, section length, driveway density, un-signalized intersection density, type of median and pedestrian crosswalk density.

Wedagama et al. (2006), in their study, reported that pedestrian accidents in the city centre area are linked to increases in retail and community land use.

Ukkusuri et al. (2012) stated that zones with more fractions of residential land uses show a significantly lower likelihood of pedestrian accidents than zones with higher fractions of open, commercial and industrial land use.

In terms of the developing cities, Ross et al. (1994) indicated that incompatible land uses and unplanned growth may lead to a rate of conflict between pedestrians and motorised traffic.

2.4.3 Pedestrian and road side activities

Many studies indicate the effect of some roadside activities on pedestrian safety. Crossing in front of a parked vehicle, crossing from an incorrect location, unsafe entering/exiting a vehicle, walking, playing, working on the road edge and improper parking are reported to be among the factors affecting pedestrian safety (Al-omari and Obaidat, 2013; Al-Ghamdi, 2002; Derry et al., 2010; Ma et al., 2010; Harwood et al., 2008; and Hunter et al., 1996).

McGuigan (1987) conducted study in Lothian region and reported that there are significant relationships between accidents and location including adjacent roadside development.

Wedagama et al. (2006) investigated the effect of land use, junction density and population density on non-motorized crash casualties and indicated that pedestrian crashes are concentrated around the city centre area and associated with an increase in retail land uses.

Ukkusuri et al. (2012) indicated that areas which have a greater number of public transport stops and schools, which are sources of pedestrian activity, are more likely to have higher crash risks. In addition, the study reported that industrial and commercial land uses are associated with higher crash frequencies than residential land uses. This is consistent with other studies which highlighted the proportion of commercial land use as significant predictors of pedestrian injury crashes (Wier et al., 2009).

Also, land uses and roadside activities such as commercial areas, public transport hubs and stops and schools might influence pedestrians to commit crossing violations, to cross against the lights, or not at the crossing facilities (Cinnamon et al., 2011). Violators, in turn, could face higher risks of encountering traffic conflicts or crashes (Koh et al., 2014).

For example, King et al. (2009) examined the risk of injury for pedestrians committing crossing violations at Brisbane CBD, Australia. The results showed that the crash risk was approximately eight times greater per crossing violation event than that of a legal crossing.

Quistberg et al. (2015) examined the relationship between bus stops and pedestrians' crashes in urban Lima, Peru. The study reported that the presence of a bus stop in an intersection was threefold more likely to result in a pedestrian crash. Also, Harwood et al. (2008) found that the presence of bus stops within 300 m have a significant relationship to vehicle pedestrian accidents. Zegeer et al. (1985) performed analysis on pedestrian crashes for signalised intersections across 15 U.S. cities. After controlling for other variables, bus stops were found to be correlated to accident frequencies. However, these studies used the presence of bus stop as an indicator, while bus stops could have different bus stoppings frequencies.

Hauer and Mohammedshah (2004) performed analysis to estimate crash frequencies on undivided four-lane roads in urban areas. The study findings showed that some variables correlated with crash frequencies, including: access points (e.g. commercial driveways and other driveways, stop-controlled, signalized intersections, intersections,), and the presence of parking. The contribution of the number of commercial driveways and other driveways towards the crash risk ranked as second after the traffic flow.

Al-omari and Obaidat (2013) reported in their study that a significant share of pedestrian accidents occurred while walking on sidewalks, crossing in front of a parked vehicle, playing on the road, walking on the road, crossing at a zebra crossing, getting in or out of a vehicle, working on the road or pushing or towing a hand cart.

Derry et al. (2010) found in a study in Ghana that among the risk factors associated with pedestrian accidents are roadside activities including jaywalking and street-vendors.

Hosseinpour et al. (2014) indicated that roadside activities such as parking, bus stopping and loading were positively associated with crash injury severity. The study reported that the potential of fatal crashes was 34% in road sections with high roadside activities and not protected from through traffic. However, the level of interaction and intensities of roadside activities in this study was categorised only into two categories: low or non-interaction and high interaction. Thus, no methodological approach or empirical bases to measure roadside activities were suggested. This is may be due to the large number of segments (448) which were included in the study.

Verzosa and Miles (2016) examined the factors associated the severity of traffic accidents involving pedestrians in Manila, in the Philippines. The study indicated that the high flow of heavy vehicles along with the street vendors who share the road led to a hazardous environment for pedestrians. Also, the study found that a large share of fatal crashes occurs close to public transport stations and stops.

In the case of developing countries, some studies reported the effect of roadside activities such as, parking, vendors, bus stops and irregular walking or crossing pedestrians on the traffic speed and capacity of urban roads (Chiguma, 2007; Hidayati et al., 2012; Chand and Chandra, 2014; and Pal and Roy, 2016).

In conclusion, the traffic volume, number and type of pedestrian crossings, intersecting traffic volume, pedestrian crossing and walking along the road flow, public transport, land-use type, vehicle parking, demographic factors, traffic mixture and road layout and design are normally associated with pedestrian safety. These factors could be of more importance in urban commercialised areas where dense road side and pedestrian activities take place and could influence the potential of pedestrian-vehicle crashes in the designated areas. As the high level of pedestrian and roadside activities are generated by intense land development, therefore, it is felt necessary to model the pedestrian crash risk in terms of the scale of these activities in addition to traffic characteristics. Despite the high number of these activities, there is comparatively less literature about them and will be further examined to develop models suitable for application in urban areas with high road and roadside activities.

2.5 Pedestrian Safety Modelling

Quantifying pedestrian safety is equally important to defining contributing factors to pedestrian crashes. Many traditional pedestrian and traffic safety analyses have used crash data measures (i.e. fatalities, serious and slight) as the unit for analysis. The accident data was normally collected or estimated using different approaches. Another approach for assessing pedestrian safety is to use other measures of effectiveness which are non-crash based such as (Harwood et al., 2008)**:**

- Measures using pedestrian behaviour changes (including increased using of pedestrian crosswalks, decrease crossing during red signal and other measures.)
- Measures using drivers' behaviour changes (including reduction in traffic speed, increased number of motorists' yields to pedestrians and other measures)
- Other safety measures of effectiveness (including manoeuvre avoidance, pedestrian-vehicles' conflicts and other measures).

The aim of crash models is to assist with making decisions, prioritising, and the allocation of resources effectively. For traffic planning purposes, safety models that measure the expected traffic safety levels for a given area can be categorised into predictive/analytical models and risk models.

2.5.1 Predictive/Analytical models

These models have used mathematical approaches to describe how changes in independent factors affect dependant factors. This is often used to describe the mathematical relationships when there are a number of factors in combination and are difficult to control experimentally. Such models are normally used to measure the influence of roadway designs and treatments options as an alternative decision-making tool to before and after studies based on historical accidents' data records. Another practical use of crash prediction models is site ranking which is used to identify the hazardous road sections with higher risk. Site ranking is important to underlying problems and for improving the safety of road networks (Wang et al., 2011; Wang et al., 2016).

The literature suggests that the possibilities for generalisation for many models are restricted due to lack of flexibility and theoretical foundation (Jonsson, 2005). Therefore, a site-specific investigation is needed.

Road crash data are often classified according to severities (e.g. fatal, serious or slight) or types (e.g. head-on, vehicle-vehicle, vehicle-pedestrian or cyclist). Therefore, crash models may be used to model crash types or severities totally or separately. One common approach is to apply a crash frequency model to model the different crash categories separately. For example, Ukkusuri et al. (2012) developed a pedestrian crash rate models to investigate the link of roadway design in addition to the effect of land use and spatial aggregation on pedestrian safety for New York City. Results showed that the possibility of a vehicle-pedestrian crash increased with the roadway width and number of lanes. This is suggesting that narrowing roadways may in turn reduce pedestrian crashes.

Abdel-Aty and Radwan (2000) used Negative Binomial modelling to investigate the frequency of crash occurrence. The model showed the significance of the traffic volume; curvature degree; median; lane and shoulder widths; area type, in terms of urban or rural; and the length of sections on the crash frequency potential. Also, several Negative Binomial models were developed accounting for the demographic characteristics of drivers (gender and age). The results reported that narrow lane and shoulder widths, speeding, heavy vehicle volume, more lanes, urban road sections and less median width were associated with a higher crash frequency.

Agarwal (2011) studied pedestrian safety at intersections using simulation and surrogate safety measures instead of accidents' data and stated that the most suitable form for predictive models are negative binomial regression and generalised linear models.

Hall (1986) used generalised linear modelling techniques to relate traffic accident rates by types to pedestrian and traffic volumes, the geometric features and control characteristics of the urban intersections to develop models estimating the effects of such factors on safety. As a result, many geometric features were identified to be correlated with crashes frequencies. The developed models included the number of lanes, the width of the intersection`s approach, opposing roadways offset, gradients, sight distance and presence of median.

Harwood et al. (2008) developed a model to estimate pedestrian safety according to available site features and/or proposed treatments on urban and sub-urban roads. The model was developed for three- and four-legged signalised intersections and multi accidents modification factors were created. Also, the analysis established that intersections with more lanes were associated with higher pedestrian crashes rates.

Lyon and Persuad (2002), Leden (2002) and Zeeger et al. (2005) developed safety predictive models to estimate pedestrian crashes using the negative binomial regression approach taking the following form:

$\lambda_i = exp\ (\beta_0 + \beta_1\ ADT + \beta_2\ PedVol + \beta_3\ X_3 \ldots \beta_n\ X_n)$

Where:

λ_i is the number of expected pedestrian crashes,

$\beta_0 \ldots \beta_n$ are coefficients to be estimated; ADT is annual daily traffic; Ped Vol is the annual average daily pedestrian volume; and $X_3 \ldots X_n$ represent other site-specific feature such as speed limits, number of lanes, type of median, left turn volume percentage and the presence or absence of crosswalks.

In spite of the models' functional form similarity for these three studies, Zeeger et al. (2005) looked at intersection and non-intersection sites, whereas Lyon and Persaud (2002) and Leden (2002) focused on intersection sites.

Hauer and Mohammedshah (2004) also used a negative binomial model to estimate crash frequencies on undivided four-lane roads in urban areas using the following independent variables: traffic volume, horizontal and vertical alignments, lane and shoulder width, percentage of heavy vehicles, the presence of parking and left turn, and side access including intersections, commercialised and other side streets. The study findings showed that the variables that have correlation with crash frequencies include: traffic volume, access points (commercial driveways), and the speed limits.

Al-Twaijri et al. (2011) investigated the factors influencing the severity outcome of injury crashes in Riyadh city using a mixed logit model and a standard multinomial logit model (MNL). The results indicated that nationality, age, speed, lighting and road surface conditions are associated with fatal crash probability.

Pande and Abdel-Aty (2009) investigated the geometric design and traffic parameters correlated with severe crashes on multilane arterials. The study found that pedestrian and single vehicle crashes show no significant relationship with traffic volume; while for severe rear end crashes, traffic volume, time of day and week, pavement condition, horizontal curvature and median type were significant parameters.

Verzosa and Miles (2016) employed binomial logistic regression to investigate the factors that increase the severity of traffic accidents involving pedestrians in Manila, in the Philippines. The study reported that the majority pedestrian fatalities occurred in high traffic volume, at high speed, on multi-lane roads where particular surrounding land uses generate a mix of heavy vehicles and pedestrian flow. The analysis also found that the fatal crashes occur close to public transport stations.

Another model approach is to estimate the number of crashes by different severity levels separately. As crash frequency and severity are both important measures to identify the road segment risk, several studies have attempted to integrate both measures in modelling safety. In this context, two phase techniques were used where the number of crashes can be predicted using counts models and discrete choice methods can be applied to estimate the crash severity. The most common approach is to utilise regression frequency models (e.g. Poisson, Negative Binomial, Zero inflated Poison models) to model the number of crashes; and the severity of crashes can be predicted using a logistic model, Multinomial logit model, or probit model (e.g. Carson and Mannering, 2001; Lee and Mannering, 2002; Yamamoto et al., 2008; Kim et al., 2008; Milton et al., 2008; Wang et al., 2011; Qi et al., 2013).

Some other studies have recently utilized multivariate modelling techniques which predict crash frequency by type or severity simultaneously, rather than separately. Examples of these models include: multivariate negative binomial, multivariate Poisson and multivariate Poisson log-normal (El-Basyouny and Sayed, 2009; Ye et al., 2009; Anastasopoulos et al., 2012).

Despite the effectiveness of multivariate techniques in predicting the crash frequency and severity simultaneously, a major concern for this approach is the difficulty in predicting crash frequencies by type or severity correctly, where the crash data is featured by a small size or low sample mean. This limitation can be encountered as the total crash frequency needs to be separated into different categories. Another limitation is that they require the same set of factors to predict all response categories; whereas some severity levels are not necessarily influenced by the same factors (Wang et al., 2011).

2.5.2 Risk models

Another approach to assess safety is modelling road user behaviour and their perception with regard to crash risk or danger and traffic safety.

These modelling approaches investigate the risk factors that explain the behaviour of the individual road user and assess the safety based on the potential of risk reduction due to different interventions and countermeasures (Archer, 2005).

Christopoulou and Latinopoulou (2012) developed a model to estimate the level of service for pedestrians in Greek urban areas, using the perceived quality of service of pedestrian movement to reflect the contribution of traffic parameters, geometry, environment, sidewalk pavement and pedestrian movement factors with a questionnaire survey.

A study of Landis et al. (2001) provides a measure of the level of service for pedestrians within a roadway environment. The model was developed to objectively quantify pedestrians' perception of safety in urban roadways between intersections, through analysis of 1250 observations, from inviting 75 participants to walk on selected road sections in Florida. Lateral separation between pedestrians and traffic, traffic volume, effect of traffic speed, motor vehicle mix and driveway access frequency and volume were used as explanatory variables in this study. The model was developed using a stepwise multi-variable regression analysis.

Another study conducted by Jensen (2007) used objective quantification of pedestrian stated satisfaction with road segments in Denmark to determine the effect of existing traffic characteristics, geometric conditions and other factors. This model was developed by asking 407 participants to rate on six-point scale video clips from different roadway sites recorded by pedestrians walking along the road. The significant variables were speed and traffic volume; the volumes of pedestrians; urban land uses; the widths and types of pedestrian facilities; rural landscapes; the widths and numbers of the road lanes; cyclists and parked cars; and the presence of a bus stops, median and trees. This model was developed using cumulative logit regression of the stated ratings to estimate the level of service.

Kang, Xiong and Mannering (2013) also used the video clips' technique and asked a sample of 114 respondents to assess the level of service for a walking pedestrian using a scale of A-F in Chinese urban areas. The pedestrian perceptions were found to be influenced by many factors including; pedestrian flow; the presence of a barrier to separate the sidewalk pavement from vehicle traffic; sidewalk width; the presence parking vehicles; the cyclists flow; the speed of the cyclist; whether or not cyclists were riding against the pedestrians' flow; the presence of businesses along the walking pavement; weather conditions; and age of the participant. This study used the random parameters ordered probability approach.

Carter et al. (2006) developed a pedestrian safety model to proactively prioritize intersection crosswalks and approaches in an urban environment using data from 67 intersection approaches in US cities. This model was constructed using subjective ratings of video clips by pedestrian experts. The included variables in this model are: the traffic volume of the main street, 85^{th} percentile vehicle speed, type of intersection control (stop sign or signal), area type and the number of lanes.

Baltes and Chu (2002) proposed a methodology to assess the level of service of pedestrians crossing streets at midblock locations in urban areas in Florida. A collection of stated pedestrians' perceived level of service and selected site characteristics were analysed. The results indicated that the turning movements and signal spacing, width of painted medians, the cycle length and presence of pedestrian signals are significantly associated with perceived level of service by pedestrians for midblock crossings. This model was developed using the ordinary least-squares regression method.

Asadi-Shekari et al. (2015) conceptualized the pedestrian safety index (PSI) to evaluate the infrastructure facilities along roadway segments for pedestrians by comparing the standard conditions to the existing. This study proposes a scale system method to rate roadways to evaluate the safety facilities on an urban road. The final formula included 24 indicators covering traffic speed, geometrical design factors and roadside environment, reviewed from 20 guidelines developed in different countries.

Nagraj and Vedagiri (2013) developed a method to estimate the pedestrian level of service for crosswalks at signalised intersection with mixed traffic conditions in Mumbai, India. This study conducted a questionnaire to collect the perceived pedestrians' level of service for signalized intersection crosswalks and the significant influencing factors concluded by the study were: through traffic, turning traffic and pedestrian delay and. A stepwise regression approach was used for model development.

Wang and Liao (2012) used a structural equation model to develop a model relating the pedestrian's degree of satisfaction with a non-signalized intersection's safety with other factors including: design characteristics of a non-signalized intersection, pedestrian crossing and driving behaviours, traffic flow characteristics and environmental conditions.

Cheng et al. (2014) introduced the concept of a pedestrian safety conflict index and established a safety level system to assess pedestrian crossing safety, based on the study of the mechanism of the left-turn vehicle and pedestrian conflict at signalized intersections in Changchun, China. This study used multiple regression analysis and concluded that pedestrian volume, left-turn traffic volume and ratio of disadvantaged pedestrians to the total crossing population are among the most affecting factors.

2.6 The iRAP Model

In 2001, the concept of star rating score was developed and then extended in 2006 to be used in low and middle-income countries under the International Road Assessment Program (iRAP, 2016) to tackle the devastating economic and social cost of road crashes. This methodology provides a comprehensive approach, with the objective of assessing road safety provided by the road infrastructure for each road user category, to develop tools and cost-effective countermeasures. The iRAP methodology employs the star rating concept to reflect the effect of road design, multiple countermeasures and treatments, operating traffic speed and external flow influence in a quantitative rating score.

The main aim of the iRAP star rating approach is to estimate a score for each road segment and can be compared with scores of other segments.

The star rating score is the calculated outcome of the approach and transformed into an equivalent star rating to express the safety provided by the road infrastructure in a simple way.

The star rating can be calculated using road inspection data, giving an objective measure for the safety level provided by the road for each road users: vehicle occupants, motorcyclists, cyclists and pedestrians. The star rating score has five levels ranging from a five-star road which is the safest, to a one-star reflecting the least safe road. Importantly, another advantage is that the star ratings system can be calculated without reference to recorded crash data, which is often unavailable in the case of low and middle-income countries.

A star sating score (SRS) is calculated for each road user separately including vehicle occupants, pedestrians, motorcyclists and cyclists and for each 100-metre road segment, using the following function (iRAP, 2016):

$$SRS = \Sigma \text{ Crash Type Scores}$$

Where:

SRS represents the quantified relative risk of fatality and serious injury for a road user; and

Crash Type Scores = Likelihood x Severity x Operating speed x External flow influence

Where:

- Likelihood is a quantitative result of multiplying road attribute factors contributing to the possibility of an accident will be initiated.
- Severity represents the contribution of specific road features to the severity of an accident
- Operating speed represents factors that account for rate of changes in risk with speed
- External flow influence refers to the level to which a road user`s risk of being subject to accident is dependent on another road user.

The star rating methodology included crash types for each road user which are account for high share of fatalities and serious injuries as shown in Table 2.3.

Table 2.5 Crash types included in the SRS methodology (iRAP methodology, 2016)

Vehicle occupants	Motorcyclists	Bicyclists	Pedestrians
• Run-off road • Head-on • Intersections • Access points	• Run-off road • Head-on • Intersections • Access points • Moving along the road	• Travelling along road • Intersections • Run-off road	• Walking along road • Crossing road

For pedestrians, two crash types are considered: walking along the road and crossing the road. However, crossing the road is further divided for calculation purposes into two directions: the inspected road and the side road.

The iRAP methodology provides an approach to prepare an improvement procedure (Safer Roads Investment Plans) and present an economic appraisal for the needed infrastructure countermeasures to upgrade road safety. The concept enables users to estimate the risk of crashes in each 100-metre road segments; and a further estimate of serious casualties and fatalities can be made in each segment under the prevailing condition. The approach is finally used to prioritise and optimise the potential countermeasures within the available budget.

The economic appraisal can be done by comparing the economic benefits of preventing potential human casualties due to implementing one or more of 93 potential countermeasures with their implementation cost.

However, this approach is set to assess road safety in inter-urban roads and needs to be further developed to enable assessment of urban roads (Lynam, 2012).

To this end, this research will seek to develop a pedestrian safety model with the view to enhance the iRAP assessment process. The subsequent chapters show how this has been achieved.

Chapter 3

METHODOLOGY

3.1. Introduction

The primary aim of this study is to enhance the safety of pedestrians moving in environments with high roadside activities in urban areas. One of the objectives to achieve this is to develop a new model estimating the safety for pedestrians moving in such environments and to capture and estimate the significant risk factors contributing to pedestrian crashes.

The literature review presented in Chapter 2, found several safety models for investigating the effect of traffic speed, volume characteristics, and geometric design on pedestrians' safety. However some other factors could affect pedestrian safety in environments with high roadside activities in urban areas: traffic speed variation; significant pedestrian crossing violations (i.e. those who cross from non-specified crossings or during the red interval); type and intensity of roadside activities, including bus stoppings, parking and un- parking events (Al-omari and Obaidat, 2013; Ukkusuri et al., 2012; Box, 2010; Derry et al., 2010; Ma et al., 2010; Zheng et al., 2010; Harwood et al., 2008; Wedagama et al., 2006; Hauer, 2004; Al-Ghamdi, 2002; Garber and Ehrhart, 2000; Taylor et al.,2000; Hunter et al., 1996). Despite the high number of these activities it appears that there are no models to measure their effect on pedestrian safety.

Therefore, this study seeks to develop models for pedestrian safety accounting for volume and types of roadside activities, traffic speed variation, number of bus stops and land uses to estimate their effect on crash risks. It is envisaged that the model may be integrated in the iRAP star rating system with the view to enhance the (iRAP) star rating model framework.

The suggested model may provide traffic engineers and road safety managers with the ability to rate the road environment according to pedestrians' crash risk; and also in the process of evaluating existing roads, prioritize the roads in need of safety interventions, design new roads, or redesign existing roads where the existing iRAP is used (iRAP, 2016).

The following sections outline how the above approach may be achieved.

3.2. Overall Research Approach

The phenomenon under research is the interaction between specific actors, namely; pedestrians and drivers of motorised vehicles in a specific environment which is the urban roads and as such is focussed on agency. The research aims to determine the agency occurring between the characteristics of the environment and the incidence of conflicts between pedestrians and the drivers of motorised vehicles which result in injury or death. Since this is a real world interaction the research is based in observation of real world interactions and analysis of real world data.

The literature review showed studies suggesting evaluation methods and models addressing different factors for assessing pedestrian safety, including traffic characteristics, roadway design and land use and demographic features. Also, many previous studies on vehicle-pedestrian crash occurrence and prediction have been conducted at intersections and street segments separately. Moreover, previous research has investigated the link between land uses and pedestrian crash frequency without considering the effect of pedestrian and roadside activities' intensity: such as bus stoppings, parking, pedestrian crossing and violations' volume, loading and un-parking events and the traffic speed variation (Zegeer et al., 1985; Brude and Larson, 1993; Wedagama et al., 2006; Wier et al., 2009; Ukkusuri et al., 2012). For that reason, and based on literature review there appears to be a need to develop models suitable for application in urban areas with high road and roadside activities.

To this end, this study seeks to conceptualize a pedestrian safety by the following steps:

- Conducting literature reviews to explore existing research into the types of interaction under investigation and understand what is already known which is already presented in chapter 2.
- Selecting sites for the study in Birmingham city guided by the considered variables.
- Analysis of actual crash statistics to understand the types of crashes occurring in the selected study area and their frequency.
- Conducting a pilot study observation of roads in the selected sites to investigate what types of interactions are happening.
- Design of the field observations based on the above.
- Data collection and analysis.
- Development of statistical model using different regression tools.
- Testing and validation against actual data.
- Application of the estimated model through integration with iRAP and testing the results.

3.3 Theoretical Background

In order to investigate the significant contributing factors influencing pedestrian safety, the first part of this study will develop crash models for pedestrians moving in urban areas with high road and roadside activities. The first part involves selecting study sites and the second part performing the modelling process. Different generalised regression models will be considered to estimate the link between the risk of pedestrian crashes at the study sites and the studied contributory factors; while a probability model will be utilised to measure the relationships of the potential contributory factors on the severity of pedestrian crash outcomes. The logistic regression method will be used to model the crash severity as a binary of slight outcome or fatal or serious outcome. The reason behind the use of statistical models is their ability to examine the significance of the factors and to assess their contribution to crash frequency and severity (Chiou and Fu, 2013).

Validation of the developed models and assessment of their predictive ability will be made based on three criteria: goodness-of-fit assessment; visual comparison between the predicted and observed pedestrian crash events: and a real field validation test.

The following phase will constitute the application of the proposed model through three steps. Initially, the developed model will be used to assess the potential contribution of road design elements, traffic volume and pedestrians' volume and activities factors to likelihood of a crash occurrence or the severity level of a crash through sensitivity analyses. The second step is to incorporate the developed crash risk factors to enhance the (iRAP) by integrating the estimated risk factors into the iRAP star rating scoring function for pedestrians. Finally, the third step is to verify the performance of the suggested enhanced (iRAP) approach, the output scores, star rating, predicted and actual crash frequency data for pedestrians.

Both application and validation of the enhanced (iRAP) modelling approach will be covered and further discussed in Chapters 4 and 9.

3.4 Modelling Approach

Previous studies have showed a clear link between the built environment represented by land use characteristics, demographics, road network and traffic characteristics and pedestrian collision risk and severity (Wier et al., 2009; Ukkusuri et al., 2012).

The type and intensity of the surrounding land use are significant contributors in vehicle-pedestrian crashes. Subsequently, the pedestrian-vehicle crashes constituted a higher share in urban areas, especially settings of higher road and roadside activities and commercial tracts, where pedestrians' activities, traffic volume and traffic speed variations are higher compared to other area settings such as residential areas (Clifton and Fults, 2007; Zegger et al., 2002; Zheng et al., 2010).

Conventionally, pedestrian violations, roadside activities and traffic speed variations are not explicitly considered in pedestrian crash models; instead other physical and socio-economic characteristics are taken as direct inputs, such as land use type, population characteristics like age and income, road network and travel characteristics (Ukkusuri et al., 2012; Wier et al., 2009; Wedagama et al., 2006). To this effect, Wier et al. (2009) suggested a conceptual framework to show how the built environment, land use characteristics, employee and resident population size, travel behaviours and population characteristics may predict vehicle-pedestrian crashes as shown in Figure 3.1.

This study will concentrate on factors of pedestrian and roadside activities and estimate their contribution to vehicle-pedestrian traffic crash risk particularly in urban areas with high pedestrian and roadside activities using a statistical modelling approach to measure the variations of crash risk intensities across a road section. The research will be conducted in the designated environment for the time and budget constraint and driven by the aim of the study and literature review.

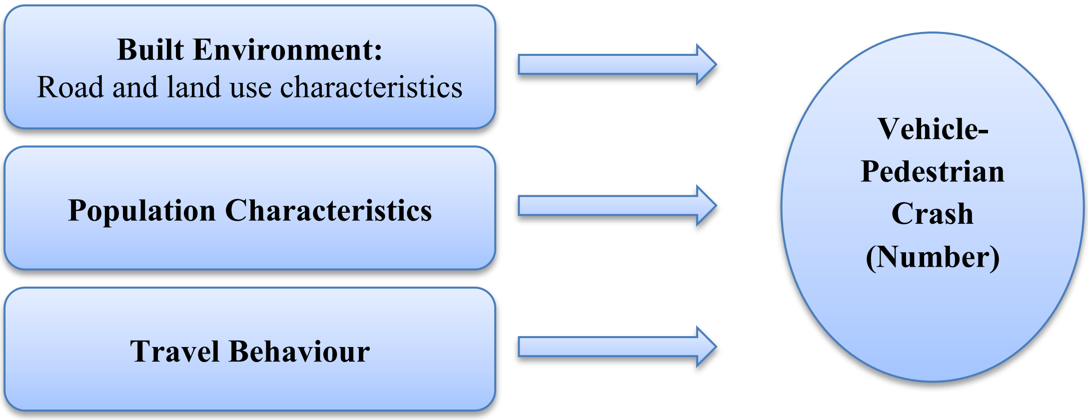


Figure 3.1 Area-level model of vehicle-pedestrian injury collisions (Wier et al., 2009)

As the high level of pedestrian and roadside activities are generated by intense land development, therefore, it is felt necessary to model the pedestrian crash risk in terms of the scale of these activities in addition to traffic characteristics.

As shown in Figure 3.2, this study suggests a framework that takes into account attributes contributing to pedestrian-vehicle crashes in urban areas where significant pedestrian and roadside activities take place. The range factors that could influence the potential of crash events includes traffic volume, number and type of pedestrian crossings, intersecting traffic volume, pedestrian crossing and walking along the road flow, public transport, land-use type, vehicle parking, demographic factors such age and income, traffic mixture and road layout and design.

Given the difficulties in measuring pedestrian and road side activities, the study will devise a method to account for these inputs using a sampling procedure to measure: pedestrian movements, crossing and violations, loading, parking and un-parking vehicle activities, besides the other inputs in the selected road sections, by conducting a field inventory and observations.

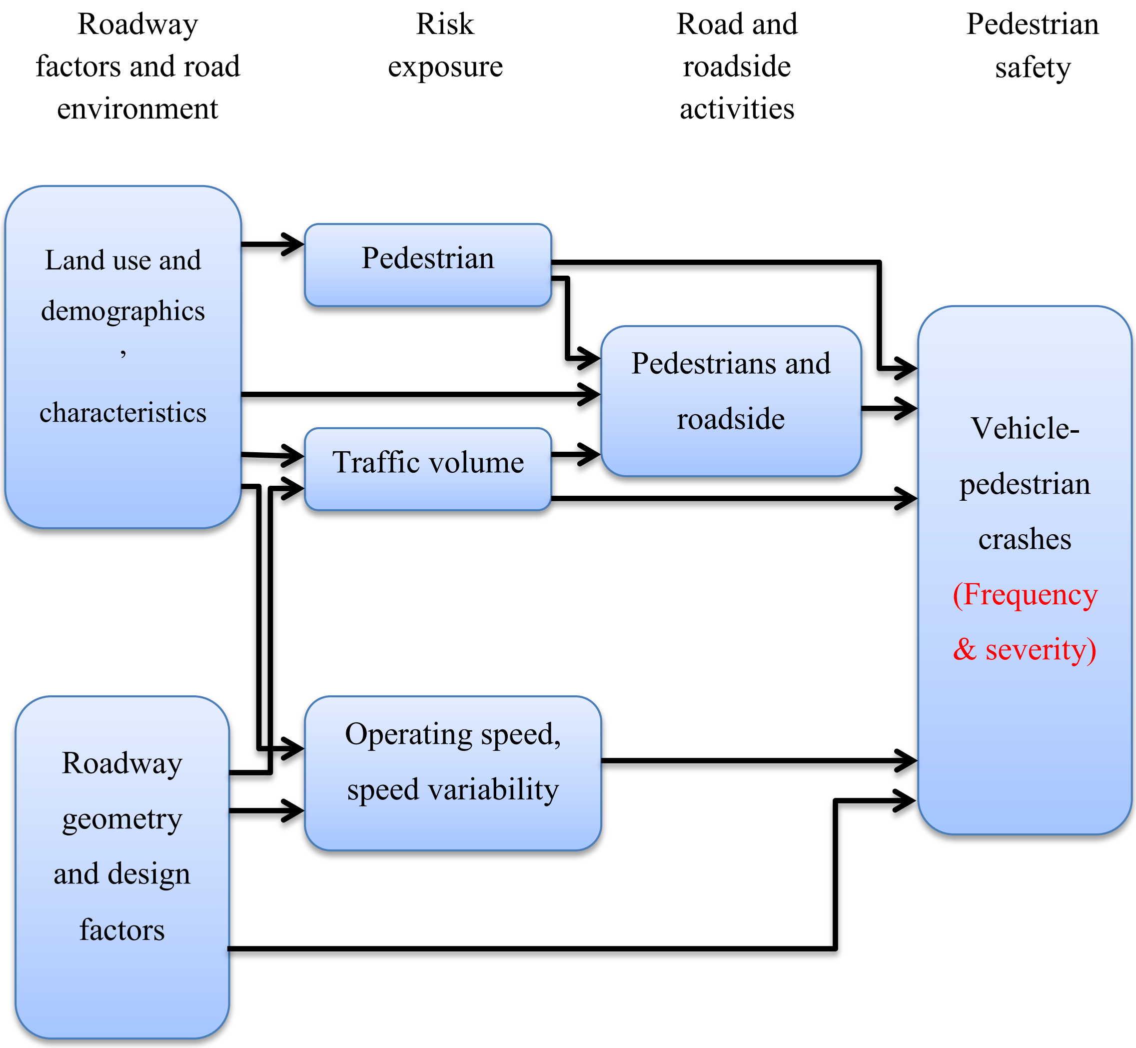


Figure 3.2 Attributes contributing to pedestrian crashes

After assessing the pedestrians' crash risk using the above approach, the iRAP model may be enhanced to improve its ability to assess pedestrian safety in urban areas by utilizing the estimated contribution of risk factors to the overall pedestrian crash potential. The overall process suggested in this research is shown in Figure 3.3.

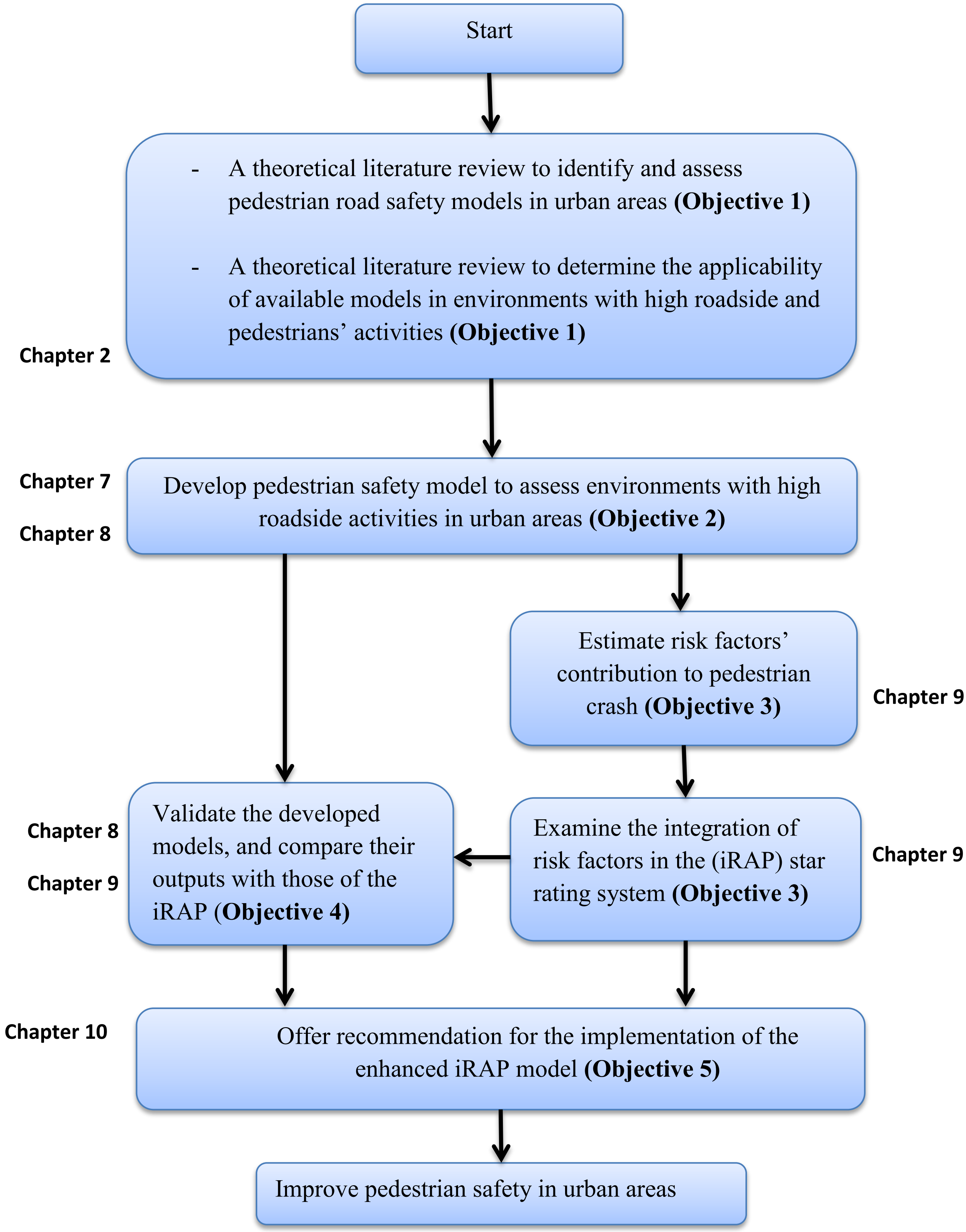


Figure 3.3 The overall research process of this study

Chapter 4

iRAP MODEL FOR PEDESTRIANS

4.1 Introduction

During 2001, a star rating process to assess road safety with regard to road design features was developed for the European Road Assessment Programme (EuroRAP). This EuroRAP used a scoring system to rate the risk of roads for car occupants and the risk scores' values assign a star rating for each road section. This concept applied to several Europeans countries and was then adopted by Australia (AusRAP) and the United States (usRAP) (Lynam 2012).

The main aim of the iRAP star rating approach is to estimate a score for each road segment which can be compared with the scores of other segments. Furthermore, the calculated scores are used to rate road sections with one to five stars.

The iRAP star rating combines both crash likelihood and crash protection scores. Also, the iRAP system developed a star rating scoring for not only car occupants, as is the case with EuroRAP, but also for motorcyclists and other vulnerable road users including cyclists and pedestrians. Furthermore, iRAP presents an approach to predict casualty crash rates along the considered roads based on the calculated star rating score. However, the iRAP support the assessment of pedestrian safety in interurban areas using measures of different inputs of road attributes.

In 2006, the concept of the star rating score was further developed and extended to coordinate activities among EuroRAP, AusRAP and usRAP; and to be used in low and middle-income countries under the international road assessment program (iRAP) to tackle the devastating economic and social cost of road crashes. This methodology provides a comprehensive approach with the objective of assessing road safety provided by road infrastructures for each road user category, and to develop tools and cost-effective countermeasures. The iRAP methodology employs the star rating concept to reflect the effect of road design, multiple countermeasures and treatments, operating traffic speed and external flow influence in a quantitative rating score (iRAP, 2016; Harwood et al., 2010).

4.2 Need for Enhancement

The current iRAP models were developed for inter-urban areas; whereas pedestrians tend to travel more in urban areas where traffic crash hazards are higher than in inter-urban settings (Lynam, 2012; WHO, 2013b). Therefore, there is a need to extend the current iRAP star rating system to include risks factors which are common in urban areas, to improve its assessment ability in an urban setting.

This study seeks to develop new models to capture the most appropriate and influential risk factors associated with pedestrian safety in urban areas, with a focus on roadside activities. All significant findings of the model will then be examined for potential integration with the iRAP star rating system, with the ultimate view to introduce a more comprehensive tool for assessing pedestrians` safety in environments with high roadside activities in urban areas. The pedestrians' risks may be mitigated by means of better road design, treatment of roadsides, delineation and road marking, road network condition, traffic flows separation, median types, frequency and design of junctions.

The iRAP star rating system was chosen to integrate the developed models in this study for the following characteristics:

- The iRAP scoring process has been developed to assess the risk for different types of road users including pedestrians.
- The iRAP star rating methodology aims to capture the risk for pedestrians that relates to the infrastructure design on road sections.
- The multiplicative feature of the models presents a clear process for risk calculation; and simple adjustment of the risk rate scoring can be achieved to estimate the potential reduction in risk.
- The iRAP star rating methodology can be used to produce risk maps documenting the risk level in a road network and in a similar way, to estimate crash numbers per section, and the potential benefits from road improvement plans.
- The iRAP model has established an international reputation in safety assessment programs and is now active in around 70 countries throughout Europe; North, South and Central America; Asia and Africa, including a wide range of low-income and middle-income countries.

4.3 Description of the iRAP Star Rating Concept

The iRAP concept involves three main activities for safety assessment and mapping of roadway systems for the defined four road users separately. These are:

- Star ratings based on inspection data of roads to examine the level of safety provided for road users; including vehicle occupants, cyclists, motorcyclists and pedestrians, from crashes and from severe outcomes when crashes occur.

- Safer roads investment plans by considering more than 90 road treatment options to improve star ratings for road sections.
- Risk mapping by using crash data if available to produce risk maps documenting the risk of death and injury crashes and show where the risk is higher.
- Performance tracking to monitor changes in the safety performance of the roadway network and measure the effect of ongoing safety improvement programs.

The following sections focus on the iRAP methodology to assess pedestrian safety and star rating.

4.3.1 Structure of the original iRAP model

The iRAP star rating methodology included crash types for each road user which account for a high share of fatalities and serious injuries, as shown in Chapter 2 (see Table 2.3). For the pedestrian star rating, two crash types are included in the process:

- Star rating scores for pedestrian movements along the road for both sides (driver side and passenger side)
- Star rating scores for pedestrian movements across the road (main inspected road and side roads).

Consequently, the iRAP star rating score (SRS) is calculated for pedestrians for each 100 metre road segment, using the following function (iRAP, 2016):

$$SRS = \Sigma \text{ Crash Type Scores}$$

Therefore,

Pedestrians SRS = (along score-driver side + along score-passenger side)/2 + (crossing score-inspected road + crossing score-side road)

Then, every crash score is calculated as expressed in Equation 2.2

Crash type scores = likelihood x severity x operating speed x external flow influence

Where likelihood, severity, operating speed and external flow influence can be calculated by multiplying the associated risk factors to account for the contribution of specific road features for the risk and severity of accidents.

4.3.2 Factors included in iRAP pedestrians` model

To calculate the star rating score for pedestrians, multiple road attribute risk factors are used as shown in Table 4.1.

Table 4.1 Pedestrian star rating score factors (iRAP, 2016)

Crash type	Component	Risk factor
Along score (driver and passenger sides calculated separately)*	Likelihood	Sidewalk Curvature Quality of curve Sight distance Lane width Delineation Grade Road condition Speed management / traffic calming

		Vehicle parking Shoulder rumble strips Street lighting School zone warning
	Severity	Sidewalk
	Operating speed External flow influence	
Crossing score (inspected road)	Likelihood	Number of lanes Median type Pedestrian crossing Pedestrian crossing quality Intersection type Intersection quality Pedestrian fencing Skid resistance / grip Street lighting Sight distance Vehicle parking Speed management / traffic calming School zone warning
	Severity	Pedestrian crossing

	Operating speed External flow influence	
Crossing score (side road)	Likelihood	Number of lanes (side road) Median type (side road) Pedestrian crossing (side road) Pedestrian crossing quality (side road) Intersection type Intersection quality Pedestrian fence Skid resistance / grip Street lighting Sight distance Vehicle parking Speed management / traffic calming School zone warning
	Severity	Pedestrian crossing (side road)
	Operating speed External flow influence	

* Along score for driver and for passenger sides calculated separately then the average value counted and used for the final score calculations

Each road attribute is related to a relative risk factor in terms of crash risk likelihood and severity. To produce a star rating score for each road section, the iRAP used all the risk factors associated with road attributes given in Table 4.1. For each road attribute, there is a set of risk factors according to its category, indicating the influence on pedestrian crash risk and severity. Examples of the road attributes' categories and their associated risk factor are shown in Table 4.2.

Table 4.2 Risk factors by road attribute category for pedestrians

Road attribute	Category	Crash type - Risk factor	
		Crossing score	Along score
Speed Management/ Traffic calming	Not present	1.25	1.25
	Present	1.0	1.0
Sight distance	Adequate	1.42	1.42
	Poor	1.0	1.0
Street lighting	Not present	1.25	1.25
	Present	1.0	1.0
Pedestrian fencing	Full length	0.0	-
	At pedestrian crossing	1.0	-
	None	1.25	-
Lane width	Wide (>= 3.25m)	-	1.0
	Medium (>= 2.75 m to < 3.25 m)	-	1.2
	Narrow (>= 0 m to < 2.75 m)	-	1.5

All the road attributes for each 100-metre road segment are gathered by site visits and reference photos to code the road attribute categories. An example of a road attribute reference photo is shown in Figure 4.1. An example of the current iRAP methodology calculation process for pedestrians can be seen in Appendix A.

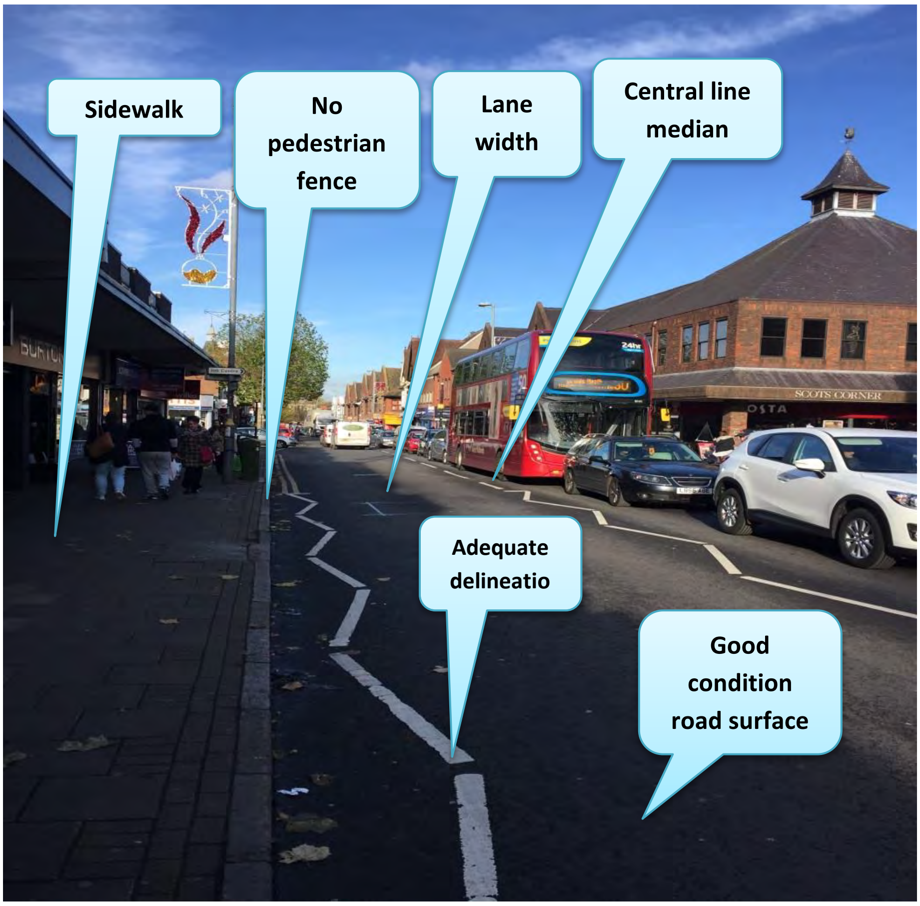


Figure 4.1 Road attributes' reference photograph

4.4 Methodology to Enhance iRAP for Pedestrians

The iRAP star rating score (SRS) is calculated separately for each pedestrian's crash type category, including along score, crossing score (inspected road), and crossing score (side road) for each 100-metre road segment, by multiplying the relative risk values for each road attribute factor related to likelihood and severity. The relative risk values of road attribute factors used in the iRAP methodology are also known as crash modification factors (CMFs) (iRAP methodology, 2013). The relative risk values were chosen to reflect the effect of road features and for most factors the optimum value is 1; the higher or lower values reflect the increased or decreased risk of potential crashes.

Therefore, the approach to enhance the existing iRAP star rating system, to improve its ability to assess the safety of pedestrians moving in road environments with high roadside activities in urban areas is to develop risk factors, measuring the relative risks of the common risk factors in urban areas and improve its assessment ability in an urban setting. The study will investigate the influence of road side activities including bus stoppings and parking activities,

Figure 4.2 Urban road with high pedestrian and roadside activities

pedestrian activities, traffic mix and speed variability and measure their effect on the likelihood and the severity of pedestrian related crashes. Figure 4.2 shows an example of urban roads with high roadside and pedestrians activities which are not accounted for in the current iRAP methodology.

This approach results in obtaining pedestrian risk factors based on the estimated models. Once crash risk factors have been obtained and quantified, integration and assessment of the chosen factors into the international road assessment programme (iRAP) will follow. This approach is consistent with iRAP methodology to assess the crash risk for different road users, as the iRAP star rating score is estimated by multiplication of the relative risk factors related to likelihood and severity of the defined road user`s crash types. The results of the risk factors for pedestrians will present the relative effect and will be impeded in the iRAP model using the flexible multiplicative form to improve the accuracy of along and crossing scores.

Figure 4.3 shows the iRAP pedestrian star rating score calculation methodology. The methodology being proposed will be demonstrated and tested in Chapter 9.

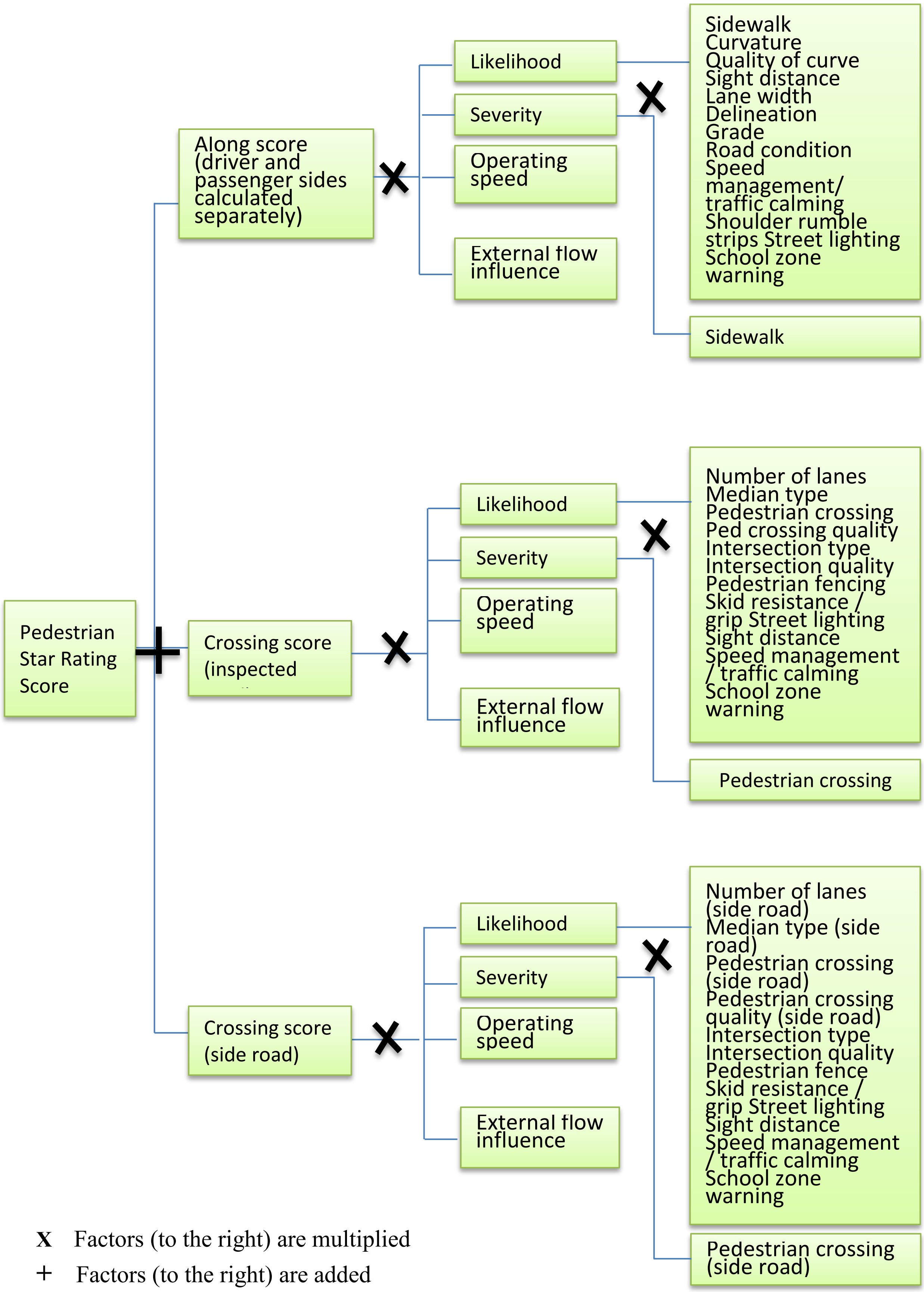


Figure 4.3 iRAP pedestrian star rating score calculation methodology (iRAP, 2016)

Chapter 5

MODEL VARIABLES

To implement the aim of this study in a systematic manner, the basis of the analysis of pedestrian safety while moving in such road environments should be defined. This chapter presents how this was achieved through the clarifications of the data requirements and it is structured in two sections. The first section introduces the new aspects to be considered in pedestrian road safety modelling. The second section presents the data requirements for the modelling process and gives the methods used for data collection.

5.1 New Aspects

One of the objectives to achieve in this study is to develop a new model for pedestrian safety and estimate the significant risk factors contributing to pedestrian crashes.

Regarding the included factors, and as stated in Chapter 2, in addition to traffic speed, volume characteristics and geometric design, some other factors may also affect pedestrian safety; such as traffic speed variation, significant pedestrian crossing violations, bus stoppings, parking and percentage of heavy vehicles in traffic flow.

Consequently, range variables that may influence the potential of crash events as investigated in literature review were categorized into: traffic volumes; traffic operating speeds and variations; pedestrian-vehicle crash data; roadway geometric design and characteristics; pedestrian volume; road and roadside activities.

5.2 Data Requirements

The development of the suggested safety models required a range of data to be collected. This section presents the methodology used to collect the data used in this study. It covers the site selection criteria, methods and the consulted sources to obtain the adequate data on traffic volumes, traffic operating speeds and variations, pedestrian-vehicle crash data, roadway geometric design and characteristics, pedestrian volume, road and roadside activities for each site. However, some of the required data were not available in most of the study areas, particularly that for pedestrians. There were challenges in measuring pedestrians' volume and activities and operating speed because of time and budget constraints.

5.2.1 Site selection

The site selection process for the data collection was mainly guided by the considered variables. Only sites which could provide the required variables and conditions were considered.

The road segments' selection was determined by the presence of a wide range of commercial, educational and residential surrounding tracts, pedestrian and roadside activities, traffic flow conditions and the percentage of heavy vehicle, bus stops, on-street parking, and parking manoeuvres. A preliminary reconnaissance survey was conducted in Birmingham city to identify the suitable road segments for potential inclusion in the study and to examine the potential interactions with traffic in the selected study environment. The preliminary visits were conducted to check the movement of pedestrian along the streets or while crossing and how they interact with traffic flow.

The study sites were selected to satisfy the following factors:

- The presence of a wide range of commercial, educational or residential tracts in the vicinity.
- No significant change to the roadway configuration occurred during the last eight years.
- The availability of active bus stops, street parking activities and pedestrian movements.
- Traffic volume of sufficient size, distributed in both directions with no limitations to heavy vehicles passing through to allow for a traffic mix.
- The need to be located in flat or gentle terrain (grade 0% to < 7.5%) with straight or gentle curving and adequate sight distance along the road segment.
- The need for all traffic characteristics and roadside' activities data to be collected in normal weather conditions and un-congested traffic flows and no roadworks on the site.
- The total number of travel lanes were one or two in each travel direction.
- For the segments, most of their lengths were undivided with a centre line or hatching and no physical median present.
- The provision of roads of generally good condition and sidewalk pavement along the road segments.

In effect, the less variation in physical configuration for the selected study sites, the better control and focus on the study's variable of interest. Accordingly, twelve road sections meeting the above criteria were chosen for model development, in addition to three other sections to provide data for testing the developed models. The selected road segments were within a 0.8-1.2 km length range; this was because it was practically difficult to find long road sections with high pedestrian and roadside activities. Table 5.1 shows the selected road sections for the purpose of the study.

Table 5.1 Selected road sections for the study

No	**Name of the Road**	**Covered Length (km)**
1	Coventry Rd N-East bound	0.9
2	Coventry Rd S-West bound	1.1
3	Hagley Rd -East bound	1
4	Hagley Rd West bound	1
5	Moseley Rd -S	0.8
6	Moseley Rd -N	0.8
7	Pershore Rd -N	0.9
8	Pershore Rd -S	0.9
9	Soho Rd -E	0.8
10	Soho Rd -W	0.8
11	Stratford Rd-N	1.2
12	Stratford Rd-S	1.6

5.2.2 Pilot study and preliminary reconnaissance survey

The variables' selection for the study's data collection were mainly guided by the research objectives, the previous literature as detailed in Chapter 2 and the study approach as shown in Figure 3.2. These variables were categorized into:

- Traffic related data (Traffic volume, operating speeds and variations)
- Pedestrian related data (pedestrian-vehicle crash data; pedestrian volume; pedestrian and roadside activities)
- Roadway geometric design and characteristics;.

However, some of the required data for modeling process were not available, particularly that for pedestrian flow and roadside activities. Therefore, the study developed a methodology to conduct a field inventory and observations to create datasets of the key parameters required to support the model development. In order develop and design the field study, a preliminary reconnaissance observation and pilot study were planned and conducted to examine the movment of pedestrian along the streets and while crossing and also the potential interactions with the traffic flow. This was an important stage in the research as the pilot study provided insights for the planning of the field study and contributed to:

- Observing and identifying the relevant factors of pedestrian and roadside activities
- Evaluating the approaches and methods for field data collection
- Identifying any potential problem areas in the research and the data requirments
- Measuring the time and resources needed for the field data collection
- Providing training and helping the resercher to be familiar with procedures and study methods

During the planing phase of the pilot study and preliminary observation, it was decieded to perform a pilot study on three from the selected sites. The road sections were considered as units of (100 metre) to provide the counts in this study and therfore it was necessarey to explore the suitable survey methods to cover the data collection over each counting station. The site visits were conducted during weekdays, when the weather allowed, during different times of the day. During the visitis, the field safety and the potential viewing positions were determined. The viewing positions were selected to be out of the way of the traffic and pedestrians, and not impeded by infrastructure.

Both a mobile observation and stationary counting techniques were considered. The locations covered different types of roadside activities with heavy pedestrian and traffic flow. The movement of pedestrian is different from other road users as they are able to move freely without being dedicated to lanes. Therfore, the preliminary site observations were designed to invistigate the movement of pedestrians along the streets or while crossing, and to check the potential interactions with traffic.

By means of both the stationary and patrolling observer methods the pedestrian and roadside activities along 100-metre section were noted and the observations identified different traffic and activities conditions and interactions that could show potential influence on pedestrian safety in such environment.

The results of the pilot study and priliminary site observations indicated a range of factors that may influnce the pedestrian safety. The frequent activities that could be considered for the data collection and model development may include pedestrian crossing violations, parking/un-parking manoeuvres and loading/unloading activities, pedestrian flow (walking along and crossing), number of bus stoppings and intersecting traffic volume.

In addition to identifying the potential key factors, the pilot study provided the experience and information to design the data collection procedures and methods. Accordingly, the mobile observation technique was considered for roadside parking while the stationary counting technique was considered for other pedestrian and roadside activities. For the stationary counting, the length of the road sections (100 metres) was found short enough to be surveyed by one surveyor. For bus stopping activities, the site visits showed that the data can be preferably collected using referencing photographs for each bus stop service timetables to count the number of bus stopping for each route. The actual circumstances encountered during the pilot study were used to refine the design of the study and data collection procedures and helped to reduce the unforeseen problems. Finally, the pilot study and the preliminary site observations offered feedback on the actual conditions which are predominant on roads similar to those of the pilot study and for the time period when the data was collected. The feedback was reflected on the final selection of study variables, position of surveys, set-up for road section survey, time of surveys and number of surveys.

As a result, the methodology used for the final variables' selection was based on comparing the factors from literature review and the obsereved factors as described above.

Consequently, a range of independent factors that may influence the potential of crash events in the model developed in this work includes:

- Traffic volume measured as annual average daily traffic (AADT),
- Traffic mean speed (V),
- Coefficient of speed variation (CV),
- Pedestrian crossing violations (PCV),
- Number of bus stoppings (BS),
- Pedestrian walking along (PAL),
- Pedestrian crossing volume (PC),
- Percentage of heavy vehicle traffic (HV),
- Intersecting traffic volume (INT),
- Number of intersecting side roads (SR),
- Roadside parking (P) and
- The number of Parking, un-parking, boarding and alight activities events (PBA).

Therefore, the generic statistical regression models sought are as follows:

$$Pedestrian\,risk = f\ (AADT, V, CV, PCV, BS, PAL, PC, HV, INT, SR, P, PBA)$$

With respect to the study objectives, the focus was more on traffic characteristics and pedestrian activities. However, roadway geometric design and attributes were also considered. The inclusion of road attributes such as road width, median arrangement, number of lanes, presence and type of pedestrian crossing facilities and sidewalk pavements was only necessary due to the difficulty of finding identical sites in terms of design to be selected for the study. Otherwise, the primary intention was to measure the potential effect of traffic characteristics and pedestrian activities on pedestrian safety; and to control, as much as possible, the effect of roadway geometric design and attributes. Therefore, the techniques and procedures followed to collect the required data in correspondence to the variables of this study are organized in sections as follows:

- Pedestrian-vehicle crash data
- Roadway segment characteristics
- Traffic volumes
- Speed data collection procedure
- Pedestrian and roadside activities' data collection procedure.

It should be appreciated that for a full scale research without the resource and time constraints of this study, the pilot study itself should be enhanced and preferably supported by a statistically defined sampling methodology. This would be a quantitative approach as opposed to the more empirically-driven qualitative method used in this research. It would potentially result in a wider coverage of conditions but this study appears to have the characteristics of empirical methods.

5.2.3 Pedestrian-vehicle crash data

Vehicle-pedestrian crash data collection was undertaken to examine their relationship with the considered variables. The iRAP and HSM recommend a period of five years to be included (iRAP, 2016; AASHTO, 2010). In this study the decision was to include an eight-year period to include a sufficient frequency of traffic crashes and the crash data for the study sites are available and well maintained; furthermore, this would reduce the crash frequency variability from year to year (Hosseinpour et al., 2014). For this time span, every pedestrian crash recorded for the study sites was considered. A pedestrian crash was defined as a crash involving at least a pedestrian and vehicle. Also, all pedestrian crash data including fatal, injury and slight crashes along the road segments were collected, including that at junctions.

As pedestrians are more fragile and the least physically protected road user (Downing et al., 2000; Eluru et al., 2008; Zhang et al., 2014), and causality injuries and fatalities are too scarce to be used for model development, the study utilized all pedestrian severity categories (i.e. fatal, serious, slight). In addition, all pedestrian crash data was then collected and extracted as map points along 100-metre sections for all road segments considered and expressed in accident/year/section for modelling purposes.

5.2.4 Segment characteristics

The first step for the segment data collection was to confirm that the selected sites meet the study site selection criteria. Then, the road attributes at each road segment were recorded to describe the physical layout. The considered sites' attributes were as follows:

- Number and width of lanes: this variable identifies the number of lanes in the road, particularly two (one each direction) or four (two each direction). However, some road sections are designed with three lanes, two in one direction and one in the other. Also, the width of lanes and road segments in total has been measured.

- Land use: this is to describe the type of land uses and buildings surrounding the study segments and is categorized into educational, commercial, residential, industrial, and other uses.
- Junctions: for each road section, the junctions have been registered in terms of their type and number. Each junction was registered with regard to the number of arms, whether it was signalized or not, and their turning traffic movement arrangement. Also, the turning traffic volume (i.e. intersecting traffic volume) for each junction was registered as a model variable. Where the turning traffic volume data were not available from existing sources, the iRAP estimate approach was used depending on the junction arms' configuration, road connections and development along the intersection road (iRAP, 2016).
- Pedestrian crossing facilities: for each road section, all pedestrian crossing facilities have been registered regardless of whether they are at a junction. Pedestrian crossing facilities included signalised, zebra crossing, un-signalised raised marked/unmarked crossing, or refuge only.
- Other road attributes: for each road section other road attributes were recorded including sidewalk pavement provision, delineation, median type, curvature and grade, road and junction condition; in addition to all other attributes as required in the iRAP system, to use their associated risk factors to calculate the star rating as explained in Chapter 4 (see Table 4.1).

5.2.5 Traffic volumes

The traffic volume data for the selected road segments was obtained from the literature (Department for Transport, 2016a).

Two variables were considered traffic volume and traffic composition. The traffic volume counts data (AADT) was firstly extracted from a database (Department for Transport, 2016a). The percentage of heavy vehicles was then calculated for each segment by dividing the number of heavy vehicles by the total traffic volume passing through.

5.2.6 Speed data collection procedure

Three speed variables were utilised: posted speed limits, operating speed and the coefficients of speed variation (CV). The posted speed limit represents the maximum legal speed allowed; while the operating speed represents the statistical average of the real speed data at any road segment. Also, this research adopted the coefficients of speed variation (CV) as an additional variable to describe the deviation of speeds from the average. This variable was estimated by dividing the standard deviation over the average of the traffic speed at any given road link. Elvik et al., (2004) highlighted that the strong correlation between speed mean and variance possibly make it difficult to separate the effects of speed variance and mean speed on crash events. To address this point, this study suggested using this factor to measure the effect of speed variability with respect to mean speed.

In order to obtain speed data, different sources were used including the existing literature and if necessary conducting field surveys for traffic speed data measurements for the selected sites.

Whenever it was applicable, speed surveys were undertaken in accordance with the technical advice note on **Vehicle Speed Measurement on All Purpose Roads DMRB Vol.5 Sec.1, TA 22/81** (DMRB, 1981).

The measurements considered the following rules:

- For each site, the choice of location to measure traffic speed needed to be level grade and straight.
- The speed measurement locations need to be well away positioned from traffic signals. The location may have been near minor intersections or driveways. No measurement of speeds of vehicles was counted when another vehicle was turning.
- Using concealed position while taking measurements to avoid attracting the attention of drivers or affect their speeds.
- For each location, measurements of 100 vehicles` speeds need to be counted to estimate the operating speed average and variability.
- The measurements excluded counts of speeds for vehicles being delayed by another vehicle.
- The selection of vehicles for measurements needs to be performed in an unbiased way, so that the selected vehicles for speed measurement are representative of the stream.

5.2.7 Pedestrians and roadside activities estimation

Data about pedestrian flows and activities could be sparse or not available despite their significant impact on safety assessment. Local knowledge and estimates could be used for model development. Such an approach may be providing less reliable results than those based on real-time observations.

Therefore, the study presented a methodology to conduct a field inventory and observations to create datasets of the key parameters required to support the model development. The data for pedestrian movement and roadside activities were collected using sampling count observations.

As a general approach, for some of the activities including parking/un-parking manoeuvres and loading/unloading activities, pedestrian volume, and pedestrian violations' data, one section of 100 metre was considered adequate to provide the counts for each road segment of similar characteristics and activities' intensities. Therefore, for each of the 12 selected study sites, at least one road section (100 metres) was taken to provide the estimation for the activities. Another section also considered if the road segments appeared to have two significantly different parts in terms of activities and land use intensities. The counting stations' number and locations for each of the selected road segments were identified during the preliminary survey.

To ensure better estimations of pedestrian and roadside activities using the sample count observations, extra attention was taken to ensure that the selected road sections were showing steady and high density of activities. The surveys were undertaken in conventional periods (not during schools and other holidays) when the weather allowed and weekend counts were avoided to minimise variations. Also, and based on the adjacent land uses, the activities were predominantly commuters and shoppers where the variation expected to be less than in roads serving leisure and tourist flows.

Because data for the pedestrian and roadside activities for all of the 100-m sections were unavailable, count observations was used to collect the needed data from sample sections manually and could not be conducted for all 100-m sections due to resources constraint.

Although this approach has some limitations, such as the pedestrian and roadside activities may vary along the different sections, it was felt it provided a sample of pedestrian activities which were acceptable estimates of the prevailing conditions and possibly the best that can be reasonably collected. Also, it was assumed that the real-time based counts may provide more representative data of the uncontrolled condition of the sites. To improve the data set used for the model development, other data collection methods such as manual or automated counts using video cameras could have been used to address the accuracy (bias) and precision (repeatability) of the data.

5.2.8 Pedestrians and roadside activities data collection procedure

For the purpose of this study, the pedestrian and roadside activities considered for modelling included parking on roadsides, bus stopping activities, parking and un-parking to boarding and alight passengers, pedestrian walking along and crossing volumes, and pedestrian violations by crossing out of allocated facilities or times.

Given the lack of pedestrian volume and roadside activities' data, the study conducted a field inventory and observations to create datasets to support the modelling process and provide the following:

- **Parking**

This factor reflects the increased risk to pedestrians associated with parked vehicles at the roadside. For the purpose of this data set, a mobile observation was considered by means of a patrolling observer moving along the selected road segments during weekdays to note the parking activities for each 100-metre section. This included street parking spaces, parked vehicles on the roadside or within 2 m of the edge of the road (iRAP, 2016). The observer then selected one of three options: no side parking, one-side parking or two-sides parking for each section.

- **Bus stoppings**

The number of bus stops has been used as a parameter to estimate pedestrian-vehicle crashes. However, as bus stops might serve different routes and bus frequencies, therefore; this study used the number of bus stopping activities per time unit and per road segment length instead of counting bus stops per road section.

To collect this data set, a preliminary study was conducted for the selected road segments. The study included the following:

- Google map navigation for each section to identify the bus stops` locations.
- Road site visit inspections and referencing photographs for each bus stop service; timetables to count the number of bus stopping for each route.
- Reviewing the websites of the bus services' calendar for each road section to count the number of bus stopping for each route.

Depending on the satellite map navigations and site visits, the number of bus stops for each 100-metre road section was located. Then the bus stopping activities per 100-metre section per hour were calculated by reviewing the referencing photographs for each bus stop, service timetables and the websites of the bus services. For this study an average number of stoppings per hour per 100 metres was used after reviewing the services' calendar during weekdays.

- **Parking and un-parking, boarding-alight activities**

The parking/un-parking manoeuvres together with the boarding/alight passengers taking place in a given road section were recorded.

For the purpose of this data set, as the length of the road sections (100 metres) was short enough to be surveyed by one surveyor, a stationary observation was considered. The observer was positioned at the selected road section where viewing was clear and not impeded by curves or infrastructure. The viewing position was selected to be out of the way of the traffic and pedestrians. Before the observation took place, a site visits were conducted to determine the field safety and potential positions with clear sight. The counts were conducted during weekdays for all road segments when the weather allowed and in 15-minute increments to take four increments to provide 1-hour counts for each road section.

The surveyor measured the number of activities per section per hour with regard to vehicles stopping to load or unload passengers and vehicles performing parking or un-parking manoeuvres. The survey set-up is shown in Figure 5.1 below:

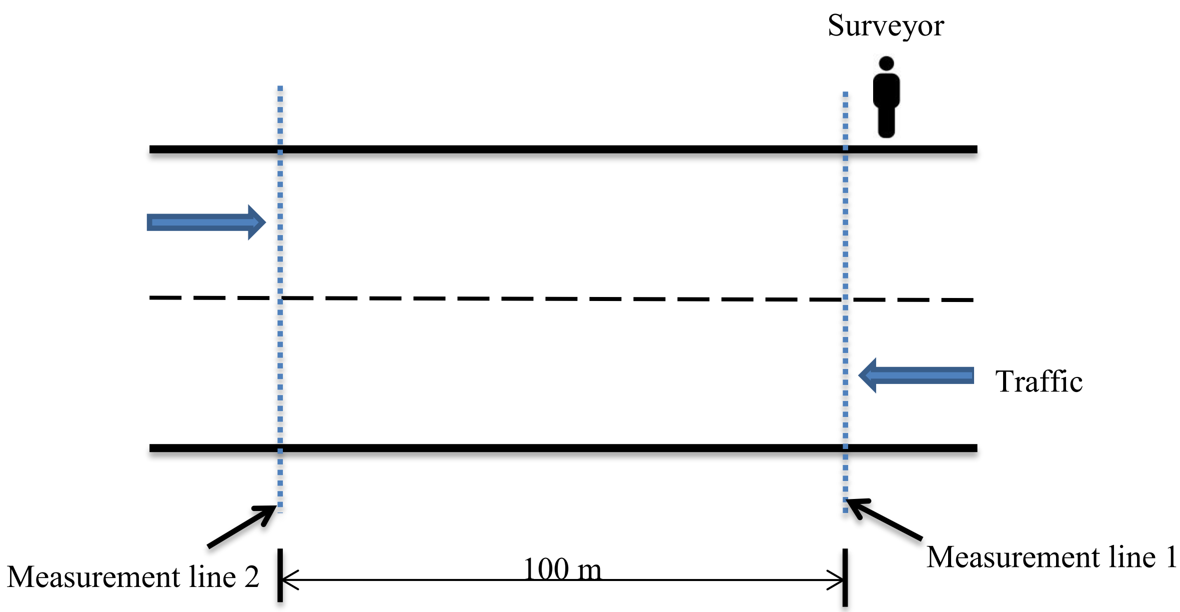


Figure 5.1 Set-up for road section survey

- **Pedestrian crossing and alongside volume**

The number of pedestrians moving along and crossing the selected road segments has been counted in-15 minute increments (Turner et al., 2006; Minge et al., 2015). Four increments were taken to provide 1-hour counts for each road section. This was done by undertaking a manual sample of actual pedestrian counts across the selected roads for assessment. The counts included pedestrians walking along both sides and pedestrians crossing the road.

This data set was gathered following the same measures taken for counting parking and un-parking activities, with the safety and position recommendations and conducted during weekdays for all road segments, when the weather allows to minimise its influence. The counts were carried out by two observers. One counted the crossing pedestrians in the selected road section; while the other counted the pedestrians walking along the road by defining a measurement line across the road and counting all walking pedestrians passing this line in both directions.

To standardise the number of pedestrians crossing the road segments, the count estimates were related to the road segment length, as a pedestrian crossing per hour per km by multiplying by 10, to convert the 100-metre segments taken for observation to km units.

- **Pedestrian violations**

The number of pedestrian crossing violations was selected as one of the study variables. This is defined as the number of pedestrians crossing the road segment freely either by not using the pedestrian crossing facilities or crossing out of the allocated time interval (i.e. crossing during the red signal). Many reasons could make pedestrians choose to cross freely and consequently encounter higher risks of traffic conflicts or crashes.

For the purpose of this data set, these counts took place in conjunction with parking and un-parking, loading-unloading activities' data collection during the same time intervals.

The same observer was positioned at the selected road section to perform the counts using the counting form shown in Figure 5.2 to record all pedestrian crossing violations in the selected 100-m section, in addition to the parking and un-parking and loading-unloading activities' data set.

For the purpose of measuring parking/un-parking manoeuvres and loading/unloading activities, pedestrian volume, and pedestrian violations' data, one section of 100 metres was considered to provide the counts for each road segment of similar characteristics and activities' intensities. In other words, for each road segment of the selected study sites (12 road segments), at least one road section (100 metres) was taken to provide the counts for the activities and pedestrian volume; another section was considered if the road segments appeared to have two significantly different parts in terms of activities and land use intensities due to time and budget constraints. The counting stations' number and locations for each of the selected road segments were identified during the preliminary survey. Although this approach has some limitations, such as the pedestrian and roadside activities may varying along the different sections, it was felt it provided a sample of pedestrian activities and volume data which were acceptable estimates of the prevailing conditions.

Road segment: Date: Start time: End time:				
Time / Events	15 minute-intervals			
Intervals	First (15 min.)	Second (15 min)	Third (15 min)	Fourth (15 min)
Parking /Loading events				
Un-parking/Unloading events				
Pedestrian violation events				

Figure 5.2 Road side activities' counting form

5.3 Summary

In this chapter the data requirements for the analysis of pedestrian safety were presented. The modelling variables were proposed and the methods and criteria for pedestrian-vehicle crashes, roadway segment characteristics, traffic volumes, speed, pedestrians and roadside activities' data collection were discussed. The field data collection was proposed to gather the required data when necessary.

Chapter 6

DATA GATHERING

Following the identification of the appropriate modelling factors and the data collection strategy, this chapter presents the data upon which the modelling and testing procedures followed in this research were based under the following headings:

1. Collision data statistics
2. Segment characteristics
3. Traffic volumes
4. Speed data collection
5. Pedestrian and roadside activities' data collection

6.1 Collision Data Statistics

6.1.1 Vehicle-pedestrian crash data collection

The pedestrian accident data set maintained and published by the Department for Transport was used. It provides the circumstances of road accidents on public roads that are reported to the police. The data set is available in different forms. The crash data can be easily accessed through maps and therefore, the application CrashMap (DfT, 2016b) has been utilized for the purpose of this study. The CrashMap offers a road crash map visualising all reported road accidents with geographical and circumstantial details. The search for accidents is facilitated by using the map or postcodes and the search can be refined to show data according to various attributes. The information provided includes the time of crash, the number of vehicles involved, the number of casualties, and also the pedestrian crash index in terms of the three severity levels:

- Fatal
- Serious
- Slight

Using the CrashMap, all pedestrian collision data was collected per 100-metre section for all road segments. For the modelling process, the collision data was calculated in terms of the annual crash rate measured in accident/year/km by dividing the total number of pedestrian crashes by the total number of years (eight) and the segment length.

The summary statistics of pedestrian-vehicle collisions are provided in Table 6.1 in terms of their frequency and severity outcome.

Table 6.1 Historical pedestrian-vehicle collision data for studied road segments

No	Name of the Road	Crash Records (2009-2016)				Crash/year/km
		Fatal	Serious	Slight	Total	
1	Coventry Rd N-East bound	0	12	18	30	4
2	Coventry Rd S-West bound	0	5	19	24	3
3	Hagley Rd-East bound	0	5	10	15	2
4	Hagley Rd West bound	2	9	12	23	3
5	Moseley Rd-S	2	8	23	33	5
6	Moseley Rd-N	1	2	4	7	1
7	Pershore Rd-N	0	6	14	20	3
8	Pershore Rd-S	0	4	14	18	3
9	Soho Rd -E	2	8	12	22	3
10	Soho Rd -W	0	12	47	59	9
11	Stratford Rd-N	0	13	24	37	4
12	Stratford Rd-S	2	13	17	32	3

6.1.2 Crash frequency distribution

All pedestrian crash data through an 8-years period were investigated for each 100-meter section of the road sites considered in this study. For modelling purposes, the crash frequency distribution per 100-meter were examined. Figure 6.1 shows the distribution of the number of crashes per100-meter section. The crash distribution is skewed to the right and does not follow normal distribution.

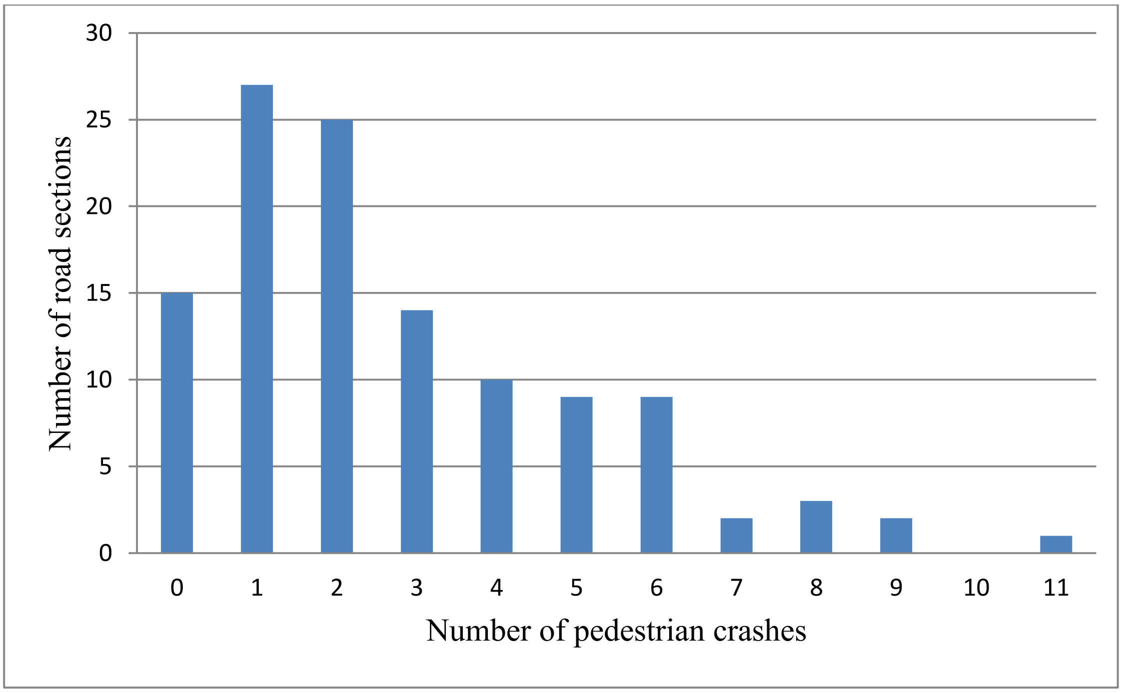


Figure 6.1 Number of pedestrian crashes per road section

As seen in the Figure 6-1, road sections with one recorded crash are the most frequent constituting 27 sections. Road sections with two crashes represent the second highest frequency followed by 15 sections with no recorded crashes.

As the selection of appropriate statistical modelling approach depends mainly on the data type, and because the crash frequency distribution does not seem to follow normal distribution, therefore; Techniques of Generalized Linear Modelling may be considered. Also, the presence of road sections with zero crash counts indicates the need to include and test modelling approaches to deal with road sections of zero crashes.

Figure 6.2 and Figure 6.3 show the distribution of crashes by severity level per 100-meter section. Both figures skewed to the right showing agreement with Figure 6.1.

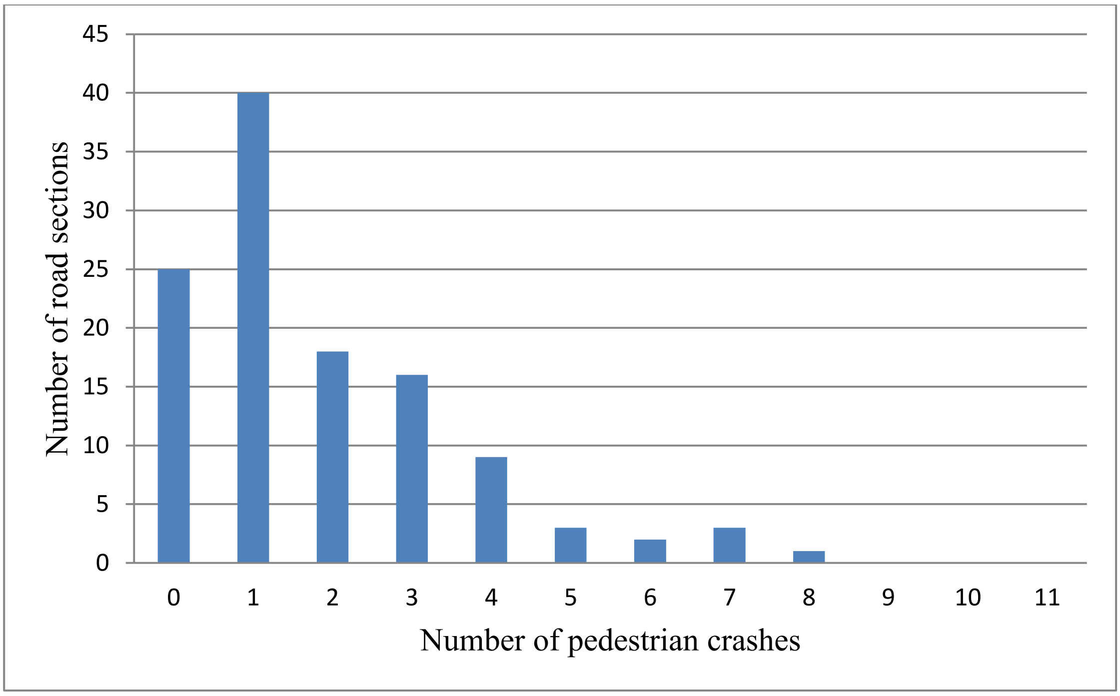


Figure 6.2 Number of slight pedestrian crashes per road

The crash distribution for each day of the week was investigated and shown in Figure 6.4. Working week days show similar trends with the highest crash events recorded on Friday, while Weekend days show lower crash dates. Overall, weekdays show higher crash events with Friday being the highest while weekend days show a lower rates where the lowest crash rates recorded on Sundays. Therefore, data collection needs to be conducted during weekdays.

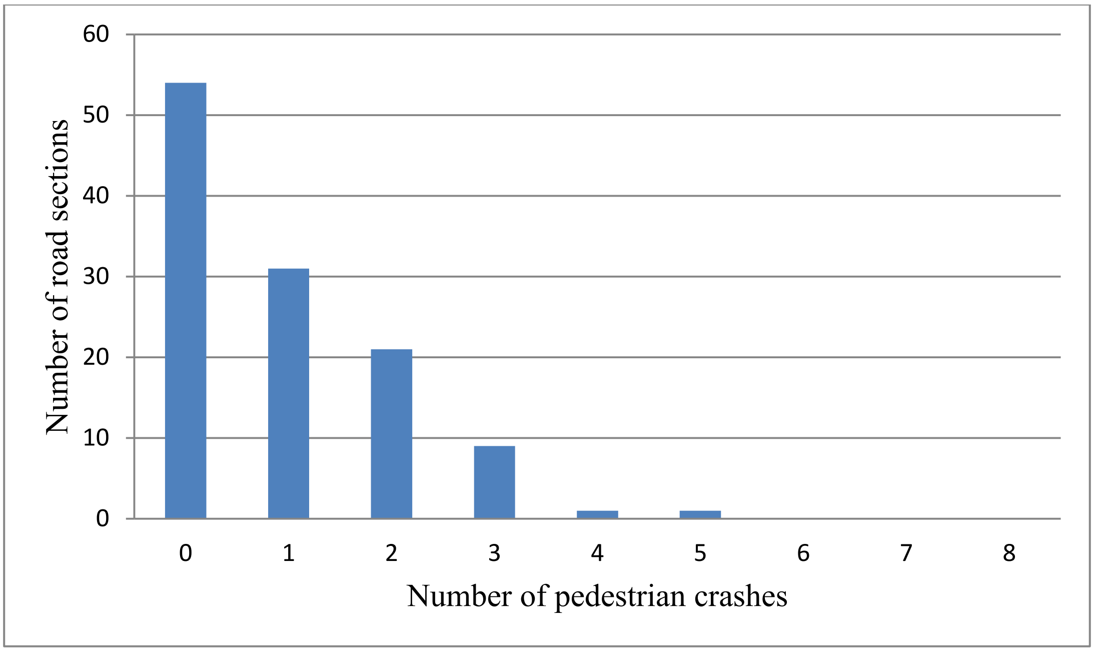


Figure 6.3 Number of serious and fatal pedestrian crashes per road section

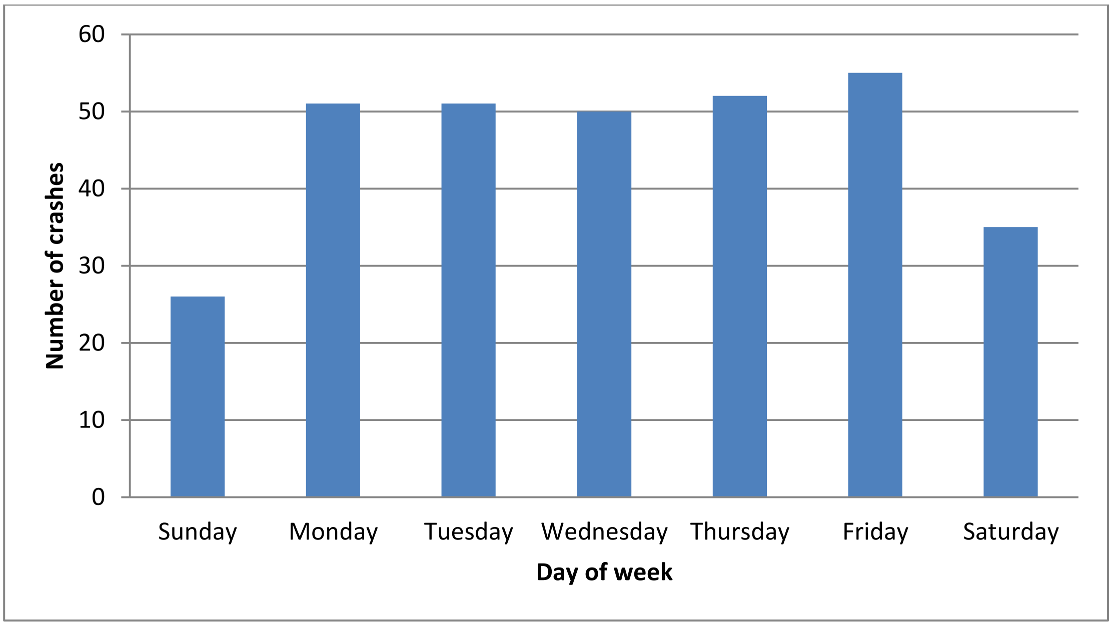


Figure 6.4 The pedestrian crashes distribution for each day of the week

6.1.3 Time of Day crash distribution

The crash frequencies by time of day are shown in Figure 6.5. Different crash rates can be seen during the hours of the day and this is may be attributed to the changes of pedestrian activities and traffic flow and other factors that are discussed in chapter two which all influence the potential of crash events. The crash rates significantly low after midnight until reaching the morning peak at 8 am. The crash rates then drop at 9 am, after which there is increase and fluctuation until 2 pm, then another peak which runs from then until 7pm, with the highest crash rate recorded at 3pm. After 7 pm the crash rates rapidly decreases. When examining all crashes distribution during the hours of the day, the highest number was recorded during the evening peak at 3 pm.

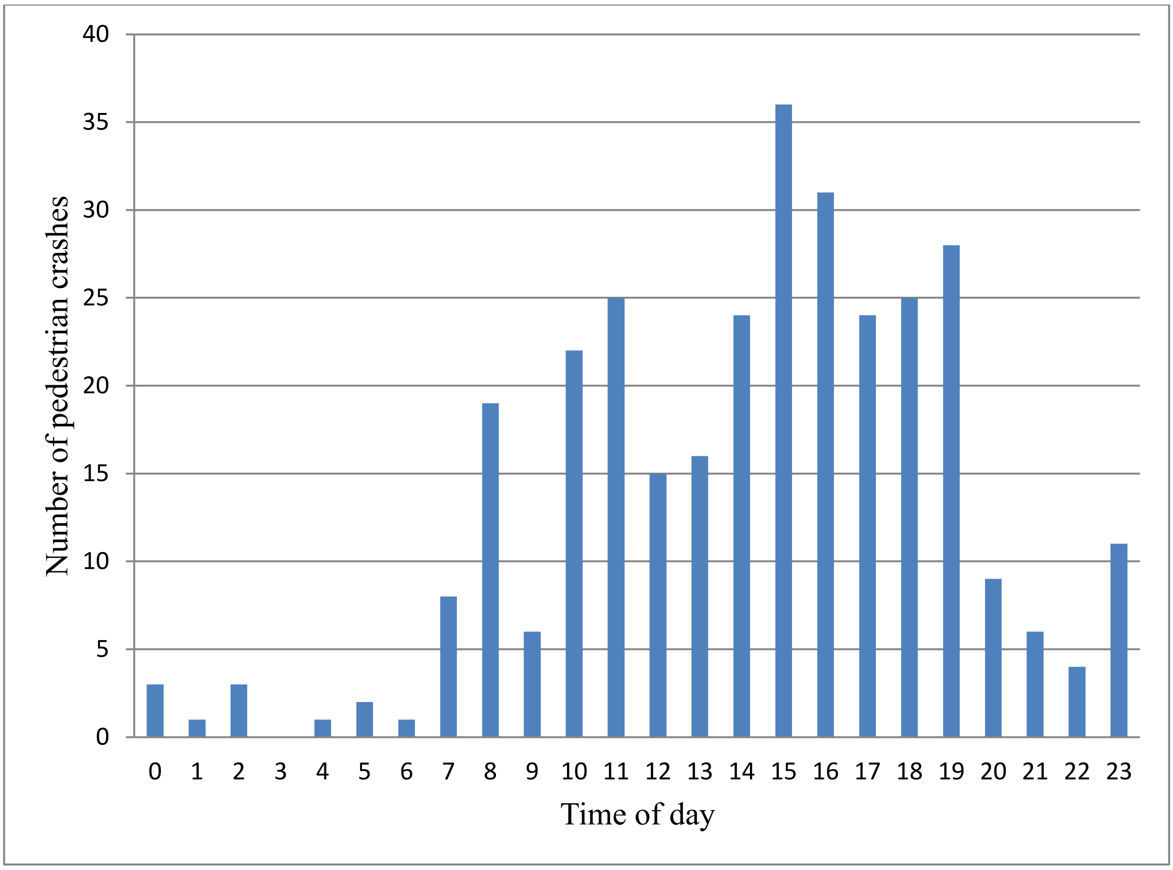


Figure 6.5 Pedestrian crashes frequencies by time of day

The crash distribution during the time of day can be used to select the time period for data collection for the study. In order to select a representative interval for data collection, the time period between 12 - 4 pm during weekdays will be initially selected to conduct data counts subject to further tests.

Table 6.2 compares between the crashes characteristics by time of day. In the selected time interval 12-4 pm, 26.3% of crashes were serious or fatal and 73.7% were slight crashes. During all the day, 27.1% of crashes were serious or fatal and 72.9% were slight crashes which are close to the proportions of the time interval 12-4 pm. In terms of gender of involved pedestrians, 58.3% of the crashes involved males during the selected interval while 41.7%involved females. When counting the gender percentages during all day, similar trends can be seen where male and female represent 60.6% and 39.4% respectively.

With regard to age groups, the highest number of crashes occurred among pedestrians age group 16-35, followed by ≤ 15, six, and 36-55 years. The percentage of crashes among pedestrians between 16 and 35 years accounting for 32.3%during the time interval 12-4 pm and 33.0% during all day. The proportions illustrated in Table 6.2 revealed similar percentages between time interval 12-4 pm and all day. Therefore, the time interval 12-4pm will be selected to conduct data collection in the selected road sites.

To enhance the clarity of the data presented, all the data shown in Figures 6.1-6.5 is also presented in Tables B.1 to B.3 in appendix B

Table 6.2 Pedestrian crashes characteristics by time of day

Variable		Time	
		All day	12:00-4:00 pm
Severity	Slight	73.7%	72.9%
	Serious and fatal	26.3%	27.1%
Gender	Male	60.6%	58.3%
	Female	39.4%	41.7%
Vehicle type	Passenger car	82.6%	76.8%
	Van /Bus	14.3%	19.5%
	Motorcycle	3.1%	3.7%
Age	≤ 15	23.0%	21.5%
	16-35	33.0%	32.3%
	36-55	22.3%	19.4%
	56-75	15.4%	17.2%
	Above 75	6.3%	9.7%

6.2 Segment Characteristics

Following the site selection step and having identified that the selected sites meet their selection criteria, the design and geometric characteristics for the selected segment were collected. For this purpose, an inventory survey was conducted along the selected road segments.

Following the inventory and site visits, a range of road attributes were recorded for each 100-metre section. An example of the road attributes collected for each road section can be found in Appendix C and E.

6.3 Traffic Volumes

The traffic volume data was obtained from the Department for Transport (DoT, 2016a). The traffic volume in AADT (annual average daily traffic) data maintained and published by the Department for Transport, provides street-level traffic counts for every junction-to-junction link on public roads for every year. These data sets are available in spread-sheet form and as an interactive map. According to the Department for Transport, all presented data is taken from annual traffic counts where each site is counted on one day for a twelve-hour period. For this study, the traffic volumes (AADT) for 2016 were utilized as it was the most recently available data.

Using the interactive map tool, the selected road segments were inspected to check for all available count points along each of them. The interactive map tool offers a geographical presentation of traffic count points and can be easily used to refine the selected area. The information gathered included the traffic counts for the period 2000-2016; road segment length covered by count point; and detailed traffic counts according to vehicle categories. Two variables were obtained: traffic volumes and traffic composition. The traffic volume counts data was firstly extracted directly from the interactive map tool and its associated spread-sheets; while the percentage of heavy vehicles was calculated for each count point by dividing the number of heavy vehicles by the total traffic volume passing through.

For each of the chosen road segments between 1 and 4 counting stations were found. The traffic volume for each 100-metre section was then obtained from this information; if the considered section lay between two counting stations, the average traffic volume was considered for the analysis.

Table 6.3 shows the traffic volume counts and heavy vehicle counts and their percentages. Based on the results, the AADT flow ranges between 18223 and 36851 vehicles per day; while the percentage of heavy vehicle ranges between 1.77% and 2.93% for the study sites.

Table 6.3 Traffic volumes' statistics

No	Name of the Road	AADT	Heavy Vehicle	Heavy Vehicle Percentage (HV)
1	Coventry Rd N-East bound	18223	323	1.77%
2	Coventry Rd S-West bound	18223	323	1.77%
3	Hagley Rd -East bound	36622	557	1.52%
4	Hagley Rd West bound	36851	568	1.54%
5	Moseley Rd -S	14506	313	2.16%
6	Moseley Rd-N	15426	299	1.94%
7	Pershore Rd-N	18754	549	2.93%
8	Pershore Rd-S	19183	409	2.13%
9	Soho Rd -E	16629	414	2.49%
10	Soho Rd -W	17178	419	2.44%
11	Stratford Rd-N	18107	284	1.57%
12	Stratford Rd-S	19247	287	1.49%

6.4 Speed Data Collection

For this study, the data for three speed variables were collected: posted speed limits, operating speed and the coefficients of speed variation (CV).

Posted speed limits' data were collected from the field survey; while operating speeds and the coefficient of speed variation were calculated using traffic speed data measurements for the selected sites.

Traffic speed data were obtained from the Department for Transport and the Birmingham City Council Transport Unit (Birmingham city council, 2016). The speed data for Birmingham city were provided for the years 2015-2016. The mapping files include almost every road in Birmingham city broken into links. The road segments finish at every junction and different speed measurements were provided for each link. The data was provided in the form of travel times for each link and then the traveling mean speed for any given link was calculated by dividing the travel times by the length of the specified link. The mean speed will be used in this study as the 85th percentile data were unavailable to calculate the operating speed. Also, the standard deviation of travel times was provided and used to calculate the coefficient of speed variation. Using the mapping files, the study sites can be detected and sectioned into 100-metre sections. The speed variables may then be calculated for every 100-metre road section. Some sections lie in two or more speed readings as the 100 metre road sections considered in this study contain junctions. In this case, a weighted average for the speed variables has been calculated according to the share of each speed reading's link length. Also for most of the links, the speed data was provided for each direction and for this case the directional average was considered. Table 6.4 shows the speed variables for the selected road segments.

Table 6.4 Speed variables for the selected road segments (mph)

No	Name of the Road	Speed Limit	Mean Speed			Coefficient of Speed Variation (CV)		
			Min.	Max	Mean	Min.	Max.	Mean
1	Coventry Rd N-East bound	20	14.4	23.5	17.6	0.57	1.11	0.80
2	Coventry Rd S-West bound	20/30*	11.6	25.9	19.0	0.63	1.11	0.80
3	Hagley Rd-East bound	30	21.0	33.9	26.6	0.86	1.18	1.10
4	Hagley Rd-West bound	30	33.9	45.0	40.3	1.18	1.98	1.60
5	Moseley Rd-S	20/30*	13.2	23.9	18.3	0.76	1.46	1.10
6	Moseley Rd-N	20/30*	27.3	37.7	31.4	0.78	1.32	1.10
7	Pershore Rd-N	20/30*	21.7	30.0	25.9	0.86	1.63	1.00
8	Pershore Rd-S	20/30*	15.8	32.1	25.9	0.72	1.49	1.20
9	Soho Rd -E	30	16.7	41.8	24.4	0.86	1.57	1.10
10	Soho Rd -W	30	12.6	17.2	14.4	0.83	1.21	1.10
11	Stratford Rd-N	30	11.9	37.1	21.1	1.05	1.87	1.00
12	Stratford Rd-S	30	19.6	47.2	35.8	0.56	2.30	1.40

* Sites with variable speed limits of 20 and 30 mph along the road

6.5 Pedestrian and Roadside Activities Data Collection

Data were collected about pedestrian and roadside activities, including; parking on roadsides; bus stopping activities; parking and un-parking to load and unload passengers; pedestrian walking along and crossing volumes; and pedestrian violations, which refers to pedestrians crossing out of allocated facilities or times.

The pedestrian volume and roadside activities for the study area were gathered by conducting a field inventory and observations along the selected sites. The data collection activities were conducted for the chosen sites. All counts were conducted during weekdays in October 2016 between 12-4 p.m. for all road segments when the weather allowed. Before actual field data collections, a one-hour pilot survey was conducted in addition to site visits to check and select the data collection positions. Table 6.5 shows the statistics of roadside and pedestrian activities for the selected road segments.

6.6 Summary

In this chapter the collected data statistics were presented. For the analysis of pedestrian safety, the required data were collected per 100-metre section for all considered sites. The presented data included collision data statistics, design and geometric characteristics, traffic volumes and composition, posted speed limits, operating speed, the coefficients of speed variation, pedestrian and roadside activities.

Table 6.5 Roadside activities and pedestrian statistics for the selected road segments

No	Name of the Road	Side Parking	Bus** Stoppings	Parking and Un-parking*	Pedestrian Crossing*	Pedestrian Alongside*	Pedestrian Violations *
1	Coventry Rd N-East bound	No/one/ two	0-30	402	814	1118	173
2	Coventry Rd S-West bound	No/one/ two	0-38	206	396	587	127
3	Hagley Rd - East bound	No side parking	0-52	93	116	186	66
4	Hagley Rd- West bound	No side parking	0-54	93	116	186	66
5	Moseley Rd-S	No/one side	0-26	340	635	1532	186
6	Moseley Rd-N	No/one side	0-42	162	386	937	139
7	Pershore Rd-N	No/one/t wo	0-38	195	204	679	69
8	Pershore Rd-S	No/one/ side	0-54	177	183	608	64
9	Soho Rd-E	No/one side	0-28	265	327	750	204
10	Soho Rd-W	No/one/t wo	0-33	402	440	1092	257
11	Stratford Rd-N	No/one/ two	0-45	450	222	648	204
12	Stratford Rd-S	No/one side	0-30	81	143	290	64

* Average

** Min-Max

Chapter 7

SYSTEMATIC MODEL SELECTION

Following the selection of the factors affecting the pedestrians' safety and the presentation of the data that facilitate the model development, this chapter gives the development of the crash *risk* and crash *severity* models. Firstly, this chapter presents the approach to measure the degree of correlation between the independent variables (that is the factors of pedestrians' safety). Secondly, it presents the selected statistical regression methods for the modelling of the *crash risk*; and the measures used in evaluating, comparing and selecting the modelling method. Finally, the chapter gives the selected statistical regression methods for modelling the *crash severity*.

7.1 Modelling of the New Aspects

To investigate the significance of the contributory factors to pedestrian safety, a statistical approach was used to measure the variations of crash risk intensities across the selected road sections. The first step was to measure the degree of correlation between the independent variables. Then, different generalised regression models (GLM) were considered to estimate the link between the risk of pedestrian crashes at the study sites and the studied contributory factors; while a probability model was utilised to measure the relationships between the potential contributory factors and the severity of the pedestrian crash outcomes.

The developed model was used then to examine the potential contribution of road design, road infrastructure elements, traffic and pedestrian volume and activities to the likelihood of a crash occurrence or the severity level of a crash. Figure 7.1 summarises the model estimation and selection process.

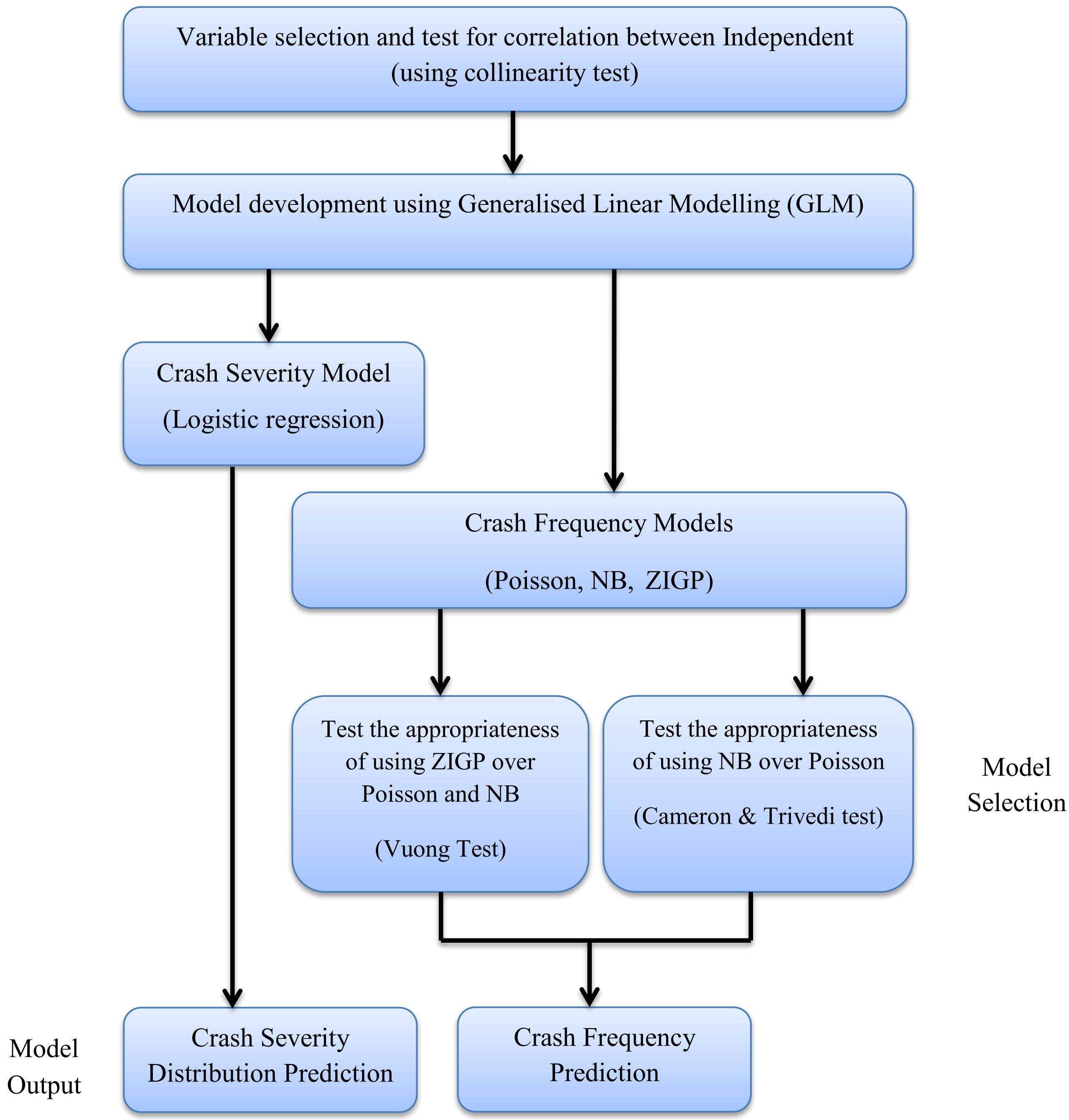


Figure 7.1 Models' estimation and selection process

7.2 Correlation between Independent Variables

The selected independent variables for modelling were checked in terms of collinearity before being included in the study. Collinearity phenomena among independent variables are often encountered with observational data. This happens when one independent variable or more is strongly correlated with others. The collinearity and multi-collinearity cause high standard errors of the estimated parameters of the collinear variables. Also, collinearity can increase the variability of the estimated dependent variable, causing unstable estimates where, in some cases, minor changes in the data produce wide changes in the parameters' estimates (Washington et al., 2011). For this purpose, this study utilised the variance inflation factor (VIF) as an indicator of the degree of collinearity or multi-collinearity of the independent variable with the other independent variables. The variance inflation factor (VIF) examines the degree of collinearity between the independent variables using the correlation coefficient R^2 (i.e. the variance in one variable that's explained by independent variable) to indicate the proportion of variance in the specified independent variable that is associated with the other in the regression model. The tolerance for any an independent variable is:

$$Tolerance = (1 - R^2) \qquad 7.1$$

The tolerance represents the proportion of variance between two independent variables. The variance inflation factor (VIF) is the reciprocal of tolerance:

$$VIF = 1/Tolerance \qquad 7.2$$

Collinearity is considered high at a VIF of 5 or 10, which correspond to tolerance values of 0.2 or 0.1 respectively (O'Brien, 2007). In case any variables show serious collinearity, one or more remedial measures need to be considered; such as eliminate or combine the independent variable that is highly correlated with the other.

7.3 Correlation between the Independent Variables and Final Selection Procedure

In order to measure the degree of correlation between the independent variables which were considered for the modelling process, i.e. traffic volume (AADT), traffic operating speed (V), coefficient of speed variance (CV), number of pedestrian crossing violations (PCV), the number of parking and un-parking and boarding-alight activities (PBA), number of pedestrian crossings (PC), number of pedestrians walking along the road (PL), number of bus stoppings (BS), percent of heavy vehicle traffic (HV), intersecting traffic volume (INT) and roadside parking (P), collinearity tests were used, using the *R2* and the variation inflation factor (VIF) (see section 7.2, Eq. (7-1) and (7-2)) as an indicator of collinearity between independent variables. Table 7.1 shows the correlation (R^2 and VIF) between all the considered independent variables to be used in the model. As explained in section 7.2, if $R^2 \geq 0.80$ (VIF ≥ 5), one of the variables should be omitted or a further remedial measure should be considered, such as combining the correlated variables.

The results shows that all VIF values are below 5, which indicate that no variables need to be excluded and no further remedial measures are needed. The highest R^2 in Table 7.1 occurred between the number of parking and un-parking and boarding-alight activities (PBA) and pedestrian violation crossings (2.922); however, the value of the VIF did not exceed the indicated threshold.

Table 7.1 Correlation between independent variables

Variable	Traffic mean speed (V)		Coefficient of Speed (CV)		PBA		Peds. crossing violations (PCV)		No. of bus stoppings (BS)		Pedestrian along Volume (PAL)		AADT		Intersecting AADT		Peds. Crossing volume (PC)		HOV%	
	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF	R^2	VIF
Traffic mean speed (V)	1																			
Coefficient of Speed (CV)	0.318	1.466	1																	
PBA	0.444	1.797	0.125	1.142	1															
PCV	0.364	1.571	0.091	1.100	0.658	2.922	1													
BS	0.000	1.000	0.000	1.000	0.002	1.002	0.000	1.000	1											
PAL	0.232	1.303	0.028	1.029	0.399	1.665	0.206	1.260	0.000	1.000	1									
AADT	0.166	1.200	0.077	1.084	0.252	1.337	0.220	1.282	0.005	1.005	0.376	1.602	1							
Intersecting AADT	0.088	1.097	0.044	1.046	0.148	1.174	0.110	1.123	0.008	1.008	0.112	1.126	0.042	1.044	1					
PC	0.277	1.383	0.094	1.104	0.256	1.344	0.452	1.823	0.002	1.002	0.267	1.365	0.161	1.192	0.236	1.309	1			
HOV%	0.095	1.106	0.011	1.011	0.057	1.060	0.057	1.060	0.000	1.000	0.159	1.189	0.189	1.233	0.001	1.001	0.03	1.031	1	
Side roads (SD)	0.068	1.073	0.030	1.031	0.082	1.089	0.100	1.112	0.005	1.005	0.040	1.042	0.008	1.008	0.237	1.311	0.06	1.063	0.002	1.0

7.4 Road Sections Selected

Due to time and budget constraints, the present study considered 12 road segments of lengths 0.8-1.2 km. For the model development, these road segments were further segmented into 100-metre sections because the road attributes may change significantly over urban areas. It was felt that the 100-meter section could be considered as the unit of the analysis and the basis for data collection and subsequently model development. However, this hypothesis required further investigation and statistical validation.

7.4.1 Differences between road sections

To establish that the 100-meter sections represented different road conditions albeit being adjacent within each site, a preliminary reconnaissance survey was carried out. This showed a wide range of:

1. Commercial, educational and residential land uses,
2. Pedestrian and roadside activities,
3. Traffic flow conditions and percentages of heavy vehicle, bus stops, on-street parking, and parking manoeuvres.

Thereafter, a further qualitative investigation was conducted to check that each of the 100-m sections within each individual site showed variations in the sections` attributes.

The investigation revealed that adjacent 100-meter sections within each road segment are characterised by differences as follows:

- Changes in the roadway configuration especially with regard to pedestrian crossing facilities and number of intersecting side roads (see Appendix C).

- Different crash rates (see Chapter 5 and Appendix C).
- Significant variations in speed variables (i.e. traffic mean speed and the speed variance, see Table 6.4 and Appendix C).
- The variability of active bus stops and the number of bus stopping activities between sections (see Appendix C).

In addition, the presence of pedestrian crossing facilities on a 100-meter section may influence the number of the crossing points at the adjacent sections within each road segment. However, in this research the variable of pedestrian crossing violations counts concerned the crossing violations activities over the length of the entire 100-m section rather than in relation to specific crossing points. This is because it was observed that pedestrian crossing violations including those who cross from non-specified crossings or during the red interval occurred at any point along the 100-m section regardless of whether there were pedestrian crossing facilities. Arguably, this approach may have required further investigation, perhaps in conjunction with enforcement regime along the road sections.

7.4.2 Statistical tests

Following the above preliminary investigation it was felt necessary to establish if the 100-meter sections within each selected site showed statistically significant variation and therefore could be used as individual sections to provide inputs for the model development despite being adjacent to each other in certain cases. To this end, the analysis of variance (ANOVA) was identified as an appropriate test to examine the above hypothesis and subsequently conducted as it has the advantage that can be applied efficiently where the observations are classified by two or more variables as in the research reported herein (Chatfield, 1991; Washington et al., 2010). Therefore, three variables were examined including the crash accidents rates, traffic speed and bus stoppings activities. The crash rate was selected because it represents the main dependent factor for the model development. The other two variables (i.e. traffic speed and bus stoppings activities) were selected because data was available for them.

It could be argued that a comprehensive application of the above statistical test could have potentially led to different conclusions, if data for other variables considered in the model development would have been available. However, given the constraint of data unavailability and those of the resources of this PhD programme in terms of time and personnel, the statistical results were assumed:

a. Sufficient at least from the view point that they were based on a robust methodology suitable for the task in hand.
b. Acceptable from a practicable point of view and
c. Applicable to the conditions likely to occur in the Birmingham area, as a minimum.

The SPSS tool was utilised to conduct the ANOVA test (IBM, 2017). In this test, the total variation of the observations is partitioned into two components, one measuring the variability between the group means, and the other measuring the variation within each group. These two components are compared by means of the F-test (Washington et al., 2010) which takes the form below:

$$F = \frac{\text{Variation among groups}}{\text{Variation within groups}} \qquad 7.3$$

The test is a method of testing the null hypothesis that there are no significant differences between the group of tested subjects and its result gives information about the significance p-value for F-test. In the case of a statistically significant F test, the hypothesis is rejected and the conclusion is that there is evidence of variation within the group of road sections.

The confidence interval is shown as a percentage, and the selected level for the test to calculate the coefficient probability p-value is 0.05 where:

$$\mathrm{p-value} = 1 - 0.95 = 0.05 \text{ (95\% confidence)}$$

The total variation of the study road sample was tested at two levels s as follows:

- **Variability between the 100-meter sections at each site**

Three tests were undertaken as follows:

1. The first test examined the crashes over an 8-year period for all 100 m sections within each site.
2. The second test examined the mean speed and its standard deviation in each site.
3. The third test examined the variability of the bus stoppings activities. For this test, the bus stopping activities per 100 m section per hour were calculated following the same procedure as in Chapter 5.

For this step the test took the form below:

$$\mathrm{F} = \frac{\text{comparing between } 100 - \text{meter sections in one site}}{\text{comparing variations in one } 100 - \text{meter section}} \qquad 7.4$$

Specifically, the null hypothesis for this test assumes that that there are no significant differences between the group of 100 m road sections and its results give information about the p-value for F-test.

Table 7.2 shows these results for each group of 100 m sections. The results show that the p-values are less than 0.05 (95% confidence) for all sites, indicating a significant difference between the groups of 100 m sections at each site. In other words it is statistically proven that each 100 m section represents different road conditions within each road site.

Table 7.2 The analysis of variance (ANOVA) test for road segments.

No	Name of the Road	Number of 100-meter sections	p-value for F Statistic		
			Crash rates	Bus stoppings activities	Traffic mean speed
1	Hagley Rd-East bound	10	< 0.001	< 0.001	< 0.001
2	Hagley Rd-West bound	10	0.001	< 0.001	< 0.001
3	Moseley Rd-S	8	< 0.001	< 0.001	< 0.001
4	Moseley Rd-N	8	< 0.001	< 0.001	< 0.001
5	Soho Rd -E	8	0.003	< 0.001	< 0.001
6	Soho Rd -W	8	0.045	< 0.001	< 0.001
7	Coventry Rd S-West bound	10	0.029	< 0.001	< 0.001
8	Coventry Rd N-East bound	9	< 0.001	< 0.001	< 0.001
9	Stratford Rd-S	16	< 0.001	< 0.001	< 0.001
10	Stratford Rd-N	12	< 0.001	< 0.001	< 0.001
11	Pershore Rd-N	9	0.049	< 0.001	< 0.001
12	Pershore Rd-S	9	0.005	< 0.001	< 0.001

- **Variability between the 12 sites**

Following the same statistical procedure as above, the differences among the 12 sites were examined. The tests examined and compared the crash rates, the speed variables and the bus stopping over the selected 12 sites. The null hypothesis in this case assumed that the 12 selected sites are similar and there are no significant differences between them. If, the test returned a statistically significant result, this assumption would be rejected indicating significant differences between the sites.

The test took the form below:

$$F = \frac{\text{comparing between sites (12 roads)}}{\text{comparing between 100} - \text{meter sections in one site}} \qquad 7.5$$

The test indicated statistically significant results with p-value of less than 0.001 for crash rates and speed variables and a p-value of 0.004 for bus stoppings activities, all of which are significantly less than 0.05 indicating significant differences between different sites.

To this end, it may be concluded the above results provide strong statistical evidence to establish that the 100-meter sections within each site represent different conditions. Therefore, it can be argued that the sample size is representative of the conditions likely to occur in Birmingham, and potentially in similar large cities in the UK.

However, the developed models should still be considered as proof-of-concept seeking to establish the relationship between potential risk variables and pedestrians crash rates and to evaluate the effectiveness of the introduced variables to predict the safety of pedestrians. As future work, more data could be collected from more sites or cities with similar characteristics to enable a more accurate and representative model.

7.5 Accident Modelling

As stated in Chapter 2, statistical regression methods (GLM) were widely used in previous studies for modelling of road crash events and therefore such methods have been used in this study because of their appropriateness to deal with count data and their ability to model simultaneous effects of multiple variables, including mixtures of continuous, count and categorical data (Myers et al., 2002; Lord and Mannering, 2010). Choosing the appropriate statistical modelling approach depends mainly on the data types. For continuous data, a standard linear regression model may be used; while for count data other regression models are recommended. Count data consist of non-negative integers and refers to the number of times an event occurs (Cameron and Trivedi, 1998).

Count data is encountered frequently in transportation and examples include the number of vehicles in a queue, the number of drivers` route changes per week and the number of accidents observed on road segments per time.

Using standard least squares regression to model count data is considered inappropriate. This is due to the fact that such models can predict negative values and also yield predicted values that are non-integers, which are both considered inconsistent with count variables (Washington et al., 2003).

Generalised Linear Modelling (GLM) is a flexible extension alternative applied to model simultaneous effects of multiple variables, including mixtures of continuous, count and categorical data (Myers et al., 2002).

There are a number of methods used to model count data, as seen in Chapter 2. The most common methods of these are Poisson, Negative binomial and Zero-inflated Poisson regression. All three methods were considered in this study and are described below.

7.6 Crash Frequency and Severity Models

Crash frequency models are mathematical functions formulated to predict the number of accidents as a function of other independent variables. Techniques of Generalized Linear Modelling were frequently used by previous studies due to the fact that some of the independent variables are count data (Myers et al., 2002; Lord and Mannering, 2010).

Generalized linear models, were first formulated by Nelder and Wedderburn (1972) and explained comprehensively by McCullagh and Nelder (1989). The GLM is a flexible extension of linear regression that allows for response variables with other than normal distribution and non-linearity to be considered in the model form. Also, GLM generalises linear regression by allowing the magnitude of the variance of each measurement to be a function of its predicted value and has the basic form:

$$P(y_i) = \mathrm{g}^{-1} X_i \beta \qquad 7.6$$

Where $P(y_i)$ is the expected value of Y; g is the link function; Xi is the i^{th} predictor of the model; and β is a vector of unknown parameters. Also, GLM usually assumes that the dependant variable follows one of the exponential family distributions.

Examples of the exponential distributions include Poisson, Normal and Gamma distributions.

The unknown parameters β, which counts for the independent variable coefficients, are typically estimated with the maximum likelihood techniques (Wood, 2006).

In the linear regression analysis, the method of the least squares is typically used to estimate the model coefficients. This is done by minimising the average squared differences predicted and observed values without counting for the form of the distribution of the errors. This is the best fit for the normal distribution or if the distribution of the errors is unknown. However, for other distribution forms such as Poisson, Binomial and logistic regressions, the method of the maximum likelihood is utilised to estimate the model coefficients. The procedure of estimating coefficients of the unknown parameters for statistical models in this method selects the values of the parameters that maximize the model likelihood function.

The traffic accidents are likely to follow Poisson distribution as they are rare events. Therefore, Poisson models were widely used in crash modelling (Geyer et al., 2006; Ye et al., 2013). However, Poisson models cannot deal with the over-dispersion, which is frequently observed in road crash data (Cameron and Trivedi, 1990). The Negative Binomial model is considered capable of handling this phenomenon (Poch and Mannering, 1996). Other studies suggested using the Zero-Inflated Generalized Poisson model because crash data often includes an unusually high proportion of zero counts in road segments (Chin and Quddus, 2003).

For modelling crash frequency, three alternatives will be tested including Poisson, Negative Binomial (NB) and Zero-Inflated Generalized Poisson (ZIGP) models.

7.6.1 Poisson model

Among the three GLM modelling methods (Poisson, NB and ZIGP), pedestrian crash data will be modelled by assuming a Poisson distribution. Poisson regression is considered a useful tool to start with because of its:

- relative ease of computation and its popularity in accidents modelling studies
- flexibility to model count data supported by discrete, non-negative integer-distribution characteristics
- flexibility to model more distributional forms.

Using Poisson regression, the probability of road section *i* having y_i crashes per year is given by:

$$P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{yi}}{y_i!} \qquad 7.7$$

Where the probability of crashes occurring in road section i is denoted by $P(yi)$; λ_i^y is the expected number of crashes on road section *i* in the time unit. To estimate λ_i^y, usually the log-linear relationship is used:

$$\lambda_i = exp(\beta' X_i) \qquad 7.8$$

Where X_i is a vector of explanatory variables;

β is a vector of coefficients to be estimated from the data.

This model was estimated then using maximum likelihood methods with the function given as:

$$L(\beta) = \prod_i . \frac{exp(-\exp(\beta' X_i))\, exp(\beta' X_i)^{yi}}{y_i!} \qquad 7.9$$

7.6.2 Negative Binomial model

The Poisson regression model assumes the mean and variance to be equal. However, over-dispersion could be encountered if the variance is greater than the mean (Washington et al., 2003; Lord et al., 2005). This commonly occurs in the analyses of traffic accidents' modelling since the crash events are rare and can significantly vary between road sections. Therefore, the use of the Poisson modelling method in such a case results in biased estimates of variables' coefficients (Fu, 2015).

Alternatively, a Negative Binomial model may be considered as a suitable option to model crash frequency due its ability to deal with over-dispersion (Washington et al., 2003; Poch and Mannering, 1996). It is another tool of the generalized linear models (GLMs) family and is widely used in crash modelling. The (NB) is an extension of the Poisson regression model and accounts for over-dispersion by adding an error term ε_i to the expected number of crashes λ_i^y such that equation $\lambda_i = exp(\beta' X_i)$ becomes for each observation i,

$$\lambda_i = exp(\beta'^{X_i} + \varepsilon_i) \quad 7.10$$

This model is estimated then using the maximum likelihood method by the given function:

$$P(y_i) = \frac{\Gamma\left(\left(1/\alpha\right) + y_i\right)}{\Gamma\left(1/\alpha\right) y_i!} \left(\frac{1/\alpha}{\left(1/\alpha + \lambda_i\right)}\right)^{1/\alpha} \left(\frac{\lambda_i}{\left(1/\alpha + \lambda_i\right)}\right)^{y_i}, \quad 7.11$$

Where Γ is a gamma function. The likelihood function:

$$L(\lambda_i) = \prod_i . \frac{\Gamma\left(\left(1/\alpha\right) + y_i\right)}{\Gamma\left(1/\alpha\right) y_i!} \left(\frac{1/\alpha}{\left(1/\alpha + \lambda_i\right)}\right)^{1/\alpha} \left(\frac{\lambda_i}{\left(1/\alpha + \lambda_i\right)}\right)^{y_i} \quad 7.12$$

7.6.3 Zero-Inflated Generalized Poisson model

In fact, another limitation may be encountered if crash data is under-dispersed, where the mean of crash events is greater than the variance. In modelling of crash events for a road section, some sections record zero accidents, however the likelihood of accidents occurring is still present. In the case of excess zero counts being present in data, the Zero-Inflated Generalized Poisson (ZIGP) model could be an appropriate option (Shankar et al. 1997; Washington et al., 2003).

The ZIGP model is another extension of the generalized linear models (GLMs) family and to address the issue of excess zero counts, a separate model that counts for excess zeros may be used. The ZIGP model has two parts; Poisson regression and logit regression model to predict zeros independently, assuming:

$$y_i = 0 \; with \; probability \;\; p_i + (1 - P_i)EXP(-\lambda_i) \qquad 7.13$$

$$y_i = y \; with \; probability \;\; \frac{(1 - P_i)EXP(-\lambda_i)\lambda_i^y}{y!} \qquad 7.14$$

This model is estimated then using the maximum likelihood methods with the function given as:

$$y_i = 0 \; with \; probability \;\; p_i + (1 - P_i)\left[\frac{\frac{1}{\alpha}}{\left(\frac{1}{\alpha}\right) + \lambda_i}\right]^{1/\alpha} \qquad 7.15$$

$$y_i = y \; with \; probability(1 - P_i)\left[\frac{\Gamma\left(\left(\frac{1}{\alpha}\right) + y\right) u_i^{1/\alpha}(1 - u_i)^y}{\Gamma\left(\frac{1}{\alpha}\right) y!}\right], y = 1,2,3, \dots \qquad 7.16$$

7.6.4 Model selection

The selection among the prescribed three models was made based on two steps,

Step 1: test the appropriateness of using Negative Binomial over the Poisson model

To make a selection decision between the Poisson and the NB models, the test of over-dispersion will be used (Cameron and Trivedi, 1990).

The process tests the assumption of the mean is equal to the variance under the Poisson distribution($(y - E[y_i])^2 - E[y_i]$. Thus, the null hypothesis,

$$H_0: var(y_i) = E(y_i) \qquad 7.17$$

The alternative hypothesis is:

$$H_A: var(y_i) = E(y_i) + \propto g(E(y_i)) \qquad 7.18$$

Where $g(E(y_i))$ is a function of expected counts. To perform this test, a linear regression model is estimated where Z_i is dependent on W_i given:

$$Z_i = \frac{\left(y_i - E(y_i)\right)^2 - y_i}{E(y_i)\sqrt{2}} \qquad 7.18$$

$$W_i = \frac{\mathrm{g}\left(E(y_i)\right)}{\sqrt{2}}. \qquad 7.20$$

After running the regression($Zi = bWi$), if b is significant in any of the case, then the null hypothesis H_0 is rejected in favour of function g.

Step 2: test the appropriateness of using ZIGP over the NB and Poisson models

To test the appropriateness of using ZIGP over the previous models, the test proposed by Vuong (1989) is utilised. The Vuong test compares the zero-inflated model with an ordinary Poisson and NB regression model. The test static calculated for each observation i as:

$$m_i = LN\left(\frac{f_1(y_i|X_i)}{f_2(y_i|X_i)}\right), \qquad 7.21$$

Where $f_1(y_i|\boldsymbol{X_i})$ is the probability function of model 1; and $f_2(y_i|\boldsymbol{X_i})$ is the probability function of model 2.

Using the above, Vuong's statistic to test model 1 against model 2 is:

$$V = \frac{\sqrt{n}\left[\left(\frac{1}{n}\right)\sum_{i=1}^{n} m_i\right]}{\sqrt{\left(\frac{1}{n}\right)\sum_{i=1}^{n}(m_i - \overline{m})^2}} = \frac{\sqrt{n}(\overline{m})}{S_m}, \qquad 7.22$$

Where $\overline{m}$ is the mean; *n* is the sample size; and S_m is the standard deviation. Vuong's value (V) is considered asymptotically to follow normal distribution, and the absolute value $|V|$is lower than $V_{critical}$ (1.96 for a 95% confidence level); the test does not support the selection process. Large positive V greater than $V_{critical}$ supports the section of the model 1 over model 2; whereas large negative values favour model 2.

7.6.5 Crash severity model

Besides pedestrian crash frequency modelling, the study attempts to test the effect of the selected variables on the crash severity level in terms of two main levels: slight crash and fatal or serious injury crash. Thus, the severity degree is a binary outcome. Logistic regression is considered an appropriate approach to predict a binary dependent outcome as a function of explanatory variables. The objective of logistic regression, similar to other regression techniques, is to identify the best fitting formula describing the relationship between a set of independent variables and binary outcomes (Washington et al, 2003).

The logistic regression can be understood simply as finding the β coefficients best fit:

$$y = \begin{cases} \beta_0 + \beta_1 x + \varepsilon & > 0 \\ 0 & else \end{cases} \qquad 7.23$$

The standard logistic function can take any real input t, where the output takes values between 0 and 1 and is interpreted as a probability function as:

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}} \qquad 7.24$$

Mathematically, t a multiple linear regression function is defined as:

$$t = B_0 + B_i . X_i , \qquad 7.25$$

Where B_0 is a constant and B_i are the model estimable coefficients for independent variables (X_i, i=1,….,n).

For modelling purposes, the observed outcomes are of two possible types; 0 and 1; which represent slight crash and fatal or serious injury crash respectively.

Accordingly, the probability function of vehicle-pedestrian crash results for fatal or serious injury may be written as

$$P(Fatal/Serious) = \frac{1}{1 + e^{-(B_0 + B_i\,.X_i)}} \qquad 7.26$$

Also, the probability function of vehicle-pedestrian crash results slight outcome can be written as

$$P(Slight) = 1 - P(Fatal/Serious) = 1 - \frac{1}{1 + e^{-(B_0 + B_i\,.X_i)}} = \frac{e^{-(B_0 + B_i\,.X_i)}}{1 + e^{-(B_0 + B_i\,.X_i)}} \qquad 7.27$$

Chapter 8

MODELS DEVELOPMENT

This chapter was structured to address the development of a new model estimating the safety of pedestrians moving in environments with high roadside activities in urban areas. This is driven by a question concerning the safety effects of such an environment on pedestrians. Another analytical question to be answered includes the effects of traffic and roadway characteristics on potential pedestrian-vehicle crashes. Specifically, a range of independent factors that may influence the potential of pedestrian-vehicle crash events include: traffic volume (ADT), traffic operating speed (V), coefficient of speed variance (CV), number of pedestrian crossing violations (PCV), the number of parking-un-parking and boarding-alight activities (PBA), number of pedestrian crossings (PC), number of bus stoppings (BS), percentage of heavy vehicle traffic (HV), intersecting traffic volume (INT) and roadside parking (P).

To answer these questions, the study conducted different analyses using the collected data and developed models using different regression approaches. The first step focused on the estimation of the coefficients of the independent variables. The second step was to select the final independent variables that are found to significantly explain the dependent variable to be included in the final model and eliminate the rest. Finally, it presents the measures used for testing and validation of the developed model.

8.1 Model building

This section presents the estimation process of the selected statistical regression methods for modelling; and the furthermore, measures used in evaluating, comparing and selecting the modelling method. Finally, it gives the model's coefficients selection and estimations.

8.1.1 Frequency model estimation

The objective of the analysis was to develop a new model to predict site-specific pedestrian-vehicle crash risks. As indicated in Chapter 7 (section 7.5), three methods were used to estimate the crash frequency model and logistic regression was used to estimate the crash severity model. The statistical packages for the Social Sciences (SPSS) and the R Project for statistical computing were utilized to estimate the coefficients of the included parameters in the pedestrian-vehicle crash prediction model.

To provide a more suitable model, the natural log transformation was applied to both traffic volume and intersecting traffic volume, as indicated in previous studies (Oh et al., 2003; El-Basyouny and Sayed, 2009; Wier , 2009; Miranda-Moreno et al., 2011; Hosseinpour, 2014).

Some of the road segments had zero intersecting traffic volume. In this case the zero was replaced by 1 to avoid dropping these segments from the analysis and eventually the magnitude of ln (1) = 0; and by applying the exponential when using the Poisson or NB model, the final outcome EXP (ln (1)) = 1.

In addition, the factors of intersecting traffic volume and number of side roads intersecting the analysed segment were included separately in the model estimation. This is due to the fact that intersecting volume and side roads are related, since in the event of no side road intersecting the considered road segment means no intersecting traffic flow at that segment. For that reason, the model was estimated using only the intersecting traffic. However, another model estimation was conducted, using the number of side roads instead of intersecting traffic volume as an alternative, in case no information was available to estimate the intersecting traffic volume. The Poisson and NB model estimation results are shown in Table 8.1

Table 8.1 Statistical properties of models -Poisson and NB models

Variables		Coefficients	Poisson Model		Negative Binomial Model	
			Estimated coefficient	p-value	Estimated coefficient	p-value
Constant		β_0	-9.848	0.005	-11.183	0.074
Traffic mean speed		β_1	-0.022	0.032	-0.027	0.162
Parking	No side parking	β_2	-0.611	0.004	-0.722	0.079
	One-side parking		-0.441	0.011	-0.442	0.213
	Two-side parking		0	•	•	•
Coefficient of speed		β_3	0.461	0.043	0.58	0.167
Parking-un-parking and boarding-alight (PBA)		β_4	-0.002	0.047	-0.002	0.222
Pedestrian crossing violations		β_5	0.008	< 0.001	0.008	0.028
No. of bus stoppings/Hr/Sec		β_6	0.015	< 0.001	0.016	0.049
Pedestrian along Vo		β_7	0.001	0.014	0.001	0.192
Number of pedestrian crossings		β_8	•	•	•	•
Percent of HV		β_9	•	•	•	•
Ln (AADT)		β_{10}	1.011	0.003	1.161	0.06
Ln (intersecting traffic volume)		β_{11}	0.042	0.03	0.049	0.183

To select between the models, a test for the appropriateness of using Negative Binomial over the Poisson model has been conducted firstly by checking the scale parameter of over-dispersion. The over-dispersion scale may be calculated by dividing the Pearson function value by the number of degrees of freedom. If the scale parameter exceeds 1 by a significant amount, it indicates over-dispersion in the data and in such a case, the NB method is appropriate for the model building. If the scale parameter ratio is larger than 2, then the use of the NB model may be required. If the scale parameter is closer to 1, it indicates that over-dispersion is not likely to be affecting the model fitting and the Poisson method provides an alternative method (Hardin and Hilbe, 2001). For this study, the calculated value of the scale parameter ratio was (1.011) which does not exceed the threshold of 2; and therefore the Poisson model was selected over the NB model.

Another considered method was the Zero-inflated Poisson (ZIP) because there are some road segments with zero crash records that need to be included in the model fitting. The Zero-inflated Poisson model estimation results are shown in Table 8.2

To test the appropriateness of using the ZIP over the previous models, the test proposed by Vuong (1989) was utilised as indicated in Chapter 7. The Vuong test compares the zero-inflated model with an ordinary Poisson or NB model, to estimate if one model fits better than another. Vuong's value (V) to be compared to $z - Statistic$ where large positive values of V greater than $V_{critical}$ favour model 1 over model 2, whereas negative values support model 2.

For this study, a test of a pair of models (Model 1-Poisson, Model2 –ZIP) was conducted using the Vuong-Test and the result was as follows:

Vuong Test-Statistic (2.466) giving that Model 1 > Model 2, with p-value 0.007; therefore the Poisson model (Model 1) is selected over the ZIP model (Model 2).

Table 8.2 Statistical properties of models -Zero-inflated Poisson model

Variables		Coefficients	Estimated coefficient	Std. Error	z value	p-value
Constant		β_0	-3.269	3.089	-1.13	0.258
Traffic mean speed		β_1	-0.029	0.010	-3.019	0.003
Parking	No side parking	β_2	.	.	.	.
	One-side parking		.	.	.	.
	Two-side parking		.	.	.	.
Coefficient of speed		β_3	0.467	0.211	2.219	0.027
Parking-un-parking and boarding-alight (PBA)		β_4	0.000	0.001	-0.066	0.947
Pedestrian crossing violations		β_5	0.004	0.001	3.427	0.001
No. of bus stoppings/Hr/Sec		β_6	0.019	0.004	4.602	< 0.001
Pedestrian along Vo		β_7	.	.	.	.
Number of pedestrian crossings		β_8	.	.	.	.
Percent of HV		β_9	7.519	15.147	0.496	0.620
Ln (AADT)		β_{10}	0.341	0.292	1.168	0.243
Ln (intersecting traffic volume)		β_{11}	0.034	0.019	1.777	0.076

After conducting the tests, the Poisson model shown in Table 8-1 was selected over the NB and the ZIP models. Another Poisson model was developed, using as the independent variable the number of intersecting side roads (SR) instead of intersecting traffic volume. Table 8.3 shows the estimated model coefficients.

Table 8.3 Statistical properties Poisson Model using number of intersecting roads

Variables		Coefficients	Estimated coefficient	Std. Error	z value	p-value
Constant		β_0	-6.694	3.2628	4.494	0.034
Traffic mean speed		β_1	-0.019	0.0099	3.703	0.054
Parking	No side parking	β_2	-0.533	0.2049	6.753	0.009
	One-side parking		-0.367	0.1683	4.742	0.029
	Two-side parking		•	•	•	•
Coefficient of speed		β_3	0.532	0.2231	5.685	0.017
Parking-un-parking and boarding-alight (PBA)		β_4	•	•	•	•
Pedestrian crossing violations		β_5	0.005	0.0015	12.536	< 0.001
No. of bus stoppings/Hr/Sec		β_6	0.013	0.004	10.933	0.001
Pedestrian along Vo		β_7	0	0.0002	2.808	0.094
Number of pedestrian crossings		β_8	•	•	•	•
Percent of HV		β_9	•	•	•	•
Ln (AADT)		β_{10}	0.685	0.324	4.475	0.034
Number of side roads		β_{11}	0.361	0.0771	21.956	< 0.001

The study felt that using the intersecting traffic volume variable for analysis could be more accurate because the intersecting side roads could have varied characteristics in terms of configuration, directions, width and number of lanes.

Therefore, the main model was estimated using the variable of the intersecting traffic volume. In order to estimate the risk factor associated with the number of intersecting side roads, another model was fitted utilizing the number of intersecting side roads instead of the intersecting traffic flow.

8.1.2 Estimated coefficients

For the best fitting of the Poisson model using the intersecting traffic volume in Table 8.1 and the Poisson model using the number of side roads shown in Table 8.3, only statistically significant independent variables were included (p-value < 0.05). Therefore, after an initial analysis any variable found insignificant was dropped before the model re-run. If more than one insignificant variable is present, the variable with highest P-value is dropped firstly before conducting another re-run.

According to estimation results of the models (Table 8.1 and Table 8.3), all significant independent variables are almost the same with relatively similar values; however, variables of the Poisson model with the intersecting traffic volume incorporate more significant effects.

The percentage of HV was not significant and excluded from the modelling process. This is may be because the percentage of HV was relatively low and has small range between 1.49% and 2.93% for all included sections in the study. The number of pedestrian crossings was also found not to be significant. This is may be because another indicator included counting the number of pedestrian crossing violations.

Traffic speed was found to be negatively associated with the crash frequency and showed only minor effect, possibly indicating less conflicts, more careful driving and better enforcement for roads with higher speed. This was also found by Al-Kaaf and Abdel-Aty (2015). Furthermore, previous studies highlighted the significant effects of traffic speed on crash severity level rather than frequency (Wier et al., 2009).

The model fitting indicated that road segments with one-sided or two-sided parking show a higher risk of pedestrian-vehicle crash compared with an absence of road-side parking. A number of studies have found that the parking variable correlated with crash frequencies, including Box (2002); Greibe (2003); Hauer and Mohammedshah, (2004); and Al-omari and Obaidat (2013).

The coefficient of speed variance affects crash frequency; suggesting that speed variation in a road segment increases accident frequency. For instance, Hunter et al. (1996); Taylor et al. (2000); Al-Ghamdi (2002); Koushki and Ali (2003); Elvik (2009); Derry et al. (2010); Pulugurtha and Sambhara (2011); and Al-Omari and Obaidat (2013), all of which have reported that the accident risk increased with higher speed variation.

In terms of roadside activities, the parking and un-parking activities at the roadside have negative effects on crash frequency; implying the crash frequency is slightly lower at segments with high parking and un-parking activities. This is possibly explained by the fact that more parking and un-parking manoeuvres make both pedestrians and drivers more aware and alert before committing to any movement and may lead to careful driving.

Pedestrian crossing violations and volume was found to have a statistically significant effect on accidents in this study. Previously, Brude and Larson (1993), Zegeer et al. (1985) and

Zeeger et al. (2005) found that the increasing of pedestrian crossing volume was associated with higher crash frequencies.

Bus stopping activities per hour per road segment were found to influence crash frequency, suggesting that a higher number of crashes occurs on segments accommodating more bus stoppings. Some previous studies investigated the effect of bus stops rather than the number of real bus stoppings activities and found a positive and significant relation, including Harwood et al. (2008); Miranda-Moreno (2011); and Ukkusuri et al.(2012).

The AADT was found to be a significant variable for developed models and as the traffic flow increased, crash numbers increased. The reason for this may be that the higher traffic flow leads to more interactions between pedestrians and vehicles and possibly increases the exposure of pedestrians to risk. Previous studies with similar findings includes Garber and White (1996), Wang and Abdel-Aty, (2008), and Wier (2009). For intersecting traffic volume, Table 8.1 shows higher intersecting traffic volume is associated with higher number of crash events. Similarly, Lyon and Persaud (2002), and Leden (2002) indicated the effect of turning traffic volume on crash risk.

For the model in Table 8.3, the number of intersecting side roads was found to be statistically significant and their number leads to an increase the crash risk. This is likely because of the interacting traffic environment and intersections bring a higher risk for vulnerable road users with the interactions and possibly increase the exposure of the pedestrians to risk. Also, this finding was indicated by Garber and White (1996) and Greibe (2003).

To summarise, the results show that the variables found to be significant in the Poisson model of Table 8.1 (intersecting traffic volume) were: traffic mean speed (V), roadside parking (P), coefficient of speed variance (CV), the number of parking-un-parking and boarding-alight activities (PBA), number of pedestrian crossing violations (PCV), number of bus stoppings (BS), pedestrian walking along (PAL), traffic volume (AADT) and intersecting traffic volume (INT). As such, the variables found to be significant in the Poisson model in Table 8.3 (number of intersecting side roads (SR)) were essentially the same but the number of parking-un-parking and boarding-alight activities (PBA) were not considered and the number of intersecting side roads (SR) instead of the intersecting traffic volume (INT) were considered.

Consequently, the best fitting Poisson model (Table 8.2) capturing the probability of a pedestrian-vehicle crash is as follows:

$$P(y_i) = 4.61 \times 10^{-5} \times AADT^{1.011} \times INT^{0.042} \times e^{\sum \beta 2 - 0.022V + 0.461CV - 0.002PBA + 0.008PCV + 0.015BS + 0.001PAL} \quad 8.1$$

Where *β2* = -0.611 if no side parking in road segment; -0.441 if one side parking or zero for two side parking.

When using the number of intersecting side roads (SR) instead of the intersecting traffic volume (INT), the Poisson model (Table 8.3) is given by:

$$P(y_i) = 12.39 \times 10^{-4} \times AADT^{0.685} \times e^{\sum \beta 2 - 0.019V + 0.532CV + 0.005PCV + 0.013BS + 0.361SR} \quad 8.2$$

Where *β2* = -0.533 if no side parking in road segment; -0.367 if one side parking or zero for two side parking.

8.1.3 Severity Model

Apart from developing a model for the probability of the pedestrian crashes, this study developed another model describing the effect of the selected variables on the crash severity. As the severity level is a binary outcome (see Chapter 7), logistic regression was utilised to predict this set outcome as a function of explanatory variables. The statistical package for the Social Sciences (SPSS, 2016) was utilized to estimate the coefficients of the included parameters in the pedestrian-vehicle crash severity model. For modelling purposes, the observed outcome was coded as 0 and 1, which represent slight crash and fatal or serious injury crash respectively.

For model fitting, the same procedure was used as only the statistically significant independent variables were included (p-value < 0.05). Therefore, after a preliminary analysis any variable found insignificant was excluded before the model re-run. Table 8.4 shows the estimated coefficients for the severity model. It appears that only the traffic mean speed (V) has a statistically significant effect in increasing the risk of death/serious injury crash as a result of a vehicle–pedestrian crash.

Table 8.4 Pedestrian-vehicle crash severity logistic model estimation

Variables	Coefficients	Estimated coefficient	Std. error	z value	p-value
Constant	β_0	-1.212	0.418	8.413	0.004
Traffic mean speed	β_1	0.033	0.017	3.749	0.05

Consequently, the best fitting severity model using the logistic equation (7.23) capturing the probability of a pedestrian-vehicle crash resulting in fatal or serious injuries is expressed as follow:

$$P(Fatal/Serious) = \frac{1}{1 + e^{-(-1.212+0.033\,.V)}} \qquad 8.3$$

Also, the probability function of vehicle-pedestrian crash resulting in a slight outcome may be written as:

$$P(Slight) = \frac{e^{-(-1.212+0.033\,.V)}}{1 + e^{-(-1.212+0.033\,.V)}} \qquad 8.4$$

The relationship between traffic speed and the crash severity outcome has been previously established (Lin et al., 2003; Nilsson, 2004; Peden et al., 2004; Lin and Kraus, 2009). According to the existing literature review, in addition to the significant effect of speed, there are other factors identified to significantly contribute to the severity risk resulting from a vehicle–pedestrian crash including age, gender and other characteristics (Lee and Mannering, 1999; Moudon et al., 2011; Olszewski et al., 2015).

However, these factors were beyond the scope of this study and the estimated severity model included only traffic mean speed as input. Therefore, the severity model will not be considered for further application in this study.

8.2 Model Testing and Evaluation

This section presents the evaluation and validation of the developed models to examine how well the models fit the observations and perform with different data sets. Three measures of evaluation were performed including goodness of fit assessment, graphical exploration of the observed and predicted crash numbers and model field validation.

8.2.1 Goodness of fit (GOF) measures

Goodness of fit (GOF) measures are useful for comparing models and examining how well the observed data set fits the developed model. For the model development, the variables selected and tested were as in (7.3). During the model development, another evaluation was conducted by estimating the statistically significant coefficients β for each variable as in section (8.1.2), while the process of selecting the most appropriate model was performed as explained in Chapter 7 (7.5.4).

The likelihood ratio test was used to assess the model GOF. This is a common measure used to assess Poisson regression models. It provides evidence of the appropriateness of one model over another and seeks to test the hypothesis that additional variables do not improve the fit. To use the likelihood ratio test, the smaller model must be nested in the other one. That is, the smaller model can be derived from the full model by having a reduced number of model dependent variables (Washington et al. 2005). The likelihood ratio test statistic is:

$$X^2 = -2[LL(\beta_S) - LL(\beta_F)] \quad 8.5$$

Where:

$LL(\beta_S)$ is the log likelihood of the smaller model

$LL(\beta_F)$ is the log likelihood of the full model

The calculated X^2 with the degree of freedom was utilised to examine the significance of the full model for each added variable using the critical chi-square probability.

Another measure to assess GOF is to use the standardised likelihood ratio as a scale parameter by dividing the likelihood ratio on the degree of freedom. The closer the standardised likelihood ratio to one, the better fitting the model and the smaller likelihood ratio, the better fit of the model (Al-Matawah, 2008).

Akaike information criteria (*AIC*), which is another GOF indicator, were also adopted as an additional GOF measure for determining model fit. The *AIC* was calculated as follows (Washington et al., 2011):

$$AIC = 2\,Q - 2\,LL(\beta) \qquad 8.6$$

Where: Q is the number of variables in the model; $LL(\beta)$ is the log likelihood at final convergence.

To perform the GOF measures, the model was fitted in a step-by-step procedure starting with the null model which contains only the constant value. At each following step, one variable was added and the GOF measures calculated. Table 8.5 shows the GOF for pedestrian-vehicle crash Poisson model estimation (with intersecting traffic volume). The value of the likelihood ratio and the AIC of the full model are the smallest, indicating that the full model shows a better fit. Another indication for a better fit of the full model is the standardised likelihood ratio being closer to one. The calculated p-value for the decrease in likelihood ratio with every variable added indicated a significant decrease for most of the added variables. For the variable of intersecting traffic volume (INT), the p-value for the likelihood ratio decrease is 0.11; however, adding this variable shows a decrease in both AIC and likelihood ratio. Overall, the full model shows a better fit in comparison to other nested models with reduced variables.

Table 8.5 Goodness of fit tests for pedestrian-vehicle crash Poisson model estimation (with intersecting traffic volume)

	Variable added									
	0	1	2	3	4	5	6	7	8	9
Evaluation measure	Null	BS	PCV	V	CV	INT	AADT	P	PAL	PBA
Likelihood ratio	226.4	212.17	146.99	140.51	135.04	132.50	129.35	124	121.3	117.7
DF	115	114	113	112	111	110	109	107	106	105
Likelihood ratio drop		14.201	65.183	6.483	5.471	2.541	3.146	5.314	2.72	3.579
Chi-square probability		>0.001	>0.001	0.01	0.019	0.11	0.07	0.07	0.09	0.05
St. likelihood ratio	1.968	1.861	1.301	1.255	1.217	1.205	1.187	1.16	1.14	1.121
AIC	514.35	502.15	438.97	434.49	431.01	430.47	429.33	428	427	425.7

The same procedure was followed to perform the GOF measures pedestrian-vehicle crash Poisson model estimation (with number of intersecting roads). Table 8.6 shows the GOF for the pedestrian-vehicle crash Poisson model with the number of intersecting roads.

The value of likelihood ratio and AIC of the full model are again the smallest indicating that full model shows better fit. The calculated P-Value for the decrease in likelihood ratio with every variable added indicated a significant decrease for most of the added variables. For the variables of AADT and parking (P), the p-value for likelihood ratio decrease is 0.1. On the other hand, the standardised likelihood decreases by adding the coefficients of AADT and parking (P). However, the full model outperforms the other models in AIC value.

Table 8.6 Goodness of fit tests for the pedestrian-vehicle crash Poisson model estimation (with number of intersecting roads)

	Variable added							
	0	1	2	3	4	5	6	7
Evaluation measure	Null	SR	PCV	BS	V	CV	P	AADT
Likelihood ratio	226	165.333	131.390	120.46	116.3	109.9	105.34	102.854
DF	115	114	113	112	111	110	108	107
Likelihood ratio drop		61.042	33.943	10.926	4.165	6.39	4.567	2.487
Chi-square probability		>0.001	>0.001	0.001	0.041	0.01	0.1	0.1
St. likelih. ratio	1.97	1.450	1.163	1.076	1.048	.999	.975	.961
AIC	514	455.310	423.367	414.44	412.28	407.89	407.32	406.832

8.2.2 Observed and predicted crash numbers

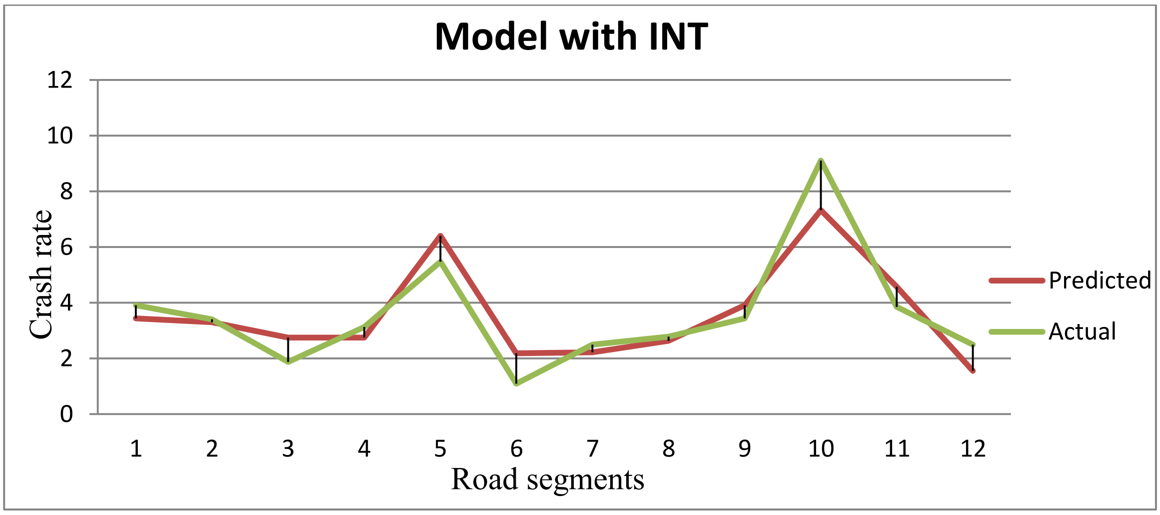
To examine the developed models based on results, the actual crash frequency observed in the field is compared with the predicted crash frequency estimated by the developed models for the same road sections and models' estimation data set. First, the predicted crash frequency was estimated by the model using the number of intersecting roads (INT) in equation 8.1 and the model using the intersecting traffic volume (SR) in equation 8.2. Secondly, all predicted and actual crash data were normalized for each road section in terms of crash rate per km per year. Then, predicted and actual crash data were compared by computing the accuracy of estimation. Table 8.7 shows the results.

The results indicated that the average percentage of agreement between the predicted and the actual pedestrian-vehicle frequency crash data was 81.3% when using the model with INT and 83.6% when using the model with SR. This result shows a good level of agreement between actual values and predicted values for both developed models in equations 8.1 and 8.2.

Table 8.7 Predicted and actual number of crash frequencies values for each road section

No.	Name of the Road	Observed	Poisson Model (INT)		Poisson Model (SR)	
			Predicted crashes	Percentage correct %	Predicted crashes	Percentage correct %
1	Coventry Rd N-East	3.91	3.44	88.0%	3.75	96.0%
2	Coventry Rd S-West	3.41	3.30	96.7%	3.64	93.8%
3	Hagely Rd -East	1.88	2.75	68.2%	2.75	68.2%
4	Hagely Rd -West	3.13	2.75	88.0%	2.88	92.0%
5	Moseley Rd -S	5.47	6.41	85.4%	6.41	85.4%
6	Moseley Rd -N	1.09	2.19	50.0%	2.03	53.8%
7	Pershore Rd -N	2.50	2.22	88.9%	2.22	88.9%
8	Pershore Rd -S	2.78	2.64	95.0%	2.64	95.0%
9	Soho Rd -E	3.44	3.91	88.0%	3.59	95.7%
10	Soho Rd -W	9.11	7.32	80.4%	7.68	84.3%
11	Stratford -N	3.85	4.58	84.1%	4.38	88.1%
12	Stratford -S	2.50	1.56	62.5%	1.56	62.5%
			Average	81.3%	Average	83.6%

The actual and predicted crash frequencies' values for both INT and SR models for the 12 road sections used to develop the models were standardized to crash rate per km per year and presented as trend lines in Figure 8.1.



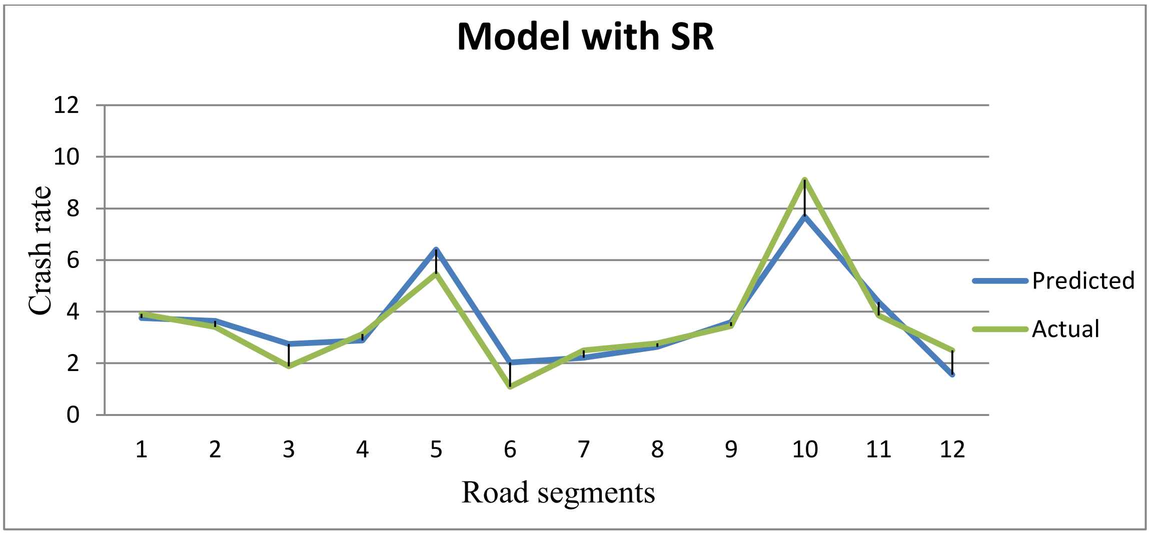


Figure 8.1 Predicted and actual number of crash frequencies' values

The graphical visualisation of both models in comparison with the actual crash rates shows good agreement and therefore the ability of the models to predict crash rates at a road section level with reasonable accuracy.

8.2.3 Field validation

In order to further test the developed models, a field validation was conducted. The performance of the suggested models was tested using data collected from three road sections following the same data collection measures as explained in Chapter 5. The summary of validation data can be seen in Appendix D (Tables D1 to D4).

The actual crash frequency observed in the validation site was compared with the predicted crash frequencies estimated by the developed models.

First, the actual crash frequency was collected and the predicted crash frequency was estimated by the INT and SR models (see equations 8.1 and equation 8.2) for each selected site. Second, all predicted and actual crash data were standardized for each road section to be measured in terms of crash rate per km per year. Then, the predicted and actual crash data were compared in terms of the accuracy of their estimation. Table 8.8 shows the comparison results.

The results indicated that the average percentage of agreement between the predicted and the actual pedestrian-vehicle crash frequencies was 87.27% when using the INT model and 81% with the SR model. These results show a good level of agreement between the actual values and the predicted values and similar trends with the results produced by the developed model.

Table 8.8 Predicted and actual number of crash frequencies values for the validation road sections

No.	Name of the Road	Observed	Poisson Model (INT)		Poisson Model (SR)	
			Predicted crashes	Percentage correct %	Predicted crashes	Percentage correct %
1	Dudley Rd	4.77	3.93	82.39%	3.85	80.71%
2	Waterloo-Cape Hill Rd	4.17	3.47	83.21%	3.44	82.49%
3	Alum Rock Rd	5.80	5.58	96.21%	4.64	80.00%

The graphical visualisation of both models in comparison with the actual crash rates show good similarity in trends and therefore the ability of the models to predict crash rates at 100-metre segments with reasonable accuracy. However, it is obvious that the models could not predict some sudden high crash rates in some road segments. This is possibly because there are other variables (e.g. human factors) which were not considered in this model building process (Abdel-Aty and Radwan, 2000).

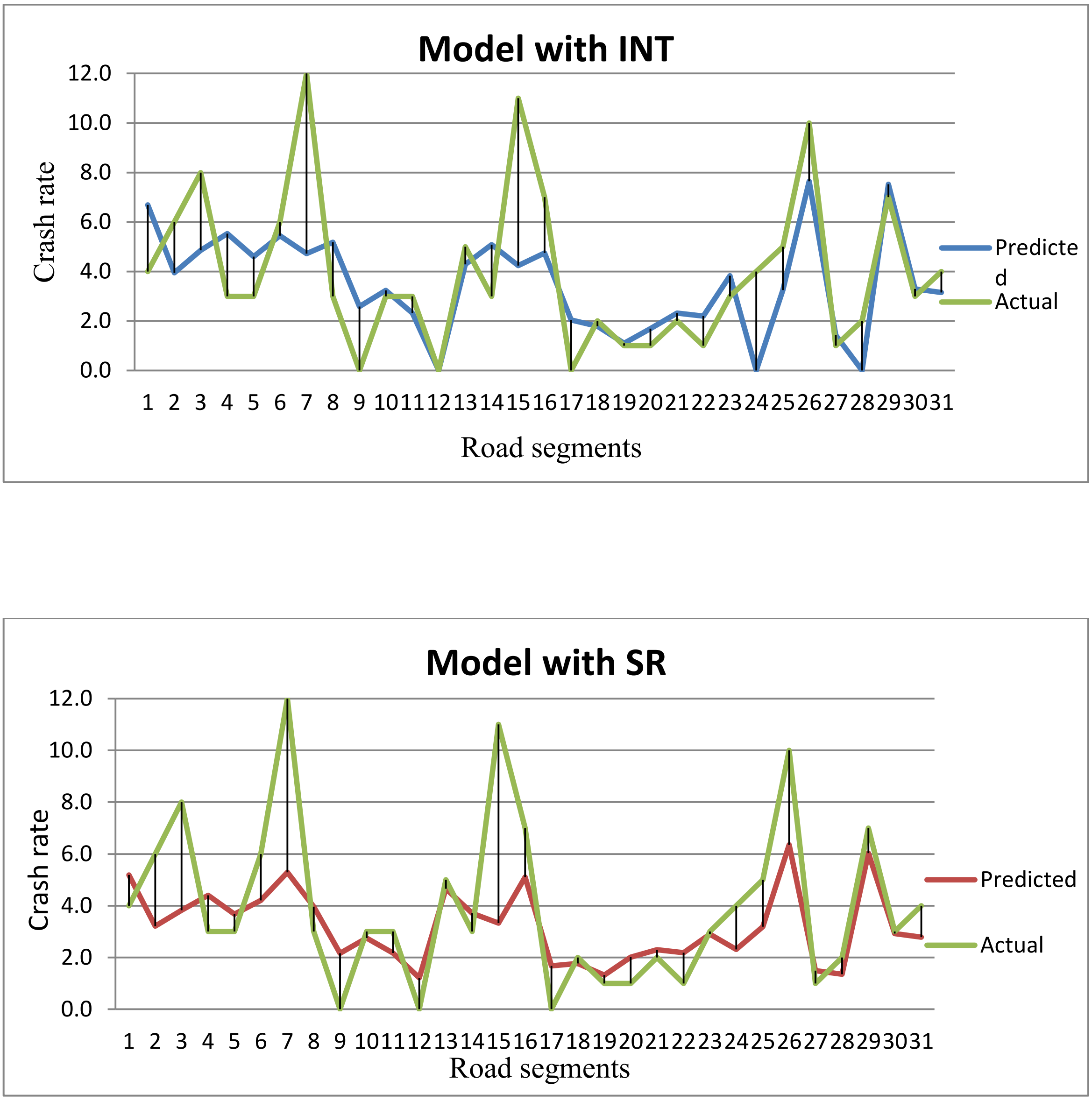


Figure 8.2 Predicted and actual number of crash frequencies' values for the validation sites

Other goodness of fit (GOF) measures were examined to investigate how well the developed models fit the data by comparing their performance with the original data set versus the verification data set. For this purpose, the mean squared prediction error (MSPE) and the mean squared error (MSE) were used.

The MSPE is defined as the sum of the squared differences between actual and predicted crash rates divided by the sample size. The MSPE is usually used to assess the error associated with a verification data set and is given by:

$$MSPE = \frac{\sum_{i=1}^{n_1} (\hat{Y}_i - Y_i)^2}{n_1} \quad 8.7$$

Where:

n_1 is the number of model validation sites; $\hat{Y}_i$ is average predicted crash rate; and Y_i is the actual crash rate observed in same model verification of a road section.

The MSE is defined as the sum of the square difference between predicted and actual crash rates divided by the sample size minus the number of model independent variables. The MSE is usually a measure of model error associated with the estimation data set and given by:

$$MSE = \frac{\sum_{i=1}^{n_2} (\hat{Y}_i - Y_i)^2}{n_2 - P} \quad 8.8$$

Where:

n_2 is the number of model estimation sites; P is the number of model independent variables; $\hat{Y}_i$ is the average predicted crash rate; and Y_i is the actual crash rate observed in same model estimation of a road section.

The developed model performance can be estimated then by comparing the MSPE and MSE values. If the MSPE higher than the MSE value, this may indicate the potential of over-fitting to the estimation data or some important variables were omitted (Oh et al., 2003)

For this study, these goodness of fit (GOF) measures were used for the entire data set (3 verification road sections and 12 model building road sections).

When using the model with INT, the MSPE for crash rate per km per year was (0.412) for the verification data and significantly lower than that of the MSE (2.71) for the model building data. Also for the model with SR, the MSPE for crash rate per km per year (0.91) for the verification data was slightly lower than the MSE (1.19) for the estimation data. Both developed models show MSPE values lower than MSE values. However, for the model with (RS) both MSPE and MSE values are similar in magnitude, indicating that the validation and estimation data set fits the model similarly.

8.2.4 Field validation-Low roadside activities

Furthermore, and in order to test the model performance in road sections with a low pedestrian activities, three road sections were selected and tested. This analysis has been conducted to provide a counterpoint and contrast for the sections of high activity. The road sections were selected following the same data collection measures as explained in Chapter 5. In effect, most of their lengths were undivided with a centre line or hatching and no physical median present. Also, the presence of active bus stops was a condition. However, the selected roads were not surrounded by commercial tracts to ensure less pedestrian and roadside activities. The summary of validation data can be seen in Appendix D (Tables D5 to Table D8).

The actual crash frequency represented in crash per year per km in the low activity sites was compared with predicted crash frequencies. Table 8.9 shows the comparison results.

The results showed that the agreement between the predicted and the actual pedestrian-vehicle crash frequencies was 73.46% when using the INT model and 68.69% with the SR model. These results show less level of agreement between the actual values and the predicted values than those found when higher pedestrian and roadside activities sections were utilised for validation (87.27% and 81% respectively). Also, the graphical relationship could not be proved between the predicted crash rates of both models and the actual crash rates.

Therefore, the developed model can give higher accuracy when used to estimate the crash frequency in high activities sections.

Table 8.9 Predicted and actual number of crash frequencies values for the validation road sections-Low roadside activities

No	Name of the Road	Observed	Poisson Model (INT)		Poisson Model (SR)	
			Predicted crashes	Percentage correct %	Predicted crashes	Percentage correct %
1	City Rd	0.88	1.18	74.46%	1.55	56.76%
2	Yardley-Sherbroune Rd	1.63	1.10	67.20%	1.58	96.86%
3	VicarageRd	0.88	1.12	78.72%	1.68	52.44%

Chapter 9

MODEL APPLICATION

9.1 Introduction

This chapter proposes an application of the developed models to enhance the current international road assessment programme (iRAP) star rating system, to improve its ability in assessing the safety of pedestrians moving in road environments with high roadside activities in urban areas. The first section covers the model's sensitivity analysis, using the developed models in this study to examine the effect of variation of the input variables on the safety outputs. The second section assesses the effect of the independent variables on crash frequency potentials by estimating the relative contribution of the variables' risk factors. The third section suggests an approach to enhance the current iRAP star rating tool for pedestrians, by incorporating the developed risk factors in the calculation mechanism of the star rating.

9.2 Sensitivity Analysis

The input variables of any model are subject to change; therefore, the impact of these potential changes on the resulting output may be investigated through sensitivity analysis. Also, sensitivity analyses can be used to communicate the quantification of the system, model development and identify the most important variables (Pannell, 1997).

For this purpose, the developed model for the prediction of the crash rate, shown in Equation 8.1, was analysed for different ranges of inputs to investigate the effect of the change of the input variables on crash probabilities. The tested variables in the sensitivity analysis included traffic volume (AADT), traffic operating speed (V), coefficient of speed variance (CV), number of pedestrian crossing violations (PCV), the number of parking-un-parking and boarding-alight activities (PBA), the number of bus stoppings (BS), intersecting traffic volume (INT) and roadside parking (P). Also, the developed model in Equation 8.2 was used to test the sensitivity towards the number of intersecting side roads (SR).

The sensitivity analysis was performed by changing the input variables to observe the corresponding changes in the output values of the crash rate. The input variables' limits were based on the observed maximum and minimum variables' values in the collected data set; and in the case of testing for specific input variables, the estimated mean was considered for the rest of the input variables. Despite its limitations, this was felt to be a reasonable approach because it was based on the data collected and avoided potential errors of extrapolations.

9.3 Data Set for Sensitivity Analysis

To conduct the sensitivity analysis, the modelling estimation and validation data set as presented in Chapters 6 and 8 was used to ensure a wider range of maximum and minimum inputs. Both data sets were merged and are summarised in Table 9.1.

9.4 Sensitivity Analysis Results

To assess the significance of independent variables, a range of selected variables was examined to estimate their relative influence on various traffic flow levels and parking situations. A selected group or pair of independent variables was tested to examine their combined effect on crash frequency probability compared to a base case which was taken as 1.0. The rest of the tabled values show the percentage of change with regard to the base case and correspond to the changes in the considered variables. The sensitivity analysis was carried out with regard to:

Table 9.1 Data set summary for sensitivity analysis

Input variables	Mean	Minimum	Maximum
Traffic mean speed	25	10	47
Coefficient of speed	1.08	0.36	2.30
Pedestrian along volume	650	173	1929
Parking-un-parking and boarding-alight	228	81	450
Pedestrian crossing violations	137	38	268
No of bus stoppings/hr/section	15	0	54
AADT	19902	12267	36851
Intersecting traffic volume	3956	0	15000
Number of side roads	1	0	3
Side-parking	1	0	2

9.4.1 Traffic flow variables and parking arrangement

The sensitivity analysis was performed to examine the influence of intersecting traffic volume (INT), traffic volume (AADT) and parking arrangements on crash frequency rate. Different values for INT, ranging from 3000-15000, with AADT levels ranging from 15000-35000 were used to predict the crash rate as shown in Table 9.2. The base case was taken as AADT equals to 15000, with an intersecting AADT of 3000 and no parking present at the roadsides to show the percentage of change in the crash rate corresponds to other values for the considered variables.

Table 9.2 The sensitivity of crash frequency to traffic flow variables and parking

AADT	Parking	Crash frequency rate (percentage of change)				
		Intersecting AADT				
		3000	6000	9000	12000	15000
15000	No parking	1.000	1.030	1.047	1.060	1.070
	One-side	1.185	1.220	1.241	1.256	1.268
	Two-side	1.842	1.896	1.929	1.952	1.971
20000	No parking	1.337	1.377	1.400	1.418	1.431
	One-side	1.585	1.632	1.660	1.680	1.696
	Two-side	2.464	2.537	2.580	2.611	2.636
25000	No parking	1.676	1.725	1.755	1.776	1.793
	One-side	1.987	2.045	2.080	2.106	2.125
	Two-side	3.087	3.178	3.233	3.272	3.303
30000	No parking	2.015	2.075	2.110	2.136	2.156
	One-side	2.388	2.459	2.501	2.531	2.555
	Two-side	3.712	3.822	3.888	3.935	3.972
35000	No parking	2.355	2.424	2.466	2.496	2.519
	One-side	2.791	2.874	2.923	2.958	2.987
	Two-side	4.338	4.466	4.543	4.598	4.642

9.4.2 Traffic flow with parking arrangement and intersecting side roads

Using the same levels of AADT and parking arrangements, the sensitivity analysis was conducted to predict the crash rate using the number of intersecting side roads (SR). For this purpose, the number of intersecting side roads' (SR) input ranged from 0 to 3 as shown in Table 9.3. The base case was taken as AADT equals to 15000 with no parking present at the roadsides and no intersecting road, to show the percentage of change in the crash rate corresponding to other values for the considered variables.

By comparing the maximum and minimum outputs in Tables 9.2 and 9.3 using the same AADT and parking arrangements, the results indicated that the predicted crash rate showed more sensitivity for the SR variable. This may due to the fact that more SR leads to more side road crossings for pedestrians, while the intersecting traffic volume does not necessarily reflect the number of side roads to cross.

Table 9.3 The sensitivity of crash frequency to number of intersecting side roads (SR)

		Crash frequency rate (percentage of change)			
AADT	Parking	Number of intersecting side roads (SR)			
		0	1	2	3
15000	No parking	1.000	1.435	2.059	2.954
	One-side	1.180	1.694	2.431	3.488
	Two-side	1.704	2.446	3.509	5.034
20000	No parking	1.222	1.753	2.515	3.608
	One-side	1.442	2.069	2.969	4.259
	Two-side	2.081	2.987	4.285	6.148
25000	No parking	1.427	2.047	2.937	4.213
	One-side	1.684	2.416	3.467	4.974
	Two-side	2.431	3.488	5.004	7.180
30000	No parking	1.620	2.324	3.333	4.783
	One-side	1.911	2.743	3.935	5.646
	Two-side	2.759	3.959	5.680	8.150
35000	No parking	1.802	2.586	3.711	5.324
	One-side	2.128	3.053	4.380	6.285
	Two-side	3.072	4.407	6.323	9.072

9.4.3 Traffic flow with pedestrian crossing violations and pedestrian along volume

To assess the influence of pedestrians' movements, the sensitivity analysis was performed to examine the effect of pedestrian crossing violations (PCV) and pedestrian along volume (PAL). Different values for PCV ranging from 50 to 250 and PAL ranging from 300 to 1500 with AADT levels from 15000 to 35000 were used to predict the crash rate as shown in Table 9-4. The base case was taken as AADT equals to 15000 with pedestrian crossing violations 50 and pedestrian along volume 300, to show the percentage of change in the crash rate corresponding to other values for the considered variables.

The predicted crash rate showed sensitivity for both variables, PCV and PAL. Based on results in Table 9.4, the predicted crash rate percentage of change ranged up to 11.63 crashes per km per year for different levels of AADT and PCV and up to 7.84 crashes per km per year for different levels of AADT and PAL.

Table 9.4 The sensitivity of crash frequency to pedestrians' variables PCV and PAL

	Crash frequency rate (percentage of change)									
AADT	Pedestrian crossing violations (PCV)					Pedestrian along volume (PAL)				
	50	100	150	200	250	300	600	900	1200	1500
15000	1.00	1.49	2.22	3.31	4.93	1.00	1.35	1.82	2.46	3.33
20000	1.33	1.99	2.97	4.43	6.60	1.34	1.81	2.44	3.30	4.45
25000	1.67	2.49	3.72	5.55	8.28	1.68	2.27	3.06	4.13	5.58
30000	2.01	3.00	4.47	6.67	9.95	2.02	2.73	3.68	4.97	6.71
35000	2.35	3.50	5.22	7.80	11.63	2.36	3.19	4.30	5.80	7.84

9.4.4 Coefficient of speed and bus stoppings with parking arrangements

The sensitivity analysis was also conducted to examine the effect of the coefficient of speed (CV) and bus stoppings (BS) on the crash frequency rate. Different values for CV ranging from 0.4 to 2 and for BS ranging from 10 to 50 with different parking arrangements along the roadside were used to predict the crash rate as shown in Table 9.5 and Table 9.6. The base case was taken as the coefficient of speed equals to 0.4 and 10 bus stoppings per hour with no parking present at the roadsides to show the percentage of change in the crash rate corresponding to other values for the considered variables.

The results of the predicted crash rate were sensitive for both variables, CV and BS. The predicted crash rate percentage of change ranged from 1 crash rate per km per year in the base case of no parking presence at the roadside and with the coefficient of speed equal to 0.4, up to 3.851 crash rates per km per year in case of parking present at both roadsides and the coefficient of speed value equal to 2. Regarding the BS variable, the predicted crash rate percentage of change ranged from 1 crash per km per year in the base case of no parking present at the roadside and 10 bus stopping activities per hour, up to 3.357 crashes per km per year in the case of parking present at both roadsides and 50 bus stopping activities per hour.

Table 9.5 The sensitivity of crash frequency to parking and coefficient of speed (CV)

	Crash frequency rate (percentage of change)				
Parking	Coefficient of speed (CV)				
	0.4	0.8	1.2	1.6	2
No parking	1.000	1.202	1.446	1.738	2.091
One-side	1.185	1.425	1.714	2.061	2.478
Two-side	1.842	2.215	2.664	3.203	3.851

Table 9.6 The sensitivity of crash frequency to parking and number of bus stoppings (BS)

	Crash frequency rate (percentage of change)				
Parking	Bus stoppings (BS)				
	10	20	30	40	50
No parking	1.000	1.162	1.350	1.568	1.822
One-side	1.186	1.377	1.600	1.859	2.160
Two-side	1.842	2.141	2.487	2.890	3.357

9.5 Contribution of Variables on Crash Risk

The sensitivity analysis sections focused on the identification of the degree of influence of a set of contributing factors on crash rates. In this section the developed models presented in section 8.1.2 were tested to assess the relative contribution of each of the independent variables on the crash frequency risk of pedestrians moving in environments with high roadside activities in urban areas. This test aimed to produce relative risk factors representing the effect of the significant variables which were found in this study to propose enhanced risk factors for the iRAP.

9.5.1 Relative effect of variables' risk factors

Risk factors are used in the iRAP star rating methodology to relate road attributes with crash frequency rates and subsequently to calculate star rating scores for different road users in road segments based on those road attributes. The iRAP methodology generally uses a range of risk factors reflecting the contribution of the selected attributes on crash risk using a multiplicative form. Therefore, to integrate the developed models with iRAP, a range of risk factors were calculated from the parameters' coefficients of the developed models using the concept of crash modification factors (CMF) to be integrated later using the multiplicative form of the iRAP.

A CMF is a multiplicative value used to calculate the expected frequency of a crash at a defined location and evaluate the effect on safety due to changes in traffic features, or in the geometric design for the same location (Gross et al., 2010; Carter et al., 2012). The CMF is calculated as a ratio by dividing the expected frequency of a crash for one condition by the expected frequency of a crash under a different condition. To calculate the ratio between expected crash rates after and before, or with and without a given change (HSM, 2010); a CMF can be calculated as:

$$CMF = \frac{\text{expected crash frequency after or with a given change}}{expected\ crash\ frequency\ before\ or\ without\ a\ given\ change} \qquad 9.1$$

Therefore, the base CMF value of each attribute is 1.0. Any attribute associated with a greater crash risk has a CMF higher than 1.0; while a value less than 1.0 indicates an attribute associated with a reduction in the crash risk.

9.5.2 Determining crash risk factors for pedestrians

The safety effect of the independent variables may be calculated using the developed models by means of developing risk factors associated with crash rates between different conditions. The concept of Equation 9.1 may be further used to develop the risk factors by calculating the ratio of crash frequencies due to a relative change in one of the dependent variables assuming all other variables remain unchanged. Therefore, risk factors may be calculated as follows:

$$Risk\ factor = \frac{expected\ crash\ frequency\ under\ diffrent\ condition}{expected\ crash\ frequency\ under\ base\ condition} \qquad 9.2$$

Following the sensitivity analysis presented in section 9.4, the base condition was taken the minimum values of the independent variables and the risk factors were calculated for each variable by considering an incremental increase, until reaching the maximum value (as found in data set) for the defined variable.

Using the developed models and Equation 9.2, risk factors could be calculated for the intersecting AADT, the number of intersecting side roads (SR), vehicle parking (P), the coefficient of speed (CV), the pedestrian crossing violations (PCV), the pedestrian along flow (PAL) and the number of bus stoppings (BS).

The relative risk factors' results for these variables are presented in Tables 9.7 to 9.13

9.5.3 Relative risk factor for intersecting road volume

The relative risk factors for the intersecting road volume were computed and presented in Table 9.7. To start, the value of the intersecting road volume was taken as zero (substituted 1 for computational purpose as indicated in section 8.1.1) and then a gradual increase was considered to calculate the risk factors. Using the previously developed models, it was found that the crash rate becomes less sensitive if the intersecting road volume exceeds 2000 per 100 m road section. Therefore, the risk factors may be divided into three categories as shown in Table 9.7.

Table 9.7 Relative risk factor for intersecting traffic volume

Intersecting Road Volume	0	≤ 2000	>2000
Crash rate (accident/yr/km)	2.7	3.7	3.8
Risk factor	1	1.3	1.5

9.5.4 Relative risk factor for the number of intersecting side roads

The relative risk factors with regard to the number of intersecting side roads were computed and are presented in Table 9.8. The range of number of the intersecting roads per 100 m road section was found to be between 0 and 3 for the sections included in the study. Therefore, the crash rates and their relative risk factors was calculated for these cases as shown in Table 9.8.

Table 9.8 Relative risk factor for number of intersecting side roads

Number of intersecting side roads	0	1	2	3
Crash rate (accident/yr/km)	1.9	2.7	3.9	5.5
Risk factor	1	1.4	2.1	3

9.5.5 Relative risk factor for parking vehicles

The relative risk factors for vehicle parking were computed and presented in Table 9.9. The vehicle parking condition per 100m road section may be as follows: no parking permitted, one-side parking or two-side parking. Therefore, the crash rates and relative risk factors were calculated for these three options.

Table 9.9 Relative risk factor for parking vehicles

Parking	No parking	One side parking	Two sides parking
Crash rate(accident/yr/km)	3.10	3.67	5.71
Risk factor	1.0	1.2	1.8

9.5.6 Relative risk factor for coefficient of speed

The relative risk factors for coefficient of speed were computed and presented in Table 9.10. The range of the speed coefficient per 100 m road section was taken as indicated in the sensitivity analysis (see section 9.4.4). Therefore, the relative risk factors for the sections included in this study, ranged between 1 and 2.1.

Table 9.10 Relative risk factor for coefficient of speed

Coefficient of speed	0.4	0.8	1.2	1.6	2
Crash rate (accident/yr/km)	2.3	2.7	3.3	3.9	4.7
Risk factor	1.0	1.2	1.4	1.7	2.1

9.5.7 Relative risk factor for pedestrian crossing violations

The relative risk factors for pedestrian crossing violations were computed and are presented in Table 9.11. The base value of the pedestrian crossing violations' value was empirically taken as a rounded mean value of 100 for all included sections; and then the effect of its gradual increase was considered using the developed models. For simplicity, the relative risk factors for this variable were calculated for three categories as indicated in Table 9.11.

Table 9.11 Relative risk factor for pedestrian crossing violations

Pedestrian crossing violations	≤100	100-200	≥200
Crash rate (accident/yr/km)	2.3	3.4	5.1
Risk factor	1.0	1.5	2.2

9.5.8 Relative risk factor for pedestrian flow- along

The relative risk factors for pedestrian flow- along were computed and are presented in Table 9.12. The base value of the pedestrian flow-along value was empirically taken as a rounded mean value of 500 for all included sections, and then the effect of its gradual increase was considered using the developed models. The relative risk factors for this variable were calculated for the cases as shown in Table 9.12 below.

Table 9.12 Relative risk factor for pedestrian flow- along

Pedestrian flow-along	≤ 500	750	1000	1250	≥1500
Crash rate(accident/yr/km)	3.1	3.42	4.4	5.64	7.25
Risk factor	1	1.1	1.4	1.8	2

9.5.9 Relative risk factor for number of bus stoppings

The relative risk factors for the number of bus stoppings were computed and are given in Table 9.13. The base value for the number of bus stoppings' value was empirically taken as a mean value of 15 for all included sections; and then the effect of its gradual increase was considered using the developed models. The relative risk factors for this variable were calculated for the categories indicated in Table 9.13 below.

Table 9.13 Relative risk factor for number of bus stoppings

Number of bus stoppings	≤15	20	30	40	≥50
Crash rate(accident/yr/km)	3.87	4.20	4.88	5.67	6.59
Risk factor	1.0	1.1	1.3	1.5	1.7

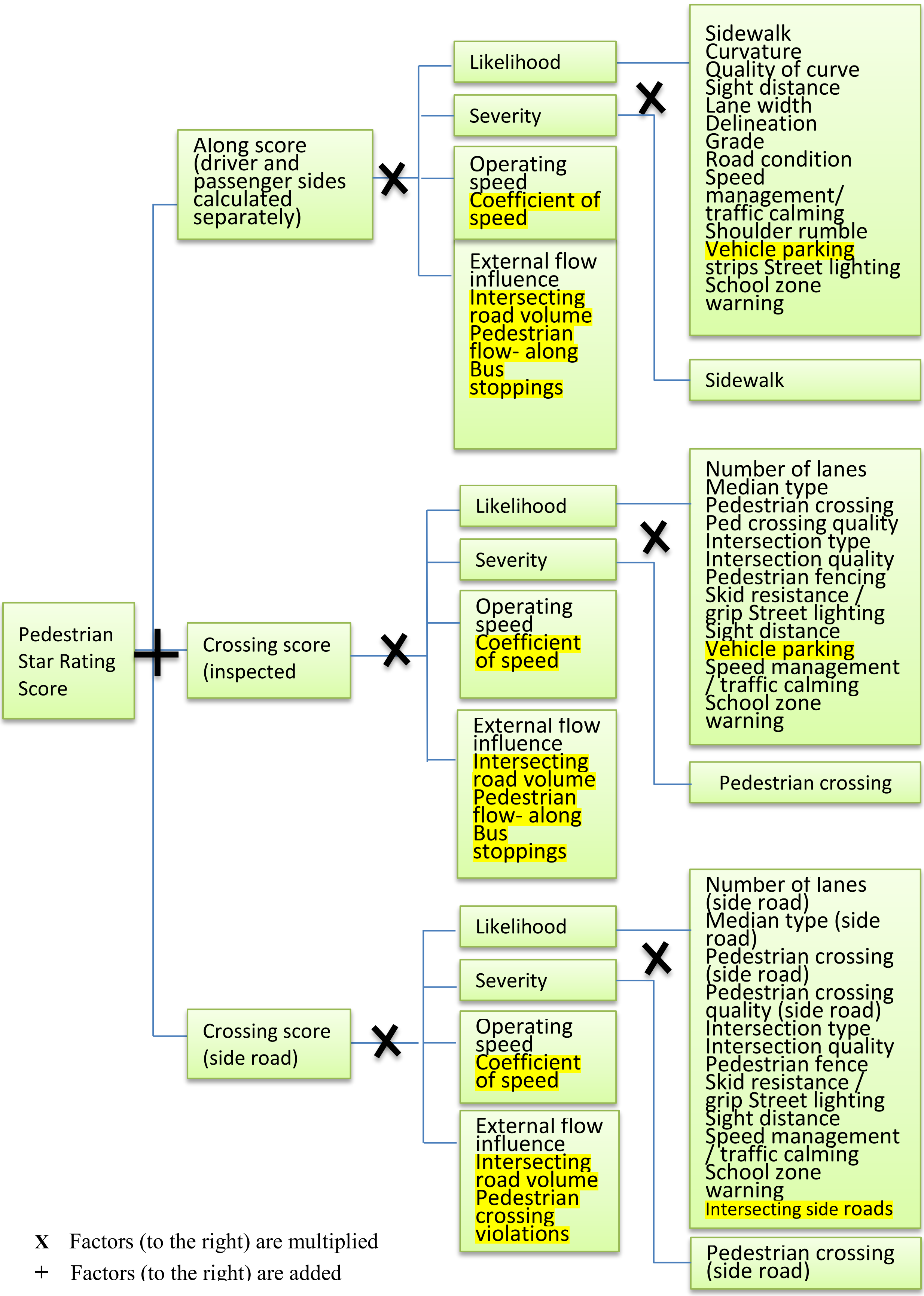
9.6 Enhancing the iRAP Star Rating System for Pedestrians

The main use of the modelling presented in the previous sections is to predict accident risks and thereafter to identify appropriate safety counter measures and improvements rather than using accident prediction models (Fu, 2015). However, the models were developed using limited road configurations and therefore can be used for limited road designs which are comparable to the considered road segments.

To apply and test the developed models as part of a fully developed road safety management system, the International Road Assessment Programme (iRAP) has been selected (see Chapter 4). Therefore, all risk factors computed using the developed models were embedded in the existing (iRAP) star rating system to introduce a comprehensive and enhanced tool for assessing pedestrians` safety in environments with high roadside activities in urban areas; which was not explicitly addressed to date. The following section elaborates further on this process.

9.6.1 Proposing an enhanced iRAP star rating methodology

The iRAP star rating score is estimated by multiplying the relative risk factors of road features related to the likelihood and severity of each road user`s crash types. Therefore, the multiplicative approach was utilised to embed the obtained risk factors in a manner consistent with iRAP methodology. Figure 9.1 shows the iRAP methodology and the newly suggested risk factors for integration within the iRAP system as opposed to the original model which was given in Chapter 4 (Figure 4.1).



Factors highlighted in Yellow are suggested to be embedded in the current model

Figure 9.1 iRAP pedestrian star rating score calculation methodology (iRAP, 2016) and suggested risk factors to be embedded

9.6.2 Data used

To assess the proposed methodology, a comparison has been conducted between the existing iRAP and the suggested model, enhanced with the new risk factors calculated with the developed model in this study. To achieve this comparison, the necessary data to run the iRAP star rating model were collected (see Chapter 4). All the risk factors' data shown in Table 4.1 were collected and coded in accordance with the iRAP data specification and coding manual. These data were collected and coded for each 100-metre section along the road segments used, which were road sites used for sensitive analysis.

Table 9.14 Selected road sections for the iRAP application

No	Name of the Road	Covered Length (km)
1	Coventry Rd N-East bound	0.9
2	Coventry Rd S-West bound	1.1
3	Hagley Rd -East bound	1
4	Hagley Rd West bound	1
5	Moseley Rd -S	0.8
6	Moseley Rd -N	0.8
7	Pershore Rd -N	0.9
8	Pershore Rd -S	0.9
9	Soho Rd-E	0.8
10	Soho Rd-W	0.8
11	Stratford Rd-N	1.2
12	Stratford Rd-S	1.6
13	Dudley Rd	1.1
14	Waterloo Rd-Cape Hill Rd	0.9
15	Alum Rock Rd	1.1

The traffic characteristics and sample observations together with the road attributes' dataset for iRAP (see Chapter 4) were collected in accordance with iRAP data specification (iRAP, 2014). A sample of iRAP geometry and traffic characteristics' input are shown in Table 9.15.

Table 9.15 Geometry and traffic characteristics of road segment (sample iRAP input)

Road name	Median type	Intersection type	Number of lanes per direction	Lane width	Pedestrian crossing facilities - inspected road	Vehicle parking	Sidewalk Pavement	Speed limit	Vehicle flow (AADT)
Moseley Rd S-1	Centre line	3-leg signalised	2	Narrow*	Signalised with refuge	One side	1.0m to <3.0m	30 mph	14506
Moseley Rd S-2	Centre line	3-leg signalised	1	Narrow*	Signalised with refuge	One side	1.0m to <3.0m	30 mph	14506
Moseley Rd S-3	Centre line	3-leg unsignalised	1	Narrow*	Signalised without refuge	One side	1.0m to <3.0m	30 mph	14506
Moseley Rd S-4	Centre line	3-leg unsignalised	1	Narrow*	No facility	One side	1.0m to <3.0m	30 mph	14506
Moseley Rd S-5	Centre line	3-leg unsignalised	1	Narrow*	Signalised without refuge	One side	1.0m to <3.0m	30 mph	14506

* ($\geq$ 0 m to < 2.75 m) In accordance with the iRAP input specification

The iRAP sample input of all attributes for road segments can be found in Appendix E. Pedestrian-vehicle crash data for the selected sites were collected for the last 5 years as recommended (iRAP, 2016) and then standardized to crash rates per km per year. This is given in Table 9.16.

9.6.3 Comparison test

For testing and comparison, the crash rates for the road segments under consideration were estimated using iRAP. The iRAP model can be used to make estimates of the number of expected fatalities as a function of the star rating score (SRS) and AADT for each 100-metre segment of the evaluated road under the existing conditions. Also, iRAP estimates the number of serious injuries' crashes (SI) as a function of fatalities as follows (iRAP, 2016):

$$SI = Calculated\ fatalities\ x \frac{actual\ number\ of\ serious\ injuries\ in\ the\ jurisdiction}{actual\ number\ of\ fatalities\ in\ the\ jurisdiction} \qquad 9.3$$

Both the actual and the estimated crash values need to be consistent; therefore, the actual crash data was transformed to the fatality-crash rate to be comparable with the iRAP estimates in the following two steps:

Firstly, the crash rates per km per year for pedestrians for the selected sites were computed for each 100-metre section based on the last 5 years of crash history as follows:

$$Crashes/km/yr = \frac{\text{Total crash frequency for the selected time span}}{Section\ length\ \text{x}\ Number\ of\ years} \qquad 9.4$$

Secondly, the actual fatality-crash rate values were calculated using a similar approach to Equation 9.4 but using all crash events instead of SI crashes as follows:

$$fatality\ rate\ = crash\ rate\ x\,\frac{actual\ number\ of\ fatality\ crashes\ in\ the\ jurisdiction}{actual\ number\ of\ total\ crahes\ in\ the\ jurisdiction} \qquad 9.5$$

In order to estimate fatality –crash rates in Equation 9.5, the number of fatality crashes in all road segments were divided by the number of total crashes in all road segments.

9.6.4 Results

The results shown in Table 9.16 indicated that the percentage of agreement between the predicted and actual pedestrian-vehicle fatality-crash rate on average was 32% when using iRAP estimates and 55% when enhanced methodology was used. This result shows a significant improvement in the level of agreement between actual values and predicted crash rates.

In terms of road segments, the proposed methodology provided better assessment and causality estimation for two-thirds (11 out of 15) of the studied road segments.

Also, the relationship between the SRS and crash rates was investigated by plotting the SRS estimated by iRAP against the enhanced iRAP and actual crash rates. For the road sites included in this study, the relationship could not be proved between the SRS and crash rates for pedestrians. These findings are in agreement with previous research conducted to test the potential applicability and feasibility of instituting a U.S. road assessment program (usRAP). The study attributed this finding to the variability of crash data and non-roadway factors that may influence the crash rates (Harwood et al., 2010).

Another reason could be the difficulty in comparing the crash rates and SRS for pedestrians as they are significantly varied with the number and type of crossing facilities. The presence of a pedestrian crossing may increase the SRS for the investigated 100-metre section (Lynam, 2012).

Table 9.16 Actual and predicted iRAP and enhanced methodology crash rate comparison

No.	Name of the road	Actual crash rate	iRAP		Enhanced methodology	
			Predicted crashes rate	Percentage correct %	Predicted crashes rate	Percentage correct %
1	Moseley-N Rd	0.040	0.019	47.50%	0.066	60.61%
2	Hagley-W Rd-A456	0.045	0.036	80.00%*	0.099	45.45%
3	Coventry-N Rd	0.049	0.027	55.10%*	0.117	41.88%
4	Pershore-N Rd	0.072	0.020	27.78%	0.067	93.06%
5	Strat-S Rd	0.077	0.058	75.32%*	0.183	42.08%
6	Hagley-E Rd-A456	0.090	0.179	50.28%*	0.419	21.48%
7	Coventry-S Rd	0.100	0.011	11.00%	0.061	61.00%
8	Pershore-S Rd	0.101	0.013	12.87%	0.038	37.62%
9	Strat-N Rd	0.102	0.026	25.49%	0.22	46.36%
10	Soho-E Rd-A41	0.113	0.046	40.71%	0.166	68.07%
11	Waterloo-Cape Hill	0.129	0.023	17.83%	0.11	85.27%
12	Dudley Rd	0.135	0.012	8.89%	0.058	42.96%
13	Moseley-S Rd	0.145	0.017	11.72%	0.148	97.97%
14	Alum Rock Rd	0.164	0.014	8.54%	0.069	42.07%
15	Soho-W Rd-A41	0.291	0.015	5.15%	0.106	36.43%

*Sites where the original iRAP predicted crash rates are better than those of the new model

Furthermore, the limited road sites included in this study, due to time and budget constraints, may have imposed an effect on the results. All in all, a large number of road sites and combinations of varied roadway types and crash rates would be needed to test the embedded risk factors and to investigate their effect further.

Chapter 10

DISCUSSION

10.1 Introduction

This chapter discusses the findings of this study with regard to the methods selected for the problem at hand and the contributing factors considered; the evaluation of the success of the model by testing the correlation between the contributing factors; and the validity of the developed models, their applicability; and their potential application and integration to enhance the current iRAP system. It also covers data issues with regard to their collection process and sample size; and furthermore, gives the limitations of this study.

10.2 Importance of Pedestrians' Protection and Risk Assessment

Virtually every traveller starts and ends as a pedestrian. Therefore walking, an essential part of a non-motorised transport mode, is very important especially in urban settings. Thus, promoting walkability through infrastructure design and investment can be beneficial for the environment and health. Despite the importance of this transport mode, road infrastructure in urban areas is still causing harm to pedestrians (Miranda-Moreno et al., 2011). In order to achieve safe walkable environments, practical tools are needed by transport professionals to assess and mitigate the influence of development on pedestrian safety, including pedestrian-vehicle accidents. Many of the traffic safety factors are related strongly to the roadway design (Abdel-Aty and Radwan, 2000).

Given that the cost of an accident is strongly based on the human casualty outcome and only minor costs could result from other damages (DfT, 2016c), and because the pedestrian is more fragile and the least physically protected road user (Downing et al., 2000, Eluru et al., 2008), traffic crashes involving pedestrians could cost more than other vehicle crashes. Therefore, they need to be reduced and addressed appropriately.

Pedestrian causalities are mostly preventable and proven intervention measures are available; yet successful interventions to make walking safer and protect pedestrians need an understanding of the risk factors related to pedestrian crashes. Therefore, a better assessment of the risks faced by pedestrians is important to achieve a safer walking environment and to reduce the risks. The safer environment for pedestrians has health, economic and social benefits including physical activities, reduction in traffic noise, improved air quality, social cohesion and economic growth (WHO, 2013b).

10.3 The model development

In spite of the importance of roadside activities in urban areas, it appears that there are no sufficient studies which have given attention to them and investigated their influence on pedestrians' safety.

As a consequence, this study developed a model to assess pedestrian safety and capture the effect of pedestrian and roadside activities' intensity. The study concluded that the number of bus stoppings per unit of time, parking, pedestrian crossing and violations' volume, the traffic speed variation, the number of intersecting side roads, in addition to through and intersecting traffic volume, were among the significant risk factors related to the pedestrian crash risk. The traffic operating speed was found to be the major risk factor determining the crash severity outcome. Furthermore, all significant findings of the developed model were embedded in the iRAP system to improve its assessment ability in urban settings and capture the risk effect of roadside activities. The importance of its improved assessment ability focuses on the identification of hazardous locations for pedestrians in urban areas. Identifying the safety problem is essential to introduce and design engineering interventions to promote safety. Also, this identification can be used to prioritise the road network for improvement programmes and to ensure cost-effectiveness. Ultimately, making walking safer and protecting pedestrians is important to reduce the social and economic burden of crashes and to encourage people to walk as a travel option.

10.3.1 The modelling method selection

Choosing the appropriate statistical modelling approach may depend on the data availability. For the model development in this study, the pedestrian-vehicle crash frequency was used as the dependent variable to predict site-specific pedestrian-vehicle crash risks. The number of accidents observed on road sections per time is considered as count data consisting of non-negative integers (i.e. the number of accidents is integer and cannot be a negative) (Cameron and Trivedi, 1998).

As stated in Chapter 2, statistical regression methods (GLM) were used because of their appropriateness to deal with count data and their ability to model simultaneous effects of multiple variables, including mixtures of continuous, count and categorical data (Myers et al., 2002; Lord and Mannering, 2010).

There are a number of generalised regression models (GLM) to model count data, the most common methods of which are Poisson, Negative Binomial and Zero-inflated Poisson regression. All the three methods were considered in the modelling, while a probability model was utilised to measure the relationships of the potential contributory factors on the severity of pedestrian crash outcomes.

After conducting the tests, the Poisson model was identified as better in terms of capturing the probability of the pedestrian-vehicle crash risk and was therefore selected over the NB and ZIGP models. This is because the over-dispersion scale parameter ratio was (1.011) indicating no significant over-dispersion as the value is closer to 1. This is attributed to the proximity between the values of crashes mean and standard deviation.

Among the strengths of the selected approach is that of providing realistic results, as no negative results are allowed by the form of the model. Also, the Poisson regression may provide an appropriate option to analyse crash data as crashes are infrequent events and tend to be non-normally distributed and skewed.

In terms of weaknesses, crash data may violate the assumption that the mean and variance of the data on which the model is based are equal. Therefore, special consideration for this point needs to be taken when using this approach.

For application purposes, this approach implies that a unit change in any included variable leads to a percentage change in the number of the dependent variable. This feature is consistent with the multiplicative form of the iRAP model; it also facilitates the calculation of the risk factor effects on the crash frequency outcome and can be embedded in the iRAP model individually. In addition, there is potential for the developed model to further improve by including a larger representative data set and a wider range of roadway configurations.

10.3.2 Contributing factors

Before including the selected independent factors in the study, the degree of correlation between each pair was examined. The results indicated that no variables needed to be excluded. Other factors that could be included, if a larger data set was available, are the percentage of heavy vehicles, the number and width of lanes, the crossing facilities and other geometric features. In this study, the number and width of lanes in addition to crossing facilities were included in the initial analysis and then omitted as they were found to be insignificant. This is possibly because of the similarity of these features along the selected sites. Also, the percentage of heavy vehicles was not significant and excluded from the analysis later; possibly because this percentage in the traffic composition was relatively low and ranged between 1.49% and 2.93% for all included sections in the study.

The traffic speed was found to have minor effect in the pedestrian crash frequency risk and a significant effect on the crash severity. These results were consistent with findings reported in previous studies (Wier et al., 2009; Al Kaaf, 2015). The severity outcome model has shown that the potential of a serious or fatal pedestrian crash is 59 percent of all crash events at sections at a speed of 30 mph. This can be a subject for further investigation as it provides a direct link between pedestrian crash severity and traffic operating speed, rather than impact speed, for an individual vehicle involved in a crash. Also, the comparisons between speed limits and the actual operating speed revealed that in some road segments with a reduced speed limit of 20 mph, some part of the road section had an actual operating speed higher than the speed limit. This may indicate the need for additional enforcement. The investigated road segments with reduced speed limits just have repeated speed limit signs and no further measures. As physical measures may not be appropriate for these roads, as they accommodate high traffic volumes, speed cameras could be a more appropriate and effective enforcement to reduce the speed limit violations.

The coefficient of speed variation was found to influence the crash frequency significantly. The finding appeared to be consistent with previous studies, however higher speed variation rates were found at selected sites; this is likely because pedestrian and roadside activities in addition to congestion may lead to higher speed differences. This significant effect may occur because the higher variability of traffic speed leads to shorter gaps in passing traffic and decreasing crossing opportunities. Also, the traffic movement becomes more complex and unpredictable. The importance of both speed factors can be used in deciding the speed reduction policies and measures; some measures might reduce the traffic operating speed but neglect the effect of speed variations, or conversely, reduce the speed variations and neglect the traffic operating speed.

The presence of vehicles parking on one or both sides and the higher number of bus stopping activities have been found to have significant effect and lead to an increase in the pedestrians' crash risk. This could be attributed to the risk of crossing in front of parked vehicles causing restricted visibility. The presence of parked vehicles or buses could block the sight distance, particularly where pedestrians try to cross the road; thus, affecting the required time for drivers to identify hazards and to react in order to avoid them.

Other variables that were found to be significant and where their increase leads to higher pedestrian-vehicle crashes, included traffic flow, intersecting traffic flow and the number of side roads. The reason being that higher traffic flow leads to more interaction with pedestrians and therefore higher a crash risk.

The findings of this study indicated that the volume of pedestrian and roadside activities influence pedestrian safety. The data for these variables were counted over a selected time intervals representative of all day. However, longer counts are recommended to cover a wider range of the independent variables' values.

10.4 Success of the New Model

This study developed a systematic methodology to assess the developed models. The evaluation of their success has been assessed firstly, by testing the correlation between the independent variables included in the model and secondly, by evaluating and validating the developed models with regard to whether they model the crash frequency appropriately.

10.4.1 Correlation between independent variable

The purpose of this test was to measure the degree of correlation between the independent variables which were considered in the modelling process. The variables were checked in terms of collinearity. Using the variance inflation factor (VIF), the degree of collinearity (or multi-collinearity) of each independent variable with the other independent variables was measured. The results show that all VIF values were below 5 (see Table 7.1) indicating that the considered independent variables were not strongly correlated and therefore there was no need to exclude variables (O'Brien, 2007).

10.4.2 Evaluation and validation of the developed models

The purpose of the evaluation and validation of the developed models was to examine how well the models fit the observations and perform using a different data set. Three evaluation measures were considered including the goodness of fit assessment, a graphical exploration of the observed versus the predicted crash numbers and a field validation. The results were as follows.

- **Goodness of fit assessment**

In this step, selected tests were conducted to assess the goodness of fit (GOF) and examine how well the observed estimation data set fits the outputs from the developed models. The likelihood ratio, the standardised likelihood ratio and the Akaike information criteria (*AIC*) were used as GOF measures (Washington et al., 2005; Al-Matawah, 2008; Washington et al., 2011). To apply these measures, the model was fitted initially using only the constant and excluding all variables. Following that, the model was then fitted by adding one variable at a time as shown in Table 8.5. In this process, the smaller models with fewer variables were considered as nested in the full model. The results indicated that the full model (which included all significant selected variables) shows a better fit in comparison to the other nested models with less variables. These findings suggest that the full model is more accurate and represents better risk estimates.

- **Observed and predicted crash numbers**

The purpose of comparing the observed with the predicted crash numbers was to examine the developed models based on actual data and graphically investigate the relation between the observed and the predicted crash numbers. This is conducted by estimating the percentage of the accuracy of estimation, using the predicted crash frequency calculated by the developed models, against the actual crash frequency data.

The comparison results indicated that there was a very good agreement between the predicted and actual pedestrian-vehicle frequency crash data, which was 81.3% when using the model with intersecting traffic flow and 83.6% when using the model with the number of intersecting side roads. Also, the graphical visualisations for both developed models in comparison with the actual crash rate values showed good similarity in the data trends.

Regarding the prediction accuracy for each road section, the results indicated a lower prediction accuracy at sections with less observed crashes while a higher accuracy rate observed in the case of sections with more crash events. The lowest prediction accuracy rate was 50 % for Moseley road where the lowest crash rate occurred, and the prediction accuracy rate was 80.4 % for Soho road –W where the highest crash rate occurred. This finding suggests that the developed model can give better results when used to estimate the crash frequency in high activities sections where higher crash events can be observed.

- **Model field validation**

The purpose of the field validation was to further test the developed models by examining the performance of the suggested models using different field data, through a comparison of the predicted crash frequency with the actual crash frequency observed in the validation road sections. The results with regard to the evaluation measures were as follows:

1- There was a good agreement between the predicted and actual pedestrian-vehicle frequency crash data; which was 87.27% when using the intersecting traffic flow model and 81% for the intersecting side roads model. The data visualisation test revealed similar trends.

2- In the case of less pedestrian and roadside activities road sections, the results show a lower level of agreement between the actual values and the predicted values than those found at sections with more pedestrian and roadside activities. The agreement between the predicted and actual pedestrian-vehicle frequency crash data was 73.46% when using the intersecting traffic flow model and 68.69% for the intersecting side roads model.

3- The mean squared prediction error (MSPE) and the mean squared error (MSE) were used to investigate how well the developed models fit the data set, by comparing the performance when using the estimation and the verification data sets. Both developed models show MSPE values lower than MSE values, indicating acceptable goodness of fit with regard to the validation data set.

The above evidence demonstrated the prediction ability of the developed models and their reliability and potential for assessing and ranking the risk of pedestrian-vehicle crash frequencies for environments with high roadside activities in urban areas. Accordingly, the developed models may be used to prioritise roads in need of improvement interventions.

However, a wider range of representative road sections and data sets could be considered for further assessment of the developed models.

10.5 Applicability

10.5.1 Application of the developed models

The study developed analytical models that assess pedestrian safety, by accounting for risk factors related to pedestrian and roadside activities' intensity and estimating their effects on vehicle-pedestrian crash likelihood and severity. The developed models appear to be suitable for application in urban areas with high road and roadside activities according to the evaluation and validation tests.

The data sets for the models' fitting were collected from Birmingham' urban roads in the UK. However, the ultimate objective was to test the transferability of the model to developing countries. There are differences between the developed and developing countries and the latter have more severe problems with pedestrian road safety (WHO, 2015). Thus, the roads were selected in such a manner so that their conditions may be comparable and the models can therefore be transferred.

It is also felt that the investigation of the proposed risk factors and examination of their effect in developing countries might give a better insight into pedestrian road safety, especially in road links accommodating high roadside activities. Therefore, it would be desirable to further test and possibly calibrate the findings of this study in developing countries. This could be performed by selecting representative roads accommodating high pedestrian and roadside activities which are to some extent comparable with UK roads, in terms of road configuration, to test the model development and the effect of the variables.

10.5.2 Potential application to enhance the existing iRAP system

Another potential application proposed by this study is to use the developed models to enhance the current iRAP system in its ability to assess the safety of pedestrians.

The developed models were used to develop risk factors which were integrated with the existing iRAP system using its multiplicative model form. The enhanced model included risk factors associated with road links with high roadside activities and it was demonstrated that the safety assessment ability was improved for the road segments considered.

Generally, the comparisons between the proposed models and the existing iRAP indicated that the percentage of agreement between the predicted and actual pedestrian-vehicle fatality-crash rate increased from 32% (when using iRAP estimates) to 55% when the proposed methodology was used. In terms of road segments, the proposed methodology provided better assessment and causality estimation for two-thirds (11 out of 15) of the studied road segments.

For further investigation, the studied road segments may be separated into two categories, segments with actual crash rate less than 0.1 (6 segments) and segments with actual crash rate equal or more than 0.1. When testing the prediction ability for segments with lower actual crash rates (less than 0.1), the existing iRAP showed better accuracy for 4 out of 6 segments. However, when testing the prediction ability for segments with higher actual crash rate (equal or more than 0.1), the proposed methodology provided better assessment and causality estimation for all road segments (9). Also, the iRAP prediction accuracy for segments with actual crash rate equal or more than 0.1 reduced to 16% in comparison with 58% prediction accuracy for the proposed methodology.

This result suggested that iRAP crash estimation ability reduced when used in dense urban roads where higher pedestrian crashes occurring, whereas the proposed methodology can give better results when used to estimate the crash frequency in high activities sections where higher crash events can observed. Therefore, the proposed models may be used to assess roads in terms of the pedestrian-vehicle crash frequency risk in urban environment with high pedestrian and roadside activities where high pedestrian crash events observed.

In terms of the iRAP star rating score, the relationship could not be demonstrated between star rating score (SRS) and crash rates for pedestrians. This could be attributed to the variability of the crash data and the non-roadway factors, such as different road users' behaviours, which may influence the crash rates. However, it could be tentatively suggested that the full integration of the newly developed models could offer an iRAP pedestrian modelling technique which is more sensitive and more accurate than the current one.

Also, the iRAP SRS for pedestrians is significantly varied with the number and type of pedestrian crossing facilities. The presence of crossing facilities could decrease the star rating score (SRS) for the considered 100-metre section and possibly may affect the high crash risk caused by other risk factors.

The multiplicative nature of the iRAP facilitates the integration of new factors. The iRAP model improvement is based on a modular approach and therefore could be transferred and further enhanced to describe new conditions. It is believed that the integration of the risk factors proposed in this study into the iRAP methodology will improve its risk prediction ability. However, further investigation in this regard is needed to reflect these improvements in the iRAP SRS outputs. A large number of road sites and combinations of varied roadway types and crash rates would be necessary to test the embedded risk factors and to investigate further the application of the methodological framework presented in this study. Other data sets with varied road configurations in terms of number and widths of lanes, crossing facilities, median arrangements and a wider range of considered variables could be useful to ensure the findings.

10.6 Data Issues

10.6.1 Data collection

The present study used 12 road segments of 0.8-1.2 km lengths, where extensive pedestrian and roadside activities were observed. The selected segments may seem relatively short. This is because it was practically difficult to find long road sections with high pedestrian and roadside activities. Yet, despite the short lengths, these road segments experience high crash rates and therefore were suitable for the model development and testing.

The data for the model development was collected with regard to specific road segments of a length of 100 metre. The data included crash data, traffic volume, segments' characteristics, traffic operating speed, speed variance, speed limit and pedestrian and roadside activities. Data were also collected about pedestrians and roadside activities; including parking on roadsides, bus stopping activities, parking and un-parking to load and unload passengers, pedestrians walking along flow and crossing volumes, and pedestrian violations, which refers to pedestrians crossing out of allocated facilities or times.

Given that the pedestrian volumes and roadside activities for the study sections were unavailable, it was essential to conduct a field inventory and make observations along the selected road segments. The data for pedestrian movement, crossing and violations and parking and un-parking manoeuvring activities were collected using a sampling counts procedure.

The enhanced methodology proposed introducing additional risk factors to the existing iRAP and therefore, a cost-effective method for related data collection is needed. The existing iRAP related road attributes data is normally acquired by conducting road surveys using cameras fitted on a survey vehicle. Traffic and pedestrian volume and speeds are collected separately for the considered road sections using local knowledge or site measurements. The data requirements for the newly introduced risk factors of the number of intersecting roads and their traffic volumes, and the pedestrian flow and parking vehicles can be collected by the same current procedure taken by iRAP. The traffic speed variance data can be collected using the same resources or measurements as those for the operating speed. The data related to the number of bus stoppings may possibly be acquired from local knowledge or sample counts. Therefore, the data collection methodology followed in this study may be considered as sustainable and also transferable to similar road conditions as it had minimal cost and did not require additional or expensive equipment.

Data about pedestrian flows and activities could be sparse or not available and their significant impact on safety assessment highlights the need to collect and maintain reliable data. Local knowledge and estimates in addition to sample counts can be combined to maintain a better data set for pedestrian activities.

The needed data to estimate the relative risk factors was aggregated into ranges empirically for simplicity. Nevertheless, the relative risk factors presented in this study were based on a simplified methodology which may need further improvements in terms of accuracy.

10.6.2 Data sample size for model coefficient estimation

The development of the models in this study used data from 12 sites, were segmented into 100-metre sections making 118 sections in total. The pedestrian crash data collected for the above sections counted 320 crash events in total. For comparison and contrast, some previous recommendations regarding the suggested sample size for crash prediction model estimation has been examined. The Highway Safety Manual (HSM) recommends that to test prediction models, 30-50 locations may be used, where locations refer to intersections or road links regardless of length. Further sample size guidelines recommended by Shirazi et al., (2016) for the highway safety manual which proposed selecting the number of data locations for prediction models with regard to the coefficient of variation, which refers to the ratio of the standard deviation to the mean of the observed crashes. This guidelines suggesting using 30 sections when coefficient of variation is less or equals to 0.6, 100 sections for the coefficient of variation value of 0.8 -1.0 and as the coefficient of variation increased more road sites recommended. However, these guidelines were recommended for the highway safety manual (HSM).

The HSM model provides a base condition crash modification factors (CMFs) which needed to be calibrated for any location using the recommended road sites sample size. The calibrated CMFs quantify the change in expected average crash frequency as a result of geometric or operational modifications to a site that differs from set base conditions. Therefore, the sample size used to estimate crash modification factors for a wide range of than 50 variables and for segments (both divided and undivided) and for intersections (signalised and stop controlled intersections) for the following facility types:

- Rural Two-Lane, Two-Way Roads
- Rural Multilane Highways
- Urban and Suburban Arterials

In effect, these recommendations are not comparable to this study as the included factors are less and only urban undivided road segments were included.

Another recommendation by Lord and Miranda-Moreno (2008) proposed the use of the crash rate mean to determine the number of required sample road sections for crash model development. The recommended guidelines suggested using 20 road sections if the crash mean is equal or less than 2.0. In this study, the crash sample mean per road section was 2.75; hence, 20 road sections or less were recommended for model development.

In addition, in this study the model was estimated using 12 road segments where each one divided into 100-meter sections. The test of homogeneity conducted as in Chapter 7 (section

7.3.2) revealed significant differences between the 100-meter sections within each group in terms of crash rates, traffic speed and bus stoppings activities.

Therefore the above recommendations, in addition to the statistical analysis of the correlation between the independent variables and the developed models' evaluation and validation, demonstrate that the quality of the data set was high, and the quantity were good and representative of a wide range actual conditions.

10.7 Implementing Pedestrian Safety Interventions

10.7.1 Role of roadway design and urban planning

Roadway designs intended for the needs of motorists are generally viewed to be at odds with the needs of pedestrians' safety (Dumbaugh and Wenhao, 2010).

Also, the roadway network design can both reduce and increase the pedestrian-vehicle crash risk depending on the extent to which facilities such as crossings and sidewalk pavements are provided.

Hence, the road network should be designed to meet the requirements of all road users. This study provides steps towards this purpose by underlining the affecting factors on pedestrian safety and proposes a modelling approach to assess the potential risk in roads; where this risk may be reduced or mitigated by sustainable engineering countermeasures.

Apart from the roadway design elements, the planning and land-use types (i.e. commercial, educational, residential, industrial or mixed) can contribute to the risk of pedestrian crash occurrence. Some land-use factors such as population and density, land-use mixes and activities' location and intensity can influence risks for pedestrians. Additionally, the extent to which pedestrians' facilities are considered and provided by planning and land-use has a significant effect on the crash risk for pedestrians.

It is recommended that roadway design and land-use practices should aim not only to accommodate the pedestrian's needs to access services such as retail shops, hospitals, schools, social meeting places and public transport, but also to improve their safety (WHO, 2013b). Therefore, a specific consideration to pedestrian safety is needed in land-use planning and road design.

10.7.2 Countermeasure selection

Better selection and planning of appropriate countermeasures and effective interventions depend on assessing the pedestrian safety problem. The objective of safety measures is to improve walking safety and also contribute to environmental, economic and social benefits. However, there are costs implications and the selection of countermeasures needs to be prioritised to address the risk factors. Safety interventions are generally categorised into engineering measures and behavioural change measures. However, a significant improvement to pedestrian road safety may be achieved by using a balanced approach of both measures (Fylan and Stradling, 2014). This study showed that a number of measures leading to a reduction in vehicle pedestrian conflicts and crashes include: pedestrian fencing, sight distance improvements, parking improvements, pedestrian refuge island, speed management (WHO, 2013b; iRAP, 2016).

Speed management is applied to make drivers reduce their speed by employing engineering, enforcement and educational measures. The engineering approach consists of a range of traffic calming physical treatments such as speed humps, raised platforms and roundabouts.

As shown by this study, special attention to speed measures needs to be taken to address traffic operating speed and speed variation (i.e. measures that could reduce the operating speed and increase the speed variations). The application of a reduced speed limit of 20 mph could be such an example, but absence of enforcement may encourage road users to show different level of compliance and consequently higher speed variation.

In addition, there are vulnerable groups within the pedestrians such as children, elderly people, disabled and other groups with special needs. Some of the vulnerable groups may have difficulties in hearing or seeing the approaching traffic, others may take longer to cross the roads due to their decreased mobility or awareness and judgment skills. Therefore, countermeasures and road design should account for the vulnerable groups and be tailored to meet their special needs' consideration. It may be necessary therefore to extend the scope of this study to include specific consideration for the most vulnerable road users.

10.7.3 Behaviour of pedestrians and drivers and the role of enforcement

The understanding of pedestrian behaviour features such as crossing behaviour, speeds, conflict and interaction with traffic is useful to understand pedestrian activities and associated risk (Zegeer and Sandit, 2009). Another important behavioural feature is the compliance of both pedestrians and drivers to road rules. Road users obey the rules only if they perceive that violations are penalised effectively. Therefore, effective enforcement measures should be applied. This is critical to pedestrian safety and to ensure the effectiveness of interventions. To ensure effective enforcement, the perceived possibility of being penalised for violating road rules and regulations should be high. Also, the unwanted outcome of being caught for violation should be enough to discourage road users from committing unlawful behaviour. Pedestrian and driver behaviour features could be established using observational studies, cameras and video recording. An example of critical compliance behaviour to be investigated may include drivers' yield behaviour at pedestrian crossings and compliance with speed limits. In this context, one of the findings in this study shows speed limit violations in some road segments with a reduced speed limit of 20 mph, where a part of the road section had an actual operating speed higher than the speed limit. The enforcement measure could be one or a mix of physical measures, police enforcement programs and speed radars. For a high degree of effectiveness, the enforcement measures need to be introduced with an awareness campaign, applied over a wide range, be difficult to avoid, be highly visible and unpredictable (iRAP, 2016).

10.8 Limitations of this Study

The main contribution of this study is to quantify the safety of pedestrians moving in environments with high roadside activities in urban areas. Nevertheless, the model development has certain limitations which may be summarized as follows:

- For the model development, the selected 12 road segments were further segmented into 100-metre sections because the road attributes may change significantly over urban areas. The 100-meter section considered as the unit of the analysis and the basis for data collection and subsequently model development. One limitation of this approach could be that there are no significant differences between the groups of 100 m road sections within each site. However, the qualitative investigation showed significant variations in the sections` attributes. A further statistical analysis test concluded that the 100-meter sections within each selected site showed statistically significant variation albeit being adjacent within each site. Therefore, the 100-meter sections within each site used individually to provide inputs for the model development with confidence.

- The length of the road sections (100 metres) could be considered short for data collection and subsequently model development as the pedestrian crash events could be rare. For instance, Kulmala (1995) indicated that the crash model development process can be challenging if the crash rate per site is lower than 0.2. However, these road sections experience average pedestrian crash rates of 2.75 during the counting periods, which is equivalent to a 0.34 pedestrian crash rate per year per 100-metre section. Therefore, modelling based on 100-metre lengths would seem satisfactory, as the crash rate per road section is higher than 0.2.
- The Study showed that the pedestrian and roadside activities influence pedestrian safety. However, the data for some variables which required site counts were collected over sample points for each road segment. Therefore, for each road segment of the selected study 12 sites, at least one road section (100 metres) was taken to provide the counts for the activities and pedestrian volume. Given that the pedestrian and roadside activities for the study sections were unavailable, the sampling count observations used to collect the needed data. Although this approach has some limitations, such as the pedestrian and roadside activities may vary along the different sections, it was felt it provided a sample of pedestrian activities and volume data which were acceptable estimates of the prevailing conditions. Also, such approach based on real-time observations may provide more representative data than those based on estimates and local knowledge.
- For pedestrian and roadside activities, the data collection was conducted between 12 and 4 p.m. during October. Although the pedestrian roadside activities may vary with

the hours of the day and month, the time was selected to capture average busy hours and during October where the weather is moderate and no season holidays impact the working days.

- Regarding the area characteristics and road configuration, this study investigated only road links where a high intensity of pedestrian and roadside activities were taking place. Also, specific road geometric characteristics such as different median types or horizontal and vertical alignment were not considered. Drivers' and pedestrians' behaviour may vary in different road configurations; thus, other areas and roads types should be included and investigated, if the scope of the research were expanded.
- For pedestrian crash severity outcome, only speed was found to be significant. Other factors may relate to the severity risk including age, gender and time variations but were not included as they were beyond the considered scope of this study.
- The pedestrian crash frequency model was developed using all pedestrian-vehicle crash severities; with the view to estimate risk factors for integration with the iRAP methodology, which aims to identify the risk fatalities and severe injuries but excluding slight injuries. Despite the difference in the considered crash severities between the developed model and the iRAP, it was felt that any recorded pedestrian-crash event would be important for inclusion in the analysis because the pedestrian is more fragile and the least physically protected road user (Downing et al., 2000; Eluru et al., 2008), and severe injuries and fatalities are too scarce.

Chapter 11

CONCLUSIONS AND RECOMMENDATIONS

11.1 Conclusion

Pedestrian safety is a key aspect in road safety management. Yet to date there have been a limited number of models satisfactorily addressing pedestrian movements in the road environment and subsequently suggesting measures to make roads safer for them. This study has aimed at developing such a model with emphasis on roadside activities in an urban environment. The work followed a systematic approach which was based on a representative data set; analysed the problem at hand, suggested appropriate mathematical models; and tested these models.

Furthermore, it proposed the integration of the models with iRAP, an established road safety management system, demonstrated how this may be achieved and highlighted their successes if used in practice. The integration was systematically tested and validated, and it was found that the new model provided better results compared to the original iRAP.

The main conclusions from this work may be presented as follows:

1) Model development

a. Model appropriateness

i. This study provided a reference methodology for pedestrian safety assessment in developing countries with high pedestrian and roadside activities, a diverse traffic mixture and a high share of pedestrians involved in road traffic fatalities and injuries.

ii. This study examined the influence of factors affecting pedestrian safety in environments with high roadside activities. It found that these factors were: traffic volume, traffic operating speed, coefficient of speed variance, number of pedestrian crossing violations, number of parking and un-parking activities, number of pedestrian crossings, number of bus stoppings, percentage of heavy vehicle traffic, intersecting traffic volume, roadside parking, the number of intersecting side roads and the number of pedestrians walking along the roads.

iii. The developed models may be used to assess roads in terms of the pedestrian-vehicle crash frequency risk and to prioritize them for subsequent improvement interventions.

b. Model for crash risk

i. Generalized linear modelling (GLM) techniques are suitable to model pedestrian risks subject to comparison between the mean and variance of actual crash data.

ii. The Poisson model appeared to be the most suitable for pedestrian crash frequency prediction.

iii. Two models were developed to estimate the pedestrian crash frequency risk; the first model included the variable of intersecting traffic volume and the second model included the variable of the number of intersecting side roads, in addition to other chosen variables.

iv. It was found that traffic operating speed has an insignificant effect on crash frequency risk.

v. The most influential factor in crash frequency risk was found to be the number of intersecting side roads.

vi. The coefficient of speed variation was found to influence the crash frequency significantly, suggesting that higher speed variations in a road segment increase the accident frequency.

vii. The presence of vehicle parking on one or both sides of the road, and the number of bus stopping activities were found to have a significant effect on and lead to an increase in the pedestrians' crash risk.

viii. It was found that other significant factors affecting crash risk are the traffic flow and intersecting traffic flow, together with the number of pedestrians and roadside activities. Their increase leads to a higher number of pedestrian-vehicle crashes.

ix. The number of bus stopping activities per time unit per road section should be used instead of the presence of bus stops as an input to pedestrian-vehicle crash frequency prediction models.

c. Model for crash severity

i. A probability model was developed to estimate crash severity using the logistic regression technique.

ii. The traffic operating speed was found to be the only significant factor explaining the variation in pedestrian crash severity.

iii. Using the crash severity model, it was demonstrated that the potential of serious or fatal pedestrian crashes is 59 percent of all crash events at speed of 30 mph.

d. Model testing

The proposed independent variables and the developed crash risk models were systematically tested and assessed; and the results were as follows:

i. The considered independent variables were not strongly correlated and therefore they were all included in the modelling process.

ii. The goodness of fit assessment suggested that the full model, which includes all significant variables, is more accurate and represents better risk estimates than the smaller models with less variables included.

iii. The graphical exploration of the observed versus the predicted crash numbers indicated that their agreement was good and showed good similarity in the data trends for both the developed models of the crash risk.

iv. The validation assessment for both developed models indicated that there was a very good agreement between the predicted and actual pedestrian-vehicle frequency crash data and a good similarity between the data trends was shown.

v. Both of the developed models for the crash risk demonstrated an acceptable goodness of fit for the validation data set.

e. Data used

i. Twelve representative road segments, where extensive pedestrian and roadside activities occur, were used for model development.

ii. Due to practical difficulties, the road sections considered were of relatively short lengths of 0.8-1.2 km.

iii. A simple methodology was developed for the preliminary examination of the data set gathered. It was found that:

- The selected road segments were experiencing high crash rates and therefore were suitable for the model development and testing.
- The study 's data size in terms of the number of road segments was satisfactory and larger than the minimum sample size.
- The quality of the data set was high, and the quantity was good and representative of the actual conditions.

2) Model application

a. Integration with iRAP

- The developed model for the estimation of crash risk factors of roadside activities was integrated in the iRAP star rating system.
- The integration introduced an enhanced model framework to improve the assessment of pedestrian safety in urban environments.
- The integration of the developed risk factors within the iRAP system improved the degree of agreement between the predicted and the actual crash rates.
- The proposed methodology provided better assessment and causality estimation for two-thirds (11 out of 15) of the studied road segments.

b. Data requirements

The data collection methodology followed in this study may be considered as sustainable and also transferable to similar roads for the following reasons:

i. The data for the newly introduced risk factors of number of intersecting roads and their traffic volumes, pedestrian flow and parking vehicles can be collected by the same data collection procedure currently used by iRAP.

ii. The traffic speed variance data can be collected using the same resources or measurements as those for the operating speed.

iii. The data related to the number of bus stoppings may possibly be acquired from local knowledge or sample counts.

3) Pedestrian safety improvement

i. The actual operating speed may be noticeably higher than the speed limits, even on roads with reasonable enforcement and a reduced speed limit of 20 mph.

ii. The number of pedestrian fatalities was found to be significantly higher than other vehicle crash fatalities on the studied roads.

iii. Both the traffic operating speed and the coefficient of speed variation need to be used when deciding speed reduction policies and associated measures. It is important to implement measures that reduce the traffic operating speed together with the speed variations.

iv. A specific consideration for pedestrian safety is needed in land use planning and road design.

v. The safety countermeasures and road design should account for the vulnerable groups and be tailored to meet their special needs.

vi. With regard to the risk variables, a range of countermeasures could be considered to reduce and mitigate the pedestrian casualties. These countermeasures may include physical fencing to prevent unwanted pedestrian crossing; improving the sight distance; removing obstacles, particularly where pedestrians try to cross; reviewing and improving the on-street parking and occasionally removing it; and also introducing speed management techniques and appropriate enforcement to ensure that road rules will be obeyed.

vii. Any infrastructure specific measures to increase road safety should be accompanied by an appropriate enforcement regime.

viii. To ensure effective enforcement, the perceived possibility of being penalised for violating road rules and regulations should be high.

ix. For a high degree of effectiveness, any enforcement measures need to be introduced with a wide-ranging awareness campaign and be difficult to avoid.

4) Model use

The developed models provide traffic engineers and safety managers with the ability to rank the roads according to pedestrians' crash risk scores and may be used to evaluate existing roads, prioritize roads in need of safety interventions, design new roads, or redesign existing roads. Specifically, the proposed model may be used in the following ways:

a. The pedestrian crash prediction model developed in this study provides a tool to evaluate pedestrians' safety along street segments by:

    i. Identifying the contribution of risk factors to pedestrian crash frequency;
    ii. Predicting pedestrian crash risk and estimating the effect of any change in roadside activities and their implications;
    iii. Modelling pedestrian and roadside activities' intensity and estimating their effects on pedestrians' crash risk.

b. Introducing, ultimately, appropriate regulations and enforcement programs to address pedestrian safety problems.

c. Assisting transportation planners and traffic engineers to address pedestrian safety problems through improvement of traffic conditions and infrastructure management.

d. Evaluating, ultimately, the economic consequences of pedestrian injuries and fatalities and highlight the social burden of pedestrians' crashes.

11.2 Areas for Further Studies

The experience acquired from this work and its constraints indicated that further research could be conducted to investigate the following:

1. The use of a wider range of road configurations, segments and data range to further assess, test and calibrate the developed models.
2. The effect of specific engineering countermeasures and interventions to mitigate the pedestrian crash risks related to roadside activities.
3. The effect of roadside activities in low and middle-income countries on pedestrian safety, with a view to capture the influence of different road users' behaviour and road design aspects.
4. The effect of the presence of heavier vehicles in the traffic mix on pedestrian safety, using more representative data and with emphasis on low and middle-income countries.
5. The common types and details of pedestrian crashes associated with roadside activities in low and middle-income countries.
6. The effect of age and gender of pedestrians on crash frequency and severity.
7. In-depth investigation of pedestrians' crossing behaviour and the effect of an associated distraction such as smart phones.
8. Include specific consideration for the most vulnerable groups within pedestrian road users in the developed models, such as children, the elderly, the disabled and other groups with special needs; and identify effective interventions to address their specific needs.

REFERENCES

AASHTO, 2010. *Highway Safety Manual, First Edition*. American association of state, highway transportation officials, Washington DC.

Abdel-Aty, M., Wang, X., 2006. Crash estimation at signalized intersections along corridors: analyzing spatial effect and identifying significant factors, Transportation Research Record: *Journal of the Transportation Research Board*, No. 1953, TRB, National Research Council, Washington, DC, pp. 98–111.

Abdel-Aty, M.A., Radwan, A.E., 2000. Modelling traffic accident occurrence and involvement. *Accident Analysis and Prevention* 32, 633–642.

Agarwal, N., 2011. *Estimation of pedestrian safety at intersections using simulation and Surrogate safety Measures*. University of Kentucky Doctoral Dissertations.

Al Kaaf, K., Abdel-Aty, M., 2015. Transferability and calibration of highway safety manual performance functions and development of new models for urban four-lane divided roads in Riyadh. Transportation Research Record: *Journal of the Transportation Research Board*, No. 2515, Transportation Research Board, Washington, D.C.,

Al-Ghamdi, A.S., 2002. Pedestrian–vehicle crashes and analytical techniques for stratified contingency tables. *Accident Analysis and Prevention*, 34, pp.205–214.

Al-Matawah, J.A., 2008. *An investigation of driver attitudes towards road safety in Kuwait.* PhD dissertation. University of Southampton. Southampton, UK.

Al-Omari H, B., Obaidat S, E., 2013. Analysis of pedestrian accidents in Irbid city, Jordan. *Open Transportation Journal* 7, pp 1-6.

Al-Twaijri, S., Quddus, M., Bristow, A., 2011. Factors affecting the severity of traffic crashes in Riyadh City. *Proceedings of the 90th Annual Meeting of the Transportation Research Board,* Washington, D.C.

Anastasopoulos, P.C., Shankar, V.N., Haddock, J.E., Mannering, F.L., 2012. A multivariate tobit analysis of highway accident-injury-severity rates. *Accident Analysis and Prevention* 45, 110–119.

Anderson, R., Mclean, A., Farmer, M., Lee, B., Brooks, C., 1997. Vehicle travel speeds and the incidence of fatal pedestrian crashes. *Accident Analysis and Prevention*, 29 (5), pp. 667-674.

Ansari, S., Akhdar, F., Mandoorah, M., Moutaery, K., 2000. *Causes and effects of road traffic accidents in Saudi Arabia*. Public Health, 114, pp.37–40.

Archer, J., 2005. *Indicators for traffic safety assessment and prediction and their application in micro-simulation modelling: A study of urban and suburban intersections*. Doctoral Dissertation, Royal Institute of Technology, Department of Infrastructure, Division for Transport and Logistics, Centre for Transport Research.

Asadi-Shekari, Z., Moeinaddini, M., Zaly Shah, M., 2015. Pedestrian safety index for evaluating street facilities in urban areas. *Safety Science*, 74 , pp. 1–14

Ashton, S.J., Pedder, J.B., MacKay, G.M., 1977. Pedestrian injuries and the car exterior. *SAE Paper* 770092, 1977.

Baltes, M., Chu, X., 2002. Pedestrian level of service for midblock street crossings. *Transportation Research Record*, 1818, 125–133.

Birmingham City Council, 2016. *Transport policy. Average speed data*. Birmingham, UK.

Box, P., 2002. Angle parking issues revisited. *ITE Journal. Institute of Transportation Engineers*, Washington, D.C., pp. 36-47.

Brude, U., Larson, J., 1993. Models for predicting accidents at junctions where pedestrians and cyclists are involved. How well do they fit? *Accident Analysis and Prevention* 25, pp. 499-509

Cameron, A.C., Trivedi, P.K., 1990. Regression-based tests for overdispersion in the poisson model. *Journal of Econometrics* 46, pp. 347-364.

Cameron, A.C., Trivedi, P.K., 1998. *Regression analysis of count data*. Cambridge University Press, Cambridge.

Carson, J., Mannering, F., 2001. The effect of ice warning signs on accident frequencies and severities. *Accident Analysis and Prevention* 33, 99–109.

Carter, D., Srinivasan, R., Gross, F., and Council, F., 2012. *Recommended protocols for developing crash modification factors.* American Association of State Highway and Transportation Officials (AASHTO), NCHRP 20-7 (314), Final Report.

Carter, D.L., Hunter, W.W., Zegeer, C.V., Stewart, J.R., Huang, H.F., 2006. *Pedestrian and bicyclist intersection safety indices: Final report*. FHWA-HRT-06-125. U.S. Department of Transportation.

Chand, Sai, Chandra, S., 2014. Impact of bus stop on urban traffic characteristics –a review of recent findings. *Journal of Society for Transportation and Traffic Studies (JSTS),* 5(2), pp.57–72.

Chapman, R., 1973. The concept of exposure. Accident *Analysis and Prevention* 5, 95-110.

Chapman, R., 1978. *Accidents on urban arterial roads*. TRL report (LR838). Transport Research Laboratory.

Chatfield, C., 1983. *Statistics for Technology*. Chapman and Hall/CRC, London, England.

Cheng, W, Zhang, N., Wei, L., Jianfeng , X., 2014. Modeling and application of pedestrian safety conflict index at signalized intersections. *Discrete Dynamics in Nature and Society*,Volume 2014, 1–6.

Chiguma, M.L., 2007. *Analysis of side friction impacts on urban road links. Case study, Dar-es-Salaam.* Doctoral Thesis, Royal Institute of Technology, Stockholm, Sweden.

Chin, H.C., Quddus, M.A., 2003. Modeling count data with excess zeroes: An empirical application to traffic accidents. *Sociological Methods and Research* 32, 90–116.

Chiou, Y.C., Fu, C., 2013. Modeling crash frequency and severity using multinomial-generalized Poisson model with error components. *Accident Analysis and Prevention* 50, 73–82.

Christopoulou, P., Pitsiava-Latinopoulou, M., 2012. Development of a Model for the Estimation of Pedestrian Level of Service in Greek Urban Areas. *Procedia - Social and Behavioral Sciences*, 48, pp.1691–1701.

Cinnamon, J., Schuurman, N., and Hameed, S. M., 2011. Pedestrian injury and human behaviour: Observing road-rule violations at high-incident intersections. *PLoS One.* Vol 6: e21063.

Clifton, K.J., Kreamer-Fults, K., 2007. An examination of the environmental attributes associated with pedestrian-vehicular crashes near public schools. *Accident Analysis and Prevention* 39, 708–715.

Derry, J., Ebel, B.E., Mock, C.N., Afukaar, F., Donkor, P., 2010. Pedestrians' injury patterns in Ghana. *Accident Analysis and Prevention* 42, 1080–1088.

DFT 2016a. *Department for Transport, Traffic Counts.* Available at: http://www.DfT.gov.uk/traffic-counts/area.php/ [Accessed 10 Aug. 2016].

DFT, 2016b. *Department for Transport. CrashMap.* Available at: http://www.crashmap.co.uk/(Last accessed 12/03/2017).

DFT, 2016c. *Department for Transport. Reported road casualties in Great Britain 2016 annual report.*

DMRB, 1981. *Design Manual of Roads and Bridges*, TA 22/81. Vehicle speed measurement on all purpose roads, London, UK.

Downing, A., Jacobs, G., Thomas, A., Sharples, J., Silcock, D., Lottum, C., Walker, R., Ross, A., 2000. *Review of road safety in urban areas. Project report PR/INT/200/00*. TRL, Berkshire.

Dumbaugh, E., Li, W., 2011. Designing for the safety of pedestrians, cyclists, and motorists in urban environments. *Journal of the American Planning Association*, Vol. 77, No. 1, pp. 69-88.

Ekman, L., 1996. *On the treatment of flow in traffic safety analysis: a non-parametric approach applied on vulnerable road users*. Department of Traffic Planning and Engineering, University of Lund, Bulletin 136, Lund, Sweden.

El-Basyouny, K., Sayed, T., 2009. Collision prediction models using multivariate Poisson-lognormal regression. *Accident Analysis and Prevention* 41, 820–828.

Eluru, N., Bhat, C.R., Hensher, D.A., 2008. A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes. *Accident Analysis and Prevention* 40, 1033–1054

Elvik R., 2009. The non-linearity of risk and the promotion of environmentally sustainable transport. *Accident Analysis and Prevention* 41, pp. 849-855

Elvik, R., 2014.*Speed and road safety – new models.* Institute of Transport Economics, TOI report 1296, Oslo, Norway.

Elvik, R., Christensen, P., Amundsen, A., 2004. *Speed and road accidents. An evaluation of the power model.* Institute of Transport Economics, TOI report 740, Oslo, Norway.

Elvik, R., Hoye, A., Vaa, T., Sorensen, M., 2009. *The handbook of road safety measures,* Emerald Group Publishing Limited, second edition.

Finch, DJ, Kompfner, P, Lockwood, CR, Maycock, G., 1994. *Speed, speed limits and accidents.* TRL Project Report 58, Transport Research Laboratory, Crowthorne, UK.

Frith, J., Patterson, L., 2001. Speed variation, absolute speed and their contribution to safety, with special reference to the work of Solomon. *Proceedings of the IPENZ Transportation Group Technical Conference:* 1-8. Wellington, NZ: Institute of Professional Engineers New Zealand.

Fu, C., 2015. *Modeling spatial-temporal dependencies of crash frequency and severity: multinomial-generalized poisson models*. Ph.D. dissertation. Department of Transportation and Logistics Management, National Chiao Tung University, Taipei, Taiwan.

Fylan, F., Stradling, S., 2014. Behavioural change techniques used in road safety interventions for young people. *European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée* 64, pp. 123-129.

Garber, J., Ehrhart, A., 2000. *The effect of speed, flow, and geometric characteristics on crash rates for different types of virginia highways*. Virginia Transportation Council.

Garber, N., White, T., 1996. Guidelines for commercial driveway spacing on urban and suburban arterial roads. *Conference Proceedings. 2nd National Access Management Conference.* Vail, Colorado.

Garber, N.J., Gadiraju, R., 1989. Speed variance and its influence on accidents. *Transportation Research Record*, 1213, pp. 64-71, Washington D.C., USA.

Geyer, J., Raford, N., Pham, T., Ragland, D.R., 2006. Safety in numbers. Data from Oakland, California. *Transportation Research Record: Journal of the Transportation*. Research Board, No. 1982, Washington D.C., pp. 150-154.

Greibe, P., 2003. Accident prediction models for urban roads. *Accident Analysis and Prevention* 35, pp. 273-285

Gross, F., Bhagwant Persaud, B., Lyon, C., 2010. *A guide to developing quality crash modification factors.* Federal Highway Administration, Washington, D.C. Report No. FHWA -SA-10-032.

Haddon, W. Jr. 1980. *Advances in the epidemiology of injuries as a basis for public policy.* Public Health Report, 95 : 411–421

Hall, R., 1986. *Accidents at four-arm single carriageway urban traffic signals*. Transport and Road Research Laboratory, Crowthorne, UK.

Hardin, J., Hilbe, J., 2001.*Generalized linear models and extensions*. Stata Press, College Station, TX.

Harwood, D. , Torbic, D. , Gilmore, D. , Bokenkroger, C. , Dunn, J. , Zeeger, C., et al., 2008. *Pedestrian safety prediction methodology*. NCHRP Document 129: Phase III. Transportation Research Board, Washington, DC.

Harwood, D., Bauer, K.M.,Gilmore, D.K., Souleyrette, R., Hans, Z.N., 2010. U.S. *Road Assessment Program (usRAP) Pilot Program—Phase III*. AAA Foundation for Traffic Safety. Washington, DC

Hauer, E., Council, F., Mohammedshah, Y., 2004. Safety models for urban four-lane undivided road segments. *Transportation Research Record: Journal of the Transportation* Research Board 1897, 96–105.

Hidayati, N., Liu, R. & Montgomery, F., 2012. The impact of school safety zone and roadside activities on speed behaviour: the Indonesian case. *Procedia - Social and Behavioral Science*s, 54, pp.1339–1349.

Hosseinpour ,M., Yahaya,A, Sadullah,A., 2014. Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: Case studies from Malaysian federal roads. *Accident Analysis and Prevention* 62, pp 209– 222.

Hunter, W, Stutts, C, Pein, E, Cox, L., 1996. *Pedestrian and bicycle crash types of the early 1990's*. FHWA-RD-95-163. Federal Highway Administration.

iRAP, 2016. International road assessment program - methodology fact sheet. Available at: http://www.irap.org/ (Last accessed 12/02/2016).

Jacobs, G., Thomas, A., Astrop, A., 1999. *Estimating global road fatalities*. Department for International development, TRL report 445, London.

Jacobsen, PL, 2003. Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury Prevention*, 9, pp. 205-209

Jensen, S. U., 2007. Pedestrian and bicyclist level of service on roadway segments. *Transportation Research Record*, 2031, 43–51.

Jonsson, T., 2005. *Predictive models for accidents on urban links. A focus on vulnerable road users.* Ph.D. Dissertation. Department of Technology and Society, Lund Institute of Technology. Lund University, Lund.

Jonsson, T., 2013. Safety performance models for pedestrians and bicyclists. *Proceedings of the 16th International Conference Road Safety on Four Continents*, Beijing, China. 15-17 May.

Kang, L., Xiong, Y., & Mannering, F. L., 2013. Statistical analysis of pedestrian perceptions of sidewalk level of service in the presence of bicycles. *Transportation Research Part A: Policy and Practice*, 53, 10–21.

Kim, J. K., Ulfarsson, G., Shankar, V., Kim, S., 2008. Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic Logit analysis. *Accident Analysis and Prevention* 40, 1695-1702

Kim, K., Brunner, I., Yamashita, E., 2006. Influence of land use, population, employment, and economic activity on accidents. *Transportation Research Record* 1953, pp 56–64.

King, M.J., Soole, D., Ghafourian, A., 2009. Illegal pedestrian crossing at signalised intersections: Incidence and relative risk. *Accident Analysis and Prevention* 41, 485–490

Koh, P.P., Wong, Y.D., Chandrasekar, P., 2014. Safety evaluation of pedestrian behaviour and violations at signalised pedestrian crossings. *Safety Science* 70, 143–152.

Koushki, A, Ali, A., 2003. An analysis of pedestrian accidents in Kuwait. *Kuwait Journal of Science and Engineering*, 30(1), pp.139–151.

Kröyer, H.R.G., 2015. Accidents *between pedestrians, bicyclists and motorised vehicles: accident risk and injury severity.* Ph.D. Dissertation. Department of Technology and Society, Lund Institute of Technology. Lund University, Lund, Sweden.

Kulmala, R., 1995. *Safety at rural three- and four-arm junctions: Development and applications of accident prediction models*. VTT Publications, Technical Research Centre of Finland, Espoo.

Landis, B. W., Vattikuti, V. R., Ottenberg, R. M., McLeod, D. S., Guttenplan, M., 2001. Modeling the roadside walking environment: A pedestrian level of service. *Transportation Research Record* 1773, pp.82-88

Lave, C.A., 1985. Speeding, coordination, and the 55-mph limit. *The American Economic Review*, 75, pp.1159–1164.

Lawson, S.D., 1985. *Aspects of the identification, analysis and treatment of hazardous road locations.* PhD Thesis, University of Birmingham. Birmingham, UK.

Lawson, S.D., 1986. Descriptions and predictions of accidents on urban radial routes. *Traffic Engineering and Control*, 27, pp. 310-319.

Leden, L., 2002. Pedestrian risk decrease with pedestrian flow. A case study based on data from signalized intersections in Hamilton, Ontario. *Accident Analysis and Prevention* 34, pp.457–464.

Lee, C., Abdel-Aty, M., 2005. Comprehensive analysis of vehicle-pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention* 37, 775–786.

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. *Accident Analysis and Prevention* 34, 149–161.

Lee, J., Mannering, F., 1999. *Analysis of roadside accident frequency and severity and roadside safety management*. Olympia. Washington State Department of Transportation.

Lin, M.R., Chang, S.H., Pai, L., Keyl, P.M., 2003. A longitudinal study of risk factors for motorcycle crashes among junior college students in Taiwan. *Accident Analysis and Prevention* 35, 243–252.

Lin, M.R., Kraus, J.F., 2009. A review of risk factors and patterns of motorcycle injuries. *Accident Analysis and Prevention* 41, 710–722.

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.

Lord, D., Miranda-Moreno, L.F., 2008. Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: A Bayesian perspective. *Safety Science* 46, 751–770.

Lord, D., Washington, S.P., Ivan, J.N., 2005. Poisson, poisson-gamma and zero-inflated regression models of motor vehicle crashes: Balancing statistical fit and theory. *Accident Analysis and Prevention* 37, 35–46.

Loukaitou-Sideris, A., Liggett, R., Sung, H.G., 2007. Death on the crosswalk: A study of pedestrian-automobile collisions in Los Angeles. *Journal of Planning Education and Research* 26, pp 338–351.

Lynam, D., 2012. *Development of Risk Models for the Road Assessment Programme.* RAP504.12 and TRL Report CPR1293, Published by iRAP and TRL.

Lyon, C., Persaud, B., 2002. Pedestrian collision models for urban intersections. *Transportation Research Record* 1818, pp. 102-107.

Ma, M., Yan, X., Abdel-Aty, M., Huang, H., Wang, X., 2010. Safety analysis of urban arterials under mixed-traffic patterns in Beijing. *Transportation Research Record: Journal of the Transportation Research Board* 2193, pp.105–115.

Mccullagh, P., Nelder, J. A., 1989. *Generalized linear models*. Second Edition. Chapman and Hall, London.

McGuigan, D., 1982. Non-junction accident rates and their use in “black-spot” identification. *Traffic Engineering and Control*, 23 (2), pp. 60-65.

McGuigan, D., 1987. *An examination of relationships between road accidents and traffic flow.* Ph.D. Dissertation. Newcastle University, Newcastle.

McLean, A., Anderson, R., Farmer, M., Lee, B., Brooks, C., 1994. *Vehicle travel speeds and the incidence of fatal pedestrian collisions; vol. I.* The University of Adelaide, CR146

Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the Mixed Logit model: An exploratory empirical analysis. *Accident Analysis and Prevention* 40, 260–266.

Minge, E., Falero, C., Lindsey, G., Petesch, M., 2015. *Bicycle and pedestrian data collection manual.* Minneapolis, Minnesota: Department of Transportation.

Miranda-Moreno, L. ,Morency, P. ,El-Geneidy, A., 2011.The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, 43, pp. 1624-1634.

Moudon, A.V., Lin, L., Jiao, J., Hurvitz, P., Reeves, P., 2011. The risk of pedestrian injury and fatality in collisions with motor vehicles, a social ecological study of state routes and city streets in King County, Washington. *Accident Analysis and Prevention* 43, pp. 11-24.

Munawar, A., 2011. Speed and capacity for urban roads, Indonesian experience. *Procedia - Social and Behavioral Sciences* 16, pp.382–387.

Myers, R.H. , Montgomery, D.C., Vining, G.G., 2002. *Generalized linear models: With applications in engineering and the sciences*. Wiley Publishing Co., New York.

Nagraj, R., Vedagiri, P., 2013. Modeling pedestrian delay and level of service at signalized intersection crosswalks under mixed traffic conditions. *Transportation Research Record: Journal of the Transportation Research Board* 2394: 70–76.

Nelder, J.A., Wedderburn, R.W., 1972. Generalized linear models. *Journal of the Royal Statistical Society A*, 135: 370–384.

Nilsson, G., 2004. *Traffic safety dimensions and the power model to describe the effect of speed on safety*. Lund bulletin 221, PhD thesis, Lund Institute of Technology.

O'BRIEN, R. M., 2007. A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, 41, 673-690.

Ogendi, J., Odero, W., Mitullah, W., Khayesi, M., 2013. Pattern of pedestrian injuries in the city of Nairobi: implications for urban safety planning. Journal of urban health: b*ulletin of the New York Academy of Medicine,* 90(5), pp.849–56.

Oh, J., Lyon, C., Washington, S.P., Persaud, B.N.,Bared, J., 2003. Validation of the FHWA crash models for rural intersections: lessons learned. *Transportation Research Record*, 1840 , pp. 41-49.

Olszewski, P., Szagała, P., Wolański, M., Zielińska, A., 2015. Pedestrian fatality risk in accidents at unsignalized zebra crosswalks in Poland. *Accident Analysis and Prevention* 84, pp. 83-91.

Pal, S., Roy, S.K., 2016. Impact of roadside friction on travel speed and LOS of rural highways in India. *Transportation in Developing Economies* 2, 9.

Pande, A., Abdel-Aty, M., 2009. A novel approach for analyzing severe crash patterns on multilane highways. *Accident Analysis and Prevention* 41, 985–994.

Pannell, D.J., 1997. Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. *Agricultural Economics* 16: 139-152.

Peden, M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A., Jarawan, E., Mathers, C., 2004. *World report on road traffic injury prevention*. WHO, Geneva, Switzerland.

Petch, R.O., Henson, R.R., 2000. Child road safety in the urban environment. *Journal of Transport Geography* 8, pp 197–211.

Poch, M., Mannering, F., 1996. Negative Binomial analysis of intersection-accident frequencies. *Journal of Transportation Engineering* 122, 105–113.

Priyantha Wedagama, D.M., Bird, R.N., Metcalfe, A.V., 2006. The influence of urban land-use on non-motorised transport casualties. *Accident Analysis and Prevention* 38, pp 1049–1057.

Pulugurtha, S, Sambhara, R., 2011. Pedestrian crash estimation models for signalized intersections. *Accident analysis and prevention* 43, pp.439–46.

Qi, Y., Srinivasan, R., Teng, H., Baker, R., 2013. Analysis of the frequency and severity of rear-end crashes in work zones. *Traffic Injury Prevention*, 14 :61–72.

Quddus, M., 2013. Exploring the relationship between average speed, speed variation, and accident rates using spatial statistical models and GIS. *Journal of Transportation Safety and Security*, 5(1), pp.27-45.

Quistberg, D.A., Koepsell, T.D., Johnston, B.D., Boyle, L.N., Miranda, J.J., Ebel, B.E., 2015. Bus stops and pedestrian–motor vehicle collisions in lima, peru: A matched case–control study. *Injury Prevention* 21, e15–e22.

Ross, A., Baguley, C., Hills, B., McDonald, M., Silcock, D., 1994. *Towards safer roads in developing countries: A guide for planners and engineers.* Transportation Research Laboratory, Overseas Development administration.

Sawalha, Z., Sayed, T., 2001. Evaluating safety of urban arterial roadways. *Journal of Transportation Engineering* 2, pp.151–158.

Shankar, V., Mannering, F., Barfield, W., 1995. Effect of roadway geometrics and environmental factors on rural accident frequencies. Accident Analysis and Prevention 27(3), 371-389.

Shankar, V., Milton, J., Mannering, F.L., 1997. Modeling accident frequency as zero-altered probability processes: an empirical inquiry. *Accident Analysis and Prevention* 29, pp. 829-837.

Shirazi, M., Lord, D., Geedipally, S.R., 2016. Sample-size guidelines for recalibrating crash prediction models: Recommendations for the highway safety manual. *Accident Analysis and Prevention* 93, 160–168.

Solomon, D., 1964. *Accidents on main rural highways related to speed, driver and vehicle.* US Department of Commerce, Bureau of Public Roads, Washington D.C.

Taylor, M., Lynam, D., Baruya, A., 2000. *The effect of drivers' speed on the frequency of road accidents*. TRL report 421. Transport Research Laboratory.

Tiwari, G., 2001. Pedestrian infrastructure in the city transport system: a case study of Delhi. *World Transport Policy & Practice*, 7(4), pp.13–18.

Turner, B., Smith, G., 2013. *Safe system infrastructure: Implementation issues in low and middle income countries.* ARRB Group Ltd, ARR 383.

Turner, S.A., Roozenburg, A.P., Francis, T., 2006. *Predicting accident rates for cyclists and pedestrians.* Land Transport New Zealand Research Report 289. 180 pp.

Ukkusuri, S., Miranda-Moreno, L., Ramadurai, G., Isa-Tavarez, J., 2012. The role of built environment on pedestrian crash frequency. *Safety Science*, 50, pp.1141–1151.

Verzosa, N., Miles, R., 2016. Severity of road crashes involving pedestrians in Metro Manila, Philippines. *Accident Analysis and Prevention* 94, 216–226.

Vuong, Q., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57 (1989), pp. 307-333.

Wang, C., Quddus, M., Ison, S., 2013. The effect of traffic and road characteristics on road safety: A review and future research direction. *Safety Science*, 57, pp.264–275.

Wang, C., Quddus, M.A., Ison, S.G., 2011. Predicting accident frequency at their severity levels and its application in site ranking using a two-stage mixed multivariate model. *Accident Analysis and Prevention* 43, 1979–1990.

Wang, K., Liao, M., 2012. *Modeling Pedestrian Degree of Satisfaction Based on Structure Equation Model*. ICLEM: pp. 442-448.

Wang, X., Abdel-Aty, M., 2008. Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models. *Accident Analysis and Prevention*, 40, pp. 1674-1682

Wang, X., Yang, J., Lee, C., Ji, Z., You, S., 2016. Macro-level safety analysis of pedestrian crashes in Shanghai, China. *Accident Analysis and Prevention* 96, 12–21.

Washington, S., Karlaftis, M., Mannering, F., 2003. *Statistical and econometric methods for transportation data analysis.* Chapman and Hall/CRC, Boca Raton, FL.

Washington, S., Karlaftis, M., Mannering, F., 2010. *Statistical and econometric methods for transportation data analysis,* Second Edition. Chapman and Hall/CRC, Boca Raton, FL.

Wedagama, M, Bird, N, Metcalfe, V., 2006. The influence of urban land-use on non-motorised transport casualties. *Accident Analysis and Prevention* 38, pp.1049–1057.

WHO (World Health Organization), 2013a. *Global status report on road safety 2013*. WHO, Geneva, Switzerland.

WHO (World Health Organization), 2013b. Pedestrian Safety: a road safety manual for decision-makers and practitioners. WHO, Geneva, Switzerland.

WHO (World Health Organization), 2015. Global status report on road safety 2015. WHO, Geneva, Switzerland.

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 and Prevention* 41, pp. 137–45.

Wood, S.N., 2006. *Generalized additive models: an introduction with R*. Chapman and Hall/CRC, Boca Raton, Florida.

Wramborg, P. 2005. A New approach to a safe and sustainable road structure and street design for urban areas. Paper presented at Road Safety on Four Continents Conference, Warsaw, Poland.

Yamamoto, T., Hashiji, J., Shankar, V., 2008. Underreporting in traffic accident data, bias in parameters and the structure of injury severity models. *Accident Analysis and Prevention* 40, 1320–1329.

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, pp. 443–452.

Ye, Z., Zhang, Y., Lord, D., 2013. Goodness-of-fit testing for accident models with low means. *Accident Analysis and Prevention* 61, pp. 78-86.

Zeeger, C. , Stewart, J. , Huang, H. , Lagerway, P. , Feaganes, J. & Campbell, B., 2005. *Safety effects of marked versus unmarked crosswalks at uncontrolled locations. Final report and recommended guidelines.* FHWA-RT-04-100. Federal Highway Administration. Washington DC., U.S. Department of Transportation.

Zegeer, C., Sandt, L., Scully, M., 2009. *How to develop a pedestrian safety action plan.* Washington, Federal Highway Administration, report number FHWA-SA-05–12.

Zegeer, C., Opiela, K., Cynecki, M., 1985. *Pedestrian signalization alternatives*. Report No. FHWA-RD-83-102, Federal Highway Administration, Washington, DC., U.S. Department of Transportation.

Zegger, C.V., Seiderman, C., Lagerwey, P., Cynecki, M., Ronkin, M., & Schneider, R., 2002. *Pedestrian facilities users guide-providing safety and mobility*. FHWA-RD-01-102. Federal Highway Administration. Washington, DC., U.S. Department of Transportation.

Zhang, G., Yau, K.K.W., Zhang, X., 2014. Analyzing fault and severity in pedestrian-motor vehicle accidents in China. *Accident Analysis and Prevention* 73, pp. 141–150.

Zheng, Z., Ahn, S., Monsere, C.M., 2010. Impact of traffic oscillations on freeway crash occurrences. *Accident Analysis and Prevention* 42, pp. 626–636.

APPENDICES

APPENDIX A

AN EXAMPLE OF THE EXISTING IRAP METHODOLOGY CALCULATION PROCESS

This appendix demonstrates an example of iRAP star rating score (SRS) calculation for pedestrians (see Chapter 4). This example based on 100-metre segment of road shown in the image below from Soho road. In this study, the calculation has been conducted for all the 148 segments similarly using ViDA, iRAP online platform to calculate star rating scores (SRS), casualty estimation and create safer road investment plans.

The calculation showed that the star rating score (SRS) for the sample segment is 8.86. The star rating is determined by the SRS value; subsequently, the star rating is (4) as it is the respondent star rating to the band (5 to < 15).

Figure A.1 100-metre segment from Soho road

Pedestrian Star Rating Scores

Along (driver side)

Type of risk factor	Category	Risk factor	Score
Road attribute (likelihood)			
Sidewalk - driver-side	Non-physical separation 1.0m to <3.0m	0.09	
Curvature	Straight or gently curving	1.00	
Quality of curve	Not applicable	1.00	
Sight distance	Adequate	1.00	
Lane width	Narrow (≥ 0m to < 2.75m)	1.10	
Delineation	Adequate	1.00	
Grade	0% to < 7.5%	1.00	
Road condition	Good	1.00	
Speed management / traffic calming	Not present	1.25	
Vehicle parking	One side	1.2	
Shoulder rumble strips	Not present	1.25	
Skid resistance / grip	Sealed - adequate	1.00	
Street lighting	Present	1.00	
School zone warning	Not applicable (no school at the location)	1.00	
Product of road attribute (likelihood) risk factors			0.1856
Road attribute (severity)			
Sidewalk - driver-side	Non-physical separation 1.0m to <3.0m	0.09	
Product of road attribute (severity) risk factors			0.09
External flow influence	17178 15000 vehicles per day		0.032
Operating speed	<30km/h		0.02
Along (driver side) Star Rating Score			0

Along (passenger side)

Type of risk factor	Category	Risk factor	Score
Road attribute (likelihood)			
Sidewalk - passenger -side	Non-physical separation 1.0m to <3.0m	0.09	
Curvature	Straight or gently curving	1.00	
Quality of curve	Not applicable	1.00	
Sight distance	Adequate	1.00	
Lane width	Narrow (≥ 0m to < 2.75m)	1.10	
Delineation	Adequate	1.00	
Grade	0% to < 7.5%	1.00	
Road condition	Good	1.00	
Speed management / traffic calming	Not present	1.25	
Vehicle parking	One side	1.2	
Shoulder rumble strips	Not present	1.25	
Skid resistance / grip	Sealed - adequate	1.00	
Street lighting	Present	1.00	
School zone warning	Not applicable (no school at the location)	1.00	
Product of road attribute (likelihood) risk factors			0.185 625
Road attribute (severity)			
Sidewalk - passenger -side	Non-physical separation 1.0m to <3.0m	0.09	
Product of road attribute (severity) risk factors			0.09
External flow influence	17178 15000 vehicles per day		0.033
Operating speed	<30km/h		0.02
Along (driver side) Star Rating Score			0

Crossing (inspected road)

Type of risk factor	Category	Risk factor	Score
Road attribute (likelihood)			
Number of lanes	Two	2.80	
Median type	Centre line	3.00	
Pedestrian crossing facilities - inspected road	No facility	6.7	
Pedestrian crossing quality	Adequate	1.00	
Intersection type	3-leg unsignalised with no protected turn lane	1.1	
Intersection quality	Adequate	1.00	
Pedestrian fencing	None	1.25	
Skid resistance / grip	Sealed - adequate	1.00	
Street lighting	Present	1.00	
Sight distance	Adequate	1.00	
Vehicle parking	One side	1.2	
Speed management / traffic calming	present	1.00	
Product of road attribute (likelihood) risk factors			92.9
Road attribute (severity)			
Pedestrian crossing facilities - inspected road	Signalised without refuge	90.00	
Product of road attribute (severity) risk factors			90
External flow influence	17178 15000 vehicles per day		0.033
Operating speed	<30km/h		0.02
Crossing (inspected road) Star Rating Score			5.52

Crossing (side road)

Type of risk factor	Category	Risk factor	Score
Road attribute (likelihood)			
Number of lanes	One (two side roads)	2.80	
Median type	Centre line	3.00	
Pedestrian crossing facilities - side road	No facility	6.70	
Pedestrian crossing quality	Not applicable	1.00	
Intersection type	3-leg unsignalised with no protected turn lane	1.1	
Intersection quality	Adequate	1.00	
Pedestrian fencing	None	1.25	
Skid resistance / grip	Sealed - adequate	1.00	
Street lighting	Present	1.00	
Sight distance	Adequate	1.00	
Vehicle parking	One side	1.2	
Speed management / traffic calming	present	1.00	
Product of road attribute (likelihood) risk factors			92.9
Road attribute (severity)			
Pedestrian crossing facilities - side road	No facility	90.00	
Product of road attribute (severity) risk factors			90
External flow influence	5000 vehicles per day		0.02
Operating speed	<30km/h		0.02
Crossing (side road) Star Rating Score			3.34

Pedestrian Star Rating Score and Star Rating

Crash type	**Star Rating Score**	**Star Rating**
Along (driver side)	0	
Along (passenger side)	0	
Along (average of driver and passenger sides)	0	
Crossing (inspected road)	5.52	
Crossing (side road)	3.34	
Total Score / Star Rating	8.86	4

APPENDIX B

THE DATA TABLES TO CLARIFY THE CRASH FREQUENCIES DISTRIBUTIONS

Table B.1 Crash frequencies per road section (Data for Figures 6.1-6.3)

Pedestrian crashes frequencies per section	Number of the road sections*		
	Figure 6.1 (all pedestrian crashes)	Figure 6.2 (slight pedestrian crashes)	Figure 6.3 (serious and fatal pedestrian crashes)
0	15	25	55
1	27	40	30
2	25	18	21
3	14	17	10
4	12	9	1
5	10	2	0
6	6	2	0
7	2	3	0
8	3	1	0
9	2	0	0
10	0	0	0
11	1	0	0
	320	214	106

* The number of crashes calculated by multiplying the pedestrian crashes frequency by the number of road sections

Table B.2 The pedestrian crashes distribution for each day of the week (Data for Figure 6.4)

Day of week	Number of pedestrian crashes
Sunday	26
Monday	51
Tuesday	51
Wednesday	50
Thursday	52
Friday	55
Saturday	35
Total	320

Table B.3 Pedestrian crashes frequencies by time of day (Data for Figure 6.5)

Time of day	Number of pedestrian crashes
0	3
1	1
2	3
3	0
4	1
5	2
6	1
7	8
8	19
9	6
10	22
11	25
12	15
13	16
14	24
15	36
16	31
17	24
18	25
19	28
20	9
21	6
22	4
23	11
Total	320

APPENDIX C

THE SELECTED SITES DATA

Road section	Speed (mean)	CV	Bus Stops	Bus Stoppings	AADT	Pedestrian activities			No. of Lanes	Lane width	HoV %	Intersecting AADT	Road Parking	No. of Side roads	Median type*	No. of crashes	Pedestrian crossing facilities*
						C*	A*	V*									
HAGL-E-1	23	1.16	2	52	36622	132	214	76	2	N	1.52	0	No parking	0	C	2	N
HAGL-E-2	23	1.16	0	0	36622	132	214	76	2	N	1.52	0	No parking	1	C	2	N
HAGL-E-3	23	1.16	0	24	36622	132	214	76	2	N	1.52	0	No parking	1	C	2	S
HAGL-E-4	23	1.16	1	0	36622	132	214	76	2	N	1.52	0	No parking	2	C	3	R
HAGL-E-5	21	1.12	1	26	36622	106	167	60	2	N	1.52	7500	No parking	0	C	0	R
HAGL-E-6	25	0.99	0	50	36622	106	167	60	2	N	1.52	3000	No parking	1	C	3	R
HAGL-E-7	30	0.86	1	0	36622	106	167	60	2	N	1.52	0	No parking	1	C	2	S
HAGL-E-8	32	0.9	0	0	36622	106	167	60	2	N	1.52	7500	No parking	2	C	1	N
HAGL-E-9	33	1.01	1	26	36622	106	167	60	2	N	1.52	3000	No parking	1	C	0	N
HAGL-E-10	34	1.18	0	0	36622	106	167	60	2	M	1.52	0	No parking	0	C	0	R
HAGL-W-1	34	1.18	1	54	36851	98	139	55	2	N	1.54	3000	No parking	2	C	5	N

HAGL-W-2	37	1.42	1	0	36851	98	139	55	2	N	1.54	7500	No parking	2	C	1	N
HAGL-W-3	37	1.42	2	30	36851	98	139	55	2	N	1.54	0	No parking	0	C	3	N
HAGL-W-4	37	1.42	1	30	36851	98	139	55	2	N	1.54	0	No parking	0	C	1	N
HAGL-W-5	45	1.19	0	0	36851	98	139	55	2	N	1.54	3000	No parking	1	C	3	SR
HAGL-W-6	44	1.59	1	18	36851	98	139	55	2	N	1.54	7500	No parking	2	C	2	N
HAGL-W-7	42	1.98	1	18	36851	157	187	91	2	N	1.54	0	No parking	1	C	2	N
HAGL-W-8	42	1.98	0	0	36851	157	187	91	2	N	1.54	0	No parking	0	C	2	S
HAGL-W-9	42	1.98	0	0	36851	157	187	91	2	N	1.54	0	No parking	0	C	0	N
HAGL-W-10	43	1.54	1	38	36851	157	187	91	2	N	1.54	7500	No parking	2	C	6	N
MOS-S-1	13	1.14	2	24	14506	801	1779	217	2	N	2.16	7500	One side	1	C	5	SR
MOS-S-2	15	1.1	2	26	14506	801	1779	217	1	N	2.16	7500	One side	2	C	6	SR
MOS-S-3	15	1.36	0	0	14506	801	1779	217	1	N	2.16	7500	One side	2	C	5	S
MOS-S-4	15	1.46	1	20	14506	801	1779	217	1	N	2.16	7500	One side	2	C	6	N
MOS-S-5	24	0.76	1	24	14506	801	1779	217	1	N	2.16	7500	One side	2	C	6	S
MOS-S-6	23	0.89	1	26	14506	801	1779	217	1	N	2.16	3000	One side	2	C	5	N

MOS-S-7	23	1	1	0	14506	138	790	92	1	N	2.16	3000	One side	1	C	2	S
MOS-S-8	18	1.2	1	24	14506	138	790	92	1	N	2.16	3000	No parking	1	C	0	N
MOS-N-1	30	1.32	1	0	14506	152	673	89	1	N	2.16	3000	No parking	1	C	0	SR
MOS-N-2	30	1.32	0	0	14506	152	673	89	1	N	2.16	0	No parking	0	C	0	N
MOS-N-3	30	1.32	1	22	15426	152	673	89	1	N	1.94	500	No parking	0	C	0	N
MOS-N-4	38	1.01	2	42	15426	152	673	89	1	N	1.94	3000	No parking	1	C	0	N
MOS-N-5	37	1.21	0	0	15426	152	673	89	1	N	1.94	500	No parking	1	C	1	N
MOS-N-6	29	0.89	0	24	15426	776	1378	223	1	N	1.94	7500	One side	1	C	1	S
MOS-N-7	30	0.8	1	27	15426	776	1378	223	1	N	1.94	500	One side	0	C	1	N
MOS-N-8	27	0.78	1	38	15426	776	1378	223	2	N	1.94	3000	One side	1	C	4	SR
SOHO-E-1	28	1.27	1	14	16629	327	750	204	2	N	2.49	0	No parking	0	C	2	N
SOHO-E-2	42	0.86	1	17	16629	327	750	204	2	N	2.49	3000	No parking	1	C	2	N
SOHO-E-3	24	1.57	1	15	16629	327	750	204	2	N	2.49	3000	No parking	1	C	3	N
SOHO-E-4	18	1.25	1	0	16629	327	750	204	2	N	2.49	12500	One side	2	C	6	S
SOHO-E-5	20	1.08	0	17	16629	327	750	204	2	N	2.49	3000	One side	1	C	4	S

SOHO-E-6	23	0.88	2	11	16629	327	750	204	1	N	2.49	7500	No parking	2	C	1	N
SOHO-E-7	23	0.98	1	28	16629	327	750	204	1	N	2.49	3000	One side	1	H	3	N
SOHO-E-8	17	0.94	0	0	16629	327	750	204	1	N	2.49	500	One side	1	C	1	S
SOHO-W-1	14	1.01	0	0	17178	440	1092	257	1	N	2.49	3000	Two side	2	C	8	S
SOHO-W-2	13	1.03	2	33	17178	440	1092	257	1	N	2.49	0	Two side	0	C	9	N
SOHO-W-3	15	1.14	1	14	17178	440	1092	257	1	N	2.49	7500	One side	2	H	9	S
SOHO-W-4	14	1.11	0	21	17178	440	1092	257	1	N	2.44	500	One side	1	H	4	S
SOHO-W-5	14	1.11	0	0	17178	440	1092	257	1	N	2.44	0	One side	0	H	2	N
SOHO-W-6	17	0.83	1	14	17178	440	1092	257	1	N	2.44	12500	No parking	2	C	8	SR
SOHO-W-7	16	1	1	31	17178	440	1092	257	2	N	2.44	7500	One side	3	C	8	N
SOHO-W-8	13	1.21	0	14	17178	440	1092	257	2	N	2.44	7500	One side	2	C	11	S
COVT-N-1	21	0.88	2	30	18223	814	1188	173	1	N	1.77	15000	Two side	1	C	6	S
COVT-N-2	15	0.68	0	0	18223	814	1188	173	1	N	1.77	12500	Two side	1	C	5	S
COVT-N-3	16	0.67	0	8	18223	814	1188	173	1	N	1.77	7500	Two side	3	C	6	N
COVT-N-4	19	0.79	1	0	18223	814	1188	173	1	N	1.77	15000	One side	0	C	2	SR

COVT-N-5	18	0.57	1	10	18223	814	1188	173	1	N	1.77	12500	One side	1	C	1	R
COVT-N-6	23	0.63	1	10	18223	814	1188	173	1	N	1.77	7500	One side	1	C	1	N
COVT-N-7	14	1.11	0	0	18223	814	1188	173	2	N	1.77	7500	One side	2	C	3	R
COVT-N-8	14	1.11	1	12	18223	814	1188	173	1	N	1.77	15000	One side	0	C	1	SR
COVT-N-9	14	1.11	1	12	18223	814	1188	173	1	N	1.77	0	Two side	0	C	5	S
COVT-S-1	12	1.11	0	12	18223	331	508	115	1	N	1.77	7500	Two side	2	C	4	S
COVT-S-2	15	0.9	2	12	18223	331	508	115	1	N	1.77	3000	One side	1	C	2	N
COVT-S-3	20	0.69	0	0	18223	331	508	115	1	N	1.77	3000	Two side	1	C	1	N
COVT-S-4	22	0.63	1	11	18223	331	508	115	1	N	1.77	0	Two side	0	C	0	SR
COVT-S-5	26	0.69	1	12	18223	331	508	115	1	N	1.77	7500	One side	2	C	2	N
COVT-S-6	23	0.69	1	14	18223	331	508	115	1	N	1.77	7500	Two side	2	C	4	N
COVT-S-7	20	0.77	0	18	18223	331	508	115	1	N	1.77	7500	Two side	2	C	5	S
COVT-S-8	21	0.74	1	0	18223	547	770	154	1	N	1.77	7500	One side	1	C	1	N
COVT-S-9	18	0.88	1	38	18223	547	770	154	2	N	1.77	3000	Two side	2	C	4	N
COVT-S-10	18	0.88	1	0	18223	547	770	154	2	N	1.77	7500	No parking	1	C	2	SR

STRAF-S-1	46	1.93	2	30	18794	171	348	71	1	N	1.45	500	No parking	1	C	4	N
STRAF-S-2	46	2.3	0	16	18794	171	348	71	1	N	1.45	0	No parking	0	C	2	N
STRAF-S-3	40	1.71	2	12	18794	171	348	71	1	N	1.45	7500	No parking	1	C	1	N
STRAF-S-4	35	1.23	0	16	18794	171	348	71	2	N	1.45	500	No parking	1	C	3	S
STRAF-S-5	25	0.81	2	13	18794	171	348	71	2	N	1.45	500	No parking	0	C	1	N
STRAF-S-6	20	1.48	0	20	18794	171	348	71	2	N	1.45	7500	No parking	2	C	6	S
STRAF-S-7	32	1.17	1	14	18794	171	348	71	2	N	1.45	500	No parking	1	C	2	N
STRAF-S-8	28	0.56	0	0	18794	171	348	71	1	N	1.45	7500	No parking	1	C	1	N
STRAF-S-9	31	0.67	1	14	18794	114	231	57	1	N	1.45	3000	One side	1	C	2	S
STRAF-S-10	33	0.79	0	16	18794	114	231	57	1	N	1.45	3000	One side	1	C	4	N
STRAF-S-11	27	0.82	1	14	18794	114	231	57	1	N	1.45	3000	One side	1	C	2	S
STRAF-S-12	44	2.01	0	0	18794	114	231	57	1	N	1.45	7500	No parking	1	C	1	N
STRAF-S-13	47	2.08	0	0	20606	114	231	57	1	N	1.61	3000	No parking	1	C	1	N
STRAF-S-14	43	1.84	2	22	20606	114	231	57	1	N	1.61	3000	No parking	1	C	1	N
STRAF-S-15	38	1.66	0	0	20606	114	231	57	1	N	1.61	0	No parking	0	C	0	N

STRAF-S-16	38	1.66	1	13	20606	114	231	57	2	N	1.61	0	No parking	0	C	1	N
STRAF-N-1	37	1.87	1	22	20606	256	792	229	2	N	1.61	7500	One side	1	C	3	N
STRAF-N-2	24	0.96	0	10	20606	256	792	229	2	N	1.61	3000	Two side	1	C	5	S
STRAF-N-3	19	0.94	1	14	20606	256	792	229	1	N	1.61	0	One side	1	C	7	N
STRAF-N-4	13	1.15	1	10	20606	256	792	229	1	N	1.61	7500	No parking	1	C	1	SR
STRAF-N-5	13	1.15	1	20	20606	256	792	229	2	N	1.61	3000	Two side	1	C	3	S
STRAF-N-6	15	0.8	0	0	20606	256	792	229	1	N	1.61	500	Two side	1	C	1	N
STRAF-N-7	12	1.05	0	0	15608	188	503	178	1	N	1.52	3000	Two side	1	C	4	S
STRAF-N-8	19	0.89	2	45	15608	188	503	178	1	N	1.52	7500	Two side	2	C	7	N
STRAF-N-9	24	0.86	0	0	15608	188	503	178	1	N	1.52	3000	Two side	1	C	3	R
STRAF-N-10	25	0.68	0	23	15608	188	503	178	1	N	1.52	3000	One side	1	C	1	N
STRAF-N-11	25	0.8	1	0	15608	188	503	178	1	N	1.52	3000	No parking	1	C	0	S
STRAF-N-12	26	0.85	0	22	15608	188	503	178	1	N	1.52	7500	One side	2	C	2	R
PERSH-N-1	22	1.05	0	16	18754	266	756	84	1	N	2.93	7500	No parking	2	C	3	N
PERSH-N-2	22	1.02	0	16	18754	266	756	84	1	N	2.93	3000	Two side	1	C	2	S

PERSH-N-3	24	0.86	1	0	18754	266	756	84	1	N	2.93	500	One side	0	C	1	N
PERSH-N-4	24	1.09	1	30	18754	266	756	84	1	N	2.93	3000	Two side	1	C	2	N
PERSH-N-5	27	1.03	0	0	18754	266	756	84	1	N	2.93	3000	Two side	1	C	2	R
PERSH-N-6	30	0.98	1	18	18754	79	423	38	1	N	2.93	0	One side	0	C	1	N
PERSH-N-7	28	1.3	1	38	18754	79	423	38	1	N	2.93	7500	One side	2	C	4	S
PERSH-N-8	27	1.63	0	0	18754	79	423	38	1	N	2.93	3000	No parking	1	C	3	N
PERSH-N-9	30	1.24	0	0	18754	79	423	38	1	N	2.93	3000	One side	0	H	2	SR
PERSH-S-1	32	1.36	2	50	19183	101	458	43	1	N	2.13	0	One side	0	C	1	S
PERSH-S-2	32	1.36	0	0	19183	101	458	43	1	N	2.13	7500	No parking	0	C	1	S
PERSH-S-3	32	1.05	2	50	19183	101	458	43	1	N	2.13	0	No parking	0	C	0	N
PERSH-S-4	29	0.72	0	0	19183	101	458	43	1	N	2.13	500	No parking	0	C	0	S
PERSH-S-5	30	1.44	2	26	19183	287	856	89	1	N	2.13	3000	No parking	1	C	0	N
PERSH-S-6	22	1.11	0	26	19183	287	856	89	1	N	2.13	500	One side	0	C	3	N
PERSH-S-7	22	1.11	2	18	19183	287	856	89	1	N	2.13	7500	One side	1	H	2	S
PERSH-S-8	16	1.49	2	19	19183	287	856	89	2	M	2.13	3000	No parking	1	C	5	N

PERSH-S-9	18	1.05	2	54	19183	287	856	89	2	N	2.13	0	No parking	1	C	6	S

CV- Coefficient of speed variation

Pedestrian activities
C-Pedestrian crossing
A-Pedestrian alongside
V-Pedestrian violations

Lane width
N -Narrow (≥ 0m to < 2.75m)
M - Medium (≥ 2.75m to < 3.25m)

Median type
C - Centre line
H - Central hatching (>1m)

Pedestrian crossing facility
SR - Signalised with refuge
S - Signalised without refuge
R - Refuge only
N - No facility

APPENDIX D

THE SUMMARY OF VALIDATION DATA

Table D1.Selected road sections for the study-lengths and traffic volumes' statistics

No	Name of the Road	Length (km)	AADT	Heavy Vehicle	Heavy Vehicle Percentage (HV)
1	Dudley Rd	1.1	15379	186	1.21%
2	Waterloo Rd-Cape Hill Rd	0.9	12612	143	1.13%
3	Alum Rock Rd	1.1	18754	550	2.93%

Table D2.Historical pedestrian-vehicle collision data for studied road segments

No	Name of the Road	Crash Records (2009-2016)				Crash/year /km
		Fatal	Serious	Slight	Total	
1	Dudley Rd	1	10	31	42	5
2	Waterloo Rd-Cape Hill Rd	1	6	23	30	4
3	Alum Rock Rd	0	13	38	51	6

Table D3. Speed variables for the selected road segments (mph)

No	Name of the Road	Speed Limit	Operating Speed			Coefficient of Speed Variation (CV)		
			Min.	Max	Mean	Min.	Max.	Mean
1	Dudley Rd	20	14.9	24.1	19.3	0.63	1.26	0.90
2	Waterloo Rd-Cape Hill Rd	20/30*	10.8	39.9	23.3	0.36	0.9	2.04
3	Alum Rock Rd	20/30*	9.53	25	17.26	0.68	1.34	0.91

* Sites with variable speed limits of 20 and 30 mph along the road

Table D4. Roadside activities and pedestrian statistics for the selected road segments

No	Name of the Road	Parking on side	Bus** stoppings	Parking and un-parking*	Pedestrian alongside*	Pedestrian violations*
1	Dudley Rd	No/ Two	0-23	251	317	166
2	Waterloo Rd-Cape Hill Rd	No/one/ Two	0-30	214	345	190
3	Alum Rock Rd	One/ Two	0-23	203	242	153

* Average

** Min-Max

Table D5.Selected low activity road sections for the study-lengths and traffic volumes' statistics

No	Name of the Road	Length (km)	AADT	Heavy Vehicle	Heavy Vehicle Percentage (HV)
1	City Rd	1.0	9502	52	0.55%
2	Yardley Rd	1.0	8641	62	0.72%
3	Vicarage Rd	1.0	9446	128	1.36%

Table D6.Historical pedestrian-vehicle collision data for low activity road segments

No	Name of the Road	Crash Records (2009-2016)				Crash/year /km
		Fatal	Serious	Slight	Total	
1	City Rd	0	1	6	7	0.88
2	Yardley Rd	0	3	10	13	1.63
3	Vicarage Rd	0	2	5	7	0.88

Table D7.Speed variables for the low activity road segments (mph)

No	Name of the Road	Speed Limit	Operating Speed			Coefficient of Speed Variation (CV)		
			Min.	Max	Mean	Min.	Max.	Mean
1	City Rd	30	19.1	25.7	21.1	0.46	1.97	1.06
2	Yardley Rd	30	16.7	24.4	20.8	0.79	1.33	1.07
3	Vicarage Rd	20/30*	5.8	19.1	15.8	0.80	1.79	1.22

* Sites with variable speed limits of 20 and 30 mph along the road

Table D8.Roadside activities and pedestrian statistics for the low activity road segments

No	Name of the Road	Parking on side	Bus** stoppings	Parking and un-parking*	Pedestrian alongside*	Pedestrian violations*
1	City Rd	No/one/ Two	0-29	57	120	72
2	Yardley Rd	No/ One	0-17	94	188	136
3	Vicarage Rd	No/ One	0-39	64	155	85

* Average

** Min-Max

APPENDIX E

EXAMPLE OF COLLECTED DATA FOR IRAP INPUT (ONE ROAD SAMPLE)

Image reference	Road name	Section	Distance	Length	Latitude	Longitude	Landmark	Comments
Stratf.-S- 01	Strat-S Rd	Strat-S Rd	0.003	0.1	52.43084	-1.843832	Stratf Rd	
Stratf.-S- 02	Strat-S Rd	Strat-S Rd	0.103	0.1	52.43161	-1.844376	Stratf Rd	
Stratf.-S- 03	Strat-S Rd	Strat-S Rd	0.203	0.1	52.43247	-1.845012	Stratf Rd	
Stratf.-S- 04	Strat-S Rd	Strat-S Rd	0.303	0.1	52.43329	-1.845584	Stratf Rd	
Stratf.-S- 05	Strat-S Rd	Strat-S Rd	0.403	0.1	52.43415	-1.846053	Stratf Rd	
Stratf.-S- 06	Strat-S Rd	Strat-S Rd	0.503	0.1	52.435	-1.846448	Stratf Rd	
Stratf.-S- 07	Strat-S Rd	Strat-S Rd	0.603	0.1	52.43595	-1.846626	Stratf Rd	
Stratf.-S- 08	Strat-S Rd	Strat-S Rd	0.703	0.1	52.43681	-1.846852	Stratf Rd	
Stratf.-S- 09	Strat-S Rd	Strat-S Rd	0.803	0.1	52.43762	-1.847267	Stratf Rd	
Stratf.-S- 10	Strat-S Rd	Strat-S Rd	0.903	0.1	52.43849	-1.847813	Stratf Rd	
Stratf.-S- 11	Strat-S Rd	Strat-S Rd	1.003	0.1	52.43937	-1.848272	Stratf Rd	
Stratf.-S- 12	Strat-S Rd	Strat-S Rd	1.103	0.1	52.44023	-1.848708	Stratf Rd	
Stratf.-S- 13	Strat-S Rd	Strat-S Rd	1.203	0.1	52.44104	-1.849346	Stratf Rd	
Stratf.-S- 14	Strat-S Rd	Strat-S Rd	1.303	0.1	52.44181	-1.850119	Stratf Rd	
Stratf.-S- 15	Strat-S Rd	Strat-S Rd	1.403	0.1	52.44238	-1.851235	Stratf Rd	

Carriageway	Upgrade cost	Motorcycle observed flow	Bicycle observed flow	Pedestrian observed flow across the road	Pedestrian observed flow along the road driver-side	Pedestrian observed flow along the road passenger-side	Land use - driver-side	Land use - passenger-side	Area type	Speed limit	Motorcycle speed limit
3	3	2	3	6	6	6	3	6	2	5	5
3	3	1	1	6	6	6	3	3	2	5	5
3	3	1	1	6	6	6	3	3	2	5	5
3	3	1	2	6	6	6	3	3	2	5	5
3	3	4	5	6	6	6	3	3	2	5	5
3	3	2	4	6	6	6	4	4	2	5	5
3	3	1	5	6	6	6	4	4	2	5	5
3	3	2	1	6	6	6	4	4	2	5	5
3	3	2	1	6	6	6	4	3	2	5	5
3	3	2	1	6	6	6	4	3	2	5	5
3	3	2	1	6	6	6	4	4	2	5	5
3	3	2	1	6	6	6	3	3	2	5	5
3	3	2	3	6	6	6	3	3	2	5	5
3	3	1	1	6	6	6	3	3	2	5	5
3	3	1	1	6	6	6	3	3	2	5	5

Truck speed limit	Differential speed limits	Median type	Centreline rumble strips	Roadside severity - driver-side distance	Roadside severity - driver-side object	Roadside severity - passenger-side distance	Roadside severity - passenger-side object	Shoulder rumble strips	Paved shoulder - driver-side	Paved shoulder - passenger-side	Intersection type
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	12
5	1	11	1	2	12	1	12	1	4	4	8
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	10
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	4
5	1	11	1	2	12	1	12	1	4	4	12

Intersection channelisation	Intersecting road volume	Intersection quality	Property access points	Number of lanes	Lane width	Curvature	Quality of curve	Grade	Road condition	Skid resistance / grip	Delineation
1	5	1	1	1	3	1	3	1	1	1	1
1	7	3	1	1	3	1	3	1	1	1	1
1	3	1	1	1	3	1	3	1	1	1	1
1	5	1	1	2	3	1	3	1	1	1	1
1	5	1	1	2	3	1	3	1	1	1	1
1	3	1	1	2	3	1	3	1	1	1	1
1	5	1	1	2	3	1	3	1	1	1	1
1	3	1	1	1	3	1	3	1	1	1	1
1	4	1	1	1	3	1	3	1	1	1	1
1	4	1	1	1	3	1	3	1	1	1	1
1	4	1	1	1	3	1	3	1	1	1	1
1	3	1	1	1	3	1	3	1	1	1	1
1	4	1	1	1	3	1	3	1	1	1	1
1	4	1	1	1	3	1	3	1	1	1	1
1	7	3	1	1	3	1	3	1	1	1	1

Street lighting	Pedestrian crossing facilities - inspected road	Pedestrian crossing quality	Pedestrian crossing facilities - intersecting road	Pedestrian fencing	Speed management / traffic calming	Vehicle parking	Sidewalk - driver-side	Sidewalk - passenger-side	Service road	Facilities for motorised two wheelers	Facilities for bicycles
2	7	3	17	1	1	1	4	4	1	6	4
2	7	3	7	1	1	1	4	4	1	6	4
2	7	3	17	1	1	1	4	4	1	6	4
2	3	1	17	1	1	1	4	4	1	6	4
2	7	3	17	1	1	1	4	4	1	6	4
2	3	1	3	2	2	1	4	4	1	6	4
2	7	3	17	1	1	1	4	4	1	6	4
2	7	3	17	1	1	1	4	4	1	6	4
2	3	1	17	1	1	2	4	4	1	6	4
2	7	3	17	1	1	2	4	4	1	6	4
2	3	1	17	1	1	2	4	4	1	6	4
2	7	3	16	1	1	1	4	4	1	6	4
2	7	3	17	1	1	1	4	4	1	6	4
2	7	3	17	1	1	1	4	4	1	6	4
2	7	3	7	1	1	1	4	4	1	6	4

Roadworks	Sight distance	Vehicle flow (AADT	Motorcycle %	Pedestrian peak hour flow across the road	Pedestrian peak hour flow along the road driver-side	Pedestrian peak hour flow along the road passenger-side	Bicycle peak hour flow	Operating Speed (85th percentile)	Operating Speed (mean)	Roads that cars can read
1	1	18794	3	7	9	7	5	5	5	1
1	1	18794	3	7	10	7	5	5	5	1
1	1	18794	3	10	10	8	5	3	3	1
1	1	18794	3	10	10	8	5	2	2	1
1	1	18794	3	9	10	8	5	1	1	1
1	1	18794	3	7	9	7	5	1	1	1
1	1	18794	3	6	9	7	5	2	2	1
1	1	18794	3	6	9	7	5	1	1	1
1	1	18794	3	6	9	7	5	1	1	1
1	1	18794	3	6	9	7	5	2	2	1
1	1	18794	3	6	9	7	5	1	1	1
1	1	18794	3	6	9	7	5	4	4	1
1	1	20606	3	6	9	7	5	5	5	1
1	1	20606	3	6	9	7	5	4	4	1
1	1	20606	3	6	9	7	5	3	3	1

Vehicle Occupant Star Rating Policy Target	Motorcycle Star Rating Policy Target	Pedestrian Star Rating Policy Target	Bicycle Star Rating Policy Target	Annual Fatality Growth Multiplier	School zone warning	School zone crossing supervisor
6	6	6	6	1	3	2
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3
6	6	6	6	1	4	3

*The numbers in above tables representing road attributes in accordance with iRAP methodology coding specification (See example below) which can be found at https://www.irap.org/

Example of iRAP input coding specification:

Median type	Intersection type	Delineation	Pedestrian crossing - inspected road
11 - Centre line	8 - 4-leg unsignalised with no protected turn lane	2 - Poor	1 - Grade separated facility
14 - Wide centre line (0.3m to 1m)	17 - Mini roundabout	1 - Adequate	2 - Signalised with refuge
10 - Central hatching (>1m)	7 - 4-leg unsignalised with protected turn lane		3 - Signalised without refuge
8 - Continuous central turning lane	10 - 4-leg signalised with no protected turn lane		4 - Unsignalised marked crossing with refuge
9 - Flexible posts	13 - Railway Crossing - passive (signs only)		5 - Unsignalised marked crossing without a refuge
7 - Physical median width 0 to <1m	4 - 3-leg unsignalised with no protected turn lane		6 - Refuge only
6 - Physical median width 1 to <5m	9 - 4-leg signalised with protected turn lane		7 - No facility
5 - Physical median width 5 to <10m	6 - 3-leg signalised with no protected turn lane		14 - Unsignalised raised marked crossing with refuge
2 - Safety barrier - concrete	3 - 3-leg unsignalised with protected turn lane		15 - Unsignalised raised marked crossing without refuge
1 - Safety barrier - metal	5 - 3-leg signalised with protected turn lane		16 - Raised unmarked crossing with refuge
12 - Safety barrier - motorcycle friendly	14 - Railway Crossing - active (flashing lights / boom gates)		17 - Raised unmarked crossing without refuge
15 - Safety barrier - wire rope	2 - Roundabout		
4 - Physical median width 10 to <20m	1 - Merge lane		
3 - Physical median width >=20m	15 - Median crossing point - informal		
13 - One way	16 - Median crossing point - formal		
	12 - None		

APPENDIX F

Published Conference Paper

The following conference paper was prepared from this thesis during the course of this study:

Kraidi, R., Evdorides, H., 2016. A review of pedestrian safety models for urban areas in low and middle income countries. The 4th Chinese–European Workshop on Functional Pavement Design, CEW 2016, Delft, Netherlands, DOI: 10.1201/9781315643274-177.