Optimisation of speed camera locations using genetic algorithm and pattern search

by

Agnes Boscoe-Wallace

A Doctoral Thesis

Submitted in partial fulfilment of the requirements

for the award of

Doctor of Philosophy of Loughborough University

February 2017

Abstract

Road traffic accidents continue to be a public health problem and are a global issue due to the huge financial burden they place on families and society as a whole. Speed has been identified as a major contributor to the severity of traffic accidents and there is the need for better speed management if road traffic accidents are to be reduced.

Over the years various measures have been implemented to manage vehicle speeds. The use of speed cameras and vehicle activated signs in recent times has contributed to the reduction of vehicle speeds to various extents. Speed cameras use punitive measures whereas vehicle activated signs do not so their use depends on various factors. Engineers, planners and decision makers responsible for determining the best place to mount a speed camera or vehicle activated sign along a road have based their decision on experience, site characteristics and available guidelines (Department for Transport, 2007; Department for Transport, 2006; Department for Transport, 2003). These decisions can be subjective and indications are that a more formal and directed approach aimed at bringing these available guidelines together in a model will be beneficial in making the right decision as to where to place a speed camera or vehicle activated sign is to be made. The use of optimisation techniques have been applied in other areas of research but this has been clearly absent in the Transport Safety sector.

This research aims to contribute to speed reduction by developing a model to help decision makers determine the optimum location for a speed control device.

In order to achieve this, the first study involved the development of an Empirical Bayes Negative Binomial regression accident prediction model to predict the number of fatal and serious accidents combined and the number of slight accidents. The accident prediction model that was used explored the effect of certain geometric and traffic characteristics on the effect of the severity of road traffic accident numbers on selected A-roads within the Nottinghamshire and Leicestershire regions of United Kingdom. On A-roads some model variables (n=10) were found to be statistically significant for slight accidents and (n=6) for fatal and serious accidents.

The next study used the accident prediction model developed in two optimisation techniques to help predict the optimal location for speed cameras or vehicle activated

signs. Pattern Search and Genetic Algorithms were the two main types of optimisation techniques utilised in this thesis. The results show that the two methods did produce similar results in some instances but different in others. Optimised results were compared to some existing sites with speed cameras some of the results obtained from the optimisation techniques used were within proximity of about 160m. A validation method was applied to the *genetic algorithm* and *pattern search* optimisation methods. The pattern search method was found to be more consistent than the genetic algorithm method. Genetic algorithm results produced slightly different results at validation in comparison with the initial results. T-test results show a significant difference in the function values for the validated genetic algorithm (M= 607649.34, SD= 1055520.75) and the validated pattern search function values (M= 2.06, SD= 1.17) under the condition $t(79) = 5.15$, $p=0.000$.

There is a role that optimisation techniques can play in helping to determine the optimum location for a speed camera or vehicle activated sign based on a set of objectives and specified constraints. The research findings as a whole show that speed cameras and vehicle activated signs are an effective speed management tool. Their deployment however needs to be carefully considered by engineers, planners and decision makers so as to achieve the required level of effectiveness. The use of optimisation techniques which has been generally absent in the Transport Safety sector has been shown in this thesis to have the potential to contribute to improve speed management. There is however no doubt that this research will stimulate interest in this rather new but high potential area of Transport Safety.

Statement of originality

The author (Agnes Boscoe-Wallace) is wholly responsible for the work carried out and presented in this thesis. Dr. Markus Deublein allowed the author of this research to use his method for calculating the slope and radius of the road segments.

Work carried out in this thesis was funded by Loughborough University Graduate scholarship.

This statement certifies that neither the submission nor the original work contained in this thesis has been submitted for an award or other degree awarding body.

Acknowledgments

I would like to take this opportunity to express my gratitude to my supervisors Dr. Richard Frampton and Professor Andrew Morris for their kind support, advice and encouragement throughout this research. There were times I just needed some advice on how to move on and the little chats were all very useful. I would also like to thank my internal assessor Professor Pete Thomas for all the useful advice and suggestions.

My sincere gratitude to Nottinghamshire County Council, Leicestershire County Council, Leicestershire Police and Julian Hill for allowing use of their data. Thanks to Dr. Vincent Dwyer, Professor M Quddus, Dr. Markus Deublein and Dr. Martins Irhebhude for volunteering their time to offer very useful advice and knowledge.

My appreciation goes to Professor Serpil Acar, Dr. Jo Barnes, Dr. Andrew May and Professor Monney for your words of encouragement. To the other PhD students I had the opportunity to interact with, I would like to say thanks for being there to share ideas especially Carlos and Karthik. My gratitude also goes to Loughborough University for offering me a full scholarship to undertake this research without which it would not have been possible.

Special thanks to my parents for all their sacrifice towards my education and to my siblings for your support and encouragement. Special thanks to my husband and children for being understanding, supportive, encouraging and for your prayers which has made all the difference in seeing me through.

Finally but not the least I would like to thank God for giving me the strength and wisdom to undertake this research.

Dedication

To my Parents

Table of Contents

List of Figures

List of Tables

1 Introduction and overall research aims

1.1 Background to research problem

There is evidence to show that inappropriate speeding is a major contributor to the severity of road traffic accidents (Barker, 1997; Chen, Meckle and Wilson, 2002; Winnett and Wheeler, 2002; Taylor, Baruya and Kennedy, 2002; Perez et al., 2007; World Health Organisation [WHO], 2013). Road traffic speed reduction measures are essential to reducing road traffic accident severities (Crombie, 2002; Peden et al., 2004; Pilkington and Kinra, 2005) with speeding inappropriately for prevailing conditions or exceeding set speed limits being the two most common speed related factors to road traffic accidents.

In an attempt to influence and reduce driver speed various road infrastructure devices such as road humps, chicanes, rumble strips, narrowing, mini-roundabout and village gateway schemes have been used. More recently, speed control devices such as speed cameras and vehicle activated signs have been deployed along roads to help reduce driver vehicle speeds. A lot of research work has also been conducted into the effectiveness of speed cameras and vehicle activated signs with positive responses obtained in most instances.

1.2 Current approach to installing vehicle activated signs and speed cameras

The only source of advisory information identified for the installation of vehicle activated signs is Traffic Advisory Leaflet (TAL) 1/03 (Department for Transport, 2003). This leaflet states that Vehicle Activated Signs (VAS) are to be used to address problems associated with inappropriate speeds in situations where conventional signing has been ineffective. VASs however should not be used as an alternative to fixed signs. It also adds that VASs should be considered in instances where there is an accident problem associated with inappropriate speed. It should be used where standard signing has been unable to satisfactorily remedy problem. Also in instances where safety cameras and other signs are not cost effective or an appropriate solution, VASs can be used.

It is advised that an audit of existing furniture, fixed signs, road condition and road markings are undertaken to assess their standard and condition before installing a

VAS. It is advised in TAL 1/03 that a detailed accident investigation should be undertaken to both identify the dominant accident patterns and to confirm that VASs are an appropriate remedial measure. The site selection should also take into account the number of speed related accidents and in particular inappropriate speed for the conditions prevailing. Monitoring of traffic speeds can be carried out to establish that a problem of inappropriate speed exists. Traffic Advisory Leaflet 1/03 advises using speed data collected prior to the installation of a VAS to help estimate a suitable threshold speed for the sign to display the message (Department for Transport, 2003).

For warning sign VASs, TAL 1/03 (Department for Transport, 2003) advises that the speed threshold be set at the 50^{th} percentile speed measured before installation. For speed limit signs it is advised that the threshold set be dependent on the road conditions with a reasonable benchmark being the ACPO (Association of Chief Police Officers) guidelines on enforcement of 10% + 2mph. It is also necessary to ensure that permanent warning signs sited in advance of any VAS are correctly and appropriately placed in accordance with Chapter 4 of the Traffic Signs Manual (Department for Transport, 2013).

Safety cameras are noted to provide a valuable and cost-effective method for preventing, detecting and enforcing speed and traffic light offences. They have been proved to encourage modified driver behaviour improving road safety for all road users. The Department for Transport (DfT) circular 01/07 (Department for Transport, 2007) provides guidance and best practice on the deployment of speed and red-light cameras after April 2007.

Circular 01/2007 (Department for Transport, 2007) recommends that before deciding on using speed cameras, investigations should be conducted on the nature of the problem. The investigation should include current vehicle speeds, the proportion of vehicles exceeding the speed limit under free-flowing conditions, the proportion of different collision types, and the causes of the collisions. It is also essential that traffic authorities confirm that the speed limit at each proposed site is appropriate. In addition to these in order to select camera sites, a collision data analysis needs to be carried out over a minimum period (most recent 3 years or preferably 5 years) and the 85th percentile speed of vehicles within a 12 month period must be collected. Cameras must be positioned such that they are not obscured and clearly visible at all times. For 40mph or less speed limit roads a minimum visibility distance of 60 metres is recommended with a minimum of 100 metres at all other speed limit locations by TAL 1/03 (Department for Transport, 2003).

The Parliamentary Office of Science and Technology POSTnote 04/218 (Parliamentary Office of Science and Technology, 2004) provided DfT guidelines applicable to speed camera use. It states that the majority (85%) of cameras must be in areas which have a specified minimum level of death and injury within 1km in the previous three years (4 collisions resulting in death/serious injury for fixed cameras, 2 for mobile). It recommends that speeding must be shown to be a problem at the location even though crashes do not need to have been speed-related.

From the summaries provided on the guidelines for the installation of VASs and speed cameras, it is evident that the VAS guidelines provide some suggestions of the instances under which they should be placed as well as the possible investigations which should be carried out.

Current practice for installing road side speed control devices to manage vehicle speed involves engineers, planners and designers to visit the site and make an assessment and judgment on where a roadside speed control device should be placed. This is usually based on the threshold of recorded accidents, site conditions such as the terrain, visibility and other site constraints and considerations. Considering that these practices can be subjective, time consuming and also costly, optimisation techniques undoubtedly seems to be a promising and suitable alternative way forward to achieve the requirements for vehicle speed management.

The motivation for this research is to develop a model using genetic algorithms and pattern search to optimise the location of speed camera to contribute to better management of vehicle speeds. The model will focus on the infrastructural aspects of roads and other related factors and it will exclude driver behaviour. This approach will serve as a useful planning tool for road safety professionals. It can be used to predict future accident prone areas and thus can help in the recommendation of appropriate corrective measures.

1.3 Research in Context

Against this background used in implementing speed control devices across England, the following questions will be addressed;

- Which infrastructural and related parameters excluding driver behaviour contribute to road traffic accidents?
- Which parameters influence the location of speed control devices (speed cameras and vehicle activated signs)?
- Taking into consideration the above two questions, can a model be developed using an optimisation technique to assist decision makers to effectively plan the deployment of speed control devices?

1.4 Research Aim

Based on the context presented, this research aims to contribute to speed reduction by developing a model to help decision makers determine the optimum location for a speed control device.

1.5 Research Objectives

To be able to achieve this aim the following objectives have been set.

i. Identify infrastructural and other related parameters excluding driver behaviour which have the ability to influence vehicle speed through a literature review.

ii. Identify methods used in facility location problems through a literature review.

iii. Based on parameters identified in (i) above, develop an accident prediction model using A-roads and test and validate the model.

iv. Using genetic algorithm and pattern search optimisation methods, identify the parameters to be used in the optimisation model and test the model on a selection of existing and future speed control device locations along A-roads.

2 Literature review

2.1 Introduction into the wider context of road safety issues

The world's first motor cars were produced during the 1880s by a German engineer, Karl Benz with the aim of making travel faster and more enjoyable (PR Newswire, 2005). Human beings on the other hand have existed for centuries going about their daily activities. These activities sometimes require travelling which in recent times have involved the use of vehicles in one form or another. However, increased motorisation has not come without a cost to humanity resulting in many deaths and injuries from road traffic accidents.

Road traffic accidents continue to be a public health problem. About 16,000 people die each day from all types of injuries around the world. Injuries represent 12% of the global burden of disease (major diseases and injuries facing the world community) and are the third most important cause of overall mortality and the main cause of death among 1 to 40 year-olds (World Health Organisation, 2001). Road crashes dominate these injuries worldwide with about 25% of all injuries resulting in death being from road traffic accidents (Peden, McGee and Sharma, 2002).

At the inquest into the first recorded world road death in Britain in 1896 (that of Bridget Driscoll, RoadPeace, 2012) the coroner reportedly said 'This must never happen again'. In the same year, the United States of America recorded its first road accident in New York when a motor vehicle collided with a pedal cycle rider (Kane, 1981). Into the twenty first century, many people are still killed and injured on roads all around the world. The total waste of human and societal resources resulting from road traffic accidents cannot be overemphasised. Worldwide, a conservative estimate of between 20 and 50 million people are injured or disabled each year in road traffic accidents (Murray and Lopez, 1996; Murray et al., 2001; Roberts, 2005). The wide-ranging estimate is largely due to the known underreporting of road traffic accidents.

Road traffic accidents continue to be a global public health problem and indications are that this will continue to worsen unless urgent action is taken (Roberts, 2005). International organisations including the World Bank and the World Health Organisation (WHO) have all taken a keen interest in the road accident situation facing the world community.

In an effort to increase public awareness about this menace affecting communities, various initiatives have emerged over the years. The World Health Assembly in 1974 adopted Resolution WHA27.59 acknowledging road traffic accidents as a major public health problem requiring the coordinated efforts of member states (World Health Organisation, 1974). For the first time in the history of the World Health Organisation, World Health day 2004 was dedicated to road safety under the theme 'road safety is no accident'. Also, the United Nations General Assembly in Resolution A/RES/60/5 (adopted in October 2006), called for member states and the international community to designate the third Sunday in November every year as the World Day of Remembrance for road traffic victims, in recognition of road traffic victims and their families' loss and suffering (World Health Organisation, 2005). The latest campaign includes the years 2011 to 2020 being declared by the United Nations as the decade of action for road safety aimed at preventing five million road traffic deaths globally by 2020 (United Nations, 2012).

In 2002, road accident was ranked the tenth leading cause of death worldwide (Peden, McGee and Sharma, 2002). It is forecast that by the year 2020 road accidents would move up to third place in the table of leading causes of death and disability facing the world community (Murray and Lopez, 1996; Peden, McGee and Sharma, 2002; Kopits and Cropper, 2003; Peden et al., 2004) with speed being the main contributory factor in these accidents. Another publication (The World Bank, 2011) indicated that annual deaths from road traffic accidents are forecast to rise to 1.9 million by 2020 and it will be the leading health burden for children over five years old in developing countries by the year 2015. More recently, the 2013 WHO Global Status Report (World Health Organisation [WHO], 2013) on road safety also revealed that if no urgent action is taken by 2030 road traffic deaths will become the fifth leading cause of death. The report also mentioned speeding as a major road safety problem in all countries.

2.1.1 The origins of speed limits in the UK

UK road speed limits are used to define the maximum legal speed limit (it can vary) for road vehicles using public roads and are one of the ways used to control traffic speed (Safe Motoring, 2013). Speed limits are commonly displayed on nearby traffic signs or indicated by the presence of street lighting. Speed limits in miles per hour are displayed on road signs or through the use of the national speed limit (NSL) symbol. The Locomotive Act of 1865 for speed limits in Britain restricted the speed of horse-less vehicles to 4mph in open country (towns and villages) and 2mph in towns. The act necessitated the use of three drivers for each vehicle with two travelling in the vehicle and one walking ahead of the vehicle carrying a red flag. On 28 January 1896, Walter Arnold of East Peckham, Kent (UK Motorists, 2014) became the first person to be successfully charged with speeding in Great Britain. The maximum speed limit was increased to 14mph and then to 20mph in 1903. 1930 saw the end of speed limits for cars and motorcycles. In 1934, a 30mph speed limit was introduced on roads in built-up areas (roads with street lighting) and this has stayed to this day (UK Motorists, 2014; Safe Motoring, 2013). The maximum speed limit on UK roads has been 70mph since 1965. This limit is applicable to unrestricted motorways and dual-carriageways and to cars (including car-derived vans) up to 2 tonnes maximum laden weight (MLW), motorcycles, buses, coaches and minibuses up to 12 metres long and goods vehicles with MLW of 7.5 tonnes. Single carriageways carry a speed limit of 60mph. In 1999 local authorities were given the power to introduce 20mph limits without the need to obtain permission from the Secretary of State. Speed limits have traditionally been enforced by the police using speed guns, automated in-vehicle systems (mobile cameras) and automated road side cameras (UK Motorists, 2014; Safe Motoring, 2013). Speed limits on motorways and trunk roads are set by the Highways Agency with the government giving advice to traffic authorities (County, District and Borough Councils, but not Parish Councils) on setting local speed limits. These authorities determine the most suitable speed limit for their roads which depend on local factors and conditions such as history of road accidents, traffic flows, road traffic combination, nature of adjacent development and road geometry (Safe Motoring, 2013).

2.1.2 History of speed cameras and vehicle activated signs

The Road Traffic Act 1991 enabled changes to be made to the law allowing courts to receive proof of speeding from type approved cameras. This had to be accompanied by a certificate signed on behalf of the relevant police force. It enabled safety cameras (speed and red traffic light cameras) to be managed by the police forces. In 1992, the first deployment of 21 fixed speed cameras and 12 red-light cameras was carried out in West London. This continued to increase and by the year 2000 an estimated 4500 safety cameras had been deployed on British roads with most being fixed speed cameras and a smaller number being red-light and mobile cameras. One of the main impediments to the rapid implementation of safety cameras during the early years was that of resource. Benefits in the form of revenues received from speed camera fines were not received by the police forces instead it went to the Consolidated Fund of the Exchequer (Ward, 2003). In April 2000, a new system allowing fines from speed and red light cameras to pay for the costs associated with camera enforcement was piloted. The success from this trial allowed the government to continue with the system establishing the required legislation in Section 38 of the Vehicles (Crime) Act 2001. The Department for Transport in April 2007 changed the funding arrangements such that fines from cameras remained with the Treasury. It created a separate road safety fund for local safety partnerships to be used for a variety of road safety activities in addition to paying the cost or supplying and operating the cameras. In 2010, the new Coalition Government announced it will no longer fund new speed cameras demanding local authorities and the police to publish data about speed cameras, including accident and casualty figures amongst other requests (ROSPA, 2011).

In the late 1970s and early 1980s, the Transport Research Laboratory (TRL) conducted research on automatic signs that provided drivers with information relating to either following closely or excessive speed. The signs stayed unlit unless a driver exceeded a predetermined threshold which was associated with either the distance from the vehicle in front or the speed of the vehicle. Signs indicating to drivers to 'MOVE APART' were made using a back-lit message and it depended on an overhead infra-red detector to measure the separation of the close following vehicle from the vehicle in front. For signs indicating to drivers to 'SLOW DOWN' a message was created by means of a number of pinpoints of light singly provided through fibre-optic cables. Inductive loops hidden in the carriageway were used to measure the vehicle speed. Current generation of vehicle activated signs show a message (symbols and words) outlined by either fibre-optic cables (illuminated by quartz halogen lamps) or light emitting diodes (LEDs) placed on the front panel of the sign. The signs remain blank when not activated by a vehicle. The main types of signs used are either speed enforcing or warning (example, of a hazard) (Winnett and Wheeler, 2002).

2.1.3 Why do people speed

There are different views expressed about why people speed. Research (Corbett and Simon, 1999) has shown that though many drivers recognise the danger posed by high speed they think they are in control when they themselves drive fast and thus they will not be harmed. Others believe they are better than an average driver and less likely to get involved in an accident as a driver than as a passenger (Corbett and Simon, 1999). Drivers who believe they are better than an average driver also hold the view that they are less likely to have an accident in comparison to other drivers. They also believe the roads will be safer if others drove like them with fast drivers believing they are less likely to have an accident than slower drivers (Finn and Bragg, 1986; Corbett and Simon, 1999; Horswill and McKenna, 2005; Comte, Wardham and Whelan, 2000; Warner, Ozkan and Lajunen, 2010).

The reasons behind drivers speed have been grouped into five factors; inattention, thrill, time pressure, disdain of driving and ego gratification by a research study conducted by Gabany, Plummer and Grigg (1997). It is worth noting that the sample used in this study comprised college students with no other sample population to compare with. Ego gratification was highly regarded by males compared to females with younger people being more in favour of risk taking than older people. Time pressure, inattention and disdain of driving were mostly agreed upon by females. These factors were identified by the students as what they believed were the reasons for people exceeding the speed limit and not necessarily their reasons for exceeding the speed limit.

A review of some 40 studies showed a positive relationship between sensation seeking and risky driving (Jonah, 1997). Paradoxically some car manufacturers display speed as a key performance feature or a higher standard of sporting performance and it is of no surprise that the number of people watching the Formula One motor racing show keeps increasing. It was reported that the biggest innovation introduced to the Formula One racing in 2009 was derived from the KERS (Kinetic Energy Recovery System) technology for capturing and storing the car's braking energy instead of wasting it as heat (The Economist, 2009) - the idea being that racing drivers will be able to make use of the energy stored to deliver quick bursts of speed for overtaking, making the sport more entertaining. During the 2010 FIA

Formula One World Championship the total global television audience was 527 million people in comparison to 516 million people in 2009 making it one of the most watched television sports programme and clearly showing the increasing thrill speed offers to people (Formula1, 2011).

Tranter and Warn (2008) investigated the relationship between interest levels in motor racing, speeding attitudes and speeding violations on public roads. Of about 5000 questionnaires distributed randomly to households in Queanbeyan, Australia a return yield of 524 was realised. The focus of the study was on mature drivers, aged above 25 years and had 2 or more years of driving experience. The final analysis sample obtained was 478 respondents with results revealing that the level of interest in motor racing is significantly related to speeding attitudes.

In another instance (Kanellaidis, Golias and Zarifopoulos, 1995) 207 fully completed questionnaires in the absence of an interviewer were obtained from a random choice of drivers. Male and female drivers aged 18 to 68 with driving experience ranging from 1 to 42 years and education ranging from elementary to university was represented in the study. The main reasons identified by respondents for speeding include being in a hurry, desire to show off to other people, underestimation of the risk of speeding and over estimation of driving abilities.

2.1.4 The contribution of speed to road accidents

Road accidents do not just happen but result from a number of factors. These factors have been grouped as Road user, Vehicle and Road environment (O'Flaherty et. al., 1997). Of these factors, the human element of the road user contributes to about 95 percent of road traffic accidents of which speed is a causative factor.

The road user factors are classed as impairment, errors in perception, skills deficiency, and execution manner. The road environment factors involve adverse road design, unexpected obstructions, deficiency in road furniture or markings and adverse environment. Vehicle factors involve those due to a deficiency in keeping up with regular maintenance of the vehicle by the user.

The likelihood of being involved in a road accident is linked to the vehicle speed (McKenna, 2005). Most studies have indicated a clear relationship between speed and road accidents with the severity increasing as speed increases – as a general rule, it has been shown that there is a 5 percent increase in accidents resulting from a 1mile/h increase in average speed (McKenna, 2005; Taylor, Lynam and Baruya, 2000). At a given road it has been shown that crash rate increase as vehicle speed increases (Aarts and Schagen, 2006). This can possibly be further explained in relation to the common association between kinetic energy ‘K’, mass of the object ‘M’ and velocity or speed ‘v’ of the object given as $K=1/2(Mv^2)$. Assume the object in this case to be the vehicle. In any collision, kinetic energy is released due to the speeds involved and it can be argued that at high speeds the reaction time to environmental changes is reduced requiring greater stopping distance and thus leads to increased severity of a road accident.

Schemes with vertical speed deflection measures have recorded a 44 percent reduction in personal injury accidents whereas sites with safety cameras recorded a 22 percent reduction (Hirst, Mountain and Maher, 2005; Mountain, Hirst and Maher, 2005). Mean speeds, 85^{th} percentile speeds and the percentage of vehicles exceeding the speed limit are normally reduced but the difficulty usually lies in reducing the speeds of drivers who continually speed (Hirst, Mountain and Maher, 2005; Hirst, Mountain and Maher, 2005). A 1km/h reduction in vehicle speed has also been shown to lead to a 3 percent reduction in accident risk meaning not only does speed contribute to the severity of road accidents but also to the risk of being involved in the actual accident (Finch et al., 1994). Hauer (1971) developed a relationship between observed accident involvement rates on rural highways for 100 million vehicle-miles travelled and varying travel speeds. The study showed the initial reduced probability of being involved in an accident with increasing travel speed. The lowest probability point was reached in a u-shaped curve but beyond the lowest point on the curve, any more increase in vehicle travel speed resulted in an increased rate of accident involvement.

Elvik, Christensen and Amundsen (2004) carried out a number of studies trying to understand and establish the relationship between speed and road traffic accidents. This has been known as the Power model which makes use of a meta-analysis to provide approximations of how changes in speed influence road accident and road users’ injury severity numbers using a power function. A power function is a mathematical equation or function relating two variables to each other by raising values of one of the variables to a power in order to obtain values for the other variable (Elvik, Christensen and Amundsen, 2004). For example the relationship for fatal accidents was given as shown in equation 1 with an exponent of 4 proposed.

$$\frac{Fatal\ accidents\ after}{Fatal\ accidents\ before} = \frac{Speed\ after^4}{Speed\ before^4}$$

…………….Equation 1

In an attempt to evaluate the validity of the Power Model some inconsistencies were found. This required the model to be reformulated such that the various levels of accident or injury severity did not overlap but instead treated as mutually exclusive categories. Despite the limitations associated with the study it was shown that a strong and consistent statistical relationship existed between speed and road safety such that an estimated 10 percent reduction in the mean speed of traffic results in a 37.8 reduction in the number of fatalities. Elvik (2009) later provided an updated analysis of the relationship between speed and road accidents. The findings differed from the original study with regards to first the exponents which were found to differ depending on the initial speed such that two new versions of the Power model were developed. One model applied to urban and residential areas and the other to rural roads and freeways. The second finding was that the exponents were adjusted with a tendency to become smaller over time indicating that the effects on speed also become smaller. Despite these findings one thing that stood out from all these studies was that speed is a very important risk factor in accident occurrence and the severity of injury sustained. A more recent study by Elvik (2013) attempted to reanalyse the Power Model by fitting exponential functions to data points. Changes in speed and accidents were sorted in groups of 10km/h depending on the initial speed. Even though the exponential function and the power function fitted the data almost equally well, a clear-cut distinction was found between the functions especially at high speeds. Generally, the analyses indicated a stronger support to the use of an exponential function than the power function with the exponential function indicating that the effect on accidents from a given relative change in speed is greatest when initial speed is highest. These studies showed (Elvik, Christensen and Amundsen, 2004; Elvik, 2009; Elvik, 2013) that speed plays a significant role in road traffic accidents.

2.1.5 The contribution of speed to pedestrian fatality

Pedestrians are identified as the most vulnerable road users in most countries globally with a relationship existing between vehicle speeds and their rate of involvement in accidents and injury outcome (Peden et al., 2004).

In UK, the 2010 road accidents report (Department for Transport, 2010) identified pedestrians as the most vulnerable road users accounting for 405 deaths, a reduction compared to the 2009 figure of 500. The problem of pedestrians being classed as vulnerable road users has for a long time been identified in literature. Ashton and Mackay (1979) set the scene and drew light on the relationship between vehicle speeds and pedestrian injuries. It was noted in the study that the speed distribution on impact from the front of a vehicle was dependent on the severities of the injuries sustained. Taking into consideration the design of cars in the 1970s, pedestrians hit at impact speeds less than 30km/h mainly sustained slight injuries with speeds above 30km/h mainly causing non-minor injuries. At speeds ranging from 50km/h to 60km/h injuries sustained varied from survivable to fatal. Figure 1 shows the cumulative impact speed distributions for pedestrians struck by the front of cars. For all casualties, the 50th percentile impact speed varied from 20 to 25km/h with non-minor injuries having the same percentile speed hovering around 35km/h. At the 90th percentile impact speed, all casualties record a speed of 40km/h, 50km/h for non-minor injuries and 65km/h for fatalities.

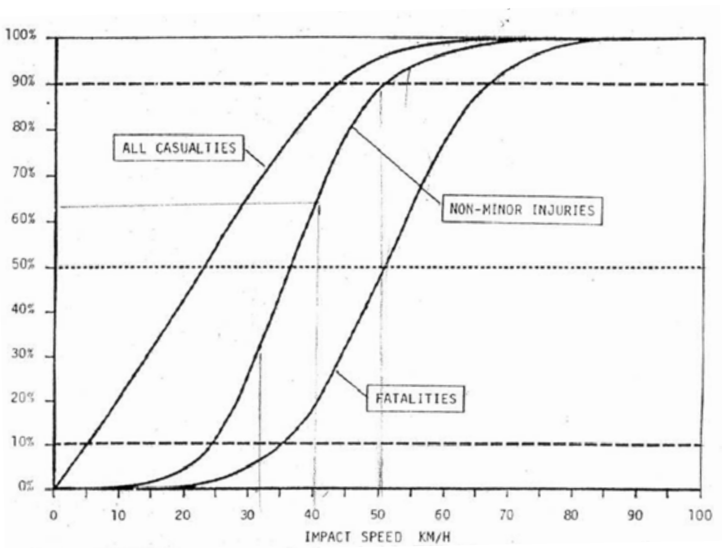


Figure 1 Cumulative impact speed distributions for pedestrians struck by the front of cars or car derivatives (Ashton and Mackay, 1979)

Some differences have been noted in other literature since the studies carried out by Ashton and Mackay (1979). Investigation by Rosén and Sander (2009) into pedestrian fatality risk as a function of car impact speed showed a much lower fatality risk than was generally shown in literature. This was explained to be a result of sample bias towards severe injury accident by earlier works. However a strong influence on impact speed was obtained since fatality risk at 50km/h was more than twice as high as the risk at 40km/h and more than five times higher than the risk at 30km/h. These results were obtained from a sample of 490 adult pedestrians aged 15 to 96 years. Adults were of particular interest in the study as there was the need to distinguish adult pedestrian fatality risk from child pedestrian fatality risk due to the anatomical and biomechanical differences observed in children and adults as reported by Tarriére (1995) (in Rosén and Sander (2009)). Figure 2 below shows the relationship between fatality risk and impact speed for adults with the dotted curve giving an indication of the 95th per cent confidence limits.

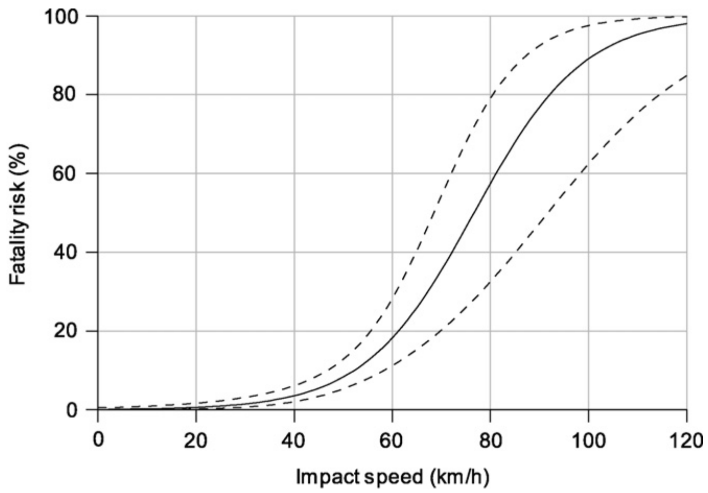


Figure 2 Fatality risk as a function of impact speed for adult pedestrians hit by the front of a passenger car (Rosén and Sander, 2009)

A review of published literature up to and including the year 2009 on pedestrian fatality risk as a function of car impact speed was carried out by Rosén, Stigson and Sander (2011). In all, 11 relevant literatures were identified. Consistently running through all the literature was the fact that fatality risk increased monotonically with the impact speed of the car but absolute risk estimates varied considerably. Another striking feature noted was that almost all the data sources used in studying pedestrian fatality risk was made up of a higher percentage of fatalities in comparison to the corresponding national pedestrian statistics introducing a phenomenon of sample bias. Pre-2000 studies used direct analyses of data which had a huge bias towards severe and fatal injuries resulting in an overestimation of fatality risks. On the other hand it was noted that post-2000 studies made use of less biased data or adjusted for bias and though it revealed a steep increase of risk with impact speed the risks estimated were much lower than had been previously reported (Rosén, Stigson and Sander, 2011). Even though the level of pedestrian fatality risk as a function of car impact speed has been shown to be much lower than previously envisaged in the older literature, there is one common theme that runs through all these literatures being that a relation exists between pedestrian fatality risk and car impact speed.

2.1.6 The contribution of speed to car occupant fatality

In Great Britain, car occupants have remained the largest road user group in each casualty severity (Department for Transport, 2014) with cars making up 80 percent of traffic. In 2014, 45 percent of road deaths were car occupants rising by 1.5 percent in comparison with 2013. There was a 7.6 percent increase in the number of car occupant fatalities in 2014. In 2014, there was a 5.2 percent increase in the number of seriously injured car occupants in comparison with the 2013 figure.

In January 1983, the use of seat belts by front-seat occupants of cars and light vans came into operation in Britain (Mackay et al., 1992) with the law aimed at reducing the national casualty figures. Mackay et al., (1992), investigated 120 fatal cases in a five year study period of cars less than six years old where at least one of the occupants being in either the front or rear of the vehicle died. In this study 83 percent of front seat occupants who used seat belts died. Results from the study indicates intrusion as a result of structural mismatch from large trucks in frontal collisions and the very high energies involved in car to car frontal collisions. The seat belt was unrelated to the mechanism of injury in 80 percent of the cases. There was frequent occurrence of abdominal injuries in fatally injured drivers as compared with passengers. Fatal head/neck injuries dominate as a major cause of death. The sample was biased towards collisions resulting in occupant injury with the study focussing on fatally injured occupants.

Miltner and Salwender (1995), analysed 319 cases of seat belt restrained front seat car occupants (234 drivers and 85 passengers) from 241 vehicles involved in a car to car head on collision in Germany. The main factors contributing to occupant injury severity were the energy equivalent speed (EES) which represents the energy transfer in an accident, the change in velocity, the maximum deformation depth and the angle of collision. The results indicated fatal injuries to be expected in EES greater than 50km/h with no occupant remaining uninjured above 60km/h. The position of occupants affected only the severity of head injury with drivers being more severely injured than passengers. These are consistent with the findings of Mackay et al., (1992). Generally, results from the study show that the severity of injury at all body locations and the total injury severity of front seat occupants is primarily determined by the dynamics of the accident and passenger tolerance to mechanical loading.

A mathematical model was developed by Buzeman, Viano and Lövsund (1998) to estimate average injury and fatality rates in frontal car-to-car crashes for changes in vehicle fleet mass, impact speed distribution and inherent vehicle protection. The results revealed a possible 27 to 35 percent improvement in frontal crashes as a result of a 10 percent increase in fatality risk parameters which reflects significant improvement in inherent vehicle protection. A 10 percent impact speed reduction was obtained from a 40 percent safety improvement. The effects of vehicle fleet mass were not strong but found to depend on the average mass ratio of the fleet. The study noted a reduction in mass range would be advantageous. With a uniform mass reduction of 20 percent, there would be a 5.4 percent increase in fatality rate. Impact speed reductions strongly improved traffic safety.

The EU Directive 96/79/EC (1996) makes provision for a high level of protection to belted occupants of motor vehicles in Europe in the event of a frontal impact by introducing frontal impact test requirements including biomechanical criteria. EuroNCAP (European New Car Assessment Programme) also plays a significant role in the impact performance of cars. The tests are based on those developed for legislation by the European Enhanced Vehicle safety Committee (EEVC) for frontal and side impact protection of occupants of cars and for the protection of pedestrians hit by the front of cars (EuroNCAP, 2004). Cars manufactured from the mid 1990s onwards have shown an overall reduction in car occupant injuries from national accident data in Great Britain (Frampton, Page and Thomas, 2006). Frampton et al., (2002) analysed national car to car crashes occurring in 1997 and 1998 and used it to estimate changes in mean casualty rates between cars manufactured from 1988 to 1992 and 1993 to 1998. The use of only two years of data was to reduce the effects of any accident reduction measures. The outcome of the study showed a reduction in the casualty rate of newer cars with an 18 percent reduction in fatalities obtained. The killed/serious injury rate was reduced by as much as 15 percent. Taking into consideration that significant change to crash safety during the evaluation period was targeted at frontal crashes, it was assumed that the benefits obtained in injury reduction could have resulted from improved frontal crash protection.

Frampton, Page and Thomas (2006) used a stratified random sampling method based on injury severity to assess accident cases for investigation. The study aimed to investigate the scope for further fatality reduction using passive safety improvements in frontal crashes. The sampling considered crashes that involved towed cars less than 7 years old at the time of the accident in selected rural and urban roads in Great Britain. The Equivalent Test Speed (ETS) was used as the crash severity measure. ETS is computed on the assumption that deformation is caused by impacting a rigid barrier with the force directed through the centre of the crush area with the vehicle not being assumed to be brought to rest (Frampton, Page and Thomas, 2006). Results from the study showed no evidence in support of increasing crash test speeds. However, at least 27 percent of fatal drivers and 39 percent of all fatal front seat passengers have potential for survival given attention to older occupant's chest injury tolerance and passenger compartment's intrusion under 60km/h. When belted, fatally injured front seat occupants in frontal crashes with no significant overrun was considered an estimated survival potential for 49 percent of drivers and 60 percent of front seat passengers with improved passive safety was observed. The study suggests that targeting unbelted occupant protection could have additional benefit. Figure 3 shows the distribution of crash severity for fatally injured belted drivers with airbags. A median ETS of 50km/h with a 32 to 65km/h interquartile range obtained. Results from the study showed 78 percent of driver deaths were obtained with an ETS below 66km/h which is consistent with the EuroNCAP crash severity. With an ETS below 60km/h, the study revealed that 68 percent died (consistent with the crash severity for EU Directive 96/79/EC). Evidence from the results obtained was not immense enough to require raising of frontal crash test speeds specifically with the added suggestion for ‘stiffening’ of vehicle structures. The sampling method used in the study gives a bias towards serious injury crashes due to about 80 percent of serious and fatal crashes are included with the investigation of all fatal crashes in the sample areas studied.

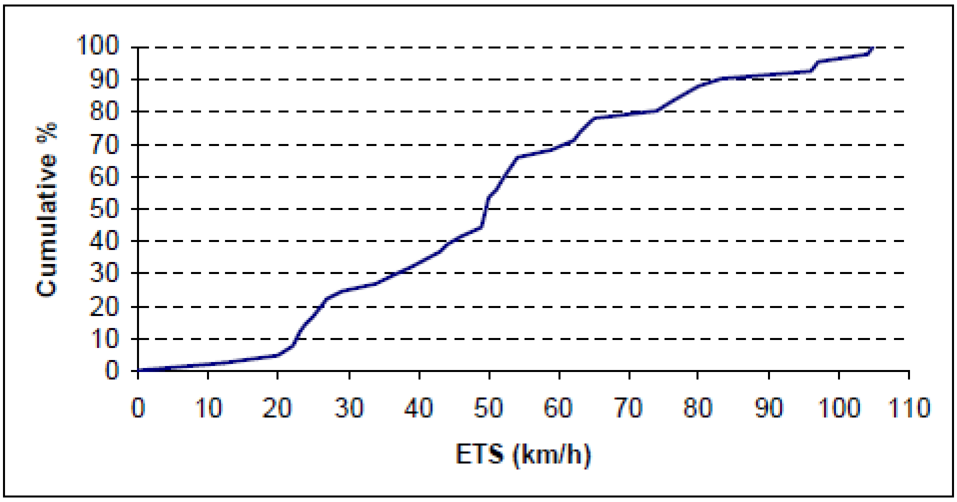


Figure 3 ETS Distribution for fatally injured belted drivers with airbags (N=41) (Frampton et al., 2006)

Frampton, Page and Thomas (2006) were able to show that frontal crash fatalities do not always arise from crash conditions that are extraordinary. Many of the frontal crash fatalities arise from crash severities that are well below those examined with crash tests and thus there is scope to improve protection for fatal injuries using passive safety.

2.1.7 Pedal cyclist as vulnerable road users

The vulnerable road user group which comprises pedestrians, pedal cyclists and motorcyclists are excessively represented in casualty numbers taking into consideration the distance travelled by these. Pedal cyclists present a noticeable difference in casualties. Even though pedal cyclists have a similar fatality rate to pedestrians in terms of deaths per billion miles travelled (around 35 to 38 deaths per billion miles travelled), their overall reported casualties is very different. Pedal cyclists record a close rate to motorcyclists for casualties of all severities at more than 6, 500 casualties per billion passenger miles (Department for Transport, 2014). Pedestrians however have a rate of 2.110 casualties per billion miles walked. With the exception of 2012 to 2013, the number of seriously injured pedal cyclists has increased every year since a low figure of 2,174 was recorded in 2004. This shows a growing issue with pedal cyclists casualties. The development of countermeasures to prevent crashes with through bicyclists on priority roads crossing a minor road at non-signalised intersections within built up urban areas was identified (Schepers et al., 2011). Speed reduction measures for drivers leaving or entering the main road was identified as the most effective measure to improve the safety of cyclists. However red coloured pavement and other markings were found to decline the safety of cyclists.

Morris et al., (2013) assessed the impact of current and upcoming Intelligent Transportation Systems (ITS) applications on the safety and mobility of vulnerable road users. The majority of cyclist accidents were found to occur in urban areas on relatively low speed limit roads. Also most cycling accidents happened at junctions/intersections.

2.2 Factors influencing vehicle speed

2.2.1 Introduction

A number of factors come to play to influence a vehicle's speed and for that matter a driver's choice of speed on the road. Various researches have been carried out in an attempt to identify these factors. Figure 4 below summarises the findings of Wahlgren (1967) on the general groups of factors influencing vehicle speed.

In this chapter a literature review was carried out on each of the factors shown in Figure 4. Some of the studies carried out on these factors are quite dated but it is important to consider them in order to have a broader perspective of investigations carried out to date on various factors identified to affect speed. Taking into consideration the availability of data and the extent of influence of these factors on road traffic accidents, it is worth stating that some of the factors identified will not be included in the accident prediction model to be developed in chapter 5.

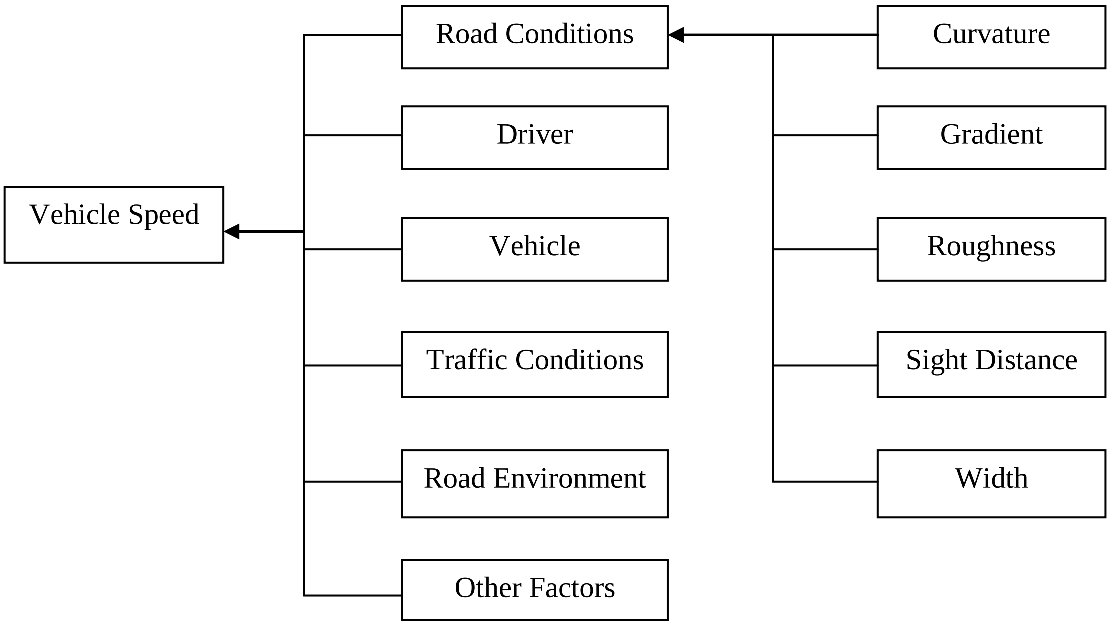


Figure 4 Factors influencing vehicle speed (Wahlgren, 1967)

In identifying the influence of risk on road traffic accident deaths, Smeed (1949) and Smeed (1972) derived a formula to calculate the number of road traffic accident deaths in a country (D) based on the number of licensed motor vehicles (N) and the population (P). The study was based on 1938 data for 20 countries and is given in equation 2 as

$$D = 0.003(NP^2)^{1/3} \qquad \text{………………Equation 2}$$

This formula into road traffic accidents was basic and did not include other factors such as the road characteristics. After these initial works, other studies have evolved taking into consideration the influence of engineering, economics and policy on road traffic accidents. The use of laws is another tool applied in countries to contribute to road safety improvement.

Bjørnskau and Elvik (1992) used a game theory model to understand the relationship between road user behaviour and police enforcement on road safety. While some researchers focus on identifying what measures are effective at reducing road traffic accidents, others have tried to establish the causative factors and establish trends in road traffic accidents.

Other researches have looked at using statistical or empirical approaches to make deductions about road accident causation seeking to unveil 'myths' about the mechanics of accident causation. In an attempt to explain the mechanism of road traffic accident causation, Elvik (2006) revealed some statistical regularities that determine the shape of the relationship between risk factors and accident occurrence referring to them as "laws of accident causation". The laws proposed by Elvik (2006) are: *'The universal law of learning'* which states that the accident rate per unit of exposure will decline as the amount of exposure increases. *'The law of rare events'* states that the more rarely a certain risk factor is encountered the larger its effect on accident rate. '*The law of complexity*' states that the more units of information per unit of time a road user must attend to, the higher becomes the probability that an error will be made and finally *'The law of cognitive capacity'* which states that the more cognitive capacity approaches its limits, the higher the accident rate. Even though this study by Elvik (2006) does not provide a lot of empirical data to support his reasoning, it does provide some useful explanation to the reasons behind factors affecting road traffic accidents and it makes way for more statistical analysis to validate these laws.

In the development of the study by Elvik (2006), Davis and Swenson (2006) applied a causal model to three freeway rear-end collisions. This model compared what happened to what would have happened had the supposed cause been absent. One of the findings from this study revealed short following headways by the colliding drivers as probable causative factors in the collisions.

2.2.2 Road Conditions

2.2.2.1 Road curvature and infrastructure

Horizontal curves have an effect on road safety and especially on two lane rural roads due to increased speed of driving along these curves (Johnston, 1982; Bhatnagar, 1994). Increasing the degree of horizontal curvature increases the number of accidents (Haywood, 1980; Johnston, 1982; McDonald, 2004) with highways having single sharp curves in combination with long tangents and flat curves creating situations which are often hazardous. Radius of curves less than 600 metres have been found to be over represented in road accidents (Choueiri and Lamm, 1987; Johnston, 1982) with enough evidence also suggesting that horizontal curve radii less than 400 metres contribute to road accidents. In the case of single vehicle accidents, there is a 34 percent increase in accident frequency per sharp curve per kilometre (McDonald, 2004). A sharp curve was defined as one marked with a chevron and/or curve warning sign. The definition given for sharp curve was vague but it did serve as a guide to the authors for assessing the data.

There is however some argument among researchers that operating speeds are affected by the characteristics of the preceding section of the road (McLean, 1979; McLean, 1981). Research has produced speed prediction models for individual curves with these models found to be simple and possess enough predictive power with R^2 values ranging from 0.75 to 0.95 (McLean, 1981).

Among the other important parameters used in regression equations for predicting operating speeds on horizontal curves are degree of curvature, curve radius, length of curve, deflection angle, cross section and superelevation. Curve radius was a widely used parameter since it is considered the most important parameter in obtaining operating speeds along horizontal curves (Bennett, 1994; Abdul-Mawjoud and Sofia, 2008). Other studies have also shown that the 85^{th} percentile speeds on horizontal curves can be predicted in combination with parameters such as curve radius, superelevation, deflection angle and length of curve (Kadiyali et al., 1981; Lamm and Choueiri, 1987; Krammes et al., 1995).

In another study based on passenger vehicles, there were 28 sites used to develop the speed prediction model and 20 other sites used to validate it for horizontal alignments of rural two lane highways in northern Iraq (Abdul-Mawjoud and Sofia, 2008). The regression models developed to predict the 85^{th} percentile speed was also based on the geometry of the curves with the approach speed on the preceding tangent being based on various gradients. This study used 4 independent variables which were 85^{th} percentile approach speed, radius, deflection angle and superelevation. With the application of data from 20 validation sites to the four individual equations developed for different ranges of gradient it was found that the mean absolute percent error in predicting the 85^{th} percentile curve speed ranged from 7 to 9 percent indicating a reasonable degree of precision of the model.

The curvature change rate of single curves (CCR_s) was identified to be highly successful at explaining the variability in operating speeds and accident rates (Lamm, Psarianos and Mailaender, 1999; Charlton and dePont, 2007) and it expresses CCR_s in gon/km[5] (gon is a measure of angle and there are 400 gon in a circle in comparison to 360 degrees or 2π radians). For multiple curves, a curve radii ratio was the measure employed and it was given as a ratio of the curve radius to the radius of the immediately preceding curve. Ratios greater than 0.8 had relatively minor effect on crash rate. For ratios less than 0.8 the crash rate increased as the ratio decreased and rose rapidly for values less than 0.2.

The road infrastructure as a parameter has been identified to impact on road traffic accidents. Shankar, Mannering and Barfield (1996) used a negative binomial model to establish the effect of weather and geometric parameters on road traffic accidents. The model suggests that effort should be taken to avoid steep grades and horizontal curves with low design speeds in areas with adverse weather. As an example the study revealed that removing all horizontal curves with design speed less than 96.5 kilometres per hour on a roadway section experiencing at least 5.1 cm of snowfall one or more days in a month can reduce the monthly accident frequency by 47.3%. Milton and Mannering (1998) used a negative binomial regression to isolate the effects of various highway geometric and traffic characteristics on annual accident frequency for the state of Washington, USA. Results from this study indicate accident frequency and exposure to accidents increase as the length of road section increases. Horizontal curves with more space between them had a tendency to increase accident frequency. An increase in horizontal curve radius decreases the number of accidents and smaller tangent lengths preceding a horizontal curve were found to lower the frequency of accidents. Also, road widths less than 3.5m were found to contribute to a reduction in accident frequency. Haynes et al. (2008) investigated the influence of road curvature on fatal crashes in New Zealand and found no evidence that frequently curved roads had more crashes than elsewhere supporting an earlier finding by Milton and Mannering (1998). The number of junctions per kilometre was found to have a negative association with crash rate and this was found to be strongly associated with urban settings. One of the limitations identified in the study by Haynes et al.(2008) was the small numbers of data (4058 fatal crashes) used for the analysis since the statistical power of the study was limited given that the data was distributed onto 73 territorial local authority roads.

Noland (2003) used data from 50 states in the US over 14 years to analyse the effects of road infrastructure improvements on traffic related fatalities and injuries while controlling for other factors known to affect safety in general. Results from the negative binomial regression model did not confirm the hypothesis that infrastructure improvement effectively reduces total fatalities and injuries. In a later study, Noland and Oh (2004)*)* were still unable to support the hypothesis that changes in road infrastructure and geometric design have benefits on road safety in terms of road fatalities and reported accidents. They however were able to deduce that an increase in the number of lanes was linked to both an increase in traffic-related accidents and fatalities. Also an increase in lane width was found to be linked with an increase in fatalities with an increase in outside shoulder width contributing to a decrease in accidents.

Navin, Zein, and Felipe, (2000), Peréz (2006) and Gomes and Cardoso (2012) revealed the association between certain types of improved road infrastructure and road traffic accidents. Navin, Zein and Felipe (2000) investigated the impact of road safety engineering on whiplash injuries in British Columbia, Canada. The study revealed that simple and affordable solutions such as enhancing signal visibility compared to the more expensive remedies such as the geometric upgrades at intersections were found to cost effectively reduce the frequency of rear end collisions. Peréz (2006), investigating the effects of four engineering treatments on road safety found that highway upgrading had a positive and significant impact on safety with the updating and improvement of traffic signing, repainting of road markings and pavement resurfacing showing no significant impact on safety. In a similar study conducted by Daniels et al. (2010) in Belgium, results indicated no effect of additional road marking on driver speed. Gomes and Cardoso (2012) revealed that the application of low cost engineering measures on a multilane road in Portugal showed a 10% reduction in the expected annual number of personal injury accidents. There was also a 70% decrease in the expected annual number of head on collisions and a 26% reduction in the expected annual frequency of accidents involving killed and seriously injured persons. Even though finding by Gomes and

Cardoso (2012) are similar to those of Navin, Zein and Felipe (2000), the former provided expected savings and not actual savings which makes it difficult to compare results from both authors.

Altering the infrastructure was another subject investigated by a number of researchers to find its effect on improving road safety. De Brabander, Nuyts, and Vereeck (2005) assessing the effects of roundabouts observed that roundabouts tend to be most effective at intersections of a main road with a high speed limit (90km/h) and an adjacent road with a lower speed limit (50km/h or 70km/h). Later, De Brabander and Vereeck (2007) in a study into the effectiveness of roundabouts in Flanders, Belgium built between 1994 and 2000 found that roundabouts were not always effective at improving road safety. Haynes et al. (2007) used geographical information systems to generate indicators of average road curvature from a road-network in England and Wales at the local authority district level. These indicators included number of bends per kilometre, the proportion of straight road lengths, the cumulative angle turned per kilometre, the ratio of road distance to straight distance and the mean angle of each bend. A negative binomial regression analysis was used to establish the relation between each of the road curvature indicators and the number of fatal, serious and slight collisions. It was found that collision numbers were negatively related to road curvature with the cumulative angle being most strongly related to fatal road crashes. A 1° per km increase was associated with approximately 0.5% reduction in crashes. A weakness identified in this study was the use of district averages which did not account for relationships associated with road type within districts. Mountain, Fawaz, and Jarrett (1996) used Generalized Linear Modelling (GLM) to develop regression estimates of expected accidents for six highway categories while improving the estimates obtained with Empirical Bayes by combining them with accident counts. It was found that accidents on highway sections were a non-linear function of exposure and minor junction frequency.

Berhanu (2004) used Poisson and negative binomial regression methods to develop accident predictive models based on data from arterial roads in Addis Ababa, Ethiopia. Findings from this study indicated a decrease in road curvature results in decreased accidents. A possible explanation is that in urban environments, traffic speed is likely to be low and this allows for more driver reaction time to reduce speed on curved sections of the road and have better control of the vehicle. The study by Berhanu (2004) also recognised that drivers tend to speed on straight sections of a road than within bends so there is the possibility of accident risk being reduced at areas of increased road curvature. The relationship between lane width and the total number of accidents on undivided roads was found to be significant with improved safety benefits of lane widening realised for two-lane and four-lane undivided roads.

From the models discussed above, it is clear that a relationship exists between vehicle speeds and road curvature with various levels of statistical significance achieved between road curvature and vehicle speeds. As mentioned by Noland (2003) and Noland and Oh (2004), the positive effect of infrastructure improvement still remains unclear and more research is needed at the local (single road segment) level and area wide level.

2.2.2.2 Gradient

Even though road gradient is known to have an effect on vehicle speeds, the effects described by researchers are varied. In one study, it was found that the 85th percentile speed V_{85} was not varied on downgrade sections when average speeds decrease (Gombard and Louah, 1986). On the other hand another study showed that average speeds significantly increase downgrade (Yagar and Van Aerde, 1983). Both studies however pointed out that speeds decrease on upgrades while other studies indicate no significant effect of gradient on speed (Reinfurt et al, 1992). These studies show that there is an effect of gradient on vehicle speeds with the level of effect varying.

2.2.2.3 Sight distance

Every driver must be able to see ahead to an appropriate distance when driving in order to be able to identify any hazards, take the appropriate action and avoid crashing into an obstacle. Sight distance is one of the fundamental design parameters that should be satisfied in every geometric design. For design purposes when discussing sight distance in United Kingdom (UK) stopping sight distances and full overtaking sight distances are considered (Highways Agency, 2002).

Speed is a factor that has been widely acclaimed to have an impact on sight distances (Harwood, Mason and Brydia, 1999). The stopping sight distance is the distance required for a driver to stop a vehicle travelling at design speed based on design conditions. Passing sight distance is described as the distance required by a vehicle driver on a two-lane road to execute a normal passing manoeuvre based on design conditions and design speed. Finally, the term decision sight distance is the distance required for a driver to detect an unexpected or difficult-to-perceive condition, recognise the condition, select an appropriate manoeuvre and complete the manoeuvre based on design conditions and design speed (Washington State Department of Transport, 2010).

The stopping distance is made up of the distance travelled during perception and reaction and the distance taken to bring the vehicle to a halt. In design, the perception and reaction distance is usually given as the distance travelled in 2.5s at the design speed. In UK, the stopping sight distance is measured from a minimum driver's eye height of between 1.05m and 2.00m to an object height of between 0.26m and 2.00m both above the surface of the road and checked in both the horizontal and vertical plane (Highways Agency, 2002*)*. In urban areas where there is likely to be more distractions from objects, the object height can be increased to 2 ft (0.61m). Stopping sight distance is influenced by both the vertical and horizontal curvature of the road. Vertically, the stopping sight distance is influenced by the presence of crest and sag curves. Computation of sight distance depends on factors such as the reaction time of the driver, vehicle speed, efficiency of brakes, frictional resistance between the tyre and the road surface and the gradient of the road (Highways Agency, 2002*)*.

A study was conducted to establish the relationship between design and operating speeds at crest vertical curves with limited sight distance. 3,500 paired speed data (speeds at control and crest sections) together with geometric data of 36 sites in 3 states in the United States of America was used (Fambro, Fitzpatrick and Russell, 2007). For the range of conditions studied it was observed that both the 85^{th} percentile speed and the mean operating speeds were above the design speeds of the crest vertical curves. Mean reductions in speed between the control and crest sections were found to increase as the available sight distance was decreased.

Leong (1968) measured free speeds at 31 locations on sections of two-lane two-way rural highway in New South Wales, Australia between 1963 and 1967. Sight distance was found to affect free speeds with an increase in sight distance causing an increase in free speeds. An increase in 2.4km/h per 100m of sight distance for mean car speeds was obtained.

Yagar (1984) estimated highway speeds as a combined function of both traffic volumes and the geometric and environmental conditions of the highway. The study was carried out on a 2-lane rural highway in Ontario, Canada and involved 6000 sample points from 37 different environmental locations consisting each of a 1500m stretch upstream. Factors used included road curvature, gradient, lane width, land use, extra lane, access, shoulder width to nearest obstruction, sight distance, centreline marking and speed limit. Results from this study revealed that the only factor found not to be clearly statistically significant was the slope of the road. This was further explained to be probably due to the limited range of that factor in the data set used. Factors found to be statistically significant on speeds were traffic volumes, direction and type of vehicles, existence of driveway access to adjacent land use, access from other highways, speed limit, existence of extra lane and grade/slope. Due to the generality of the models developed, it was advised not to use either traffic volumes or geometric/environmental properties outside the range of values for which the models were calibrated for. Consideration of other locations and slopes with the possibility of testing the model outside Ontario, Canada will be useful to gaining confidence in the results.

2.2.2.4 Road Width

Free speeds were found to be affected by shoulder and pavement widths with an increase in pavement width and shoulder width increasing free speeds. This observation was noted in a study in New South Wales, Australia of free speeds measured at 31 locations on sections of two-lane two-way rural highways between 1963 and 1967 (Leong , 1968).

Widths of road lanes and road shoulders have been found not to be directly related to design speed but do affect vehicle speeds (American Association of State Highways and Transportation Officials (AASHTO), 2001). Where lane widths have been found to influence speeds, the range of possible lane widths was found to be narrow and suggested it was better to produce relationships based on the lane and carriageway widths being considered (Lamm and Choueiri, 1987). An investigation into the possibility of reducing driving speeds by using lane delineation on roads of width

3.6, 3.0 or 2.5m in Australia showed driving speeds to be reduced on the narrowest width road and further reductions were obtained on straight sections of road that had centre marking with painted hatching (Godley, Triggs and Fildes, 2004). To avoid bias from the additional lave width used in conjunction with the control centreline, Godley, Triggs and Fildes (2004) all road centre marking main effect and interaction contrasts were averaged across only the narrow (2.5m) and medium (3.0m) lane widths.

2.2.3 Traffic Conditions

A Poisson regression model was used in analysing data for counties in four countries, (Denmark, Finland, Norway and Sweden) for the effects of factors such as randomness, exposure, weather, daylight and speed limits on road accidents (Fridstrøm et al., 1995). Randomness and exposure accounted for 80 to 90 percent of the observable variation in the data sets with the relationship between exposure and injury accidents appearing to be proportional. Traffic volume was considered an important factor which when decreased would result in a substantial reduction in accidents. In the following subsections the effects of different traffic conditions on road traffic accidents are discussed.

2.2.3.1 Traffic flow and traffic density

Traffic is one of the causative factors of road traffic accidents. Without traffic there will be no need to discuss road traffic accidents. It is thus useful to explore the characteristics of traffic and what it is about traffic that affects road traffic accidents. Typical characteristics of traffic are flow, density, congestion and speed. Since these characteristics are somewhat interlinked it is possible that an explanation into one can provide an insight into the others. The characteristics and how they affect road traffic accidents are discussed.

The relationship existing between speed, flow and density is normally expressed as $q = k\bar{v}$ where q is the traffic flow in vehicles per unit time, k is the traffic density in vehicles per length of road and $\bar{v}$ is the mean speed in distance per unit time. As traffic flow increases, with limited capacity up to a defined point traffic congestion and delay set in resulting in speed reduction. This phenomenon is supported by Hau (1992) stating that traffic density determines speed and not vice-versa. Since traffic flow is a product of traffic density and speed, Figures 5 and 6 provide some explanation into this phenomenon. The rectangular area shown under Figure 5 equates to the traffic flow and is given in vehicles per hour. As traffic density increases, the speed of vehicles are initially kept at a stable maximum speed S^{max} due to the absence of the inflow of more vehicles which will increase density and cause congestion. This maximum speed S^{max} is maintained for a period of time. As density increases on the road segment, vehicles are unable to maintain their maximum speed S^{max} and this is the point at which vehicle speeds begin to fall resulting in congestion. Speeds continue to fall to a point at which the maximum density of traffic on the road leads to zero speed and maximum congestion.

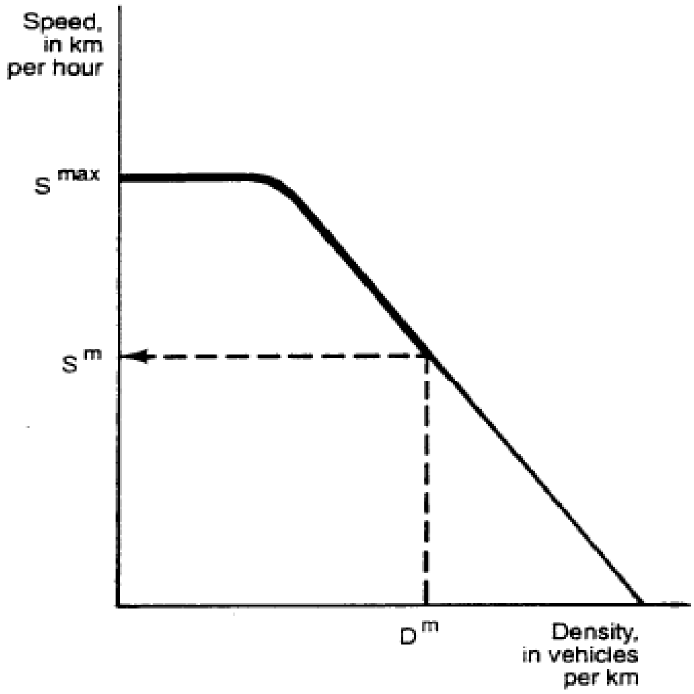


Figure 5 Speed-density curve (Hau, 1992)

Figure 6 shows a similar phenomenon whereby vehicles travel at their maximum speed S^{max} and as traffic flow increases, speeds drop to a maximum flow level of F^{max} referred to as 'Engineering capacity'. At this maximum flow into the road segment vehicle speed begins to drop, density increases and congestion sets in as shown in the lower half of the speed-flow graph reverting to zero in Figure 6.

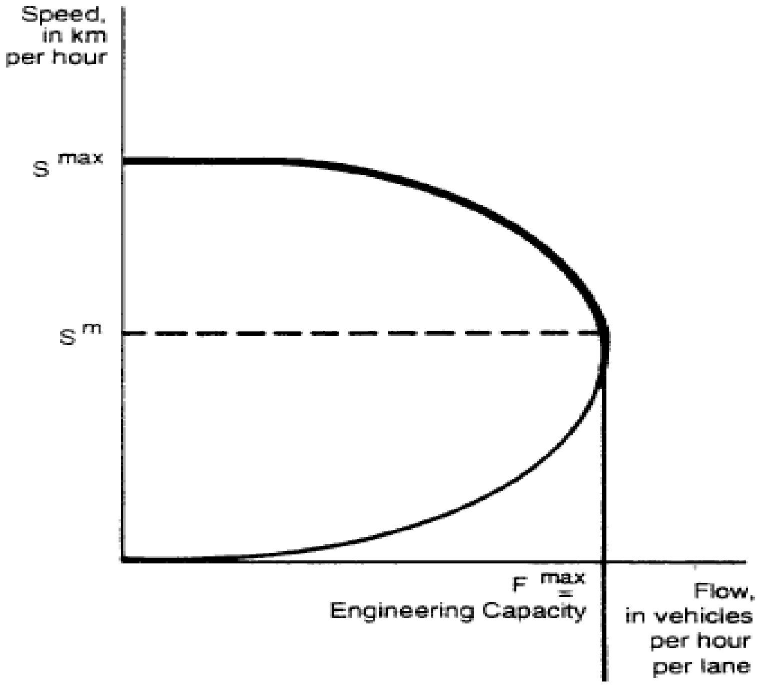


Figure 6 Speed-flow curve (Hau, 1992)

Ceder and Livneh (1978) provided some understanding into the relationship between accident and average daily traffic on interurban road sections by fitting a power function model. The total accident density which was found to increase with an increase in average daily traffic (ADT) was a combined result of a sharp increase in multi-vehicle accidents, moderate decrease in single-vehicle accidents and a negligible effect on pedestrian accidents. These findings reported by Ceder and Livneh (1978) had been noted earlier by Belmont (1953). Later on, Ceder and Livneh (1982) extended their earlier investigation (Ceder and Livneh, 1978) into the relationship between road accidents and hourly traffic flow instead of ADT using a power function model with data over an 8 year period. The reason for using hourly traffic flows was to provide a better understanding into the relation between accidents and traffic flow. A major finding from this study was the significant deterioration in single-vehicle accidents at the mid flow range (around 1000 veh/hr) which is somewhat intuitive.

Ceder (1982) followed on with another study investigating the relationship between road accidents and hourly traffic flow using single and multi-vehicle accident rates in conjunction with free-flow and congested flow conditions. Under free-flow traffic conditions, a U-shaped relationship was observed between the total accident rate and hourly flow curve. Where congested flow data was used which was often characterised by multi-vehicle accidents, accident rates were found to increase sharply with hourly flow.

Martin (2002) investigated the relation between crash incidence rates and hourly traffic volume for French inter-urban motorways and found crash incidents to be lowest at traffic flow rates of 1000 to 1500 vehicles/h. Crash incidence rates increased steadily as traffic increased on 2 and 3 lane motorways to levels of 3000 vehicles/h. The number of crashes was higher on weekdays for heavy traffic with no significant difference found between the number of daytime and night-time crashes irrespective of the traffic level. For a given level of traffic, no difference was found in crash severity by number of lanes or period in the week however, severity was found to be greater during the night when hourly traffic was light.

In a study into the relationship between traffic flow conditions and the likelihood of traffic accidents by type of crash, Golob, Recker and Alvarez (2004) were able to show that the key traffic elements affecting safety are mean traffic volume and median speed as well as temporal variations in volume and speed. Hiselius (2004), using Poisson and Negative Binomial regression models indicated the importance of considering the differences between vehicle types in estimating the effect of traffic flow on accidents. Important information can be lost if no consideration is taken. When traffic was treated as homogeneous, accident rate decreased. However when cars were independently considered accident rate was found to be constant or increased while a decrease in accidents was observed with increase in lorries. Lord, Manar and Vizioli (2005) developed a predictive model to determine the statistical relationship between crashes and hourly traffic flow characteristics. Results show that predictive models making use of traffic volume as the only explanatory variable may not provide a true understanding of accidents on freeway segments. However, models making use of vehicle density and V/C (volume over capacity) ratio gave a better description of crashes occurring in either urban or rural settings. These findings suggest that traffic volume, vehicle density and the V/C ratio directly influence the likelihood and severity of a crash. Ivan, Wang and Bernardo (2000) estimated the Poisson regression model for predicting both single and multi-vehicle highway crash rates as a function of traffic density and other land use factors. It was found that traffic intensity provided explanation into the differences in crash rates even when controlling for factors such as time of day and light conditions with differences occurring for single and multi-vehicle crashes. Due to the limited sample size of 17 sites and variable variability in the study it is advisable not to transfer findings to other sites with more investigation advised.

Golob and Recker (2001) in a study into how accidents occurring on heavily used freeways in Southern California, USA relate to traffic flow, weather and ambient lighting conditions used linear and non-linear multivariate statistical analysis. It was found that the collision type had a strong relation to median traffic speed and to temporal variations in speed within the left and interior lanes. Accident severity had an inverse relation with traffic volume. While controlling for weather and lighting conditions, severity of accidents was found to be influenced more by the volume of traffic than by speed. Levine, Kim and Nitz (1995a) analysed the spatial patterns of 1990 motor vehicle accidents in Honolulu, USA. Spatially, accidents were found to fluctuate dynamically in response to changing traffic patterns and volume.

In another study by Levine, Kim and Nitz (1995b) a spatial lag model was developed to examine the zonal relationship of motor vehicle accidents to population, employment and road characteristics. The model was independently tested for each hour of the day, weekdays and weekends. Test results from the model indicate that the predictors of accidents fluctuate according to different trip generating activities and changes considerably over the day. It is worth noting that the method employed focused on neighbourhood and area characteristics and not just on the road system.

2.2.3.2 Traffic congestion

Traffic congestion is another factor that affects road traffic accidents. Shefer (1994) hypothesised that a negative relation exists between road traffic accident fatalities and road congestion levels. The V/C ratio per unit length of road was used in defining the vehicle density in the study by Shefer (1994). The relation between road fatalities and vehicle density is shown in the bell-shaped curve in Figure 7. During stage I, with few vehicles on the road link the probability of an accident occurring is low. As the number of vehicles increase the number of fatalities also increases. For the initial part of Stage II, a steep slope is shown indicating an increasing rate of change in total fatalities. This is because during this phase high speeds of travel can be accommodated on the link relative to the allowable travel speed.

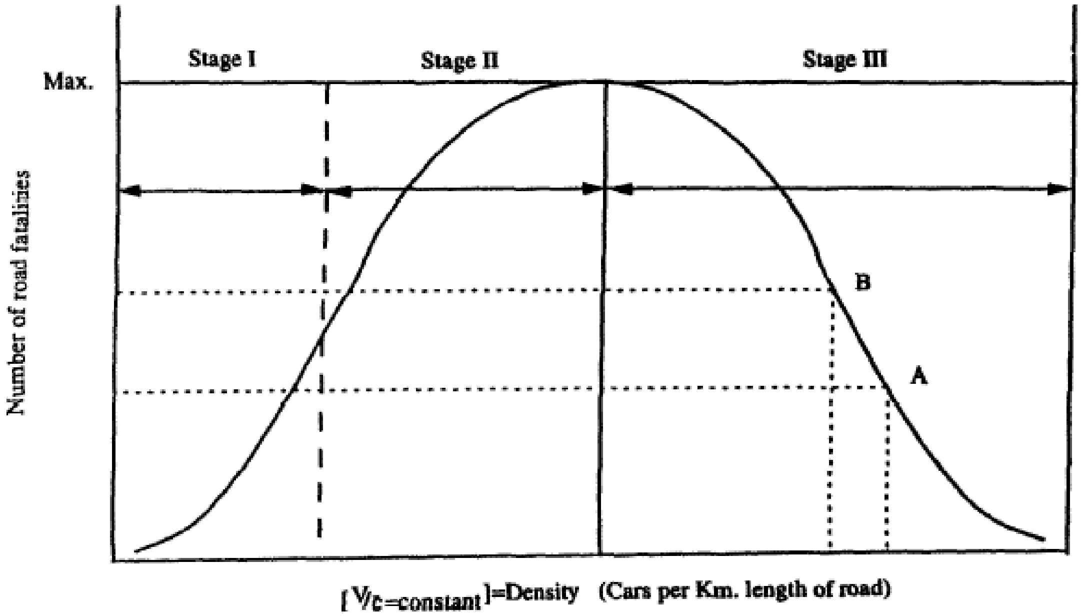


Figure 7 Relationship between road fatalities and vehicle density (Shefer, 1994)

In the second half of Stage II, it can be seen that as vehicle density increases, road congestion occurs leading to a reduction in travel speed. The slope shown is flatter meaning that with an increase in traffic density total fatalities increase but at a decreasing rate. At a certain level of vehicle density and hypothesised fatalities, fatalities begin to decrease and this is depicted in Stage III of Figure 7. This hypothesis can be supported with the reasoning that with an increase in traffic density, congestion sets in, and vehicle travel speed decreases leading to a decrease in the number of road accident fatalities. The argument being put across in Stage III of Figure 7 is in agreement with the study by Martin (2002). The study revealed that light traffic volume is a safety problem as far as the severity and frequency of road traffic accidents is concerned. However, Noland and Quddus (2005) in a spatial analysis study of congestion and safety of roads in London found little evidence to support their hypothesis that traffic congestion may result in some safety benefits. Models developed for both congested and uncongested time periods did not provide any differences to lead to any firm conclusions. Lower casualties were observed at areas with higher minor road density whilst areas with higher A-road density had more casualties. It was further explained to be the result of higher traffic levels on A-roads and lower speeds on minor roads. However, results obtained for motorway density confirmed the hypothesis that congestion may be beneficial to safety.

Uncongested periods showed an increase in serious injuries but there was no increased association with fatalities. The weakness in the proxies used for congestion was mentioned as a possible cause for the inconclusive results obtained from this study. This is because congestion can be localised to a great extent and it can be specific to a particular time of day so more precise definitions of time is required. Another reason to be attributed to the results obtained is that free-flow maximum speeds observed in the study area, London did not exceed 40mph (65km/h) which could have played a role since high speed study areas may display different results. In another study a spatial analysis approach was undertaken into the impact of traffic congestion on road accidents on the M25 London orbital motorway (Wang, Quddus and Ison, 2009). The study controlled for other factors such as curvature, gradient and number of lanes that may affect road traffic accidents. With other studies using a proxy in measuring congestion levels, Wang, Quddus and Ison (2009) sought to use a more accurate measurement for congestion. A series of Poisson based non-spatial and spatial models were used to take into consideration heterogeneity and spatial correlation effects. Results from the study indicated that traffic congestion had little or no impact on the frequency of road accidents. One limitation associated with the study was the lack of investigation into the effects of traffic congestion on road accidents on other roads connecting to the M25 motorway as there may be some spatial variations in congestion levels and on accident frequencies.

2.2.3.3 Traffic speed

Lave (1985) is reported to have said ‘Speed kills, slower is safer’ and went on to quote Schelling (1978) as saying “… the crucial element is often coordination. People need to do the right things at the right time in relation to what others are doing”. In the analysis by Lave (1985) of the effects of 55mph national maximum speed limit on motorists for six different types of high speed roads, the effects of limit-defying behaviour (speeding) and the absence of coordination (speed variance) on the fatality rate was measured using state cross-section data. No statistical relation was found between fatality rate and average speed. However a strong relationship was found with speed variance providing some explanation to variance in speed being the accident causative factor and not speed itself.

Johansson (1996) investigated the impact of lowered speed limit on the number of accident severities using Poisson and negative binomial count data models. The study showed that minor injury and vehicle damage accident numbers appeared to have reduced due to a reduction in speed limit on Swedish motorways. Aljanahi, Rhodes and Metcalfe (1999) in a study into the relationship between various measures of traffic speed during free flow conditions and accident rates for two different geographical locations in Tyne and Wear, UK and Bahrain found statistical correlation between mean speed and accident rate in Bahrain. Tyne and Wear showed a less significant statistical correlation for these parameters but instead a stronger relationship existed between accidents and the traffic speed variability. The study did not provide any further insight into what results would have been obtained if different accident severities had been considered. Various other researches have pointed at the effect speed has on the severity of accidents. Ossiander and Cummings (2002) analysing the effects of increasing speed limits on a rural freeway in the USA found speed variance not to have been affected by the increase in speed limit. Even though the fatal crash rate increased due to the increase in speed limit, the total crash rate showed little change implying that fatal crashes can result without necessarily having an increase in total crashes. One limitation found in the study by Ossiander and Cumming (2002) was that it did not consider the effect of spatial variation in vehicle speeds on the results obtained.

Other studies have also tried to establish the relationship between speed and road traffic accidents. Taylor, Baruya and Kennedy (2002) investigated if there will be a change in accidents on a given road section assuming everyone drove faster than they usually do with all other factors remaining constant. 174 rural road sections with 60mph speed limits across England were used for the analyses. Results from the study showed the frequency of all injury accidents rose quickly with the mean speed of traffic. However, the study only considered 60m/h roads so much comparison cannot be made with other classes of roads with different speed limits. In another study by Taylor, Lynam and Baruya (2000) the frequency of road traffic accidents was found to increase with traffic speed and higher speeds showed a quick increase in accident frequency. It is worth stating that results from the study relate to all injury accidents without differentiating between accidents involving different severities of injury.

2.2.4 Road Environment

The road environment must be taken into account when discussing factors that affect vehicle speeds. A study acknowledged AASHTO's (American Association of State Highways and Transportation Officials) findings of drivers selecting their speed based on the road environment instead of an assumed design speed (Fambro, Fitzpatrick and Russell, 2007). Free speeds were found to be affected by shoulder and pavement widths with an increase in pavement width and shoulder width increasing free speeds. This observation was noted in a study in New South Wales, Australia of free speeds measured at 31 locations on sections of two-lane two-way rural highways between 1963 and 1967 (Leong, 1968). The existence of driveway access to adjacent land use, access from other highways, grade/slope and the existence of an extra lane were found to have statistical significance on vehicle speeds in Ontario, Canada. This was calculated based on models developed for predicting vehicle speeds at the 10^{th}, 50^{th} and 90^{th} percentiles. The study was limited by the lack of severe gradients on high-volume roads and the use of standard road widths (Yagar, 1984).

Karlaftis and Golias (2002) showed that parameters such as geometric design and pavement condition are the most important factors having an effect on road accident rates. In the United States (US) about a third of highway fatalities result from single-vehicle run-off into roadside features (Lee and Mannering, 2002). A test on the combined effect of three roadway elements; shoulder width, existence of guardrail and roadway geometry (curvature) on driving showed a significant effect of roadway geometry on driver speeds. However, only in the presence of a guardrail did shoulder widths have a significant effect on actual speeds, lane position and on perceived safe driving speed (Ben-Bassat and Shinar, 2011). In the absence of a guard rail, the shoulder width loses a lot of its advantages and impact on driving behaviour. The results also indicated that roadway geometry can be used for reducing driving speeds but it can also have negative impacts in maintaining a stable lane position in sharp curves (Ben-Bassat and Shinar, 2011). Controlling road shoulder widths and the use of guard rails can be a safer way to tackle speed and lane position control.

Graham and Glaister (2003) developed a spatial model to investigate the role of urban environment and its associated traffic generation characteristics on road pedestrian accidents. Pedestrian casualties were noted to occur more frequently in residential areas than in business areas. Abdel-Aty (2003) investigated factors affecting the severity of driver injury at multiple locations in Central Florida, USA. The study showed that in addition to other factors, dark lighting conditions and roadway curves were found to be associated with a high probability of causing injuries during a road traffic accident. Results also indicate that wherever the crash occurs, older, drivers, male drivers and those not wearing a seat belt have a greater probability of a severe injury. In addition to these, passenger car drivers, vehicles hit at the driver’s side and those who speed encounter higher injury severity levels. Eluru, Bhat and Hensher (2008) also examined pedestrian and bicycle injury severity levels in road traffic accidents and identified darker periods of the day as a contributory factor to higher injury severity. Some of the other variables found to be critical in influencing non-motorist injury severity levels are the speed limit of the road (higher speed limits lead to higher injury severity levels) and location of crashes (those at signalised intersections are less severe than those elsewhere). This study had a limitation of investigating non-motorised injury severity in crashes with a single motorised vehicle. Even though these types of crashes may be common, a study of other crashes will be informative.

Khorashadi et al. (2005) investigated the differences between urban and rural driver injuries (both passenger-vehicle and large-truck driver injuries) in accidents involving large trucks (in excess of 10,000 pounds). The extent of the effect of variables found to be significant in both urban and rural models on driver severity outcomes varied remarkably. Results revealed the likelihood of a 68.7% reduction in severe/fatal injuries on rural area roads that had concrete median barriers. The results of the study suggested a complex relationship between road geometry, environmental conditions and the severity of driver injury during a road traffic accident.

Usman, Fu and Miranda-Moreno (2010) made use of three different models (Negative binomial model, generalised NB model and zero inflated NB models) to examine the relation between accident frequency during snow storm events and road surface conditions, visibility and other influencing factors while controlling for traffic exposure. It was shown amongst other things that poor road surface conditions was linked with higher accident frequencies on roads in Ontario, Canada. Also, air temperature and precipitation were found to have a statistically significant effect on accident frequency. These findings were confirmed in a later study by the same authors (Usman, Fu and Miranda-Moreno, 2012) showing a statistically significant effect of factors such as precipitation intensity, air temperature, wind speed etc. on winter road safety.

2.2.5 Other factors

Apart from the factors described so far in this chapter, other factors also contribute to the speed of vehicles and some of these are described in the following section.

2.2.5.1 Weather

Weather has been investigated by a few researchers as a factor having effect on speeds and accidents. This section presents the findings of some investigators.

In a study of motorists' speeds in wet weather as compared to dry conditions for two separate survey locations on the M4 motorway, South Wales in United Kingdom, one site had conventional asphalt surfacing while the other had porous asphalt wearing course (Edwards, 2002). It was found that drivers slowed down marginally in wet weather and although the speed reductions were statistically significant, it was found to be insufficient to compensate for the additional wet weather risks imposed. These marginal speed reductions were about 3 mph or 4.5 percent at the 85th percentile for mean speeds during wet conditions compared to dry weather conditions. A similar reduction of 3 percent was obtained for 85th percentile speed on porous asphalt surfacing and 4.3 percent reduction for dry conditions on the same surfacing type.

In a similar study by the same researcher, the effect of three weather conditions (fine, rain and misty conditions) on driver behaviour was compared (Edwards, 1999). Spot speeds were recorded for 200 vehicles in the outside lane of the eastbound dual 2-lane carriageway of the M4 Motorway in South Wales, United Kingdom. Speeds in poor weather conditions were compared to those in fine weather conditions. A small but significant reduction in mean speeds for both wet weather and misty conditions was realised. The speed reductions achieved were insufficient to compensate for the increased hazard posed by inclement weather. In another investigation by the same author, the relationship between road accident severity and the weather was investigated for England and Wales (Edwards, 1998). Findings indicated accident severity reduced significantly in rain compared with fine weather. Accident severity in fog showed geographical variation while accident severity in high winds was found to be inconclusive.

In Calgary and Edmonton, Canada, accident risks following rain was investigated. Data for 169 rain events and over 15,000 accidents occurring during the years 1979 to 1983 was used. It was found that the risk of an accident occurring in rainfall conditions was 70 percent higher than during normal conditions. The data suggests that accident risk returns to normal as soon as the rainfall stops notwithstanding the wet weather conditions remaining (Andrey and Yagar, 1993). Ivey et al. (1981) developed the Wet Weather Safety Index ($WWSI_e$) using multiple regression technique to provide a prediction of wet accident rates as a function of traffic, road geometry and pavement surface characteristics. This Index was applied to 68 highway segments in Texas, USA having accidents per year per mile with values ranging from zero to 40. It was noted that urban areas had much higher accident rates than rural areas but advised being careful of using the index for roads where specific remedial measures had been carried out. Edwards (1996) investigated weather related road accidents in England and Wales using spatial analysis. Results from this study showed a clear positive relationship between the presence of weather hazards and road accidents. This study however did not take into consideration other factors such as the type of manoeuvre being carried out, the age and years of driving experience etc. of the driver which can also influence an accident's occurrence.

Using negative binomial regression methods to investigate the impact of snowfall on road accident crash counts, Eisenberg and Warner (2005) used US traffic crash rates between 1975 and 2000. The study revealed that even though there were fewer fatal crashes on snowy days than on dry days, more non-fatal injury crashes and property damage-only crashes were noted on snowy days. The large sample size used in the study allowed for the controlling of potential confounders. Keay and Simmonds (2006) were able to show the presence of rainfall to be a hazard for driving in a study into the association between rainfall and road accidents in a metropolitan area in Melbourne, Australia from 1987 to 2002. These findings are consistent with other researches by Kim et al. (2007), Brijs, Karlis and Wets (2008), Koetse and Rietveld (2009) and Jung, Qin and Noyce (2010) about the significance of inclement weather on the severity of road traffic accidents with road surface condition being the major contributory factor in winter road safety. Andersson and Chapman (2011) investigated the relationship between temperature and severe road accidents in the West Midlands, UK. The study showed that reducing road slipperiness can reduce accidents by up to 12% assuming accidents remain constant over time. The study also acknowledged that with continual improvements in vehicle technology and road safety the assumption is not likely to be true. Al-Harbi et al. (2012) studied the effect of meteorological conditions on road traffic accidents in Kuwait found that temperature during the fall, spring and winter seasons were the most significant meteorological conditions causing road traffic accidents. During summer wind speed was noted as the most significant factor causing road traffic accidents.

Finally, in a review of methods proposed for measuring the added risk of road accidents in rainy weather conditions, two methods were applied with adjustments to data from Israel and the United States. Results indicated a substantial added risk of an injury accident in rainy conditions and it was found to be two or three times greater than in dry weather conditions (Brodsky and Hakkert, 1988). The above review of weather on road accidents indicates that weather does have an effect on road traffic accidents. Unfortunately, the extent of influence of weather is not one of the factors that can be altered.

2.2.5.2 Speed Restrictions

Studies suggest that the use of devices such as Vehicle Activated Signs (VASs) and Speed Cameras to enforce speed limits can result in reduced speed violations and road accidents (Corbett, 1995). VASs are used to address problems associated with inappropriate speed at locations where conventional signing has proved ineffective. Inappropriate speeds include vehicle speed on approaches to hazards such as bends or junctions. VASs are supposed to complement fixed signs and should not be an alternative to them. They are triggered to show a particular hazard or speed limit when drivers exceed a set threshold speed and can be accompanied by the message “SLOW DOWN”. They can even warn drivers of a safety camera if they are exceeding the speed limit. An example is shown in Figure 8.

Figure 8 ‘SLOW DOWN' sign

Recent VASs display messages with the help of either fibre optic cables or light emitting diodes (LEDs) on the front panel of the sign. Different colours can be used to show different parts of the sign with an automatic dimmer in place to reduce its intensity of brightness at night. It remains blank when not activated by a vehicle. Current VAS types used are speed enforcing and hazard warning signs. VASs are to be used only where road traffic accidents are caused by vehicle speeding. They are particularly useful where speed cameras and other related signs are not cost effective or appropriate solutions. VASs do not record any data for prosecution purposes (Winnett and Wheeler, 2002; Department for Transport, 2003) and can be moved from site to site. However, interventions that do not involve identification and punishment of drivers exceeding speed limits have been shown not to be responded to well by drivers (Morrison, Petticrew and Thomson, 2003). One such way to counteract this problem is through the use of speed cameras.

Speed cameras can be fixed (operated permanently from a roadside housing) or mobile (operated from a mobile vehicle parked by the side of the road). Although mobile cameras are flexible with regards to locations to which they can be deployed, they can however prove difficult to be resourced. The most common type of speed camera in use in countries is typically spot speed cameras and these have been shown to reduce vehicle speeds with up to 17 percent reduction in collisions obtained after introduction (Pilkington and Kinra, 2005; Champness, Sheehan and Folkman, 2005). Improvements realised from the deployment of speed cameras include a reduction in mean traffic speeds in comparison to posted speed limit levels with traffic speeds even declining in the absence of enforcement measures (Chen, Meckle and Wilson, 2002; Keall, Povey and Frith, 2002). Crash and casualty data also indicates a significant reduction in estimated casualties per crash indicating fallen speeds. There is some overseas evidence to further suggest the usefulness of speed cameras. In Australia and New Zealand, speed cameras have produced up to 32 percent and 14 percent reduction in urban and rural areas personal injury accidents respectively (Keall, Povey and Frith, 2001). An analysis of 10 studies of the effect of speed cameras in seven European countries found a 19 percent decrease in injury causing accidents (Elvik, 2002). Canada recorded a 9 percent reduction in road traffic accidents and a 2.8 kilometres per hour fall in mean speeds at speed camera sites (Chen, Meckle and Wilson, 2002; Jones, Sauerzapf and Haynes, 2008). A speed camera evaluation pilot project in the United Kingdom (UK) carried out in 1992 showed that camera use resulted in a 41 percent decrease in casualties killed or seriously injured and a mean speed reduction of 10 miles per hour (Gloag, 1993). The evaluation of the effects of a new enforcement deployment method by the Israeli National Traffic Police on 700km of interurban roads (Hakkert et al., 2001), revealed 85 percent of drivers violated the rural road speed limits in the absence of any penalising enforcement measure. From this review, even though it is evident that speed cameras do reduce vehicle speeds and road accidents it is worth stating that speed cameras tend to be introduced at sites where there are high numbers of speed related collisions. Considering that an increase in collision may not be following any pattern and be due to chance, any further reductions could be suggestive of normal variation ('regression to the mean' effect) (Pilkington and Kinra, 2005). Regression to the mean is a common phenomenon in statistics with possible severe consequences. It can lead to an imprecise conclusion that an effect is due to treatment when it could actually be due to chance (Morton and Torgenson, 2003).

2.2.6 Effectiveness of vehicle activated signs

Even though history of the first use of vehicle activated signs (VASs) dates back to the late 1970s with early research carried out by the Transport Research Laboratory, UK, not much literature can be found on the effectiveness of vehicle activated signs. This section reviews the identified literature on VASs.

Winnett and Wheeler (2002) provided insight into speed reductions achieved in the late 1970s and early 1980 trial of vehicle activated signs usage in the UK. These signs were described as automatic signs providing drivers with information relating to close following drivers or excessive speed. The signs were unlit until drivers exceeded a predetermined value in relation to either the distance from the vehicle in front or the speed of the vehicle. Results from these earlier studies for the close following signs indicated (i) a 30% reduction in the number of drivers following the front vehicle with a gap less than 1 second with the effect maintained up to a distance of 800m downstream and (ii) over a five year period no appreciable degradation of the sign's effectiveness was noted. For the speed reduction signs, speeds measured were found to decrease with time after the signs were installed and in some two villages, speeds of faster vehicles were reduced to a small extent.

The study by Winnett and Wheeler (2002) involved a large scale evaluation of vehicle activated signs in the UK following the installation of speed roundels at 30mph, 40mph, 50mph, speed change from 30mph to 20mph and speed change from 40mph to 30mph sites. Speed was used as a measure of the anticipated accident frequency since it takes time for accident data to build up. For the 30mph sites, mean speed reductions of between 2.6mph and 7.1mph was achieved. The proportion of vehicles exceeding 30mph was reduced by between 18 and 34 percent whilst the proportion of vehicles exceeding 35mph was reduced by about 15 to 51 percent. For the 30mph sites large changes in speed reductions were obtained probably because the monitoring sites were located close to the vehicle activated signs. At the 40mph sites, mean speed reductions of between 1.2 and 4.4mph was achieved. There was between 7 and 35 percent reduction in the proportion of vehicles exceeding 40mph and between 1 and 17 percent reduction in the proportion of vehicles exceeding 45mph. In the case of the 50mph sites, there was a fall in speeds of 4.6mph for lane 1 vehicles and 3.6mph speed fall for lane 2 vehicles. There was a 22 percent reduction in the proportion of lane 2 vehicles exceeding 50mph with a 26 percent reduction obtained for lane 1 vehicles. There was also a 31 percent reduction in the proportion of vehicles in lane 1 and 10 percent reduction in the proportion of vehicles in lane 2 exceeding 55mph. Where speed limits were reduced from 30mph to 20mph, drivers had difficulty attaining and maintaining a speed of 20mph. Reduction in mean speeds of between 4.4mph and 7.5mph was obtained for the 20mph speed limit. There was between 28 and 51 percent reduction in the proportion of vehicles exceeding the 25mph speed with between 38 and 56 percent reduction in the proportion of vehicles exceeding 30mph. For the 40mph to 30mph change in speed limit, there was between 6.5mph and 1.8mph reduction in mean speed. There was also between 13 and 60 percent reduction in the proportion of vehicles exceeding 30mph and between 50 and 80 percent reduction in the proportion of vehicles exceeding 35mph. Accident reductions ranging from 16 to 100 percent with a 58 percent reduction across the sites combined was achieved at the speed roundel locations.

Warning signs investigated by Winnett and Wheeler (2002) included junction, bend and safety camera repeater signs. For sites with junction warning signs, mean speed reductions of between 0.8 and 9.2mph was achieved with drivers observed to reduce their speeds not only after they have passed the sign but also on approach to the sign. Apart from one site, all other sites recorded between 45 percent and 100 percent reduction in accidents with a 26 percent reduction obtained in all sites combined. However the proportion of accidents involving fatal or serious injury accidents was found to have changed little. At sites with bend signs, mean speed reductions of between 2.1 and 6.9mph were achieved. For locations with safety camera repeater signs, mean speed reductions of between 0.5 and 3.7mph was achieved with a reduction in the percentage of vehicles exceeding various threshold speeds also obtained. Accident reductions of between 8 and 31 percent was realised at safety camera repeater sign locations with a 17 percent overall reduction across combined sites realised.

The reductions in speed obtained by Winnett and Wheeler (2002) are consistent with the findings of Burbridge, Eveleigh and Van Eysden (2010) in their study on vehicle activated signs in Queensland, Australia. Results from their study showed a 5 to 10km/h reduction in average speeds and 10 to 15 percent reduction in the proportion of drivers travelling in excess of 9km/h above the speed limit on approach to the signs. It was however noted that about 20 percent of the proportion of vehicles approaching the sign were travelling in excess of 9km/h above the limit but this percentage was found to reduce once the vehicle activated the sign. Speed data collected downstream of the signs showed less than 2 percent of vehicles travelled in excess of the 9km/h above the speed limit. There was an overall decrease in the 85^{th} percentile speed, decrease in the mean/average speed and a decrease in the overall number of speeding vehicles. It was noted from the study that speed reductions obtained a year after installation of the signs were similar to reductions obtained within a month of activation of the signs.

In another study in London, speed reductions were realised from the use of speed indicator devices (SID) (Walter and Broughton, 2011). SIDs are temporary vehicle activated signs used in detecting and displaying real-time vehicle speeds. Even though the maximum mean speed reduction obtained from the SIDs was 2.6mph, the overall speed reduction obtained was 1.4mph with a reduction from 57% to 45% in the proportion of vehicles exceeding the speed limit. The speed readings were recorded over a maximum period of three weeks. This study demonstrated the effectiveness of vehicle activated signs. When the SID was in operation, the proportion of drivers exceeding 30mph was significantly reduced at all sites. During the period of operation, most sites had mean speeds being reduced by about 0.2mph at a distance of 200m downstream of the device which was small but was found to be statistically significant. However at a distance of 400m downstream, mean speeds were found to increase by about 0.6mph signifying the depletion in effectiveness of the device at this distance. Since SIDs are temporary, it was noted that there was no lasting effect after removal with sites where the SID had the greatest effect while in place having small reduction in speeds remaining.

Santiago-Chaparro, Chitturi and Noyce (2012) investigated the spatial effectiveness of speed feedback signs (SFS). These are sometimes referred to as dynamic speed display or VASs. SFS are signs installed to give real-time dynamic display of the vehicular speed of a driver at a specific location. The investigation sought to find out how far upstream and downstream of the SFS speed reductions were maintained. Upstream and downstream speed of free flowing vehicles was obtained using radar units as well as video recording. Upstream of the SFS 50 percent of vehicles reduced their speed by at least 1.0mph. Downstream, 50 percent of vehicles increased their speed by 1.0mph. At a distance of between 381m and 441m upstream of the SFS, vehicle speeds were found to start reducing by at least 1.0mph. Downstream, the effectiveness of the SFS on vehicle speeds was found to diminish at distances from 91m to 183m. The most notable speed reductions were achieved at distances from 366m to 427m upstream whilst downstream, the most notable speed increase was obtained at distances from 91m to 152m. Findings from this study illustrates that once drivers pass the SFS there is a reduction in its effectiveness with recommendations being that SFS need to be placed closer to the location where it is intended to reduce vehicle speeds.

2.2.7 Effectiveness of speed cameras

Speed cameras (mobile or fixed) have been used in various countries with several studies pointing to the benefits to be attained following the introduction of mobile and fixed speed camera operations. This section discusses speed camera effectiveness.

Corbett (1995) studied drivers who use the A40, one of the first roads in England to have unmanned speed cameras installed along it. The study revealed that 54% of respondents said they had been driving more slowly at least in some places since the introduction of the cameras. Of these respondents, 8% said they slowed down 'a lot', 21% said they drove 'a bit' more slowly, 15% said they slowed down in some places but drove no differently in others with 10% saying they drove slowly in some places but faster in others. 46 percent of respondents said they drove 'a bit' or 'a lot' slower in areas where they suspected there were cameras. The outcome by Corbett (1995) was partly supported by speed monitoring data obtained by the Department for Transport. The data revealed that during the first six months of the introduction of the speed camera, average free-flow speeds fell by 10% and injury or damage only accidents also fell by 22%. Corbett (1995) expressed concern about the possibility of the effectiveness of speed cameras diminishing over time.

In Barcelona, Novoa et al. (2010) assessed the effectiveness of speed cameras in reducing crashes and the number of injured people on 50km/h speed limit arterial roads as well as the long term effectiveness on an 80km/h speed limit beltway. The study did not show any effect of fixed speed cameras on the arterial roads. This lack of effect was partly attributed to the small numbers of vehicles traversing the stretch of road, lower speed limit and the presence of traffic lights. On the beltway, a 30% reduction in injury accidents and a 26% reduction in people injured were realised. In another study in Barcelona, Pérez et al. (2007) used the same roads used by Novoa et al. (2010) but the beltway was used as the 'intervention group' and the arterial roads was the 'comparison group' and had no fixed cameras installed. In the case of Novoa et al. (2010), both arterial roads and the beltway had speed cameras installed on them. An estimated 27% reduction in the number of collisions and another 26% reduction in the number of vehicles involved in collisions were obtained by Pérez et al. (2007). Despite the fact that Pérez et al. (2007) showed that speed cameras are effective at reducing road collisions and the number of people injured in urban road accidents they were not able to provide any evidence about the effectiveness of the speed cameras for the arterial roads. With the method of comparison used by the two authors Novoa et al. (2010) and Pérez et al. (2007) being different, it makes it difficult to make any meaningful comparison about the effects of the speed cameras on arterial roads and the beltway even though both studies were conducted in Barcelona with the same set of roads. Some of the variables that are often considered to have confounding consequences on results obtained for observational before and after road safety studies comprise regression to the mean Pérez et al. (2007). This variable was controlled for in the study by using time-series analysis which allowed for seasonality and trend adjustment. Considering that the beltway characterises and functions as the fastest route connecting the city to the outlying metropolitan area, it was not suspected that a remarkable number of vehicles would divert onto other routes to avoid the speed cameras. Contrary to what was expected, vehicle-kilometres travelled on the beltway grew over the study period.

He et al. (2013) studied speed enforcement in China through the national, provincial and city initiatives using a combination of automatic detection from speed cameras (fixed and mobile) and other traffic law enforcement. At the national level there was a reduction from 17.2% in 2004 to 10.2% in 2007 in fatalities associated with speeding. One of the fears envisaged by He et al. (2013) was that with increase in motorisation and increase in highway lengths in China, fatalities from road traffic accidents may revert back to previous numbers. This fear can be allayed by the findings from another study (Shefer, 1994) that with increased motorisation come associated congestion resulting in reduced travel speeds and leading to no increase in road accidents associated with vehicle speeds.

In a study by Carnis and Blais (2013) in France, significant decreases in both fatal and non-fatal traffic injuries were realised following the deployment of a speed camera programme in 2003. Translating the benefits achieved into figures, 15,193 fatalities and 62,259 non-fatal injuries were prevented by the programme between November 2003 and December 2010. Fatality rate per 100,000 vehicles dropped by 21% whilst the decrease in non-fatal injuries revealed a decay function with a 26.2% reduction being recorded in the first month but dropped to 3.5% in December 2008 and 0.8% in December 2010. This dissipating effect is likely to be due to the initial deterrent effect posed by the speed cameras and the fading off effect with the passage of time. One other point raised in the study was the possibility of the decline in non-fatal injuries resulting from incidents outside the scope of the speed cameras. Some of these incidents include the use of mobile phones for listening, talking and texting while driving. This is another area of research that could provide some evidence about the decline in non-fatal injuries. It was noted that additional input of speed cameras did not translate into proportional decreases in fatal injuries consistent with the findings from Elvik (2011) that the preventive effect remains stable at about 21%.

A systematic review on the effectiveness of speed cameras in preventing road traffic collisions and related casualties by Pilkington and Kinra (2005) identified 92 published and unpublished papers for different countries. After reviewing the identified materials, 21 studies were found to be potentially suitable for use. On further elimination of some of the material, the review was eventually carried out using 14 studies. Results from the reviewed materials were mostly before-after studies. One study which had been introduced for about 4.6 years showed sustained longer term effects. Across all the studies there was a 5% to 69% reduction in collisions, 12% to 65% reduction in injuries and a 17% to 71% reduction in deaths within the immediate vicinity of the camera sites. A similar order of magnitude in reductions was obtained over wider geographical areas. Evidence from this study showed that speed cameras are an effective measure in reducing road traffic collisions and related casualties. A much smaller study carried out in Maryland, USA (Retting, Farmer and McCartt, 2008) analysed speed data obtained 6 months before and 6 months after the deployment of speed cameras in Montgomery County, Maryland where roads had speed limits ranging from 25mph to 35mph. The proportion of drivers travelling more than 10mph above the posted speed limit declined by about 70% at locations which had both speed cameras and warning signs, 39% at locations with warning signs alone and 16% on residential streets with no warning signs or speed cameras. The introduction of fixed speed cameras on the Loop 101 freeway in Scottsdale, Arizona was the first use on a major US highway (Retting, Kryychenko and McCartt, 2008). The 9 month trial of the speed cameras revealed an 88% reduction in the number of vehicles travelling 11mph or more above the 65mph limit, however traffic speeds were found to increase soon after the pilot study was over. At distances of 25 miles away from where the camera had been installed, huge reductions in speeding was observed providing some evidence that speed cameras can substantially reduce the speed of vehicles. The study however did not consider the effect of the speed reduction on crashes.

Shin, Washington and Van Schalkwyk (2009) investigating the same stretch of road, Loop 101 freeway in Scottsdale, Arizona in the US used the 9 month demonstration speed enforcement programme (SEP). Unlike the previous authors, Shin, Washington and Van Schalkwyk (2009) investigated the impact of the speed enforcement programme on crashes. The outcome of the study indicated that taking traffic flow into account, average speeds within the enforcement area reduced by about 9mph when the SEP was introduced. All types of crashes were found to reduce with the exception of rear-end crashes. Even though Retting, Kryychenko and McCartt (2008) observed speed reductions at up to about 25 miles from the enforcement site, Shin, Washington and Van Schalkwyk (2009) observed no statistically significant reduction in average speeds at about 40 miles away from the enforcement zone. Shin, Washington and van Schalkwyk (2009) cautioned generalising the results obtained from the study since it was from a short programme. Similarity in the predictions observed between the Empirical Bayes before-after study and the before-after study allowing for traffic flow adjustments indicates a comparatively small possibility of bias from regression to the mean when evaluating the impact of the SEP on safety in the enforcement zone (Shin, Washington and Van Schalkwyk, 2009).There was still evidence to suggest that speed cameras do help to reduce vehicle speeds and road traffic accidents with other long term programmes (Pilkington and Kinra, 2005) supporting this claim. Høye (2015), Høye (2015a), Mountain et al. (2004) and Elvik (1997) used the Empirical Bayes approach to control for regression to the mean in their evaluation of the impact of speed cameras on safety. Mountain et al. (2004) in their study showed that speed enforcement cameras on 30mph roads provide safety benefits over a distance of up to 1kilometre upstream and downstream of the camera. An average of 20 percent or 1 PIA/km/year decrease in accidents ascribed to reductions in speed over this distance was realised. A before-after empirical Bayes method was used to study the safety effects of 223 fixed speed cameras installed between 2000 and 2010 in Norway by Høye (2015). After controlling for effects of trends, volumes of traffic and speed limit changes, it was found that on road sections between 100m upstream and 1 km downstream of the speed cameras there was a 22 percent statistically significant reduction in the number of injury crashes. On longer sections of road and for killed and seriously injured (KSI), results were found to be statistically significant. For speed cameras installed in 2004 or later, a reduction in injury crashes and the number of KSI on road sections from 100m upstream up to both 1 km and 3 km downstream of the speed camera was noted. Greater effects were found for KSI than for injury crashes with the effects decreasing with increasing distance from the speed cameras. At distances of 100 metres upstream and downstream of the camera sites, crash reductions were smaller and non-significant.

Fourteen sites in Norway were investigated to establish the safety effects of section control by Høye (2015a) in a before-after study using the empirical Bayes method. A non-significant reduction by 12 percent in injury crashes was found. There was a 49 percent significant reduction in the number of killed or severely injured at section controlled sites. Results revealed that crash reduction in tunnels (mostly undersea tunnels with section control on steep downhill segments) are of the same enormity as on open roads. Injury crashes downstream of the section control sites (up to 3km in each direction) were observed to be remarkably reduced by 46 percent but the number of killed or seriously injured downstream of the section control sites was too small for any meaningful conclusion to be drawn. Section control was found to be effective at decreasing both speed and crashes, mainly serious crashes with the possibility of spill over effects (crash reductions at non-enforcement sites) more likely to arise than crash migration (Høye, 2015a).

Improvements realised from the use of speed cameras are not unique to the UK as shown earlier and in other parts of the world it includes a reduction in mean traffic speeds in comparison to posted speed limit levels with traffic speeds declining in the absence of enforcement measures (Chen, Meckle and Wilson, 2002; Keall, Povey and Frith, 2002). Crash and casualty data from these studies also indicate a significant reduction in estimated casualties per crash as well as reduced speeds (Chen, Meckle and Wilson, 2002; Keall, Povey and Frith, 2002). There is some overseas evidence to further suggest the usefulness of speed cameras. In Australia speed cameras have produced up to 41% reduction in fatal crashes (Cameron et al., 1994). In another Australian study (Newstead, 2009) there was an estimated 47% reduction in fatal to medically treated crashes with an overall 32% and 30% reduction respectively in all reported crashes including non-injury crashes for the 2 years assessed period of 2006 and the first half of 2007. An analysis of 10 studies of the effect of speed cameras in seven European countries found a 19 percent decrease in injury causing crashes (Elvik, 2002). Canada recorded a 9 percent reduction in road traffic crashes and a 2.8 kilometres per hour fall in mean speeds at speed camera sites (Chen, Meckle and Wilson, 2002; Jones, Sauerzapf and Haynes, 2008). A speed camera evaluation pilot project in the UK carried out in 1992 showed that camera use resulted in a 41 percent decrease in casualties killed or seriously injured and a mean speed reduction of 10 miles per hour (Gloag, 1993).

Table 1 provides a summary of factors identified in literature to affect vehicle speed as discussed in section 2.2.

Factor	Author and Year	Commentary
Radius of curvature	Abdul-Mawjoud and Sofia (2008);McDonald (2004); Donnell et al., (2001); Fitzpatrick et al., (2000); Bhatnagar (1994), Lamm and Choueiri (1987); (McLean (1981); McLean (1979)	curve radius being a widely used parameter since it is considered the most important parameter in obtaining operating speeds along horizontal curves and it does have effect on vehicle speed.
Road gradient	Reinfurt et al., (1992); Gombard and Louah (1986); Yagar and Van Aerde (1983)	speeds decrease on upgrades and has been shown to have effect on vehicle speed.
Environmental factors	Fambro, Fitzpatrick and Russell (2007); Lee and Mannering (2002); Yagar (1984)	The speed of vehicles as a function of the terrain with particular reference to the horizontal curvature of the road has an impact on vehicle speeds.
Road roughness.	Karlaftis and Golias (2002);	This is found to have an effect on vehicle speeds on a case by case basis.
shoulder widths	Ben-Bassat and Shinar (2011); Fambro, Fitzpatrick and Russell (2007); Leong (1968)	Increasing shoulder widths increases vehicle speeds. Combining shoulder widths with distances to lateral obstructions in the road environment affects vehicle speed.
Side friction	McLean (1979); McLean (1981)	This has been shown to have no effect on vehicle speed.

Table 1 Summary of factors influencing vehicle speed

Factor	Author and Year	Commentary
sight distance	Fambro, Fitzpatrick and Russell (2007); AASHTO (2004); Easa and Hassan (2000); Harwood, Mason and Brydia (1999); Highways Agency (2002); Yagar (1984); Leong (1968)	Significant impact at higher percentile speeds. As sight distance decreases, mean reductions in speed between the control and crest sections increase. Sight distance does have an effect on vehicle speed.
road widths	Fambro, Fitzpatrick and Russell (2007); Godley, Triggs and Fildes (2004); AASHTO (2001); Lamm and Choueiri(1987); Yagar (1984); Leong (1968)	This effect is significant at widths below a certain critical value. Narrowing roads to widths of less than 3.0m using a painted hatched road centre marking can be effective at reducing vehicle speeds. Road widths have some effect on vehicle speed.
superelevation	Abdul-Mawjoud and Sofia (2008); Krammes et a.l (1995); Bennett (1994); Lamm and Choueiri (1987); Kadiyali et al. (1981)	Only affects high speeds such as 85^{th} percentile speed of drivers. Superelevation has minimal effect on vehicle speed
Weather	Edwards (2002); Edwards (1999); Edwards (1998); Andrey and Yagar (1993); Wasielewski (1984); Brodsky and Hakkert (1988)	Marginal effect of wet weather on vehicle speeds. Small but significant reduction in mean speeds obtained for poor weather (wet weather and misty conditions) in comparison with fine weather condition. Weather does have some effect on vehicle speed.

Table 1 Summary of factors influencing vehicle speed (continued)

Factor	Author and Year	Commentary
Other factors (traffic volume, presence of driveway access to adjacent land use, speed limit and access from other highways)	Jones, Sauerzapf and Haynes (2008); Pilkington and Kinra (2005); Champness, Sheehan and Folkman (2005); Morrison, Petticrew and Thomson (2003); Karlaftis and Golias (2002); Elvik (2002); Chen, Meckle and Wilson (2002); Keall, Povey and Frith (2002); Keall, Povey and Frith (2001); Fridstrøm et al. (1995); Corbett (1995); Gloag (1993); Yagar (1984);	A combination of other factors, some not necessarily associated with the geometry of the road have been shown to have effect on vehicle speed.

Table 1 Summary of factors influencing vehicle speed (continued)

2.3 Summary

Transportation professionals are continually tasked with improving road safety. To improve road safety there is the need to understand what factors come to play to contribute to the problem of road safety. The chapter provided a detailed review of literature that identifies factors that affect vehicle speed. These factors include road conditions (curvature, gradient, roughness, sight distance and road width), vehicle, traffic conditions, road environment and other factors. Various factors affecting vehicle speeds were identified in the literature and these factors have been studied using a wide range of methods such as engineering and economics. Some of the important factors identified to affect vehicle speed are radius of curvature and gradient of the road. Others include environmental factors such as the speed of vehicles as a function of the terrain with particular reference to the horizontal curvature of the road.

The outcome of this literature review will contribute to this research since it seeks to investigate the optimum locations for speed control devices and thus the importance of identifying factors that affect vehicle speed. With the identification of these factors, the next chapter aims to find an accident prediction model that includes these factors. Where an appropriate accident prediction model cannot be found, one will be developed based on the factors identified. The accident prediction model will then be applied to an optimisation technique to optimise the location of the speed control device along roads.

3 Analysis Methodologies

3.1 Introduction

A wide range of models exist for predicting accident occurrence and severity along a road. In order to establish which model to use in this research a review of previous works was carried out. The primary aim of this thesis is to contribute to speed reduction by developing an optimisation model to help decision makers determine the optimum location for a speed control device. In view of this aim, a review on accident prediction models as well as a review of the optimisation techniques to be used in this research will be carried out. Some of the advantages and disadvantages exhibited by the accident prediction models are revealed. The identified accident prediction model will be used to further enhance the methodology of the thesis.

3.2 Accident Prediction Models

In road accident prediction models different areas of the road segment ie. specific areas, junctions, and segments of road are usually used. Models used for accident frequency tend to differ from those used in accident severity (Jones and Jorgensen, 2003; Wang, Quddus and Ison, 2011). Accident frequency models tend to establish the relationship existing between the number of accidents observed over a specific period of time and the contributory factors to the accidents for a given road segment, area etc. Accident severity models on the other hand tend to establish a relationship between the level of severity of the accident ie. fatal, serious and/or slight and information about the characteristic of individual accidents. Accident frequency and accident severity models are classed as accident prediction models. These two types of models have been used over the years to help predict and provide suitable remedial measures to reduce road traffic accidents. The discussion of accident prediction models will be grouped based on the type of modelling technique used. It is appreciated that accident prediction models have gained respectable recognition by road safety practitioners and advocates over the years dating back a few decades. However an attempt to review all this literature will prove too ambitious for this thesis and some very dated studies may not prove very useful. In identifying suitable material to be used a literature search was undertaken of journal articles, conference papers and books etc. Other materials were obtained from searching ‘Google scholar’ and ‘Google’. A combination of search key words used included ‘road geometry parameters and road accidents', 'road geometry and road accidents', 'accident prediction models' and 'accident prediction models for roads'. Different approaches to accident prediction modelling used in this chapter typically fit into these categories; multiple linear regression, multiple logistic regression, Poisson models, negative binomial models, random effect models etc..

3.2.1 Empirical Bayes (EB) models

Empirical Bayes models are used for modelling complicated systems, provide a mechanism for obtaining parameter estimates and are based on Bayesian models but they make use of alternate estimation techniques. Bayesian models are able to derive estimates for all parameters of interest in a model. Inferences from Empirical Bayes analysis tend to be frequentist (drawing conclusions from sample data by drawing emphasis on the frequency of the data) (Casella, 1992). Bayesian analysis is dependent on a prior distribution for the parameters of the model. Depending on unknown parameters which may be obtained from some second stage prior, the prior can be nonparametric or parametric. The order of parameters and priors make up a hierarchical model. The hierarchy must terminate at some point with the rest of the prior parameters assumed to be known. Instead of making this assumption, the empirical Bayes method applies the observed data to estimate the final stages parameters and then proceeds as though the prior was known (Carlin and Louis, 2000). Some of the good features of the empirical Bayes (EB) approach are that it helps to deal with the regression to mean (RTM) bias, EB estimates tend to be more precise and also the EB approach allows the estimation of the entire time series as required (Hauer, 1997). The central concept of the EM method is the reference population with each entity of the reference population said to have its own accident count k. The EM method utilises data about the mean and variance of the k's in the reference population with methods for evaluating the E$\{k\}$ and VAR$\{k\}$ given where E is the mean and VAR is the variance.

Brüde and Larsson (1988) used a variant of the Empirical Bayes method for a before and after period in predicting accidents at 1,901 3-way junctions on rural roads. The variant EB method made use of predicted numbers of accidents for each junction but it was not possible to provide evidence as to whether the average estimates of the expected number of accidents were better than results obtained using the conventional EB method. Also the conventional Bayes method provided satisfactory results whereas the population of objects eg. junctions was not too small. This model however made use of very few parameters ie. accident numbers and traffic flows and did not consider the role the characteristics of the road infrastructure played in accident prediction. Mountain, Fawaz and Jarrett (1996) developed a model to predict accidents on main roads with minor junctions where traffic counts were unavailable on the minor approaches. It was based on approximately 3800km of highway in the UK in addition to 5000 minor junctions using data for periods varying from 5 to 15 years. The roads studied were restricted to A- and B- roads outside major conurbations with motorways and C-roads being excluded from the study. It was noted that other explanatory variables such as traffic composition, pedestrian flow and minor road entry flow could have improved the fit of the models. Accidents on single-carriageways were found to be proportional to link length whilst accidents on dual-carriageways were less than proportional with the best results obtained using the empirical Bayes method.

A general modelling strategy was used to analyse and forecast road accident fatalities in Yemen using socioeconomic and cultural variables (Ameen and Naji, 2001). The model did not consider other categories of road traffic accidents as well as the road geometry characteristics. Elvik (2008) compared the EB method to the traditional, naive assumption of treating the recorded number of accidents as an unbiased estimator of the expected number of accidents. All versions of the EB estimates were found to give considerably more correct predictions of accident numbers than the traditional approach. The traditional method assumes that the recorded number of accidents is an unbiased estimator of the expected number of accidents. EB model estimates are not always accurate, however if the differences between EB estimates and the actual number of accidents are small and random they can be accepted considering that there is randomness present in accident counts (Elvik, 2008).

Summarising the concept of the Empirical Bayes approach, the prior information is obtained from a reference group of sites similar to those being assessed to compute a sample mean and variance, or from a calibrated safety performance function relating the frequency of crashes to their features. The point estimates of the expected mean and variance are them combined with the site specific crash count to get a better approximation of a site's long term expected crash numbers (Persaud et al., (2010).

3.2.2 Bayesian Hierarchical (BH), Hierarchical Bayes (HB) and Full Bayes Hierarchical and Bayesian multivariate models

The Bayesian approach to modelling data has been extensively used over the past years. The potential to forecast risks accurately even with sparse data or rare events is one of the main benefits of the Bayesian approach (Withers, 2002). Also, the potential to include prior knowledge without the limitation of classical distributional assumptions makes it possible to implement the Bayesian approach in a lot of fields. When sample sizes are comparatively small as is often the case with accident frequency analysis, a Bayesian approach can be chosen to obtain robust approximations. However with very large sample sizes, as is normally the case with severity analysis, a frequentist inference can be used since the approximated results are equivalent to the Bayesian method (Train, 2003). The Bayesian theorem stipulates that the posterior distribution is proportional to the prior distribution times the likelihood of the observed data. For large sample sizes, the prior becomes irrelevant and the maximum of the likelihood function becomes the same as the maximum and also the mean of the posterior (Train, 2003). Bayesian models also take account of both spatial dependence and uncorrelated heterogeneity. Deductions made from traditional spatial models can be deceptive as it does not reveal the underlying data generating processes accurately (Bhati, 2005). Also, with the spatial units being analysed getting smaller (such as zip-code, wards, post-code etc.), the number of observed counts in each sampled units reduces leading to the distribution of counts becoming a highly skewed (to the right) distribution as the number of spatial units with zero counts increases. To control these issues, a Bayesian method is used by researchers in spatial econometrics. Bayesian hierarchical models can be estimated using the Markov Chain Monte Carlo (MCMC) method (Carlin and Louis, 2000). Bayesian Hierarchical models are noted to be suitable for analysing area-wide traffic crash incidents. Nevertheless, some form of strong distributional assumption *a priori* must be implored for such fully parametric models. It is difficult to assume *a priori* in non-experimental situations such as traffic crash occurrence. Thus, model estimates and inferences obtained from them can be sensitive to distributional assumptions (Quddus, 2008). The use of an alternative semi-parametric method (the cross-entropy (CE) method) that avoids parametric distributional assumptions is proposed (Bhati, 2005). Li, Zhu and Sui (2007) used a Hierarchical Bayes approach to identify and rank roadway segments with potentially high risk for crashes to help implement preventative actions. The model however only made use of traffic volume and annual average daily vehicle miles travelled without incorporating road characteristics and severity of the crash. The study failed to answer why some road segments are riskier or what could be carried out to improve safety. HB methods analyse iteratively with multiple levels of analysis (Carlin and Louis, 2000). Whilst conventional statistical inferences derive the average parameter estimates, HB models produce parameter estimates at each analytical unit stage as well as revealing and flagging 'extra variance' (Congdon, 2001). Bayesian models are generally known to borrow information from neighbours to infer individual-level parameter estimates (Bolstad, 2007; Lee, 2004). Two stages are involved in Hierarchical Bayesian modelling. During the first stage, a likelihood model for the observed crash counts vector based on the relative risk vector of crashes is specified. The second stage then involves specifying a prior model over the space of possible relative risks (Li, Zhu and Sui, 2007).

Full Bayesian models facilitate the steady analysis of aleatory and epistemic uncertainties, non-linear dependencies amongst the indicator variables and the updating of the developed risk models established on new available data (Deublein, 2013). Aguero-Valverde and Jovanis (2006) compared a full Bayes (FB) hierarchical model to a traditional negative binomial (NB) model using annual county level crash frequency. Covariates included socio-demographics, weather conditions, transportation infrastructure and amount of travel. Total precipitation was found to be significant in the NB model but not in the FB. In the FB model, spatial correlation, time trend, and space-time interactions were found significant for injury crashes. Highly significant variables in the NB models were found to be also significant in the FB models. However, variables found to be marginally significant in the NB models were generally found to be non-significant in the FB models. Song et al. (2006) also used Full Bayes models to estimate crash rates for different crash types. MacNab (2004) used a Bayesian model to analyse variations in accidents and injury from data in British Columbia, Canada. A broad range of socioeconomic data was used with no traffic or road geometric data included in the analyses. One advantage highlighted in this study over conventional methods was its ability to account for extra variation over space.

Deublein et al. (2013) developed a Bayesian hierarchical model for the Austrian rural motorway network to predict the expected number of road accidents between risk indicating variables and the model response variables. The model was tested against a dataset that was excluded from the initial model development. When model predictions were compared with actual observations, for injury accidents and light injuries relatively high correlations were obtained with r=0.73 and r=0.67 respectively. However for severely injured road users and the number of fatalities, the correlation coefficients between model predictions and real observations were found to be relatively low with r values being r=0.50 and r=0.31 respectively. These low r values can be attributed to the low recorded numbers of these categories of road accidents. This model is useful in the way it is generic since it can be modified for use on different road types and also additional risk indicating variables can be added as required. The model has credibility as far as the accident predictive power is concerned with correlation coefficients of up to r=0.73 obtained between the predicted and the actual numbers of accidents.

A Full Bayesian model was developed and applied to crash data from Korean expressways to evaluate the safety benefits of decreasing the posted speed limit (Park, Park and Lomax, 2010). The model was tailored to a before-after study with a comparison group making use of only AADT, segment length and number of lanes as variables. Results from the model indicated that a more precise crash prediction and safety effectiveness estimate can be obtained in comparison with those from univariate models. There was an increase in correlation across crashes of different types and severities with an anticipated increase in the gain in precision of the multivariate approach. The method was however described as complex.

Wang, Quddus and Ison (2011) used a two-stage mixed multivariate model which combined both accident frequency and severity models to estimate accident frequency. A Bayesian spatial model and a mixed logit model were used. The method was able to predict low frequency accidents and was more advantageous over the traditional frequency based methods. The proposed method made use of detailed road accident data which may not be available in certain countries. The proposed method however was found to be flexible in that it allows the researcher to choose the appropriate model at each stage based on for example the sample size. Despite the advantages portrayed in this method, no validation of the method using other models or data was carried out.

Summarising, in the full Bayes approach the prior information is obtained from a model of a reference population as in the Empirical Bayes approach. However, instead of a point estimation of the predicted mean and variance done in the Empirical Bayes approach, there is the generation of a distribution of likely values. To get the approximation of long term anticipated frequency of crashes, the distribution of likely values is then merged with the site specific amount of accidents. There is precision in the calculation of the variance by applying a prior distribution instead of a point estimate (Persaud et al., (2010).

3.2.3 Negative Binomial (NB) models

Count models such as negative binomial (NB) regression models are normally aimed at establishing a relationship between area-wide traffic crashes and the contributing factors (Quddus, 2008). NB models are noted to be able to take account of the effect of unobserved heterogeneity (due to omitted variables in the model) among neighbourhoods however they may not take into account spatial correlation areas. NB accident frequency models are not able to distinguish between sections of roadway that are truly safe (near zero accident likelihood) from those that are unsafe but have zero accidents occurring along them during a time of observation. Due to the prevalence of zero accident observations the coefficient estimates produced can be biased (Shankar, Milton and Mannering, 1997). The NB model generalises the Poisson model with the introduction of an individual, unobserved effect into the conditional mean, thereby relaxing the equidispersion assumption (Graham and Glaister, 2003). Where over-dispersion exists, and is found to be moderate or high NB can be investigated (Poisson when the data is not significantly over dispersed and negative binomial when it is). The NB approach to modelling is regarded as an extension of the Poisson regression methodology with the variance being different from the mean (Abdel-Aty and Radwan, 2000). Shankar, Milton and Mannering (1997) and Abdel-Aty and Radwan (2000) used Negative binomial models to estimate the effect of various highway geometric parameters and traffic characteristics on the frequency of accidents over a length of road. Even though various parameters were observed to be significant determinants of accident frequency, all the parameters were not present in one model and were considered separately. This may be an attempt to reduce the phenomenon of multi-collinearity whereby the presence of several parameters in a model may have a relation with each other. This increases the standard errors of the coefficients and makes the coefficients less significant. Noland and Quddus (2005) used the NB model in their examination of how congestion affects traffic safety. No conclusive results were obtained from this study even though small differences were observed for the model between congested and uncongested time periods. Since the model focussed on congestion, geometric parameters were not considered. Graham and Glaister (2003) were able to show that the characteristics of the local environment have a powerful influence on pedestrian casualties. The model was however unable to provide information about the relationship existing between variables within the model. Noland and Oh (2004) in their study using the NB model could not confirm that changes in road infrastructure and geometric design are beneficial to road safety.

Whilst Poisson models assume the mean and variance are identical, the Negative Binomial model generalises the Poisson model with the introduction of an individual effect into the conditional mean allowing for a relaxation in the assumption of equal dispersion (Graham and Glaister, 2003). Shanker, Mannering and Barfield (1995) and Noland and Quddus (2004) argue for the use of NB models in the representation of vehicle crash data since it is a count distribution having a variance greater than the mean. Hiselius (2004) used a Poisson and negative binomial regression models to analyse the relationship between accident frequency and traffic flow on rural roads in Sweden. Even though the results indicated that the expected number of accidents per hour and kilometre, and the traffic flow differed considerably depending on whether different traffic modes were considered it also became apparent that other parameters such as speed should have been taken into account. This was a very basic model which did not consider a lot of variables that are often deemed important in accident analyses problems. Vieira Gomes (2013) used a NB model to estimate accident frequencies for road segments in Portugal. The model for accidents provided a good fit in comparison to the model for pedestrians which had relatively low fit quality due to the small sample size of data involved. The accident model failed to split the accidents by level of severity and no validation of the results was carried out.

The negative binomial (Poisson-gamma) regression model, an extension of the Poisson model aims to overcome possible over dispersion in the data and probably it is the most commonly used model for crash frequencies but with some limitations. These limitations include the model's inability to handle under dispersed data and dispersion parameter estimation problems when the data are characterised by low sample mean values and small sample sizes (Lord, 2006).

3.2.4 Poisson Regression (PR) models

Ivan, Wang and Bernardo (2000) provided a model to predict both single and multi-vehicle highway crash rates as a function of traffic density, land use, light conditions and time of day. Even though some observations were made about the effect of the parameters selected on road safety, mention was not made of road geometric parameters. Also the limited size in the sample used warrants some care being taken in transferring the results to other sites.

One of the limitations associated with the Poisson regression model is the variance of the accident data being equated to the mean. Accident frequency data is mostly over dispersed and depicted by a variance being greater than the mean. Poisson regression models have been found to be unable to handle over and under dispersion and can be affected by low sample means thereby producing biased results in small samples.

3.2.5 Random Effect Negative Binomial (RENB) and Random Parameters Negative Binomial Regression models

Correlations amongst variables are sometimes expected and these arise from either spatial considerations, temporal considerations or a combination of both. In an attempt to account for such correlations, random-effects models and fixed effects models are considered (Lord and Mannering, 2010). Chin and Quddus (2003) used the RENB model to investigate the relationship between accident occurrence and the geometric, traffic and control characteristics of signalised intersections in Singapore. The RENB model treats data as time series cross-section panel and is able to deal with the spatial and temporal effects in the data. Shankar et al., (1998) showed that RENB models may be suitable for road accident studies given that geometric and traffic variables are likely to have location-specific effects. RENB models allow for the introduction of a random location-specific effects term in the relationship between the expected numbers of accidents and the covariates of an observation unit in a given time period. In instances where the Poisson regression and Negative Binomial models have failed to take into account location-specific effects and/or serial correlation in time of the accident counts, the RENB is recommended for use in capturing all unobserved heterogeneity (Chin and Quddus, 2003). Despite the RENB being able to identify factors that may influence accident frequencies there remains the concern about its suitability to predict accidents.

Random-parameters negative binomial models have been used to explore and better understand the factors affecting the frequency of accidents on road segments over a period of time. This type of model has been shown to enable one to account and correct for heterogeneity that could arise from a number of factors relating to road geometrics, pavement and traffic characteristics, driver behaviour, socio economic factors and other factors (Anastasopoulos and Mannering, 2009). Using 5 years of vehicle data from Indiana, USA a random parameters negative binomial regression model was developed by Anastasopoulos and Mannering (2009) to gain new knowledge into the extent to which various factors impact on accident frequencies. Factors relating to pavement condition, road geometry and traffic were used. Findings from the study indicated that disregarding the possibility of random parameters when estimating count-data models can result in significant different marginal effects of the factors that affect accident frequencies. Random parameter models are viewed as an extension of random-effects models and instead of only influencing the intercept of the model, random-parameter models enable each estimated parameter of the model to vary across each individual observation in the dataset (Milton, Shankar and Mannering, 2008). With each observation having its own parameters, the final model often provides a statistical fit that is significantly better than a traditional fixed parameter model. Despite this advantage shown, random-parameter models have been found to be very complex to estimate; their predictive power may not be particularly significant and also it may not be possible to transfer model results to other datasets since the results are distinct to observations (Lord and Mannering, 2010).

3.2.6 Multivariate Poisson-lognormal (MVPLN) regression, Univariate Poisson regression and Poisson lognormal (PLN) models

Poisson log-normal models assume that the intensity parameter of a Poisson process follows a lognormal distribution in a sample of observations (Stewart, 1994). The lognormal distribution can model skewed distributions and it has simple parameters that are understandable in the context of normal distributions however, zero counts are problematic since the log of zero is negative infinity so normally adjustments such as adding 0.5 to zero counts may be required (Stewart, 1994).

MVPLN allows for a more general correlation structure as well as overdispersion. MVPLN provides the opportunity to incorporate the association across collision severity levels and their influence on safety analyses. The MVPLN model apart from being able to account for over-dispersion possess a reasonably general correlation structure which enables for different covariance terms and the possibility of negative correlations (El-Basyouny and Sayed, 2009). It is able to handle more than two collision categories and the MVPLN model's unknown parameters have to be estimated and a Bayesian approach can be used for the estimation. El-Basyouny and Sayed (2009) used the MVPLN model with only AADT and collision frequencies used as data variables. Estimates of the extra Poisson variation parameters were considerably smaller using the MVPLN implying a higher precision. It was estimated that the MVPLN model was more than twice as precise as the univariate Poisson log-normal (PLN) model. The better precision values obtained was attributed to the fact that the MVPLN takes into account the correlation between the latent variables representing property damage only (PDO) and injuries plus fatalities (I + F). Univariate Poisson regression models are unable to account for correlations at different levels of severity for a specific segment of roadway (El-Basyouny and Sayed, 2009). In univariate Poisson regression models traffic crash counts at different levels of severity are estimated separately. In the model by Ma, Kockelman and Damien (2008), crash data in addition to road design features, traffic intensity and geometric parameters were used. The model dealt with the effects of individual geometric design features as being independent of each other ignoring the possibility of a relationship between them.

El-Basyouny and Sayed (2009) used a PLN model for urban arterial roads in Vancouver, Canada and found a number of variables including segment length, AADT, number of lanes and a few other factors influencing the frequency of accidents. With a single data set used in that study, the authors suggested that further research with varying datasets will aid in confirming the results obtained. Another model by El-Basyouny and Sayed (2009) was done using signalised intersections in the city of Edmonton. Two severity levels; property damage only (PDO) and injuries plus fatalities (I+F) were considered using a single data set. The model was compared with an independent univariate PLN model with respect to model inference, goodness-of-fit, identification of hazardous locations and precision of expected collision frequency. Results indicated that some hazardous locations could be overlooked if the analysis is restricted to a univariate model. This model mainly used AADT as model parameters and was noted to be preferred by practitioners to models containing several covariates due to the ease of calibration (Lord, Guikema and Geedipally, 2008). AADT-only models may experience an omitted variable bias as the unobserved heterogeneity from other factors known to influence collision frequency (example, number of lanes, signal-control timing, speed limits, etc) ends up in the correlation structure and affects the estimated correlation (El-Basyouny and Sayed, 2009).

The Poisson-lognormal model, though it offers more flexibility than the negative binomial/Poisson-gamma models, is limited in the complexity of the model estimation and the effects of small sample sizes and low sample mean values (Lord and Mannering, 2010).

3.2.7 Zero Inflated (ZI) models

Zero-inflated models are normally applied to crash data containing a lot of zero data than would be expected for use in Poisson or negative binomial (NB) models (Lord, Washington and Ivan, 2007). They possess much greater flexibility in revealing processes influencing accident frequencies on observed roadway sections with zero accidents and those with observed accident occurrences (Shankar, Milton and Mannering, 1997). Even though zero-inflated models display an improved statistical fit to most crash data, Lord, Washington and Ivan, (2005) argue that the inherent assumption of a dual state process underlying the development of the model appears inconsistent with crash data. ZI models assume that the phenomenon being studied takes the form of a dual-state process; a true-zero and a non-zero state. Also, because the zero or safe state has a long term mean equal to zero, the model crash data generating process cannot be reflected properly.

3.2.8 Generalised Linear models (GLM)

Generalised Linear Models (GLM) is a big family of models of which Linear models belong to. GLMs are generalisations of the Linear Regression Model (LRM). LRMs assume the dependent variable is continuous, normally distributed with constant variance and is a linear function of a set of independent variables. The independent variables tend to be categorical, continuous or a combination of both (Dunteman and Ho, 2006). Alternatively, the response variable is taken as part of the exponential family of distributions which include Normal, Poisson, gamma, inverse Gaussian, binomial, exponential and other distributions. A multivariate method such as GLMs accommodates for correlations and enables interaction effects to be explored (Zou, Zhang and Lord, 2013). GLMs statistical computations also assist in choosing significant variables as well as in validating the model assumptions. The limitations exhibited in LRMs led to the development of GLMs by Nelder and Wedderburn in 1972.

GLMs allow the prediction of the conditional mean or some function of the conditional mean of a dependent variable as a linear function of a set of independent variables or covariates. This implies that for each subject or observation the expected value or some function of the expected value of the dependent variable is subject to the value of the independent variables or covariates (Dunteman and Ho, 2006).

GLMs are typically made up of three parts; a response variable distribution (occasionally referred to as the error structure), a linear predictor that takes into account the regressor variables or covariates and a link function connecting the linear predictor to the natural mean of the response variable (Myers, Montgomery and Vining, 2001).

Greibe (2003) used GLM to predict the expected number of accidents at urban junctions and road links. The model was used in the identification of factors affecting road safety and ‘black spots’. Vehicle traffic flow was found to be the most powerful variable for the models developed. A lot of road geometric and non-geometric variables were incorporated into the model which provided a sound blend of data variables. However, a major problem encountered in this model was the strong internal correlation that existed within the data. This correlation made it difficult for the safety effects from a single explanatory variable to be estimated since it may be affected by other variables in the model (Greibe, 2003). In another model by Cafiso et al. (2010) three models were proposed out of the 19 models developed and ranked. These models were developed using 14 variables belonging to four main groups. However the three proposed models did not contain all fourteen variables with AADT being the only common variable in all models.

3.2.9 Classification and Regression Tree (CART) models

The CART model is a data mining technique applied in business, industry and engineering and it does not require any pre-defined underlying relationship between the dependent variable often referred to as the target and the independent variable (predictors) (Chang and Wang, 2006). CART is known to be good at handling prediction and classification problems. A CART model was developed using road accident data from Taipei, Taiwan to establish the relationship between injury severity and driver/vehicle characteristics, highway environmental variables and accident variables. Some advantages and disadvantages of using the CART model were highlighted. There is no need to specify a functional form in a CART model unlike a regression model in which a mis-specification of the model can result in an erroneous estimated relationship between the dependent and independents variables as well as the model predictions. In regression analysis, outliers are known to present a serious problem with an adverse effect on the coefficient estimates. In contrast, CART models have outliers isolated into a node resulting in no effect on splitting (risk factors) (Chang and Wang, 2006). CART deals with large data sets containing a large number of explanatory variables and can produce beneficial results from using a few important variables. The main disadvantages associated with CART models include the lack of provision of a probability level or confidence interval for the risk factors (splitters). CART models also have difficulty in conducting elasticity or sensitivity analysis and are also very unstable. They are normally used to identify important variables and then some other flexible modelling technique is used to develop the final model.

3.2.10 Other models

Integer valued autoregressive (INAR) Poisson models are applicable for the analysis of time series count data since these models have the characteristics of Poisson regression and are able to deal with serial correlation providing an alternative to the real-valued time series models (Quddus, 2008). Quddus (2008) used INAR models for the time series analysis of traffic accidents in Great Britain using both disaggregate and aggregate time series data. The performance of the INAR was found to be good in comparison to Negative Binomial (NB) models. The INAR model was found to be able to control both properties of time series data. It was further suggested that INAR models should be considered when developing accident prediction models for serially correlated time series count data, particularly during instances where the time interval between successive observations is short, such as a day, a week, or a month rather than a year. As with most road accident data, INAR also has a limitation in dealing with over-dispersion.

Mixed logit models have been used in modelling accident severities at various levels and is noted to be able to account for the differential (from one road segment to the next) effects that variables have on the numbers of injury severity (Milton, Shankar and Mannering, 2008). In traditional multinomial logit models the error term (unobserved effects) are assumed to be an independently and identically distributed extreme value. It is however necessary to take into account the likelihood of shared unobservables between injury outcomes in functions that determine the injury proportions on individual roadway segments. The alternate severity outcomes are presumed to be independent with a model specification error resulting if they are not in traditional multinomial logit models. The mixed logit permits for a more general error-correlation structure, while eliminating the requirement for creating a priori assumption about the structure of shared observables (such as nested structures) in preventing this error term (Milton, Shankar and Mannering, 2008). It has been noted that the difficulties associated with modelling accidents at various severities have resulted in most road safety researchers opting for accident modelling based on frequencies. Milton, Shankar and Mannering (2008) used mixed logit which offered the flexibility of capturing segment-specific heterogeneity that can arise from a number of factors related to the roadway characteristics, environmental factors, driver behaviour, vehicle types and interactions among these factors. The inter-relationship between the factors used in the model were not revealed even though consideration of various factors provided a better understanding of the complicated interaction of factors that come to play with road safety.

3.3 Summary table of reviewed accident prediction models

A summary of the accident prediction models reviewed in this chapter is provided in Table 2.

Method	**Authors and year**	**Area studied**	**Characteristics of model**	**Results/ Comments**
Empirical Bayes (EB) (variant)	Bude and Larsson (1988)	Junctions of the same type/individual junction	1,901 3-way junctions on rural roads. Data used included number of junctions and existing number of accidents.	Model predicts accidents at junctions based on existing record of accidents. Not possible to show better results or significant difference was achieved for the variant EB as compared to conventional EB method.
	Mountain, Fawaz and Jarrett (1996)	Link and minor junction accidents	3800km of highway including 5000 minor junctions studied. Link length and traffic flow are the main explanatory variables used in the model.	Very basic model not including a lot of parameters. The EB method results in unbiased estimates of the estimated treated effect when sites are selected on the basis of high accident frequency.
Negative binomial	Shankar, Milton and Mannering (1997)	Principal arterial roads in Washington State, USA	4386km of highway and 11757 accident records were used. Model included variables such as shoulder width, horizontal and vertical curve information, traffic volume and speed data. In this model, the effects of individual variables were presented.	Model allows for over dispersion and takes account of the variance being more than the mean. A single model containing all the individual variables was not presented.
Negative binomial	Abdel-Aty and Radwan (2000)	Principal arterial roads in Florida, USA	1606 accidents over a 3 year period were used. Model revealed the significance of the AADT, degree of horizontal curvature, lane, shoulder and median widths, urban/rural, section lengths and frequency of accident occurrence.	Heavy traffic volume, speeding, narrow lane width, larger number of lanes, urban roadway sections, narrow shoulder width and reduced median width increased likelihood of accident involvement.
Negative binomial	Noland and Oh (2004)	Roads in the state of Illinois, USA	Analyses focussed on whether various changes in road network infrastructure and geometric design is associated with changes in road fatalities and reported accidents. Model hypothesises that improved infrastructure geometric design is beneficial to road safety.	Results could not confirm hypothesis. Increased number of lanes associated with increased traffic-related accidents and fatalities. Increased lane widths associated with increased fatalities.

Table 2 Summary table of reviewed accident prediction models

Method	Authors and year	Area studied	Characteristics of model	Results/ Comments
Negative binomial	Graham and Glaister (2003)	Pedestrian road casualties	The model examines the role of urban scale, density and land-use mix on the incidence of road pedestrian casualties.	Results indicate the local environment can have a powerful influence on pedestrian casualty incidents. Even though model includes parameters such as the road length and traffic flows, geometric characteristics of the road are not included in the model which could have an impact on pedestrian casualties.
Negative binomial	Vieira Gomes (2013)	Urban road networks in Lisbon, Portugal	The model predicts the frequency of accidents. The explanatory variables included vehicles and pedestrian traffic flow counts, and highway geometric design features.	Variables identified to increase the frequency of accidents included traffic, lane balance, average lane width etc. Accidents were not split by levels of severity. No validation checks were carried out.
Random Effect Negative Binomial (RENB)	Chin and Quddus (2003)	Signalised intersections in Singapore	52 four legged intersections in Southwestern Singapore were used accounting for 15% of such intersections. Accident data, traffic volume, geometric data in the form of approach curvature, sight distance to intersection and road width was used. Other parameters included median width, left turn length on slip roads, distance of upstream and downstream bus stop from intersection, uncontrolled left turn lane, exclusive right-turn lane, acceleration section and the presence of an overhead bridge near the intersection. Other parameters included the existence of surveillance camera, signal control and signal timing plan.	11 variables significantly affected safety at the intersections. The relatively small sample size used placed a limitation on the findings. Even though the RENB can be used to identify factors that influence total accident frequency, there still remains the question about its suitability in predicting accidents. This is because the identification of factors may provide a relationship with accidents occurring but not necessarily the causation of the accident.

Table 2 Summary table of reviewed accident prediction models (continued)

Method	**Authors and year**	**Area studied**	**Characteristics of model**	**Results/ Comments**
Multivariate Poisson-lognormal (MVPLN) regression model	El-Basyouny and Sayed (2009)	Signalized intersections in the city of Edmonton, Canada	99 signalized 4-leg intersections were studied in order to develop collision prediction models. Data on collision frequencies and traffic volumes were used. Collisions are grouped as property damage only and injuries and fatalities.	This model may suffer from omitted variables since it was purely based on AADT. This is because other non-flow variables have been shown to affect collision frequency. The technique generalizes the univariate posterior probability of excess commonly proposed and applied in literature to fit the multivariate relationship between latent variables.
Multivariate Poisson-lognormal (MVPLN) regression model	El-Basyouny and Sayed (2009)	City of Edmonton, Canada	99 signalised intersections were used for the development of collision prediction models relating to safety of urban 4-leg intersections to their traffic flow. Other data used included collision frequencies for three years and traffic volume. Two severity levels ie. Property damage only (PDO) and injuries plus fatalities(I+F) were considered.	The MVPLN model was found to be twice as precise as the univariate PLN model. The MVPLN also provided a superior fit over the univariate models. Small differences in the regression parameter estimates were noted between the univariate and multivariate models.
Bayesian Hierarchical (BH) models	Quddus (2008)	Greater London metropolitan area-wide, UK	633 census wards of the Greater London metropolitan area were used. Traffic characteristics, road characteristics and socio-demographic factors were used in the model.	Results obtained from the negative binomial (NB) models and the BH models were similar. BH models were found to be an appropriate model to analyse area-wide traffic crash occurrences. A series of relationships were developed between area-wide different traffic casualties and the contributing factors associated with the ward characteristics instead of a single model being developed.

Table 2 Summary table of reviewed accident prediction models (continued)

Method	**Authors and year**	**Area studied**	**Characteristics of model**	**Results/ Comments**
Bayesian Hierarchical (BH) models	Li, Zhu and Sui (2007)	Harris County, Texas state maintained roadways, USA	Five year crash data, where one or more vehicles were towed included in the data. AADT and Annual average daily vehicle miles travelled (VMT) are used.	Model does not include road characteristics and is unable to provide reasons as to why some road segments are riskier than others. Model does not also distinguish between accident severities.
Full Bayes (FB) hierarchical models	Aguero-Valverde and Jovanis (2006)	Pennsylvania State county roads, USA	Fatal and injury crash data as well as socio-demographics, weather conditions, transportation infrastructure and amount of travel data was used.	Models were developed for injury and fatal crashes. No evidence of spatial correlation in fatal crashes but significant correlation was found for injury crashes.
Full Bayes	Song et al. (2006)	Texas, USA	Crash data, weather variations and horizontal curve parameters were used in the model.	Higher accident risk locations were identified for some sites than others.
Full Bayes	Li et al. (2008)	Iowa State, USA	Data was analysed by comparing sites receiving an intervention during the study with sites which did not to assess the effect of the intervention on road safety. The number of crashes per month over the period 1982 to 2004 and average daily traffic was used.	Results not presented in this paper but it is apparent that the FB method is favoured.
Multilevel model	Jones and Jorgensen (2003)	Norway	Data from 16,000 fatally and seriously injured casualties involved in accidents between 1985 and 1996 analysed. Other parameters used in the analyses were age, sex, type of vehicle, characteristics of the impact, road section attributes, time of day and alcohol involvement.	Model found statistical significant residual variation in casualty outcomes between separate accidents and different geographical locations. The road geometric characteristics were absent from the model.

Table 2 Summary table of reviewed accident prediction models (continued)

Method	Authors and year	Area studied	Characteristics of model	Results/ Comments
Generalised Linear models (GLM)	Greibe (2003)	Urban roads in Denmark	Data from 1036 junctions and 142km of road links in urban areas was used.	Model for urban road links was able to describe more than 60% of the systematic variation (percentage-explained value) while models for junctions had lower values. Most powerful variable was vehicle traffic flow. Both geometric and traffic characteristics used however not all parameter values can be obtained for current research.
Bayesian model	MacNab (2004)	British Columbia, Canada	Hospital data from 83 local hospitals from 1990 to 1999 was used. Some of the data included socio-economic indicators, residential environment indicators, medical services availability and utilisation, population health, crime rates, speeding charges and seatbelt violation data.	Study revealed geographic/spatial patterns in injury ratios but did not indicate areas with exceptionally elevated injury rates.
Bayesian spatial model and mixed logit model	Wang, Quddus and Ison (2011)	M25 motorway and its surrounding major roads, London, England.	2003 to 2007 road accident data including parameters like date, time, lighting weather conditions, number of vehicles and number of casualties was used. Others include traffic flow, traffic delay and road curvature.	The two stage model was generally comparable to the MVPLN model and fixed proportion method. Even though the model was found to be able to predict low frequency accidents, there is the need for further research to validate the method with other data samples or models.
Poisson and Negative Binomial Regression models	Hiselius (2004)	Rural roads in Sweden	Data from 83 rural road sections, police reported accidents with personal casualties from 1989 to mid-1995 and hourly traffic flow was used. It was assumed that traffic flow counted at a stationary point along the road is valid for the road section.	Very basic model and it is mentioned that missing factors such as weather, road conditions, type of vehicle and driver characteristics can influence the occurrence of an accident.

Table 2 Summary table of reviewed accident prediction models (continued)

Method	**Authors and year**	**Area studied**	**Characteristics of model**	**Results/ Comments**
Classification and regression tree model	Chang and Wang (2006)	Taipei, Taiwan	2001 accident data in addition to driver/vehicle characteristics, highway/environmental variables and accident variables.	The most important variable associated with crash severity was found to be the vehicle type.
Integer-valued autoregressive (INAR) Poisson models	Quddus (2008)	Great Britain	Annual road traffic fatalities between 1950 and 2005 used. Explanatory variables used included seat belt wearing law, new legislation on safety and veh-km (billion). Aggregate and disaggregate time series data were considered.	The performance of the INAR model in comparison with the real-valued models show both models to perform similarly in terms of coefficient estimates and goodness of fit for the aggregated time series data. For the disaggregated time series accident data, the INAR model was found to be better than the real-values models.
Empirical Bayes (EB)	Elvik (2008)	Norway	21,738 1-km sections on national roads, accident data and some variables associated with the number of accidents for the period 1997 – 2004 was used. Road sections remained unchanged except for ordinary road maintenance like resurfacing, renewing road markings and traffic sign replacement. The main variables included in the model were AADT, speed limit, motorway type, number of lanes and number of junctions.	The EB method provided a better prediction of accidents than the traditional approach.
Mixed logit models	Milton, Shankar and Mannering (2008)	Washington State, USA	Accident severities modelled. Data consisted of 274 road way segments with mean segment length of 2.4 miles, accident data from 1990 to 1994, weather, geometric, pavement, roadside and traffic characteristics associated with road segments.	Details of geometric parameters used were quite detailed and model also included other factors which contribute to road accidents. Estimated parameters included in the model were statistically significant.

Table 2 Summary table of reviewed accident prediction models (continued)

Method	Authors and year	Area studied	Characteristics of model	Results/ Comments
Poisson lognormal model	El-Basyouny and Sayed (2009)	Urban arterial roads in Vancouver, British Columbia, Canada	392 road segments clustered into 58 corridors, accident and traffic data was used to model accident frequencies.	Apart from the intercepts, the regression coefficients were all found to be positive signifying that factors such as segment length, AADT, crosswalks density, business land use, un-signalised intersection density and number of lanes are positively associated with the number of accidents.
Random-parameters negative binomial regression	Anastasopoulos and Mannering (2009)	Rural interstate highways in Indiana, USA	5 year accident data, 322 road segments and detailed geometric and pavement data was used in developing a model to obtain knowledge about factors influencing accident frequencies.	A number of factors relating to pavement condition, road geometry and AADT were found to significantly influence the frequency of accidents. Some of the factors were found to vary significantly across road segments.
Full Bayesian (FB) multivariate	Park, Park and Lomax (2010)	Expressways in Korea	The FB method was applied to crash data from Korean expressways in order to assess the safety benefits of decreasing the posted speed limit. Crash data from 1996 to 2006 in addition to road characteristics, AADT and speed limit change information was used. The model was developed for crash counts of different types of severity for a before-after evaluation with a comparison group.	The multivariate approach can recover the underlying correlation structure of the multivariate crash counts and can lead to a more precise safety effectiveness estimate by taking into account correlations among different crash severities or types for the estimation of the expected number of crashes.

Table 2 Summary table of reviewed accident prediction models (continued)

Method	Authors and year	Area studied	Characteristics of model	Results/ Comments
Generalised Linear Modeling (GLM)	Cafiso et al. (2010)	Two lane local rural roads in Italy	5 years road accident and traffic flow data was used in addition to computed road curvature change rate (CCR) and road side hazard rating (RSH). Other variables not influencing the homogeneity of the segments included in the model were curvature ratio, tangent ratio and average operating speed. The model was developed to estimate the expected number of accidents along the road segments.	19 models were developed and three were selected as recommended. The three were selected on the basis of practical considerations, statistical significance and goodness of fit indicators. The main variable common in all three recommended models was AADT.

Table 2 Summary table of reviewed accident prediction models (continued)

3.4 Characteristics of accident prediction models

Table 3 is a summary of some accident prediction models stating some advantages and disadvantages about them.

Model type	Advantages	Disadvantages
Poisson	Most basic model; easy to estimate.	Cannot handle over- and under-dispersion; negatively influenced by the low sample-mean and small sample size bias.
Negative binomial/ Poisson-gamma	Easy to estimate can account for over-dispersion.	Cannot handle under-dispersion; can be adversely influenced by the low sample-mean and small sample size bias.
Poisson-lognormal	More flexible than the Poisson-gamma to handle overdispersion.	Cannot handle under-dispersion; can be adversely influenced by the low sample-mean and small sample size bias (less than the Poisson-gamma), cannot estimate a varying dispersion parameter.
Zero-inflated Poisson and negative binomial	Handles datasets that have a large number of zero-crash observations.	Can create theoretical inconsistencies; zero-inflated negative binomial can be adversely influenced by the low sample-mean and small sample size bias.
Conway–Maxwell–Poisson	Can handle under- and over-dispersion or combination of both using a variable dispersion (scaling) parameter.	Could be negatively influenced by the low sample-mean and small sample size bias; no multivariate extensions available to date.
Gamma	Can handle under-dispersed data.	Dual-state model with one state having a long-term mean equal to zero.
Generalized estimating Equation	Can handle temporal correlation.	May need to determine or evaluate the type of temporal correlation a priori. Results sensitive to missing values.
Generalized additive	More flexible than the traditional generalized estimating equation models and allows non-linear variable interactions.	Relatively complex to implement; may not be easily transferable to other datasets.

Table 3 Characteristics of accident prediction models (Lord and Mannering, 2010)

Model type	Advantages	Disadvantages
Random-effects	Handles temporal and spatial correlation.	May not be easily transferable to other datasets.
Negative multinomial	Can account for over-dispersion and serial correlation; panel count data.	Cannot handle under-dispersion. Can be adversely influenced by the low sample-mean and small sample size bias.
Random-parameters	More flexible than the traditional fixed parameter models in accounting for unobserved heterogeneity.	Complex estimation process; may not be easily transferable to other datasets.
Bivariate/multivariate	Can model different crash types simultaneously; more flexible functional form than the generalised estimating equation models (can use non-linear functions).	Complex estimation process; requires formulation of correlation matrix.
Finite mixture/Markov switching	Can be used for analysing sources of dispersion in the data.	Complex estimation process; may not be easily transferable to other datasets.
Duration	By considering the time between crashes (as opposed to crash frequency directly), allows for a very in-depth analysis of data and duration effects.	Requires more detailed data than traditional crash frequency models. Time-varying explanatory variables are difficult to handle.
Hierarchical/multilevel	Can handle temporal, spatial and other correlations among groups of observations.	May not be easily transferable to other datasets. Correlation results can be difficult to interpret.
Neural network, Bayesian neural network, and support vector machine	Non-parametric approach does not require an assumption about distribution of data, flexible functional form and usually provides better statistical fit than traditional parametric models.	Complex estimation process; may not be transferable to other datasets, work as black-boxes and may not have interpretable parameters.

Table 3 characteristics of accident prediction models (Lord and Mannering, 2010)(continued)

3.5 Facility Location

Facility location problems have been defined in most literature with similar wording. In one such review of location science research, facility location problems was described as one investigating where to physically locate a set of facilities (resources) in a way so as to minimise the cost of satisfying a set of demands subject to some constraints (Hale and Moberg, 2003). Facility location problems are noted to be generally solved in three main environments; continuous spaces (spatial), discrete spaces and network spaces. Continuous spaces solve problems in a continuous space (typically one, two or three dimensional) where every location is a possible ideal place for locating a facility. Discrete spaces deal with problems in which a location must be selected from a pre-chosen set of possible locations and finally network spaces deal with problems restricted to arcs and nodes of an underlying network (Hale and Moberg, 2003). Even though there seems to be some controversy about the origin of location science, it is thought (Hale and Moberg, 2003) to be traceable back to Pierre de Fermat, Evagelistica Torricelli (he is said to have been a student of the renowned scientist Galileo) and Battista Cavallieri. Hale and Moberg (2003) went on to provide a review of location science research which documented a broad review of facility location and location science research. It however was not aimed at providing a detailed list of location science topics but rather it provided the reader with a general review of the location science research landscape revealing the diverse areas and disciplines of application.

Location science has become well known with a lot of academic disciplines making use of facility models. These include civil engineers, geographers, electrical engineers, industrial engineers and urban planners to mention a few. Due to the range of academic disciplines involved in the use of facility location models, literature on this subject area has increased over the years. Over the years, a combination of exact and heuristic methods has evolved to help solve facility location problems. Brotcorne, Laporte and Semet (2003) trace the literature for ambulance location and relocation models some 30 years back showing how various stages have evolved with the passage of time. These models initially set off with static and deterministic location problems more suited to the early stages of planning thus ignoring the stochastic considerations such as the availability of ambulances. These were then followed on by the more probabilistic models reflecting the 'server within a queuing system' operation of ambulances and finally the dynamic nature of the more recent models which reflect the practicalities of relocating ambulances to efficiently and effectively cover needs.

3.5.1 Multi-objective spatial decision making

Pareto has been well acknowledged in research (Xiao, Bennett and Armstrong, 2007) as having the first study which dealt with optimality of multi-objective problems. His analytical work was illustrated using the following multi-objective optimisation problem

$$\min \quad \mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})]^T \qquad \text{subject to } \mathbf{x} \in \mathbf{S} \qquad \text{……..Equation 3}$$

where **f** is a vector of m objective functions (f_1, f_2, …, f_m) to be minimised, the decision variable vectors are **x** and the set defining all feasible solutions is **S**.

The solution space is described as the set of all feasible alternative solutions to the objective function whilst the decision space tends to be formed when the solutions are placed in a space formed by the decision variables. In a much similar way when solutions derived from the objectives are placed in a space it is described as an objective space. Non-dominated or most commonly called Pareto optimal solutions are used to describe a subset of all feasible solutions obtained which are not dominated by any solution. It also refers to solutions found to be outside the subset obtained which are dominated by at least one solution from within the subset. The Pareto front describes the set of non-dominated solutions (Xiao, Bennett and Armstrong, 2007).

Finding solutions to optimisation problems have generally been approached using either an exact method or a heuristic method. In the former method, limitations are brought to light when the problem to be solved tends to be of a large size and difficulties are encountered (Armstrong, 2000). The latter method on the other hand has been shown (Reeves, 1993) to be able to obtain optimal or near optimal solutions though this is not always guaranteed. The aim of multi-objective problems is to find near optimal solution (a Pareto optimal solution).

Three main approaches have been developed by Miettinen (1999) to solve a multi-objective problem heuristically. The first approach desires decision makers to agree on the weighting to be given to each objective. Scalarisation technique is then applied to decision makers preferred options and then a conversion of a multi-objective problem to a single objective problem to be solved. In this approach, taking a decision on which weighting to assign proves difficult to decision makers due sometimes to conflicting interests, different background knowledge and also due to the fact that certain objectives may just not be quantifiable. Examples of these are religious interests and cultural objectives.

The second approach, which has proved popular in recent times, is an interactive process whereby the preferences from decision makers are refined and inputted into the investigation. Each iteration provides decision makers with a subset of non-dominated solutions. Based on these solutions, preferred objectives are identified by decision makers leading to the formulation of a single objective problem. Solutions derived are refined and the process repeated until decision makers obtain a solution they are content with. However the reaching of an acceptable solution is not always guaranteed (Miettinen, 1999). This approach is similar in a way to the first in the way decision makers make input to the decision variables. However, this approach provides a step forward in allowing decisions to be refined and re-input into the model in an iterative fashion. This can be described as a positive step forward in obtaining better solutions even though this is not assured.

The third approach normally referred to as posterior articulation of preferences does not need the intensive involvement of decision makers when generating alternative solutions. It rather relies on methods that can be used to obtain a variety of Pareto optimal solutions uniformly distributed on a Pareto front. Solutions obtained are provided to decision makers who look at the merits of each solution and choose the most suitable option. The main problems identified with this approach are firstly to do with difficulties in obtaining a Pareto front and secondly the outcome of a large number of solutions will cause decision makers to be inundated with these solutions and faced with the challenge of selecting a solution from a number of alternatives. On the other hand, this approach is known to provide some advantages in that the Pareto front allows the real multi-objective structure of the problem to be properly visualised leading to better decision making (Brill, 1979). Another problem visualised in the third approach is to do with its reversed method of implementation. In optimisation problems, decision makers are the initial proposers or initiators of things to be done so for them not to be intensively involved during the generation of alternative solutions looks out of place. The solution will eventually revert back to these decision makers for approval so what one may ask is why not get them well involved in the decision making process so as to save time, money and provide a well-managed solution to the problem.

The use of posterior approaches has been used in evolutionary algorithms and in multi-objective optimisation problems (Badri, Mortagy and Alsayed, 1998; Yang, Jones and Shuang-Hua, 2007). In contrast to exact methods evolutionary algorithms, classified as heuristic algorithms are known to be more efficient and produces a wide set of non-dominated solutions comparative to the Pareto front (Deb, 2001). Another area the posterior approach is being made use of is in spatial analysis with visualisation capabilities such as intelligent geographic information systems (GIS) (Sasaki et al., 2010).

Many decisions to be made in location science require a multiple number of objectives to be achieved. This is mainly due to the various stakeholders involved who have objectives, normally conflicting which inevitably must all be fulfilled in one way or the other. In an example of locating airport fire stations the main objectives to be fulfilled by decision makers will be to minimise the total setup cost of the fire stations, minimise the total loss cost of an incident and minimise the longest distances from fire station to any incident point and high risk area (Tzeng and Chen, 1999).

3.5.2 Challenges of multi-objective optimisation problems

The three main challenges (Xiao, Bennett and Armstrong, 2007) to multi-objective optimisation problems are firstly the huge amount of time required for computation due to the optimisation problems being combinatorial. Secondly, different stakeholders having different backgrounds and views about the problem to be tackled are involved in the decision making. Finally, it is sometimes difficult if not impossible to meet the needs of all stakeholders involved in the decision making. This will then require a weighting of the objectives to be carried out with a compromise agreed on in order to achieve a meaningful result to the satisfaction of all stakeholders. Due to these three main challenges, solutions developed must be efficient in respect to time and complexity, effective with regards to finding good quality solutions and interactive so as to allow decision makers to visually probe different scenarios to find the most suitable solution (Xiao, Bennett and Armstrong, 2007). This is because decision makers are not normally interested in the technicalities of decision making.

3.5.3 Applicable methods for facility location problems

Salhi and Gamal (2003), in their approach for solving a location-allocation problem used genetic algorithm with the introduction of a selection and removal stage based on groups of chromosomes instead of individual chromosomes. They also proposed the use of specific crossover and mutation operators that relied on the impact of the genes selected as well as the introduction of a new operator for injecting new chromosomes into the population as and when needed. This approach by Salhi and Gamal (2003) brought about variety within the search thereby preventing early convergence. Cheung, Langevin and Villeneuve (2001) proposed an overall reconstruction of the traditional genetic algorithm method in order to overcome its inherent slow convergence weakness. A variety of crossover operators and the genetic search scheme were probed. Application of this method has been successfully executed on a number of real life problems such as hub location problems from airline networks and location-allocation problems from the oil industry. Arostegui, Kadipasaoglu and Khumawala (2006) made a comparative study of Tabu search, simulated annealing and genetic algorithms for facility location problems whilst Peng, Xu and Qin (2008) also made use of simulated annealing for a facility location problem.

Another method used in facility location problems is ant colony optimisation. Chen and Ting (2008) combined Lagrangian heuristic and Ant colony system to solve a facility location problem. Lu and Hou (2009) only made use of ant-colony optimisation to solve a facility location problem for large-scale emergencies whilst Xiang-lin, Yun-xian and Shen (2010) used a combination of fuzzy queuing and ant colony to solve large scale emergency problems.

Analytical Hierarchy Process (AHP) has also been used for facility location problems. AHP involves the decomposition of a problem into a hierarchy of easily understandable sub-problems, each of which can be independently analysed and having the elements of the hierarchy being able to relate to any aspect of the decision problem (Erden and Coşkun, 2010). Erden and Coşkun (2010) quote Siddiqui, Everett and Vieux (1996) as being the first to combine Geographic Information Systems (GIS) and AHP for a site selection problem. They also made use of AHP and GIS for the multi-criteria site selection of fire services. This same approach was used by Wei et al. (2011) for siting a fire station. Badri (1999) on the other hand made use of a combination of AHP and goal programming for solving facility location problems.

This research aims to move a step further with the application of Geographical Information Systems (GIS) in the location allocation problem being investigated. Malczewiki (2006) reveals that GIS and multicriteria decision analysis has generated a huge amount of interest over the 15 year period (1991 to 2004) investigated. GIS has a role to play in decision making or solving problems. Malczewiki (2006) went on to quote Cowen (1988) as recognising GIS ‘as a decision support system involving the integration of spatially referenced data in a problem solving environment’. Of the 319 peer-reviewed papers identified the major areas of application were found to be in environmental planning/ecology management, transportation, waste management, hydrology and water resource, agriculture and forestry in order of decreasing application with all these areas accounting for 72.4 percent of the literature surveyed. The remaining 27.6 percent was taken up by areas of application such as natural hazard management, recreation and tourism management, housing and real estate, geology and geomorphology, industrial facility management and cartography. Most of the decision problems were to do with land suitability problems. This illustrates the wide variety of research application areas of GIS in decision making. GIS and multicriteria decision making complement each other to make the decision making process more meaningful. GIS used to be an expert oriented area but this has shifted with more disciplines embracing the use of GIS leading to decision making becoming more flexible and allowing for a lot more public engagement (Malczewiki, 2006). Malczewiki (2006) further suggests that GIS multicriteria decision analysis (MCDA) can be constructed from two perspectives ie. the techno-positivist perspective on GIS and the socio-political, participatory GIS perspective. It is worth stating that, no matter what perspective is taken it should ultimately aim to provide easy integration between the two perspectives at any point in time when required.

Li and Yeh (2005) used genetic algorithms (GAs) and GIS to effectively solve a spatial decision problem for optimally locating ‘n’ sites of a facility based on population and transportation constraints derived from a GIS. The GA method performed better when compared with simulated annealing method. Indriasari et al. (2010) on the other hand used a combination of genetic algorithms (GA), tabu search

(TS) and simulated annealing (SA) with GIS in a facility location model. All methods produced better results than the existing method in use.

Despite the varied disciplines and combinations of these optimisation techniques in use, they all have their strengths and weaknesses. Li, He and Liu (2009) identified the huge combinatorial solution space problems associated with the brute-force method for solving optimisation problems with high-dimensional spatial data. This makes it an infeasible choice to opt for. Heuristic algorithms such as the Monte-Carlo method have been found to be simple and efficient amongst most methods but they do have the potential to get stuck on local suboptimal solutions since a move can only be made if a better solution is found. Genetic algorithms (GA) on the other hand have been found by Li, He and Liu (2009) to have problems in reaching convergence for large targets since the chromosome length has limitations on how long it can be. For example, a GA will find it hard to obtain the optimal location for more than 15 targets since the length of chromosome will be too long. To further explain the chromosome length, suppose it is desired to find the shortest distance to travel to six cities called A, B, C, D, E and F, a typical chromosome to represent the order of travel one may want to try can be represented as 'DFABEC'.

Li, He and Liu (2009) went on to reveal that ant colony optimisation (ACO) has been shown to have certain advantages over GAs when considered for certain areas of application. For example in operational settings with a dynamically changing system such as pipes breaking, pumps failing etc., ACO was found to be more advantageous than GA. Also ACO can be useful in instances where sequential decisions have to be made to help construct a trial solution where choosing some component solutions puts a limitation on subsequent choices.

3.6 Genetic Algorithms

Addressing the problem of site selection by decision-makers is an enormous task considering the various options that need to be taken into account. Since a large number of decision variables are involved and the search space can be large a search through heuristic methods that have been developed to handle problems associated with huge solution space was carried out. This approach was taken to make effective and efficient the decision making process associated with mounting road side speed control devices since this has been absent from the transport safety sector. Also most of the traditional methods used to solve problems associated with site selection are known to be unable to handle huge data sets and heuristic methods have been found to perform better (Li, He and Liu, 2009).

In an attempt to identify which optimisation technique is better suited for this research, an investigation into some of the optimisation techniques available was carried out. Taking into consideration the vast majority of optimisation techniques available, time limitations associated with this research and the types of optimisations techniques relevant to this area of research the choice was restricted to heuristic algorithms by making use of Genetic Algorithms and Pattern search. Heuristic Algorithms are algorithms that sift through a set of possible solutions and produce solutions close to the best. Heuristic algorithms provide approximate solutions but are able to get results quickly and easily which tends to be beneficial to decision making processes.

3.6.1 Background to Genetic algorithms

The concept of using a population of solutions to tackle practical engineering problems was considered in great detail during the 1950s and 1960s (Coley, 1999). There was however a revolutionary turn round in the 1960s when John Holland identified Genetic Algorithms (GA). Genetic Algorithms have been found to be the best and most robust kind of evolutionary algorithms (Haupt and Haupt, 2004). GA mimics natural biological evolution and belong to the broad class of evolutionary algorithms. Evolutionary algorithms work with a population of possible solutions to a problem. It then applies the principles of 'survival of the fittest', reproduction and mutation to reproduce a better breed/solution. For each evolutionary algorithm iteration, a new generation is produced through the processes of selection and reproduction leading to a new population of individuals who are better suited to the environment in which they are placed (Zalzala and Fleming, 1997).

Genetic Algorithms have been used in a wide range of applications and these include but are not limited to, image processing, laser technology, medicine, spacecraft trajectories, water networks, architectural aspects of building design and facial recognition (Zalzala and Fleming, 1997).

3.6.2 Characteristics of Genetic algorithms

Genetic Algorithms (GA) are different from the other types of heuristic algorithms in that they work on a population of feasible solutions and are also probabilistic (stochastic) and not deterministic. In a GA population each individual represents a possible solution to the problem under investigation (Zalzala and Fleming, 1997).

Once one identifies the individuals thought to be the best contributing solutions to the problem, a combination of these individuals into new individuals is carried out and with the repeated use of the method good solutions evolve (Mathworks, 2015).

The main characteristics associated with GAs is the selection process, crossover and mutation stages and these largely affect the performance of GAs (Zalzala and Fleming, 1997; Vose and Wright, 1998; Coley, 1999).

The selection process will choose the fittest individuals, whilst the cross over combines the selected individuals into new individuals and the mutation is a way of adding or taking out some 'genes' out of an individual to obtain a healthier breed.

In order for an algorithm to be described as genetic, it must have a

- mathematical representation for the solution being sought
- method for creating the start-up population such that one determines how many individuals must represent the population.
- way in which fitness can be measured so as to select the best individuals and abandon the rest.
- genetic function and this involves the selection, cross-over and mutation process.
- number of parameters and this involves deciding in advance the population size, number of parents to select, mutation rate etc.

The following sections explains some of the basic terms/procedures used in genetic algorithms

3.6.2.1 Initialisation

An initial population of possible solutions is generated for the problem to be solved. The population size can vary from little to many thousands depending on the simplicity or complexity of the problem to be solved. Traditionally, the population is normally generated at random and this covers the possible range of solutions (the search space) (Mathworks, 2015).

3.6.2.2 Selection

The selection process involves determining the number of times a particular individual is selected for reproduction and thus the resulting offspring that will evolve from the reproduction. In selecting to breed a new population, a proportion of the existing population is used. The selection is then made by choosing the fittest individuals in the population sample who always have a better chance of being included. The selection process normally has two main processes

- determining the number of trials an individual can expect to be selected (fitness assignment) and
- converting the expected number of trials into discrete number of offspring (sampling).

It is normal practice that the best 50 percent will be selected to be used in reproducing with the remaining 50 percent discarded. Though this is a practical method, it is not commonly used. This is because although it enables the best to reproduce, distinguishing between 'good' and 'very good' is problematic (Zalzala and Fleming, 1997; Vose and Wright, 1998). A more commonly used selection method is the fitness-proportional also known as roulette wheel selection (Zalzala and Fleming, 1997).

With the roulette wheel, the probability of an individual being selected depends on the individual's fitness. The analogy of a roulette wheel is the size of an individual's slot being proportional to their fitness (Coley, 1999). In the example shown in Figure 9, the circumference of the roulette wheel represents the sum of all six individual fitness values. Individual 6 is the most fit as it occupies the largest segment whereas individual 5 is the least fit occupying the least segment of the roulette wheel.

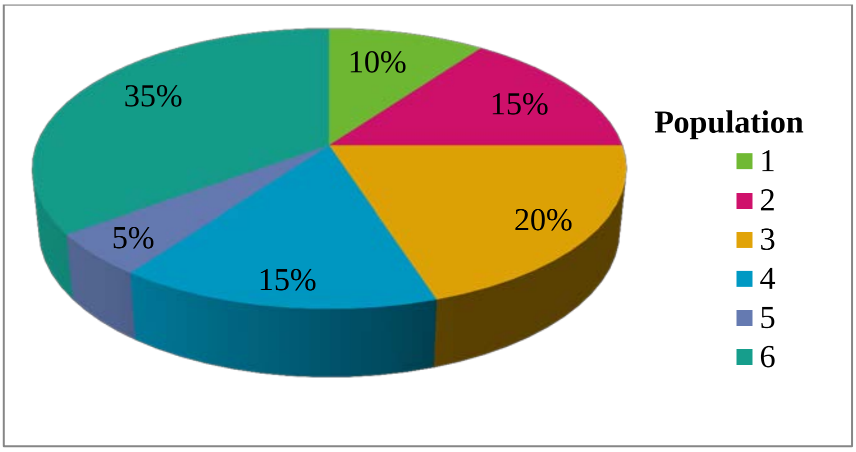


Figure 9 Roulette wheel

The wheel is spun and a figurative ball thrown in. The ball coming to rest in a particular slot is equivalent to the arc of the slot and thus to the fitness of the corresponding individual.

3.6.2.3 Crossover

Crossover is the fundamental operator for producing new chromosomes in genetic algorithm. Crossover enables new individuals to be produced having some parts of both parents' genetic make-up. It occurs between two individuals selected from the population and allows parts of their genes to be exchanged thereby forming new individuals (Vose and Wright, 1998).

Single-point crossover is the simplest form of crossover used (between one and eight crossover points are usually used in the natural world). If P_c is the probability of pairs of individuals selected undergoing crossover and R_c is a random number generated in the range 0 to 1, then the individuals can only undergo crossover if $R_c < P_c$, otherwise the pair will proceed without crossover. Typical values of P_c is from 0.4 to 0.9 with a P_c of 0.5 being the median value implying half of the population will be formed by selection and crossover with the other half being formed by selection only.

A single point crossover starts by initiating the cutting of a pair of selected strings at a random locus/point (this is chosen by selecting a random number R_L between 1 and L-1 and swapping tails to produce two child string. Assuming R_L = 4 then we have this scenario shown in Figure 10 (a). The new population consists of the same number of individuals as the original population.

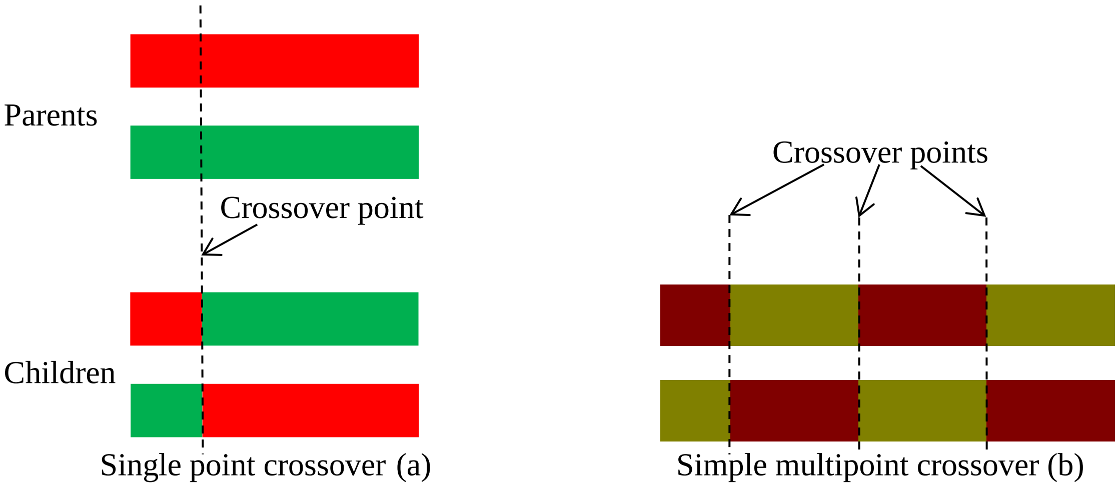


Figure 10 Simple single point crossover and Simple multipoint crossover

In the case of multipoint crossover, *n* crossover locations are chosen with k_i = {1, 2 ...., *l*-1} being the crossover points and *l* the chromosome length chosen randomly and arranged in ascending order. Both parents then exchange features between successive crossover points to produce new offspring. The first member of the series of genes occupying the first position on the chromosome at the location of the initial crossover point is left unchanged between the parents as shown in Figure 10 (b).

3.6.2.4 Mutation

The mutation process helps prevent premature convergence and it ensures that the population is thoroughly searched through, time permitting (Mathworks, 2015; Vose and Wright, 1998). This operation occurs after the crossover process and it is done through the alteration of one or more genes in a chromosome at randomly selected locations. Mutation occurs at probability rates similar to what happens in biological circles and this is usually very small (example 0.001). A 100 percent mutation implies all gene positions are altered. For example a string with this structure 1000 0001 0011 can be mutated at the second position resulting in this new string structure 1**1**00 0001 0011.

Crossover is the main operation that allows the search space to be well exploited, however crossover alone as an operation is found not to prevent a local minima convergence and this is where mutation is useful. Mutation is known to achieve better and healthier solutions than crossover only operations. It is advised not to select only crossover or mutation, with a balance between them helping to achieve good results (Coley, 1999; Vose and Wright, 1998).

The probability of mutating a gene/variable Pm, has been found to be inversely proportional to the number of variables 'n' in the chromosome. The more variables there are the smaller the probability of mutation required. The mutation rate is expressed as **m** = **1/n** (Vose and Wright, 1998).

3.6.3 Advantages and disadvantages of genetic algorithm

GAs as an optimisation and search technique is based on the principles of nature and its reproduction. Thus the 'survival of the fittest' rule in nature governs the effectiveness of genetic algorithms. Some of the main advantages of using GA is that it is able to sift through a solution very quickly enabling bad proposals to be discarded quickly and it works based on its own internal rules.

Some of the identified advantages for genetic algorithms are (Haupt and Haupt, 2004)

- It is able to optimise with both continuous and discrete variables.
- It optimises variables which have extremely complex cost surfaces.
- It does not just provide a single solution but a list of optimum variables.
- It does not require derivative information.
- It is able to effectively handle a huge number of variables.
- It can perform with numerically generated data, experimental data or analytical functions.
- It simultaneously searches from a wide sampling of the cost surface and
- It may encode variables such that the optimisation is done using the encoded variables.

Despite the above advantages, GAs cannot be said to provide the ideal solution for every problem. This is because, in problems that are not particularly difficult with few variables to solve, GA does not perform well and traditional methods of solution searching is faster and more efficient. However many real life problems are not as simplistic as one would expect and thus a quicker and more efficient problem solving approach is desired and this is where GAs comfortably fills the gap (Li and Yeh, 2005; Li, He and Liu, 2009).

3.6.4 Applications of genetic algorithms in location science

The use of genetic algorithms have been applied to areas such as locating fire stations, schools, waste disposal sites, supermarkets etc.. Other areas include health planning ie. ambulance locations, urban and regional planning, land use planning, environmental policy, land acquisition and routing problems. Examples are as follows:

Chau (2004) formulated a model on the allocation of construction facilities in GA using a mixed integer program. Lim and Kuby (2010) also used heuristic algorithms to solve a problem for the optimal locations for refuelling stations that used alternative fuels. Geroliminis et al. (2011) developed a heuristic model for the optimal deployment of many emergency response units in an urban transportation network and an application for transit mobile repair units in the city of Athens, Greece.

Yang, Jones and Yang (2007) used a combination of fuzzy multi-objective programming and genetic algorithm to determine the optimal location of fire station facilities. The three main things distinguishing this approach from existing fire station location models are firstly, the consideration of the fuzzy nature of a decision maker in the location optimisation model (this was done using the recommendations of the UK Home Office on the speed of fire engine attack to accidents in the optimal location model) (Yang, Jones and Yang, 2007). Secondly, full consideration of the demands for the facilities from areas with various fire risk categories was taken and finally the need to be practical and understandable to the decision maker was done (this involved choosing a suitable chromosome format and embedding the constraints into the fitness function of a genetic algorithm reducing the complexity of the model). One of the main advantages of this approach over existing ones is that in the use of multi-objectives and constraints, different risk categories and obstacles within specified regions were considered.

The main objectives defined for the model are to minimise the fixed cost and the total loss cost of incidents and to minimise the distance from the fire station to any incident site. The main constraints satisfied were firstly the total number of fire stations obtained must be N. The second constraint takes into consideration obstacles in a given area such that a fire station must not be built within any obstacles such as waterways and reserved areas. Thirdly there should be a reasonable distance d^r_{ab} between any two adjacent fire stations 'a' and 'b'. The decision variable S_{ij} is binary and the location is a pair of integer coordinates (x,y) taken from the National Grid. The chromosome is represented as shown in Equation 4

$[(i_1, j_1), (i_2, j_2), \ldots, (i_N, j_N)]$ where (i_a, j_a) is the coordinate of the ath fire station and $S_{i_a j_a} = 1$ ($a = 1, 2, \ldots, N$), $\quad i_a \in \{x_1, x_2, \ldots, x_k\}$ and $j_a \in \{y_1, y_2, \ldots, y_n\}$

………Equation 4

The ordinary genetic algorithm operators namely reproduction, mutation and crossover were applied to the fitness function and 30 fire station locations were identified. Most of the locations identified were similar to the actual fire stations already located (Yang, Jones and Yang, 2007).

The main difference noted between the recommended and actual fire station locations was that all actual locations were either within the city centre or in a village centre whereas some of the recommended fire station locations were between two villages or out of the city centres. The reason for this difference is because the recommended locations only targeted the various categories of risk in a given area without taking into account any social elements.

The second example of the use of multi-objective modelling for airport fire stations was considered by Tzeng and Chen (1999) who made use of a fuzzy model in combination with genetic algorithms. The model was developed to aid decision makers to help determine the optimal number and sites of fire stations at an international airport.

Binary type variables were used for the decision variables with the location model developed to optimise the number and sites of the fire stations. The main objectives of this model were to

- minimise the total setup cost of fire stations and the total loss cost of an accident. This is given mathematically as

  $\text{Min } f_1 = \Sigma\, i\, \Sigma\, j\, S_{ij} \text{ x SC} + \text{TLC x } e^{-\Sigma i \Sigma j \,\mathrm{S_{ij}}}$ ………Equation 5

Where S_{ij} is the decision variable.

- minimise the longest distance from the fire station to any point at the airport and is given by the expression

Min $f_2 = \Sigma \{(i,)\ Sij = 1, I,j\}$ {Max |x-i| + |y – j|} ………Equation 6

- minimise the longest distance from any fire station to the high risk area. The expression for this is given as

Min $f_3 = \text{Max}\ \Sigma\, i\, \Sigma\, j\ r_{ij}$ x {|x-i| + |y – j|} ………Equation 7

The first constraint imposed on the model related to the summation of S_{ij} being greater than or equal to 1 where S_{ij} is the decision variable, for a fire station located on an x,y coordinate (i, j).

$\Sigma\, i\, \Sigma\, j\ S_{ij} \geq 1$ ………Equation 8

The second constraint required a reasonable distance d^r_{ab} between any two fire stations. This distance d^l_{ab} should be such that it is not too long for fire stations to be able to support each other and the other distance d^s_{ab} should not be too short to cause an overlap of services. In formulating the fuzzy constraints d^l_{ab} and d^s_{ab} were used. The function representing the distance between any two fire stations is given as

$\mu_d(d_{ab}) = 1$ if $d_{ab} = d^r_{ab}$ …..Equation 9

$\mu_d(d_{ab}) = (d^l_{ab} – d_{ab})/(d^l_{ab} – d^r_{ab})$ if $d^l_{ab} \geq d_{ab} > d^r_{ab}$ ……Equation 10

$\mu_d(d_{ab}) = (d_{ab} – d^s_{ab})/(d^r_{ab} – d^s_{ab})$ if $d^s_{ab} \leq d_{ab} > d^r_{ab}$ …….Equation 11

$\mu_d(d_{ab}) = 0$ ….. Equation 12

Otherwise if $\{S_{ij}\}$ denote the set of $S_{ij} = 1$, with ‘a’ and ‘b’ being any two different elements taken from $\{S_{ij}\}$ then the above constraint can be given as shown in Equation 13

$|x_a - x_b| + |y_a – y_b| \approx d^r_{ab}$ …Equation 13

Equation 8 which is the first constraint relates to the minimum number of fire stations allowed such that it cannot be zero (implying no fire station) with the minimum number of fire stations desired to be greater than or equal to 1. Equations 9 to 12 relate to allowable distances. These distances are normally based on guidelines and sometimes local knowledge on what works. Distances between proposed fire stations should be such that they are not too far apart such that the desired benefits are not derived from placing the fire stations at the proposed locations. Also, the distances should not be too close such that the benefits to be derived from positioning these fire stations will be such that services will overlap each other resulting in cost inefficiencies.

Tzeng and Chen (1999) adopted a fuzzy multi-objective approach in locating fire stations at an airport for three main reasons. Firstly, the model was intended to be optimised for the number and location of fire stations at an airport and thus the input of more than one objective made a fuzzy approach suitable (Zeleny, 1982). Secondly when compared with traditional weighting methods for multi-objective optimisation models, it has shown (Bellman and Zadeh, 1970; Sakawa et al., 1997) that the fuzzy multi-objective method is simpler. Thirdly, the fuzzy multi-objective technique is more efficient in comparison to traditional methods (Chen and Hwang, 1992). Sakawa (1993) described the basic concept of fuzzy multi-objective optimisation as finding the maximal achievement level among a set of constraints with conflicting objectives. The model by Tzeng and Chen (1999) referred to as simple genetic algorithm (SGA) noted the omission of the crossover stage in genetic algorithms to prevent infeasible solutions and promote self-evolution efficiency.

The main difference between the Tzeng and Chen (1999) model in comparison to the Yang, Jones and Yang (2007) model is that in the former, risk rank was used which was statistically computed based on the accident data collected for the model location, Taipei International Airport. These accidents were categorised according to the accident frequency by type calculated. The accident frequency rate was computed for different areas and a risk rank was obtained from the inverse of the accident frequency rate computed. The latter however made use of recommendations by the UK Home Office which provides the speed of attack to fire incidents. These risks have been placed into four categories with the expected number of pumps to be made available within each risk category given. Ranges were also given for time limits for attendance as well as the distance limits to incident sites. These categories revealed the fuzzy nature of the Yang, Jones and Yang (2007) model.

Tzeng and Chen made a comparison of a GA to an enumeration method. The former was found to be more efficient with a computing time of 20.33 minutes and a comparative performance index of λ=0.75 in comparison to the latter which had a computing time of 73.57 minutes and a performance index of λ=0.78.

3.6.5 Sample calculation for Genetic Algorithm

Assume we have a population size M=20, two elite points N=2, a probability of recombination p_r=1 a probability of mutation p_m=0.02, and a number of generations K=100. The maximum number of function evaluations is $K(M-N)+N$= 1802. Manually computing 1802 evaluations will take some time and thus optimisation will prove to be a quicker and better alternative to computing values.

3.7 Pattern search

Heuristic algorithms are used in solving problems where conventional methods fall short. Like genetic algorithms, pattern search can be used to solve optimisation problems which have objective or constraint functions which are continuous, discontinuous, stochastic, do not possess derivatives or includes simulations or black box functions where some of the parameter settings have undefined values (Mathworks, 2015). Both pattern search and genetic algorithm can be customised. Pattern search optimisation was first proposed by Hooke and Jeeves (1961) with variants of this method emerging over the years. Pattern search has been applied to real life problems and to problems where the objective function is highly non-linear and discontinuous (Ackora-Prah et al, 2014). This includes the selection of an optimal portfolio of stocks (Ackora-Prah et al, 2014) and extracting parameters from various models (Al Hajri et al, 2012). Zheng and Chen (2010) used pattern search in predicting the strength and location of hazardous materials. The model was compared with other methods including genetic algorithms and it was found that pattern search can produce optimal solutions in a relatively shorter time in comparison with genetic algorithms which requires a large number of function evaluations per iteration resulting in more computation time (Zheng and Chen, 2010). Results obtained from the application of pattern search to these real life problems were found promising when compared to other methods. Pattern search is noted to have the potential for parameter estimation and system identification. Another study also identified the reduced computation time required in pattern search in comparison with genetic algorithms (Miloua et al, 2011)

Pattern search works by computing a sequence of points that approach an optimal point. During each step, the algorithm investigates a set of points called a *mesh*, around the *current point* (the point computed at the previous step of the algorithm). The formation of the mesh involves adding the current point to a scalar multiple of a set of vectors called a *pattern*. When pattern search finds a point in the mesh that refines the objective function at the current point, the new point then acts as the current point during the next step of the algorithm (Mathworks, 2015). The three main types of pattern search algorithms used are generalised pattern search (GPS) algorithm, the generating set search (GSS) algorithm and the mesh adaptive search (MADS) algorithm. These pattern search algorithms compute a sequence of points that approach an optimal point. A number of terminologies are used in *pattern search* and these are explained below.

3.7.1 Pattern

A *pattern* is a set of vectors $\{v_i\}$ that the pattern search algorithm uses to find the points to search at each iteration. The set $\{v_i\}$ is defined by the number of independent variables in the objective function, *N*, and the positive basis set (a positive basis is a positively independent set with positive span (Lewis and Torczon, 1996)).

3.7.2 Meshes

Pattern search searches through a set of points called a *mesh* at each step and also for a point that improves the objective function (Mathworks, 2015). Pattern search then forms the mesh by

- Generating a set of vectors $\{d_i\}$ by multiplying each pattern vector v_i by a scalar Δ^m. Δ^m is called the mesh size.
- Adding the $\{d_i\}$ to the current point (the point with the best objective function value found from the previous step).

  To illustrate this using the GPS algorithm assume
- The current point is [1.6   3.4]

- The pattern consists of the vectors v_1 = [1 0]; v_2 = [0 1]; v_3 = [-1 0] and v_4 = [0 -1]
- The current mesh size Δ^m is 4

The algorithm multiplies the pattern vectors by 4 and adds them to the current point. Direction is used to refer to the pattern vector that produces a mesh point.

3.7.3 Polling

The algorithm polls the points in the current mesh by calculating their objective function values during each step. With the 'Complete poll' option having a default setting 'off' in Matlab, the algorithm stops polling the mesh points as soon as it finds a point with objective function value less than that found for the current point. When this happens, the poll is described as successful and the point found serves as the current point for the next iteration (Mathworks, 2015). The algorithm only calculates the mesh points and their objective function values up to the point at which it stops the poll. At the point where the algorithm is unable to find a point that makes the objective function better, the poll is described as unsuccessful and the current point remains the same at the next iteration.

When the 'Complete poll' option is set to 'On' in Matlab, the algorithm calculates the objective function values at all the mesh points. The algorithm makes a comparison between the mesh point and the smallest objective function value to the current point. If the mesh point is found to have a smaller value than the current point then the poll is described as successful.

3.8 Choice of accident prediction model and justification and Chapter summary

Various accident prediction models have also been reviewed in the earlier sections of this chapter with a view to identifying an appropriate model to be developed and used in the optimisation model in chapter 6. Some UK road models were identified however these were rejected on the basis that they were old models, too complex for the purposes of this research, contained parameters that will not be easily obtained within the time constraints of this research and could not be validated.

In choosing an accident prediction model, some of the factors that had to be taken into account included the suitability, flexibility in obtaining data, availability of data and ability of the model to deal with confounding factors such as effects of regression to the mean.

All accident prediction models have limitations and in choosing one it was essential to also consider the objectives desired to be met from the model, advantages and disadvantages of the model, interpretation of results to be obtained and ease of transfer of the model for other use.

In this research an Empirical Bayes Negative Binomial Regression model was chosen to be developed in Chapter 5. This type of model was favoured because the primary aim of this research is not just to develop an accident prediction model but to go a step further in using the accident prediction model for optimising the location of speed control devices.

Roads selected for use in this research will be based on the high numbers of road accidents. It is possible that these high numbers may have been due to interim attributes arising from random variation in accident numbers with the possibility of the numbers returning to the net mean after a period of time (regression to the mean (RTM) effect). Similarly, the occurrence of the regression to the mean effect describes a significant effect when high crash sites are selected for treatment as is usually done. RTM depicts the statistical likelihood for high-crash tendency to decrease toward the mean in later periods of time independent of any treatment (Hauer, 1997 and Elvik, 2002). Hauer (1997) proposes the use of the empirical Bayes method to correct for the effect of the regression to the mean and this method is proposed for use in this research. In the method, the mean and variance of the expected number of accidents in a reference population is utilised to compute the revised estimate of the effect of an intervention. In this research a sufficiently large reference population of roads are used and the choice of road accident data was based on a long period of 5 years all in an attempt to help control for the effect of regression to the mean. The DETR (2001) suggests that much more than three to five years data will lead to the likelihood for changes in flow and notable changes in the network to affect accident figures. To control for RTM effects, the expected number of accidents in the before period were estimated using the Empirical Bayes method. Using this method, the mean accident frequency is determined as a weighted average of two sources of data: the observed accidents in the period before treatment, X_B and the predicted model approximation of expected accidents given the nature of the site and traffic flow levels (Hauer, 1997)

The latter sections of this chapter presented an overview of findings from various facility location problems making use of optimisation techniques, areas of application of the optimisation techniques and their strengths and weaknesses. The main purpose for carrying out the review on the use of optimisation techniques in facility location problems was to investigate its suitability, applicability and practicality to the location of speed control devices in this research. Facility location problems in the form of locating ambulance stations, schools, fire stations, hospitals etc have benefitted from the use of optimisation techniques (Wei et al., 2011; Li and Yeh, 2007; Salhi and Gamal, 2003). The use of optimisation techniques in road safety has been evidently lacking and its application in this research proves very useful. Managing vehicle speed to contribute to road traffic accident reduction involves the identification of accident prone areas from datasets. The decision making process for identifying an optimum location to place a speed control device also involves different objectives some of which include minimising the severity of road accident numbers, minimising the set-up cost of the device and minimising the maintenance costs associated with the device. Taking into consideration the fact that the cost aspects of an objective function will not affect the location of a speed control device but instead affect the number, the cost aspect of the objective function will not be included in the optimisation model. The objective function that deals with minimising the severity of road accident numbers will be optimised.

There are many studies making use of a single optimisation method or a combination of methods. As far as this research is concerned geographic information systems (GIS) in combination with two heuristic methods will be used in order to allow for meaningful comparisons to be made. Considering that earlier facility location problems made use of small datasets and simple objective function equations, the use of geographical information systems (GIS) in recent times in combination with more advanced computer programming software has facilitated the decision making process where large datasets and complex objective functions are involved (Indriasari et al., 2010). This research involves the use of large datasets and mathematical formulations that will benefit in terms of computation time savings and spatial data representation by making use of GIS and an optimisation technique.

4 Methodology and data description

4.1 Introduction to overall approach

A literature review on factors contributing to road accidents was identified in a literature review. These factors had to be selected based on their contribution to road traffic accidents, their availability and ease of collection of data to enable use in an accident prediction model. Another literature review was carried out to identify an appropriate accident prediction model to be developed using the factors identified as contributory factors to road accidents. The accident prediction model is to be used in the optimisation model. The optimisation model aims to optimise the location of road side speed control devices based on a set of objectives. The accident prediction models were tested on some independent roads data. Lastly an optimisation model which incorporates the accident prediction model was developed to identify appropriate locations to place roadside speed control devices. The optimisation model minimises the costs associated with the expected frequency of accidents along a segment of road at the required severity level. The model was then tested on independent road samples to determine the suitability as well as to validate the model.

4.2 Study Area: Nottinghamshire and Leicestershire, UK

Nottinghamshire is a county in the East Midlands of England sharing a boundary to the north-west with South Yorkshire, east with Lincolnshire, south with Leicestershire and west with Derbyshire. The established county town is Nottingham. Nottinghamshire's districts are Ashfield, Bassetlaw, Broxtowe, Gedling, Mansfield, Newark and Sherwood and Rushcliffe. The city of Nottingham was an administrative part of Nottinghamshire between 1974 and 1998 but this was made a unitary authority and only remains part of Nottinghamshire for ceremonial occasions. The estimated population of Nottinghamshire in 2011 was 785,800 with more than half of the population living in the Greater Nottingham metropolis having a population of about 650,000 (Nottingham City Council, 2011).

In a report by the Department for Transport (DfT) (2012a), exceeding the speed limit was a factor in 5 per cent of accidents and these accidents involved 14 per cent of fatalities. In 12 per cent of all accidents, at least one of 'exceeding the speed limit' and 'travelling too fast for the conditions' was reported with 25 per cent of these accidents accounting for fatalities. The Nottinghamshire County Council's 'Safer

Roads' report (Nottinghamshire County Council, 2010) recorded a 46 per cent reduction in the proportion of people killed and seriously injured in road traffic accidents as against a national target of 40 per cent. The reduction in the number of children killed and seriously injured in road traffic accidents achieved was 69 per cent against a national target of 50 per cent.

Leicestershire is a landlocked county in the Midlands of England deriving its name from the City of Leicester, long known to be the administrative centre. The City of Leicester unitary authority is currently managed separately from the rest of Leicestershire. Leicestershire is bordered to the north by Nottinghamshire, north-east by Lincolnshire, east by Rutland, south-east by Northamptonshire, south-west by Warwickshire, west by Staffordshire and north-west by Derbyshire. The population of the county is just under 1 million with more than half the population residing in Leicestershire's built-up area. The Local Government Act 1972 terminated the county borough status of Leicester city and the county status of neighbouring Rutland in 1974 changing both to administrative districts of Leicestershire. These changes were reverted on 1st April 1997 with Rutland and the City of Leicester becoming unitary authorities (Leicestershire County Council, 2011).

In Leicestershire, amongst the efforts particularly focused on improving road safety are speed management, improving safety for vulnerable road users and encouraging safer driving (Leicestershire County Council, 2011). Speed as a contributory factor to road traffic accidents in Leicestershire is provided in Table 4.

Accident contributory factor	**% of accidents involved in**
Loss of control	19
Failure to judge another person's speed and/or path	16
Travelling too fast for conditions	10
Overtaking	8
Exceeding the speed limit	5
All above factors	**58**

Table 4 Contribution of speed to accidents in Leicestershire, 2004 to 2009 (Leicestershire County Council, 2011)

Speed is still a major contributory factor accounting for 58 per cent of road traffic accidents from 2004 to 2009 in Leicestershire even though improvements have been attained for the years after 1998. ‘Travelling too fast for conditions’ and ‘exceeding the speed limit’ as contributory factors accounted for 24 per cent of road traffic accidents from 1994 to 1998 compared with results from Table 4 which shows the same contributory factors accounting for 15 per cent (Leicestershire County Council, 2011). Using the 2001 to 2004 road accident figures as a baseline for achieving the 2006 to 2011 targets, a 26 per cent reduction in the total number of killed and seriously injured casualties was attained which was in line with the set target of 26 per cent. A 40 per cent reduction in the number of children killed and seriously injured was achieved against a target of 33.3 per cent. The number of motorcyclists killed and seriously injured was reduced by 23 per cent in comparison to a reduction target of 20 per cent.

Both Nottinghamshire and Leicestershire have seen improvements in road safety. Managing vehicle speed was identified as one of the key target areas to focus on since speeding vehicles were noted (Nottinghamshire County Council, 2010; Nottingham City Council, 2011; Leicestershire County Council, 2011) to cause road traffic accidents and can be intimidating to pedestrians, cyclists and members of the communities. Leicestershire remains committed to the use of speed cameras given the effective role it has played historically in assisting to address the problem of excessive and inappropriate speed. Since speed has been identified as a contributory factor in the geographic areas of study, speed will not be isolated as the only accident contributory factor to consider instead accidents will be assessed based on the level of severity.

4.3 Accident Prediction Model Methodology

The aim of this research is to develop an optimisation model that addresses the problem of where to locate a road side speed control device such as a speed camera or a vehicle activated sign. In order to develop the model a mathematical formulation of the problem was required.

The mathematical formulation for the model involved the need for an appropriate accident prediction model. It is important to state that road traffic accidents result from a combination of a complex set of variables. Determining or controlling for all variables that are deemed to have played a role in a particular accident tends to be difficult if not impossible. The accident prediction model used in this research does not account for any human factors that may have resulted in the accident and it also does not contain all road geometry parameters. Using human factors in the classification of road traffic accidents is beyond the scope of this research. Chapter 3 has detailed information about accident prediction models and a sound justification for choosing the one used in this research is described. The accident prediction model developed in Chapter 5 relates the frequency of accidents at a specified level of severity to various road characteristics.

4.3.1 Road accident categories and terminologies

Some terminologies associated with road accidents are provided to allow for better understanding and interpretation of data. The following definitions and explanations are from the 2010 UK Road Casualties, Annual Report (Department for Transport, 2010).

An *Accident* involves personal injury occurring on a public highway (including footways) in which at least one road vehicle or a vehicle in collision with a pedestrian is involved and which becomes known to the police within 30 days of its occurrence.

Different classifications are used to describe road accident casualties as follows;

Casualty: A person killed or injured in an accident. Casualties are sub-divided into killed, seriously injured and slightly injured.

Killed: Human casualties who sustained injuries causing death less than 30 days (before 1954 about 2 months) after the accident. Confirmed suicides are excluded.

Serious injury: An injury for which a person is detained in hospital as an ‘in-patient’, or any of the following injuries whether or not they are detained in hospital: fractures, concussion, internal injuries, crushing, burns (excluding friction burns), severe cuts, severe general shock requiring medical treatment and injuries causing death 30 or more days after the accident.

Slight injury: These are injuries of a minor character such as a sprain (including neck whiplash injury), bruises or cut which are not judged to be severe or slight shock requiring roadside attention. This definition includes injuries not requiring medical treatment.

As well as casualties, the accidents are also categorised as fatal, serious and slight. A *fatal accident* is an accident in which at least one person is killed and other casualties may have serious or slight injuries. A *serious accident* is an accident in which at least one person is seriously injured but no person (other than a confirmed suicide) is killed. A *slight accident* is one in which at least one person is slightly injured but no person is killed or seriously injured.

4.3.2 Road accident data

The STATS 19 data for the UK was the main source of road accident data since it is the most up to date and reliable source of data available in the UK. Road accident data for the years 2008 to 2012 inclusive was used. The STATS 19 road accident data provides a lot of information about an accident. To allow for plotting of the road accidents onto the map, the geographical location ('x' and 'y' coordinates) of the accidents in terms of easting and northing coordinates in the British National Grid coordinate system was used.

4.3.3 Traffic flow data

The Department for Transport freely makes available AADT flow figures for most major roads and some minor roads in England and Wales on its website (Department for Transport, 2013a).Data for the years 2008 to 2012 was used. Vehicle flow data in the form of Annual Average Daily Traffic (AADT) was also obtained from Nottinghamshire County Council and Leicestershire County Council. Nottinghamshire and Leicestershire have been chosen as areas of interest in this research. This is because local councils in these regions have a history of collaboration with Loughborough University. Also these regions have VASs and speed cameras installed along some of their roads which are of interest in this research. The roads are also within modest travel distances for the researcher. The AADT data obtained was already categorised by all moving vehicles with the proportion of heavy goods vehicles (HGVs) also provided. AADT flows are equivalent to Annual Average Daily Flow (AADF) figures.

4.3.4 Road geometry data and maps

The main types of roads used in the accident prediction model were 'A' roads. This choice was made because A-Roads are close to the top end of the road classification system, no accident prediction model was identified for A-roads, speed cameras and vehicle activated signs are mounted along these roads and most of the data required for this class of road was available. A considerable amount of road accidents occur along these roads. A proportion of both high speed and moderate speed roads was considered appropriate. A-roads are classified as major roads and are intended to provide large-scale transport links within or between areas (Department for Transport, 2012).

The OS (Ordnance Survey) VectorMap Local for Nottinghamshire and Leicestershire was downloaded from EDINA digimap. These maps were provided at a scale of 1:10,000 with tile sizes being 5 x 5 km. The maps contain detailed national mapping of roads, railways vegetation, boundaries, hydrology, land areas, buildings and contours. Data is represented by points, lines, polygons and text. Road names and Department for Transport numbers are used to identify road alignment features. An OS Terrain 50 DTM map to a scale of 1:50,000 with tile sizes of 10 x 10km of the area of interest was also downloaded and superimposed on the OS VectorMap. EDINA digital maps provide contour maps, vector maps, land and height data and OS MasterMap Integrated Transport Network (ITN) map of the UK road network. The maps were in different scales and for the contour maps, the level of accuracy of the 5 metre vertical contours was ±2.5m and that of the 10 metre vertical interval contours was ±5m. Road segments were connected by links and nodes allowing for the easy identification and location of a section of road.

Junctions along the roads were physically identified using the legend from the maps used.

The road curvature represented by curve radius and the slope of the road are used in this research. The road geometry parameters were calculated using a mathematical formula used by Deublein et al (2013).

The radius of the road was calculated using a method which makes use of the x and y coordinates of three consecutive points. The slope of two vector lines between the three points is calculated followed by the calculation of the centre and radius of circle. Finally there is the smoothing of radius of radii calculated by means of average points calculated. The mathematical relation for calculating the radius is given by the as

Radius = sqrt($(x_1 - x_c)$^2 + $(y_1 - y_c)$)^2)

The slope of the road was also obtained by a mathematical formula that requires the input of a predefined set of altitude points which is the 'z' value and uses the Rise divided by Run approach. The formula is given as $((z_2 - z_1)/d)*100$. These calculations were executed in the MATLAB software.

4.3.5 Bicycle route

Bicycle route data was obtained from the England cycle route plans (Cycle-routes, 2016). For all A-roads used in the accident prediction model, the x and y coordinates for the sections of roads were compared with the cycle route plans to identify and map out the length of routes.

4.3.6 Average speed and road speed limit

Average vehicle speeds on locally managed A-roads by road name and direction of travel in England represented in miles per hour was obtained for the years 2008 to 2012 (Department for Transport, 2016). Designated road speed limits were also obtained from the STATS 19 road accidents data.

4.3.7 Model variables

In determining the variables to be used in the accident prediction model careful consideration was given to the ultimate purpose for which the accident model was being developed. Considering that the accident prediction model was to be used in the optimisation model in Chapter 6, the choice of variables was very important. An accident prediction model capable of predicting the frequency of accidents at the required level of severity was desired. In addition to this, in Chapter 2 it was essential to identify parameters that were readily available and shown in the literature to contribute to road traffic accidents. In road accident prediction models a relationship exists between the predictor variables and the dependent response variables. The predictor variables used were road direction (i.e. north-south or south north), number of lanes, slope, radius, AADT, HGV, speed limit and homogeneous segment length. These were used in predicting the dependent response variables which were the number of fatal and serious accidents combined as well as the number of slight accidents.

4.3.8 Data preparation

The data had to be in a format suitable for use in the model development. In order to achieve this all data obtained was checked to ensure it was the correct data required and this was carried out using ArcGIS which is a geographical information systems design and management software. First, EDINA digital maps comprising contour maps, vector maps, land and height data and OS MasterMap Integrated Transport Network (ITN) map of the UK road network was obtained for Nottinghamshire and Leicestershire. The OS (Ordnance Survey) VectorMap Local for Nottinghamshire and Leicestershire was also obtained from EDINA digimap. The purpose for downloading this data was to use the map as a backdrop map to other data to be subsequently added. An OS Terrain 50 DTM map of the area of interest was also obtained and superimposed on the OS VectorMap. This data was used for surface analysis of the roads to generate x and y coordinates and height values in the form of z values. Since the maps provided by EDINA contained all types of roads, there was the need to remove roads which were not of interest and only leave the roads of interest which were A-Roads.

Once the maps had been put together and checked against Google maps for any ambiguity, the road accident data was plotted onto the OS VectorMap. The STATS 19 road accidents data for the years 2008 to 2012 inclusive was used. Information in the form of speed limit, year and date of accident, number of casualties and accident severity level were available from the data. The easting and northing coordinates of road accidents for the roads of interest were plotted to identify the location along a given road segment. In plotting the road accident coordinates onto the map, it was found that some coordinates fell outside the road segment. This is because data errors may have arisen in both the road accidents data and the geographical maps since these data come from two different sources. In order to reduce any ambiguity in the plotted road accident data and to guarantee a degree of accuracy, the plotted map was compared to the already available mapped out road accidents from Nottinghamshire insight mapping (Nottingham City Council, 2013) and a crash mapping resource (Crashmap, 2014) and the results were found to tally for most cases.

Also, road accidents that fell outside the road segment lengths being considered were omitted from the analysis and these accounted for about 7% of the data. Figure 11 shows a map with road accidents along selected roads used shown.

Vehicle flow data obtained in the form of Annual Average Daily Traffic (AADT) (Department for Transport, 2013a) which had x and y coordinates in the form of easting and northing to British National Grid coordinate system was plotted onto the map. The AADT also had the direction of travel indicated so it was possible to assign the correct flow data to the correct direction of road travel. The length over which the AADT applies was also provided. It was assumed that the AADT remained the same for the segment of road starting from the count location to the next count location where it changed to a new value. Figure 12 shows a map with AADT count points for the selected roads shown. The roads data was separated by direction of travel. Data containing the x, y and z values of the roads was combined with the accident data and then with the AADT data.

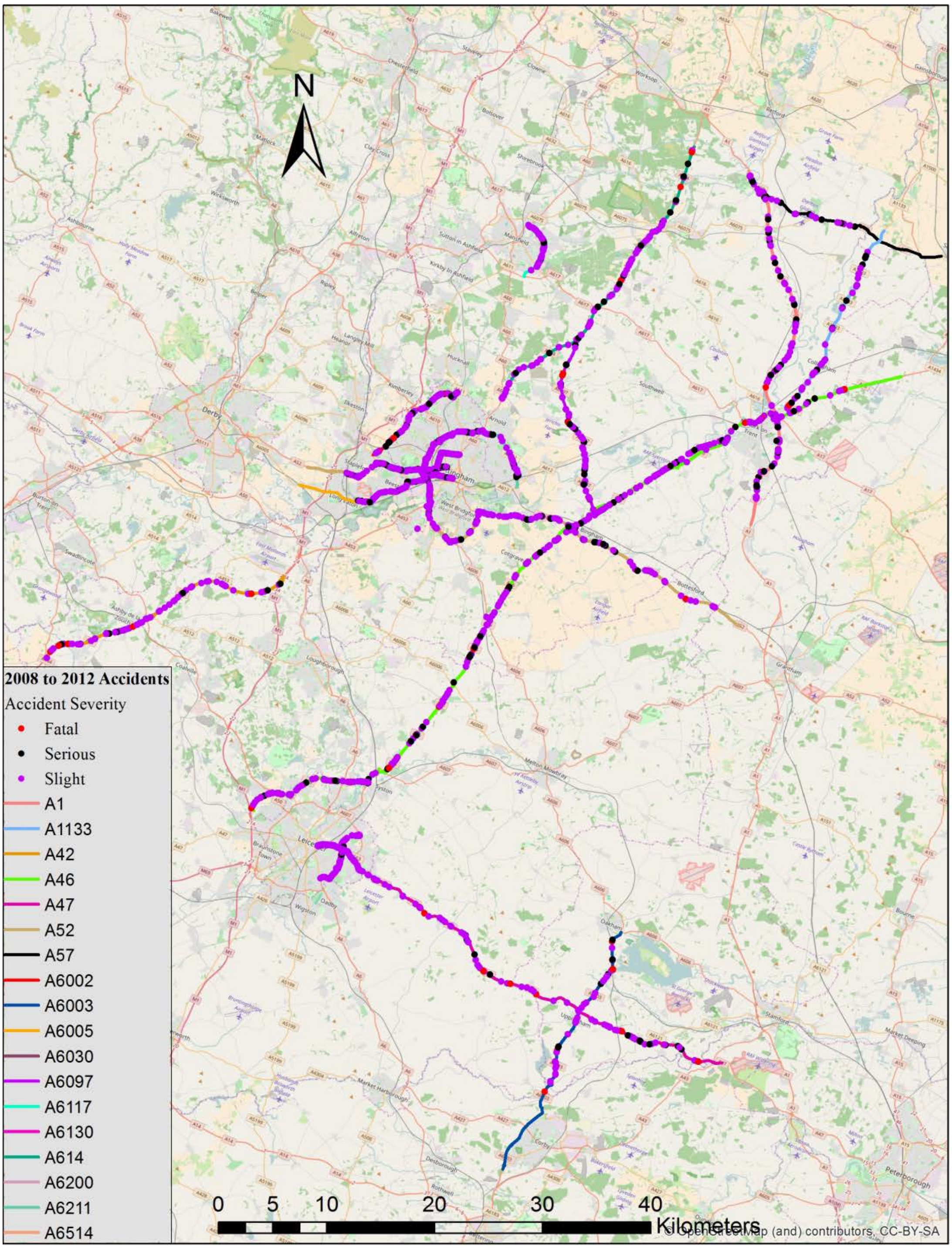


Figure 11 2008 to 2012 road accidents

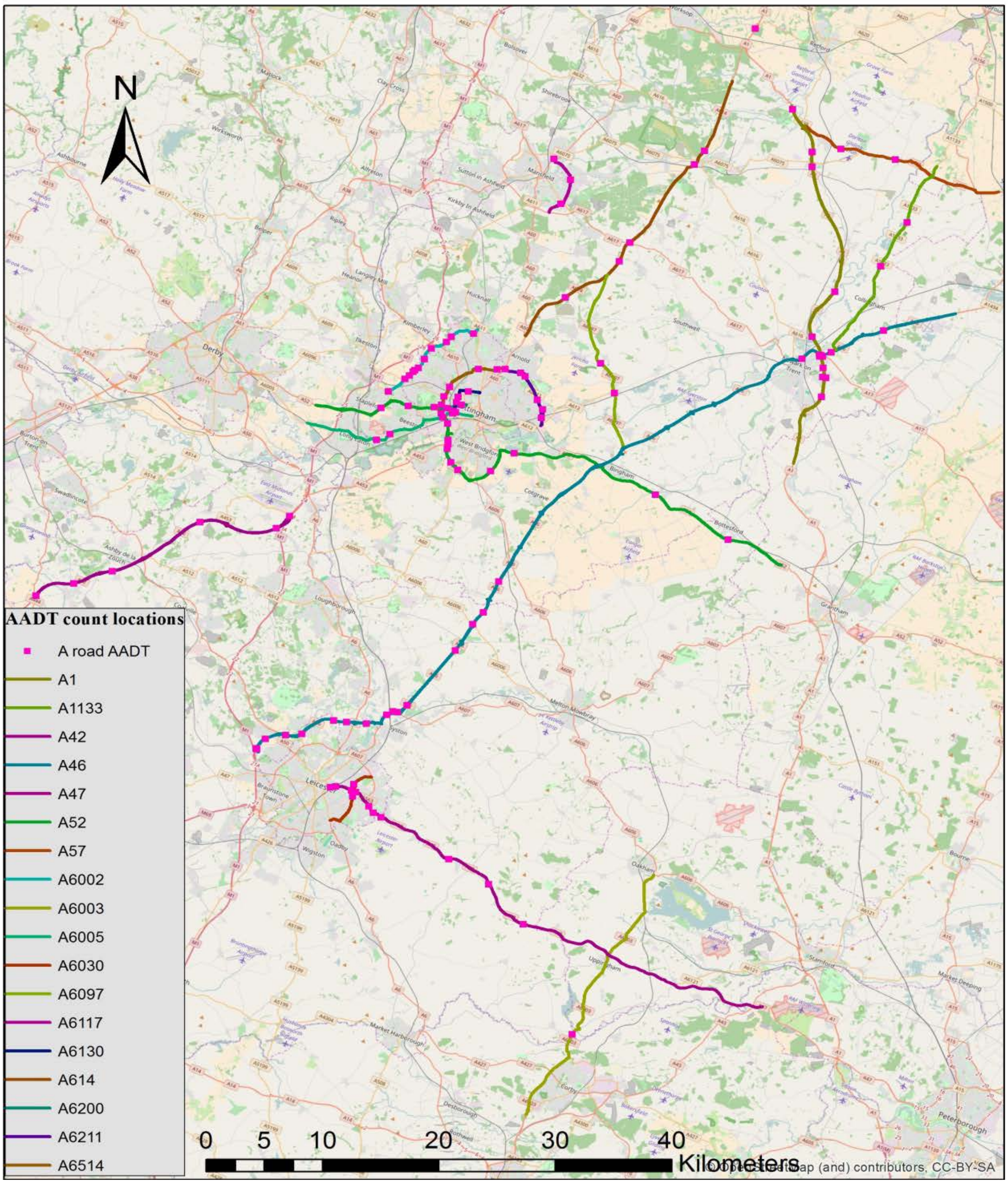


Figure 12 AADT count points

4.3.9 Site selection

The length of A-roads investigated within Nottinghamshire and Leicestershire regions in this research was approximately seven hundred and ninety kilometres long (790 km) for both directions of travel. These were randomly selected ensuring they had adequate data required for further analysis. To enable the model to be developed and tested, the dataset was split into two randomly selected road segments. The first set of 75 percent of the roads was only used for the model development and was not included in the testing of the model. The remaining 25 percent of the roads segments excluded from the model development was used to test the model. Figure 13 shows the proportion of roads used to develop and test the model.

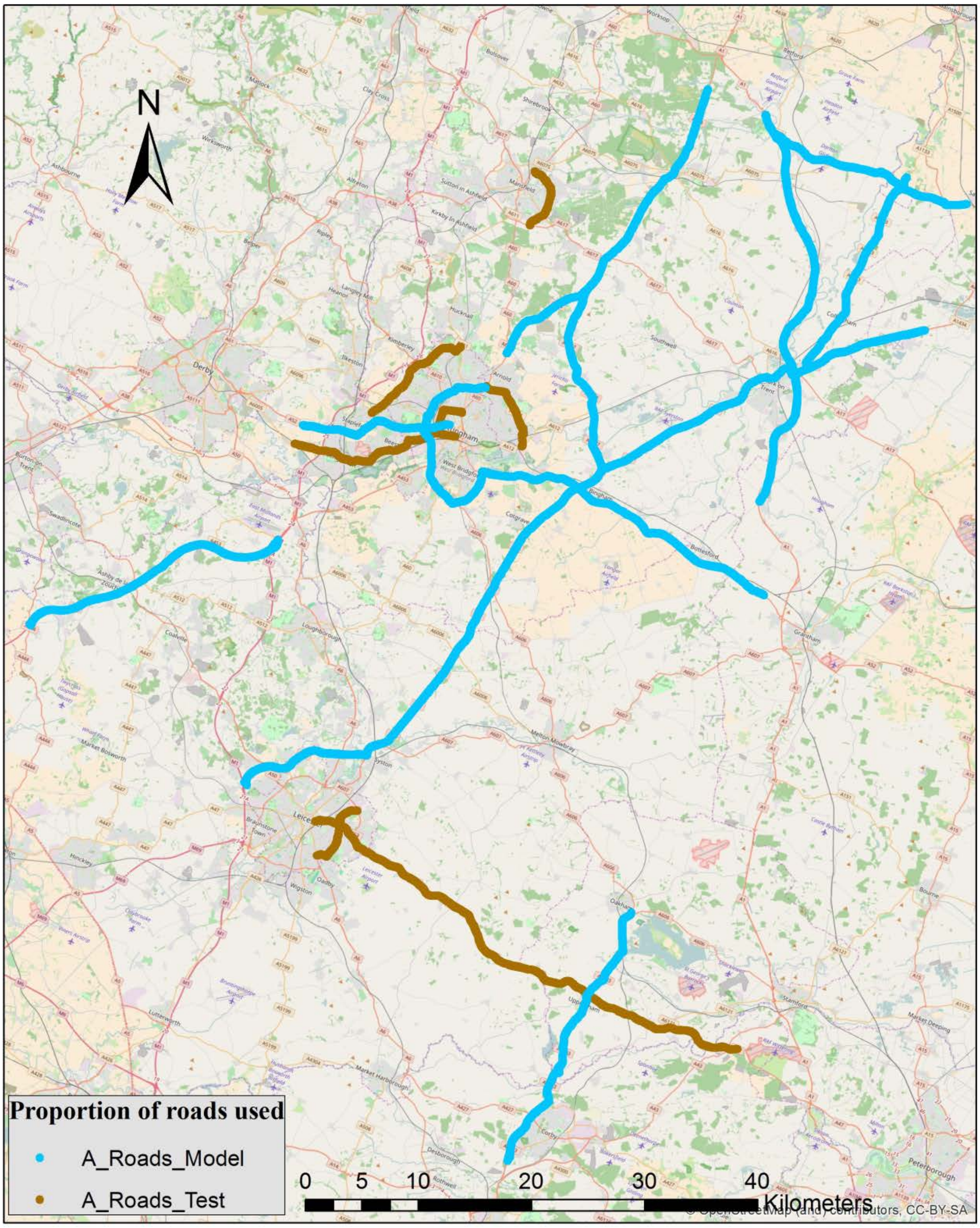


Figure 13 Roads selected in Nottinghamshire and Leicestershire

4.3.10 Model parameters

The parameters to be used in the negative binomial accident prediction model is presented in this section. The road direction, number of lanes, slope, radius, Annual Average Daily Traffic, Heavy Goods Vehicle, speed limit, homogeneous segment length, Average speed, Presence of cycle lane, length of cycle lane, presence of a junction or not, Number of junctions and number of pedestrian crossings were the main predictor variables used in the models. The various predictor variables used for developing the model are provided in Table 5

All the predictor variables given in Table 5 was used in developing the accident prediction models for fatal and serious accidents combined and a separate one for slight accidents. Predictor variables found not to be statistically significant and highly correlated in the models were removed. The models were refined by re-running them without those variables. After obtaining a suitable model with predictor variables found to be statistically significant and having no variables correlating with each other, the model was further refined using the empirical Bayes approach. This was done based on reasons given in Chapter 3 and this chapter.

Predictor variable	Symbol	Values taken	Description
Road direction [1, 2, 3, 4]	Dxn	1 = South-North 2 = North-South 3 = East-West 4 = West-East	This represents the direction of travel along the road i.e. north-south, south-north, east-west and west-east.
Number of lanes [1,2,3]	Lanes	1= single carriageway 2 = dual carriageway 3 = 3 lanes	This refers to the number of lanes per driving direction
Slope (%)	PercentSlope	-	This refers to the upward or downward longitudinal slope of the road
Radius (m)	LogRadius	-	This refers to the radius of the road segment. The logarithm of the radius was used.
AADT (veh/day)	LogAADT	-	This refers to the Annual Average Daily Traffic per driving direction. This also refers to the traffic flow and the logarithm of this variable was used.
HGV (%)	PercentHGV	-	This refers to the percentage composition of Heavy Goods Vehicle in the AADT per driving direction.
Speed limit (miles/hr)	Speed Limit	-	This refers to the speed limit of road section
Homogeneous segment length (m)	HSegLength	-	This refers to the length of homogeneous segment of road having all variables within the segment of road remaining constant.
Average Speed (miles/hr)	Avg.Speed	-	This is the average speed of vehicles along the road.
Presence of cycle lane. Yes No	PresofCR	0 - No 1 - Yes	This indicates the presence or absence of a cycle lane along the homogeneous segment considered.
Length of cycle lane (m)	CRteSUM	-	This refers to the length of cycle lane present along the homogeneous segment.
Presence of junction Yes No	Jtn	0 - No 1 - Yes	This refers to the presence of absence of junctions along the section of road considered.
Number of junctions	JtnsSum	-	This is the sum of junctions along the homogeneous segment of road considered.
Number of pedestrian crossings	PedCrossing	-	This is the sum of pedestrian crossing points along the homogeneous segment of road.

Table 5 Description of model variables

Table 6 and Table 7 provide information about the roads used.

Road Number	Number of junctions	Length of cycle lane (m)	Length of road (m)	Number of Fatal and Serious accidents	Number of slight accidents
A614	131	0	59629	29	131
A6097	132	0	33320	24	63
A6003	194	0	51192	15	38
A57	136	0	36520	8	23
A1133	130	0	38400	11	33
A6514	142	3960	15660	23	175
A6200	85	0	4720	24	82
A46	182	0	149860	81	387
A42	26	0	46782	17	71
A52	361	14280	104520	82	517
A1	60	0	64060	36	127
A6211	161	0	13840	10	44
A6130	166	0	7420	14	123
A6117	170	0	12400	17	71
A6030	155	0	12560	8	103
A6005	535	4740	30220	35	187
A6002	225	6380	20700	29	99
A47	460	0	87920	46	333

Table 6 Characteristics of A-roads investigated

Road Number	Number of lanes		Log AADT		Percent HGV		Speed limit		Average Speed		Log Radius		Slope	
	Min	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**	**Min**	**Max**
A614	1	1	3.71	4.07	3	8	30	60	39.2	40.8	1.17	6.54	-0.04	0.04
A6097	1	1	3.90	4.03	5	6	30	70	36.3	38.3	1.44	6.80	-0.04	0.04
A6003	1	1	3.50	3.56	6	6	30	60	45.9	49.6	1.14	8.09	-0.10	0.10
A57	2	2	3.53	4.06	7	10	30	60	40.1	42.1	1.34	6.02	-0.08	0.08
A1133	1	1	3.53	3.55	5	7	30	60	36.9	41.1	1.89	6.63	-0.02	0.02
A6514	2	2	4.17	4.52	2	4	30	40	16.1	18.5	1.25	7.39	-0.02	0.02
A6200	2	2	3.91	4.06	1	1	30	40	17.9	18.2	1.46	4.80	-0.06	0.06
A46	1	2	4.10	4.53	7	12	30	70	26.8	41.3	1.40	7.61	-0.03	0.04
A42	2	2	3.75	4.46	8	15	60	70	19.5	26.9	1.91	7.04	-0.04	0.04
A52	2	3	3.93	4.60	3	8	30	70	16.2	43.1	1.46	7.17	-0.07	0.06
A1	2	2	3.17	4.57	12	23	30	70	26.5	42.4	2.94	7.06	-0.04	0.04
A6211	1	1	3.62	4.02	1	3	30	40	20.4	21.1	1.21	6.45	-0.09	0.09
A6130	2	2	3.84	3.94	1	2	30	30	13.3	13.9	1.87	5.19	-0.02	0.02
A6117	2	2	3.63	3.86	2	3	30	40	22.0	23.2	1.23	5.41	-0.05	0.05
A6030	1	1	3.81	3.90	2	4	30	30	17.3	17.6	1.49	5.06	-0.05	0.05
A6005	2	2	3.78	4.29	1	2	30	40	18.6	19.9	1.22	6.19	-0.01	0.01
A6002	1	1	3.85	4.17	2	5	30	60	22.5	25.2	1.27	6.65	-0.03	0.03
A47	1	1	3.60	4.20	2	13	30	60	16.2	33.7	1.47	6.99	-0.07	0.07

Table 7 Further characteristics of A-roads investigated

4.3.11 Negative Binomial Empirical Bayes model

A Negative Binomial Generalised linear modelling approach was used to obtain regression approximations of predicted accidents for A-roads and an empirical Bayes approach used to refine the predictions.

Overdispersion arises when var(Yi) is greater than E(Yi) and for the Poisson distribution var(Yi) = E(Yi). The number of incidents Y for a given event may be Poisson with mean Z. The mean could be treated as a random variable which can be assumed to have a mean μ and index øμ in the population where E(Z) = μ and var(Z) = μ/ø resembling the Poisson distribution. A combination of these leads to the negative binomial distribution

$$\Pr(Y = y, \mu, \text{ø}) = \frac{\Gamma(y + \text{ø}\mu)\text{ø}^{\text{ø}\mu}}{y!\,\Gamma(\text{ø}\mu)(1+\text{ø})^{y+\text{ø}\mu}} \qquad y = 0, 1, 2, \ldots\ldots \qquad \ldots\ldots\text{Equation 14}$$

where E (Y) = μ and var (Y) = μ (1+ø)/ø are the mean and variance respectively.

The Negative Binomial regression model is of the form

$$\log\mu = \beta_0 + \beta_1\textbf{Dxn} + \beta_2\textbf{Lanes} + \beta_3\textbf{Slope} + \beta_4\textbf{Radius} + \beta_5\textbf{AADT} + \beta_6\textbf{HGV} + \beta_7\textbf{Speed Limit} + \beta_8\textbf{HSegLength} + \beta_9\textbf{AvgSpeed} + \beta_{10}\textbf{Jtns} + \beta_{11}\textbf{CycLength} + \beta_{12}\textbf{JtnsPresent} + \beta_{13}\textbf{PedCrossing} + \beta_{14}\textbf{PresofCR} \qquad \ldots\ldots\ldots\text{Equation 15}$$

The Empirical Bayes estimate of the total accidents in the before period M_B is computed by means of the weighted averages of the forecasts from the Negative Binomial model (μ_B) and the observed accidents (X_B) as given in equation 16

$$M_B = \alpha\mu_B + (1 - \alpha)X_B \qquad \ldots\ldots\ldots\text{Equation 16}$$

Using the forecasted accidents (μ_B) and the dispersion parameter (θ) of the predictive model given in equation 17, the weights (α) can be computed (Hirst et al., 2004).

$$\alpha = 1/(1 + \mu_B\theta) \qquad \ldots\ldots\ldots\text{Equation 17}$$

4.3.12 Software applications

Data used in the accident prediction model was prepared using ArcGIS, ArcMap Version 10.1 to enable visualisation of the data to be carried out and IBM SPSS

Statistics Version 22.0 was used to develop the model. Matlab Version R2014a was later used in the analysis of the optimisation model making use of genetic algorithms and pattern search optimisation methods.

4.4 Optimisation model Methodology

One of the most important and difficult decisions encountered by planners, engineers and designers is to identify the appropriate place to physically locate a road side speed control device in order to satisfy a set of objectives. The primary objective of the decision maker will be to reduce road accidents through the reduction of vehicle speeds. Speed cameras and vehicle activated signs as road side speed control devices are deployed at locations where speeding is a problem as well as where a specified number of road accidents are recorded. Engineers, planners and decision makers need to take a number of factors into consideration before deciding on where to place a speed camera or vehicle activated sign and these may involve satisfying multiple objectives. Some of the objectives are conflicting in nature requiring an approach such as the use of optimisation techniques that is best able to address such issues. The methodology used in developing the model is discussed in this section.

4.4.1 Roads Identification

To allow for convenience in data collection and site visits, Local Authorities that have already installed or intend installing vehicle activated signs and speed cameras in the geographical region close to Loughborough University were chosen. The Nottinghamshire and Leicestershire regions were selected for reasons given earlier in the chapter. For a site to be included in the research it had to meet some criteria as set out below;

- It must be an existing speed camera/VAS site provided by the responsible local authority.
- It must be a site proposed by the responsible local authority to have a speed camera/VAS installed in the future.
- The site must have available road accident data (STATS 19 road accident data).
- The site must have traffic flow data i.e. Annual Average Daily Traffic (AADT) or Annual Average Hourly Traffic (AAHT) to be converted to AADT or other.

Even though there will be many locations satisfying the criteria outlined above, other factors such as ease of obtaining the other data types may rule out the selection of some proposed sites. Sample sites were selected and visited by the researcher. The site visits were checked with data available from other sources such as Google earth maps. A site risk assessment sheet is provided in Appendix A.

4.4.2 Choice of optimisation model

In Chapter 2 a review of optimisation techniques was carried out. The choice of technique used will be restricted to genetic algorithms and pattern search for reasons given in section 2.4. Even though no literature has been found on the use of these techniques in locating speed cameras and vehicle activated signs, there is some literature on these techniques for locating ambulances, fire stations, schools and hospitals. These two techniques will be compared as they have proved useful in other research areas.

4.4.3 Identification of model objectives

As with any optimisation model there needs to be the identification of objectives to be achieved. A mathematical formulation of all objectives to be achieved in the model was done by expressing each of the objectives identified into a meaningful formula.

The objectives of the mathematical formulation should contribute to minimise road traffic crashes and the costs associated with deploying road side speed control devices. This section therefore focuses on developing a procedure to solve the problem of locating a road side speed control device. In summary, the main objective of this section is to create an initial search algorithm that incorporates engineering judgment. This is being done to help planners and engineers to view the whole process of locating a speed control device as practical and realistic as possible and eventually deliver good solutions.

This mathematical formulation should be inputted into the optimisation technique (genetic algorithms and pattern search) in addition to the constraints and processed for output results.

The objectives identified are to

- minimise the total set-up cost of a speed control device
- minimise the maintenance cost of a speed control device

- minimise the total lost cost of an accident

Detailed explanation and mathematical formulations for these objectives have been given below.

4.4.4 Minimise the total set-up cost of a speed control device

Minimise the total set up cost S_c of a speed control device (speed camera or VAS). This cost includes supply/purchase costs, installation costs and costs associated with relocation/diversion/bypassing of underground utilities. Every decision maker needs to establish the amount of budget available in order to know how much to invest in providing speed control devices. This objective is formulated as

$$Min\ f_1 = \sum_u \sum_v S_{uv} \times S_c \qquad \text{..........Equation 18}$$

S_{uv} is the decision variable such that if a speed device is set up on an x-y coordinate system (u, v), then $S_{uv} = 1$ if a speed device is set up otherwise $S_{uv} = 0$ and

$$\sum_u \sum_v S_{uv} \geq 1 \qquad \text{.........Equation 19}$$

4.4.5 Minimise the maintenance cost of a speed control device

Minimise the maintenance cost M_c of a speed control device (speed camera or VAS) given as

$$Min\ f_2 = \sum_u \sum_v S_{uv} \times M_c \qquad \text{..........Equation 20}$$

4.4.6 Minimise the total lost cost of the severity and frequency of an accident

Minimise the total lost cost of accidents within a length of road. This cost is dependent on the number of accidents occurring. It is assumed, few or no speed control devices will result in more road accidents within a given road section. In the absence of any speed control device, the total lost cost equates to T_{LC}. The T_{LC} can be reduced if more speed control devices are installed. However, since every decision maker has a budget to work to it is not possible to provide speed control devices beyond an optimal number. Also, beyond the optimal number of speed control devices to be provided, the set-up, operating and maintenance costs will increase. It is therefore necessary to establish a balance between set up costs, operating costs, maintenance costs and the lost costs of accidents. The expression for minimising the total lost cost of accidents is given as

$$Min\ f_3 = T_{LC} \times N \qquad \text{where } N = S_{uv} \qquad \text{...............Equation 21}$$

$$T_{LC} = exp(\beta_0 + \beta_1.\boldsymbol{Dxn_i} + \beta_2.\boldsymbol{Lanes_i} + \beta_3.\boldsymbol{Slope_i} + \beta_4.\boldsymbol{Radius_i} + \beta_5.\boldsymbol{AADT_i} + \beta_6.\boldsymbol{HGV_i} + \beta_7.\boldsymbol{SpeedLimit_i} + \beta_8.\boldsymbol{HSegLength_i} + \beta_9.\boldsymbol{AvgSpeed_i} + \beta_{10}.\boldsymbol{Jtns_i} + \beta_{11}.\boldsymbol{CycLength_i} + \beta_{12}.\boldsymbol{JtnsPresent_i} + \beta_{13}.\boldsymbol{PedCrossing_i} + \beta_{14}.\boldsymbol{PresofCR_i})$$

……………Equation 22

T_{LC} calculates the accident frequency on a link of road based on various parameters. In order to differentiate between the various kinds of accidents that occur on a road network, fatal, serious and slight accidents will be differentiated using the expression for T_{LC}.

$\boldsymbol{T_{LCfs}}$ is the total lost cost from fatal and serious accidents combined

$\boldsymbol{T_{LCsli}}$ is the total lost cost from slight accidents

Dxn is the direction of travel along a road eg. north-south

Lanes is the number of lanes per driving direction

Slope is the percentage of upwards and downwards longitudinal gradient (slope) of the road for the direction of travel being considered.

Radius is the radius of curve

AADT is the annual average daily traffic for the direction of travel being considered.

HGV is the proportion of heavy goods vehicles travelling in the section of road under consideration.

SpeedLimit is the speed limit of the road in miles/h.

HSegLength refers to the length of homogeneous segment of road being considered.

AvgSpeed is the average speed of the road

Jtns refers to the number of junctions in a homogeneous segment

CycLength refers to the cycle path length

JtnsPresent refers to the presence or absence of a junction within the homogeneous segment

PedCrossing refers to the presence or absence of a pedestrian crossing within the homogeneous segment

PresofCR refers to the presence or absence of a cycle route within the homogeneous segment

4.4.7 The objective function

The objective function relates to the accident prediction model developed in Chapter 5. This accident prediction model contains parameters such as direction of travel, number of lanes, slope, radius, AADT, percentage HGV, speed limit of the road, Average speed, number of junctions, length of cycle route and homogeneous segment length.

This study assumes that once a speed camera etc is installed, it will serve the purpose of reducing road traffic accidents within a given road link through the reduction of vehicle speed. In developing a model for locating speed control devices ie. speed cameras and vehicle activated signs (VAS), there are a number of objectives that must be satisfied.

The main objective in this research for the decision maker (DM) is to optimise the location of a speed control device by minimising the total number of accidents occurring within a road link. This is because the other objects indirectly are incorporated into the objective for minimising the number of accidents. The set up cost and maintenance costs of a speed control device do not affect the location of the speed camera as is proposed in optimising the objective of minimising accidents. Also by minimising the total number of accidents, the optimisation model is being developed to provide the best ‘x’ number of locations to place a speed control device. Depending on the budgetary allocation for the responsible authority, the required number of speed control devices can be mounted. In achieving this objective, a number of sub-objectives must be satisfied to achieve the main objective. The approach adopted in this research is similar to that of Tzeng and Chen (1999) but revised in terms of the input parameters and constraints imposed on the objective function to suit the problem at stake.

The objective function aims to minimise the functions f_1 to f_3. The level of importance assigned to each goal will have to reflect the criteria for decision making by the designated decision makers (mostly local authorities and the police in this case). The sum of objectives is given as

Minimise $Z = f_1 + f_2 + f_3$ ……………Equation 23

Minimise $Z = (\sum_u \sum_v S_{uv} \times S_c) + (\sum_u \sum_v S_{uv} \times M_c) + \sum_u \sum_v S_{uv} \times$

$$exp(\beta_0 + \beta_1.\boldsymbol{Dxn}_i + \beta_2.\boldsymbol{Lanes}_i + \beta_3.\boldsymbol{Slope}_i + \beta_4.\boldsymbol{Radius}_i + \beta_5.\boldsymbol{AADT}_i + \beta_6.\boldsymbol{HGV}_i + \beta_7.\boldsymbol{SpeedLimit}_i + \beta_8.\boldsymbol{HSegLength}_i + \beta_9.\boldsymbol{AvgSpeed}_i + \beta_{10}.\boldsymbol{Jtns}_i + \beta_{11}.\boldsymbol{CycLength}_i)$$

……….Equation 24

The optimum number of speed control devices will be determined from the third objective function which is to minimise the cost of road traffic accidents. This objective has been chosen to provide a starting point and sound justification for reducing road traffic accidents and more especially to reduce vehicle speeds. Cost has been identified to be a parameter that will not directly affect the location of the speed control device but will rather contribute to determine how many speed control devices can be mounted. The incorporation of the cost function into the model has been taken into account by setting up the model to provide the best ‘x’ number of locations to place the speed control device. The objective function can thus be represented in the following equation.

$$\text{Minimise} \quad Z = \sum_u \sum_v S_{uv} \times exp(\beta_0 + \beta_1.\boldsymbol{Dxn}_i + \beta_2.\boldsymbol{Lanes}_i + \beta_3.\boldsymbol{Slope}_i + \beta_4.\boldsymbol{Radius}_i + \beta_5.\boldsymbol{AADT}_i + \beta_6.\boldsymbol{HGV}_i + \beta_7.\boldsymbol{SpeedLimit}_i + \beta_8.\boldsymbol{HSegLength}_i + \beta_9.\boldsymbol{AvgSpeed}_i + \beta_{10}.\boldsymbol{Jtns}_i + \beta_{11}.\boldsymbol{CycLength}_i)$$

..…..Equation 25

4.4.8 Model constraints

The Highways Agency (2002) Design Manual for roads and bridges guidance to speed limits and their geometric parameter requirements was used as a guide to outlining the constraints for the objective function. Other constraints have also been included. These constraints have all been formulated mathematically.

The constraints to be considered in addition to the above mentioned objectives are set out as follows;

Firstly, there should be at least one speed control device (speed camera or VAS) along the length of road under consideration given as

$$N = \sum_u \sum_v S_{uv} \geq 1$$

…………..Equation 26

Where N is the desired number of speed control devices.

Secondly, a speed control device (speed camera or VAS) should not be located within obstacles such as driveway, property entrances, on underground utilities, water-body, on pedestrian walkway such that pedestrian movements will be disrupted or be less safe etc. This constraint is given by the following equation

$$S_{uv} = 0 \ \forall \ (u, v) \in \Psi$$

…………..Equation 27

Ψ belongs to the set of all obstacle coordinates to be avoided in order to identify a suitable location to place a speed control device (speed camera or VAS). In the unlikely event that a speed control device is located in the path of an obstacle, another solution will be randomly generated and checked to ensure it is out of an obstacle location and at a preferred location.

Thirdly, if more than one speed control device (speed camera or VAS) is to be used, an appropriate distance d_{12}^{a} must be between speed device 1 and 2. The distance between device 1 and 2 should be such that the distance d_{12}^{l} is not too long so that the effects desired is not achieved nor too short a distance d_{12}^{s} to cause an overlap of desired effects of devices. Let $\{S_{uv}\}$ denote a set of speed control devices to be set up with 1 and 2 taken from S_{uv} . Tzeng and Chen (1999) used 4 different inequality expressions containing d_{ab}, d_{ab}^{r}, d_{ab}^{l} and d_{ab}^{s} to denote the fuzzy constraints. To simplify their equations, the following expression is used to represent the constraint.

$$d_{12}^{s} \leq |u_1 - u_2| + |v_1 - v_2| \leq d_{12}^{l} \qquad \text{.............Equation 28}$$

Where *u* and *v* values are expressed in the x-y coordinate system for location 1 and 2.

Fourthly, for a given speed limit of road the Highways Agency (2002) requires certain radius and sight distances to be achieved. Sight distance measurements are not easy to obtain so the radius requirements have been used as a surrogate for achieving sight distance. Also for certain speed limits and number of lanes, there is a requirement for certain gradients to be achieved. The various geometric requirements have been mathematically formulated and used as constraints to the objective function. These constraints are given as follows;

Sum of fatal and serious accidents; sumFaSerAccd>=3
Sum of Slight accidents; sumSliAccd >=15
Summation of distance over which optimisation is executed; sumDIST >=1000m; sumDIST<=3000

Lanes >=1, Lanes <=2, Speed Limit >=20, Speed Limit <=30, Slope >=-0.06, Slope <=0.06, Radius >=360

Lanes >=1, Lanes <=2, Speed Limit >=20, Speed Limit <=30, Slope <= -0.005, Slope>= -0.06, Radius>=360

Lanes>=1, Lanes <=2, Speed Limit >30, Speed Limit <=40, Slope >=-0.06, Slope <=0.06, Radius >=510

Lanes >=1, Lanes <=2, Speed Limit >30, Speed Limit <=40, Slope <= -0.005, Slope >= -0.06, Radius >=510

Lanes >=1, Lanes <=2, Speed Limit >40, Speed Limit <=50, Slope >=-0.06, Slope <=0.06, Radius >=1020

Lanes >=1, Lanes <=2, Speed Limit >40, Speed Limit <=50, Slope <= -0.005, Slope >= -0.06, Radius >=1020

Lanes >=1, Lanes <=3, Speed Limit >50, Speed Limit <=60, Slope >=-0.06, Slope <=0.06, Radius >=1440

Lanes >=1, Lanes <=3, Speed Limit >50, Speed Limit <=60, Slope <= -0.005, Slope >= -0.06, Radius >=1440

Lanes >=1, Lanes<=4, Speed Limit >60, Speed Limit<=70, Slope >=-0.06, Slope <=0.06, Radius >=2040

Lanes>=1, Lanes <=4, Speed Limit >60, Speed Limit <=70, Slope <= -0.005, Slope >= -0.06, Radius >=2040

Lanes >=2, Lanes <=4, Speed Limit >50, Speed Limit <=60, Slope >=-0.04, Slope <=0.04, Radius>=1440

Lanes >=2, Lanes<=4, Speed Limit >50, Speed Limit <=60, Slope <= -0.005, Slope >= -0.04, Radius >=1440

Lanes >=2, Lanes <=4, Speed Limit >60, Speed Limit <=70. Slope >=-0.04, Slope <=0.04, Radius >=2040

Lanes >=2, Lanes <=4, Speed Limit >60, Speed Limit<=70, Slope <= -0.005, Slope >= -0.04, Radius >=2040

The optimisation code used in running the genetic algorithm and pattern search is given in Appendix B and Appendix C.

4.5 Application of optimisation model to selected roads

Elvik, Christensen and Amundsen (2004) in their evaluation of the Power Model stated that "if government wants to develop a road transport system in which nobody is killed or permanently injured, speed is the most important factor to regulate" and that speed limit and their enforcement are very important road safety measures.

A-roads used in the optimisation were initially split into regular 20m intervals from beginning to the end of road. The interval was chosen so as not to omit any detail in parameters along the segment of road. The easting and northings extracted from the split points along the road were used in the optimisation. Road sections for optimisation modelling were considered for lengths greater than 1000m but less than or equal to 3000m. This length was considered appropriate because for shorter lengths of roads, accident numbers were almost negligible. Another reason for choosing this length of road was because the rules and guidance for the national safety camera programme as stated in the Department for Transport (2006) provides suggestions for route lengths that need to be considered for fixed speed camera sites. In addition to these reasons, other studies on the effectiveness of speed cameras have identified distances within this range as being effective for managing speeds (Retting, Kryychenko and McCartt, 2008; Høye, 2015; Høye, 2015a).

The roads are split into homogeneous segments in which all parameters or variables within that homogeneous segment remains the same. Once a parameter changes a new homogeneous road segment begins. Accidents are assumed to remain constant over the homogeneous segment of road. The north-south direction of travel was indicated as NS, the south-north direction of travel along the road was represented as SN, the east-west direction of travel along the road was represented as EW and the west-east direction of travel represented as WE. The objective which minimises the costs associated with road traffic accidents was used to optimise the model. The optimisation model was executed such that for road segments from 1000m to 3000m in length the sum of fatal and serious accidents occurring over that stretch of road must be greater than or equal to 3 or 15 for slight accidents. Genetic Algorithms and Pattern Search were the two main optimisation techniques used in this research. These were chosen based on an initial investigation into types of optimisation techniques applied to facility location problems. These forms of optimisation techniques have been widely used in other areas of research where they have been found to adequately optimise parameters dealing with large data sets. Even though no optimisation method was identified to have been applied to a problem similar to that being investigated in this research, other areas of facility location problems were identified. Both pattern search and genetic algorithms belong to the same family of global optimisation. The choice of these two optimisation methods in this research allows for meaningful comparisons to be made. Pattern search and genetic algorithms allow customisations to the algorithms to be made by modifying options.

Unlike other optimisation methods, genetic algorithm offers the opportunity to vary options such as the crossover and mutation in order to obtain better results. This added advantage of providing variety to the search in genetic algorithms helps prevent early convergence (Salhi and Gamal, 2003). Convergence occurs when parents are unable to produce offspring who are of better quality than the parents. Also genetic algorithm was chosen because it has been found to have been applied to a lot of real life problems in facility location. (Cheung, Langevin and Villeneuve, 2001; Indriasari et al., 2010).

For genetic algorithms, the optimisation was run for 100 generations in most instances with a few run for 200 generation using a population size of 20. A crossover fraction of 0.8, a Gaussian mutation and adaptive feasible mutation was used as they are flexible and support both fine tuning of solutions and searching the domain (Heitzinger, 2002).

Road segments were considered such that they were assessed at lengths starting from the first location along the road of interest to the desired length of 3000m. The next length started at 20m away from the previous start point to the desired length of 3000m of the road. This explanation is illustrated in Figure 14.

Road number is A600NS and A600SN where NS represents North-South and SN represents South-North

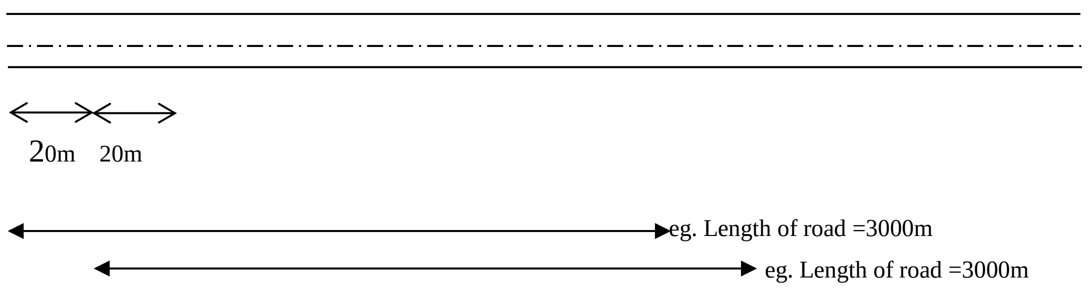


Figure 14 Sample road segments considered for optimisation

4.6 Data validation

As mentioned earlier in this chapter the main source of data was the national STATS 19 database for road accident data with vehicle flow data being obtained from the responsible Local Authorities, the Police and the Department for Transport.

The STATS 19 road accident data is the most complete, detailed and reliable single source of data that has been compiled from police completed records of injury accidents that have occurred in Great Britain. It must be borne in mind that despite every effort made to ensure that records are up to date, not all injury accidents are recorded by the police. This is because the police are not always called to the scene of an accident. However, when road accidents are compared with death registrations, very few road accident fatalities are not reported to the police (Department for Transport, 2012a; Department for Transport, 2010)). A comparison of fatalities recorded in STATS 19 and the national death registration due to land transport accident is shown in Figure 15.

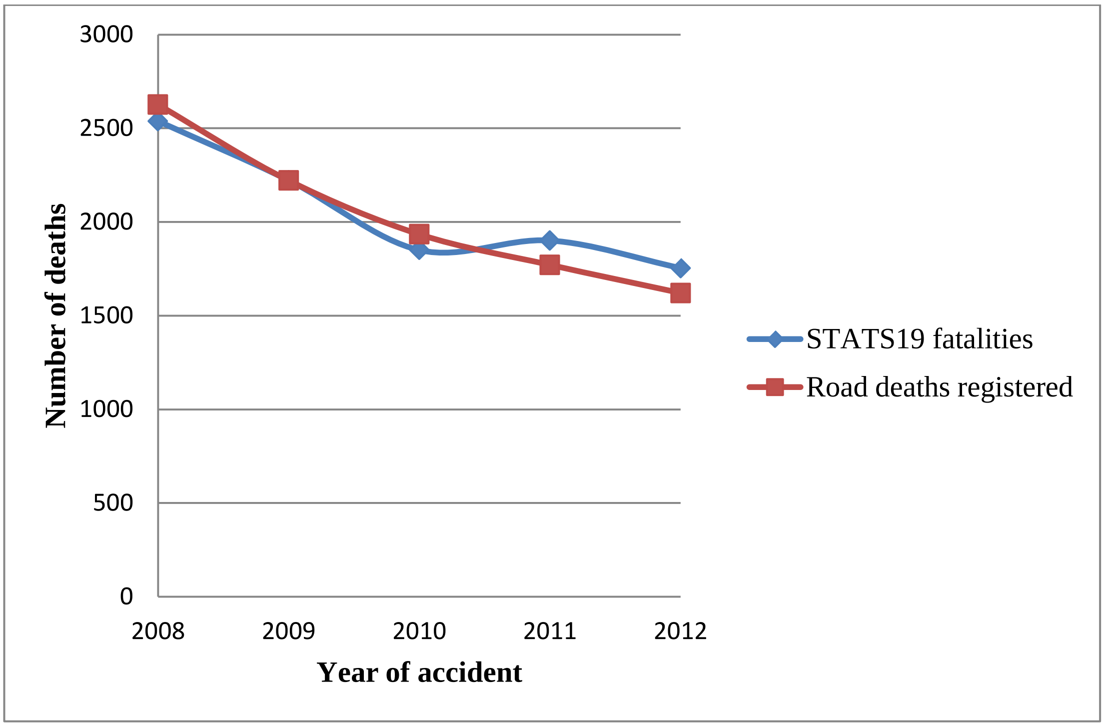


Figure 15 Comparison of STATS19 fatalities with death registration from the Office of National Statistics (2008-2012)

Differences are however observed for serious injury accidents when STATS 19 is compared with hospital admissions data from the Hospital Episode Statistics (HES) (Department for Transport, 2010; Department for Transport, 2011) suggesting there is an underreporting of road casualties in Great Britain. In 2011, there were 38.6 thousand recorded emergency admissions to hospitals in England due to road traffic accidents in comparison with 20.1 thousand serious injuries reported in STATS 19. In 2010 there were approximately 36 thousand recorded emergency admissions to hospital in England as compared to 20 thousand serious injuries recorded in STATS19. This clearly shows a gap existing in non-fatal casualties' data recorded by the police. Despite the shortfall in data from STATS19, it is still the official reliable data source for road traffic accidents in England.

In geographically mapping the various road accident severity levels to road segments, use is made of other available resources such as Crashmap.co.uk and Nottinghamshire insight mapping (Crashmap, 2014; Nottingham City Council, 2013). These resources are available for comparing with the plotted easting and northing coordinates provided in STATS 19.

For the road characteristics data, EDINA digital maps data are compared with Google maps. In checking the level of accuracy of these maps, the lengths of road segments were measured in both data sources with an average difference between the two data sources being 0.036km. The difference in values of road segments measured from the two sources was assessed using the statistical t-test and results indicate statistical insignificance at a confidence level of 95%. The use of EDINA maps, Google maps and site visits to roads in close proximity to the researcher were carried out to validate road characteristics data such as number of lanes, location of junctions and location of existing signs obtained from EDINA and Google maps. For A-roads, the vehicle speed data obtained from the local authorities was checked against the Department of Transport's online traffic count data (Department for Transport, 2013a). For non-A-roads data available from data.gov.uk is used.

5 Results (I): Accident prediction model

5.1 Results and discussion

In this chapter, an Empirical Bayes Negative Binomial Model for accident prediction was developed to establish a relationship between road traffic accidents and identified factors along selected A-roads in the United Kingdom. Results and a discussion of findings from the models are presented in this Chapter. Considering that fatal accidents do not occur that often, the frequency of these is minimal in the dataset. Modelling only fatal accidents would have resulted in the model not being significantly represented. In order to avoid this problem, *fatal and serious* accidents were combined and modelled together. Table 8 gives results of the significance of the parameter estimates for all the parameters initially used in the negative binomial model at 5 percent significance level. Parameters identified to be non-significant in the model were removed from the model and rerun.

	Fatal and Serious accidents model	Slight Accidents model
Parameter	Sig	Sig
(Intercept)	0.000	0.000
Jtn	0.000	0.000
PresofCR	0.307	0.052
Dxn	0.000	0.000
PedCrossing	0.487	0.546
JtnsSUM	0.294	0.002
Lanes	0.282	0.021
PercentHGV	0.540	0.044
Speed Limit	0.015	0.000
HSegLength	0.000	0.000
LogAADT	0.002	0.000
LogRadius	0.000	0.000
CRteSUM	0.596	0.653
Avg. Speed	0.277	0.746
Percent Slope	0.078	0.075

Table 8 Parameter estimates for initial models

5.2 Negative Binomial model results

Results of the continuous variables obtained from the rerun Negative binomial model (for both fatal and serious accidents and slight accidents) are given in Table 9 and Table 10. The minimum value, maximum value, mean and standard deviation of the variables are provided in the Tables. In Table 11, information on categorical variables used in the refined accident prediction models are shown.

The parameter estimates are given in Table 12 and Table 13 and a level of significance of 5% (< 0.05) is used in reporting the results.

		N	Minimum	Maximum	Mean	Std. Deviation
Dependent Variable	FaSer	4177	0	5	0.08	0.34
Covariate	Speed limit	4177	30	70	57.97	12.09
	HSegLength	4177	40	1220	144.73	110.33
	LogAADT	4177	3.17	4.60	4.01	0.32
	LogRadius	4177	1.14	8.09	4.24	1.26

Table 9 Continuous variable information of Fatal and Serious (FaSer) accidents

		N	Minimum	Maximum	Mean	Std. Deviation
Dependent Variable	Slight	4177	0	5	0.08	0.34
Covariate	Jtns	4177	0	7	0.38	0.75
	Lanes	4177	1	3	1.67	0.50
	Percent HGV	4177	1	23	8.21	4.44
	Speed Limit	4177	30	70	57.97	12.09
	HSegLength	4177	40	1220	144.71	110.33
	LogAADT	4177	3.17	4.60	4.01	0.32
	LogRadius	4177	1.14	8.09	4.24	1.26

Table 10 Continuous variable information of Slight accidents

Model	Variable	Category	N	Percent
Fatal and Serious	Jtn	0	3111	74.5
		1	1066	25.5
	Dxn	1	1579	37.8
		2	1553	37.2
		3	512	12.3
		4	533	12.8
Slight	Jtn	0	3110	74.5
		1	1066	25.5
	PresofCR	0	4015	96.1
		1	161	3.9
	Dxn	1	1578	37.8
		2	1553	37.2
		3	512	12.3
		4	533	12.8

Table 11 Categorical variable information of Fatal and Serious accidents and Slight accidents

Parameter	B(Estimate)	Std. Error	Sig.	Exp(B) (Odds Ratio)	95% Wald Confidence Interval	
					Lower	Upper
Intercept	-6.434	0.870	0.000	0.002	-8.140	-4.728
Jtns=0	-0.944	0.124	0.000	0.389	-1.188	-0.701
Jtns=1	0	-	-	1	-	-
Dxn=1	0.103	0.168	0.000	1.109	0.225	0.432
Dxn =2	-0.398	0.181	0.000	0.672	-0.752	-0.044
Dxn =3	-0.867	0.238	0.000	0.420	-1.331	-0.401
Dxn =4	0	-	-	1	-	-
Speed limit	-0.020	0.005	0.000	0.980	-0.029	-0.011
HSegLength	0.003	0.0004	0.000	1.003	0.002	0.003
LogAADT	1.511	0.206	0.000	4.530	1.107	1.914
LogRadius	-0.178	0.050	0.000	0.837	-0.277	-0.079
(Scale)	1					

Table 12 Parameter estimates for fatal and serious accidents on A-roads

Parameter	B(Estimate)	Std. Error	Sig.	Exp(B) (Odds Ratio)	95% Wald Confidence Interval	
					Lower	Upper
Intercept	-6.957	0.609	0.000	0.001	-8.150	-5.767
Jtns=0	-0.566	0.114	0.000	0.568	-0.790	-0.343
Jtns=1	0	-	-	1	-	-
PresofCR =0	-0.495	0.128	0.000	0.609	-0.747	-0.244
PresofCR =1	0	-	-	1	-	-
Dxn=1	0.204	0.107	0.000	1.226	0.006	0.413
Dxn =2	0.043	0.111	0.000	1.044	0.174	0.260
Dxn =3	-0.357	0.121	0.000	0.700	-0.593	-0.120
Dxn =4	0	-	-	1	-	-
JtnsSUM	0.189	0.058	0.001	1.208	0.076	0.302
Lanes	0.238	0.104	0.022	1.269	0.034	0.442
PercentHGV	-0.021	0.009	0.028	0.980	-0.039	-0.002
Speed limit	-0.022	0.003	0.000	0.978	-0.028	-0.017
HSegLength	0.002	0.0003	0.000	1.002	0.001	0.003
LogAADT	2.004	0.158	0.000	7.417	1.694	2.313
LogRadius	-0.212	0.027	0.000	0.809	-0.266	-0.158
(Scale)	1					

Table 13 Parameter estimates for slight accidents on A-road

The Odds Ratio is a relative measure of effect and enables the intervention group to be compared with the non-intervention group. It can also be explained to be a measure of the relationship between an exposure and an outcome. From Table 12 it can be interpreted that the odds of an accident occurring along a road without the presence of a junction (Jtns = 0) is 0.389 times that of a road with junctions (Jtns = 1). In Table 13, for the same parameter it can be interpreted that the odds of an accident occurring along a road without the presence of junctions (Jtns = 0) is 0.568 times that of a road with junctions (Jtns = 1). It can also be seen from Table 13 that the odds of an accident occurring along a road without a cycle route (PresofCR = 0) is 0.609 times that along a road with a cycle route (PresofCR = 1).

In Table 12 and Table 13 the 95 percent Wald Confidence Interval for the parameters shows the lower and upper limits. Values obtained for the lower and upper limits of the confidence interval suggest the Negative Binomial is more suitable for modelling the parameters as compared to the Poisson model since the limits predominantly do not include zero.

Information about the Goodness of Fit values for the Negative binomial models generated for *fatal and serious* accidents and *slight* accidents is provided in Table 14. The Value/df figure for the Pearson Chi-Square for both the *fatal and serious* accidents model (1.116) and the *slight* accidents model (2.385) are greater than 0.05 signifying that the model fits the data well. Also, the Value/df figures generated for the deviance is found to be far less than 1 for the *fatal and serious* accidents model (0.334) and closer to 1 for the *slight accidents* model (0.757) also indicating that the *slight accidents* model fits the data well.

Fatal and serious accidents model			
	Value	Df	Value/df
Deviance	1391.771	4168	0.334
Pearson Chi-Square	4652.459	4168	1.116
Slight accidents model			
	Value	Df	Value/df
Deviance	3151.852	4163	0.757
Pearson Chi-Square	9927.312	4163	2.385

Table 14 Goodness of Fit for fatal and serious accidents and slight accidents

The Omnibus Test results given in Table 15 for the *fatal and serious* accidents model and the *slight accidents* model indicate the overall models are statistically significant.

Model	Likelihood ratio Chi-square	df	Sig
Fatal and Serious accidents model	218.553	8	0.000
Slight accidents model	1002.145	12	0.000

Table 15 Omnibus Test

Presence of Junctions (Jtn)*(Table 12 and Table 13)****:*** Junctions along roads are noted in literature to influence road accident numbers and are also identified as particularly high risk locations (Clarke et. al.,1998; Golias, 1992; Mountain et al, 1996; Vorko-Jović, 2006; Greibe, 2003). The model developed showed a negative relationship between the absence of junctions and *fatal and serious* accidents. A similar observation was made for *slight* accidents with a negative relationship existing between the absence of junctions and *slight* accidents. At significance level less than 0.05 there was evidence provided by the data that the presence or absence of junctions has an effect on the number of *fatal and serious* accidents (p = 0.000) as well as *slight* accidents (p = 0.000).

Cycle Route (PresofCR) *(Table 12 and Table 13)****:*** The effect of the presence of a cycle route on *fatal and serious* accidents was found to be statistically insignificant so this parameter was removed from the model. A negative relationship however existed between the presence of a cycle route and *slight* accidents. At significance level less than 0.05 there was evidence (p = 0.000) provided by the data that the presence of a cycle route has an effect on the number of *slight* accidents. Despite this finding, the effect of cycle paths on cycle safety is unclear from published literature with the comparison of individual cycle path evaluations revealing a broad range of outcome on road traffic accidents (Phillips et. al., 2011). A meta-analysis of 14 European studies revealed that the introduction of cycle paths do not result in a remarkable overall change in cycle accident numbers (Elvik et al., 2009). The introduction of cycle paths in Sweden was linked with changes in the number of road traffic accidents varying from a 44% reduction to an 82% increase as reported in Phillips et. al (2011). A similar broad ranged variation has been published in other countries with much uncertainty about the effect of cycle paths on road traffic accidents (Kim et al., 2007; Forester, 2001; Pucher, 2001). A likely explanation for the outcome observed to date is that the success of cycle paths relies on the design and the precise context for introducing them (Phillips et. al., 2011).

A great proportion of car-bicycle accidents comprising cyclists who come from a direction inconsistent with the normal car traffic flow were noted in road traffic accidents involving cyclists (Koustanaï et al., 2008). However the geographic location where data about the accidents were obtained had very few cycle tracks.

Direction (Dxn) *(Table 12 and Table 13)**:*** For road direction 1, a positive relationship existed between the direction of travel and *fatal and serious* accidents as well as *slight* accidents. For road direction 2 a negative relationship existed between the direction of travel and *fatal and serious* accidents with a positive relationship existing between the direction of travel and *slight* accidents. For road direction 3, a negative relationship existed between the direction of travel and *fatal and serious* accidents as well as *slight* accidents.

At significance level less than 0.05 there was evidence for *fatal and serious* accidents (p = 0.000) and for *slight* accidents (p = 0.000) provided by the data that the direction of travel has an effect on the number of *fatal and serious* accidents as well as *slight* accidents.

Li, Zhu and Sui (2007) recommended differentiating directions of travel in order to assess relative crash risks since more than thirty percent of roadways investigated in their study showed statistically significant different risk values for different directions. This is because different directions of the roadway may have dissimilar risk values resulting from the characteristics of traffic (such as traffic volume), road conditions (such as road geometry), environmental conditions (such as lighting conditions, presence of hard shoulder) and other risk indicating variables. Based on a two-sample t-test, more than thirty percent of roads investigated had statistically significant different risk values for different directions (Li, Zhu and Sui, 2007). The need to distinguish directions in the analysis of relative risk can contribute to the prevention of averaging outcomes between opposite travel directions and thus produce more valid outcomes for each direction of travel.

Number of Junctions (JtnSUM) *(Table 12 and Table 13)****:*** The effect of the number of junctions on *fatal and serious* accidents was found to be statistically insignificant and this parameter was removed from the model. Results however show that the number of junctions along a segment of road has a positive effect on *slight* accidents. There was evidence provided by the data that at significance level less than 0.05 the number of junctions has a strong effect on the number of *slight* accidents ($p = 0.000$).

Lanes *(Table 12 and Table 13)****:*** Results show that the effect of the number of lanes on *fatal and serious* accidents was statistically insignificant and was thus excluded from the model. However, the number of lanes was found to have a positive effect on *slight* accidents. There was evidence provided by the data that at significance level less than 0.05 the number of lanes has an effect on the number of *slight* accidents ($p = 0.022$). This finding is consistent with the results of a study carried out by Wang, Quddus and Ison, (2011a) on major roads in the UK where the number of lanes was statistically significant and positively associated with slight injury accidents. This indicates that more slight injury accidents are likely to arise on roads with more lanes.

A number of researchers are of the view that an increase in accident rates occurs with an increase in the number of lanes (Kononov, Bailey and Allery, 2008) with other studies indicating dissimilar views. Kononov, Bailey and Allery (2008) found a 25 percent increase in accident rates between four and six-lane freeways and 40 percent between six and eight-lane freeways in Colorado. By comparing slopes of safety performance functions (SPFs) for different number of lanes in California, Colorado and Texas, it was observed that increasing the number of lanes on urban freeways initially contributes to improve safety that declines with an increase in congestion. A possible explanation to this is that with an increase in the number of lanes, there is a rise in the chances for conflicts associated with lane change. The increase in manoeuvrability relating to the availability of other lanes leads to an increase in average speed of traffic and difference in speed.

Accident rates were found to increase with an increase in the number of lanes by Garber (2000). Other studies have also indicated the association between road traffic accidents and the number of lanes. Abdel-Aty and Radwan (2000) noted that with an increase in the number of lanes on urban roadway sections, crash rates increased. Milton and Mannering (1998) also established that an increase in the number of lanes in rural Washington State resulted in more road traffic accidents. Noland and Oh (2004) were able to find that an increase in the number of lanes increased traffic-related accidents and fatalities. Golob and Recker (2001) on the other hand identified the position of lanes i.e. left and interior lanes to rather have an influence on the collision type. Council and Stewart (2000) also found that the conversion of two-lane roads to four lanes contributed to a 40 to 60 percent reduction in crashes.

HGV *(Table 12 and Table 13):* The effect of the percentage of Heavy Goods Vehicles (HGVs) on *fatal and serious* accidents was statistically insignificant and hence it was excluded from the model. For *slight* accident numbers, the percentage of Heavy Goods Vehicles was found to have a negative effect. At significance level of less than 0.05 the data provided enough evidence (p = 0.028) to suggest that the percentage HGVs has an impact on the number of *slight* accidents.

Other researches (Shankar, Milton and Mannering, 1997; Miaou, 1994; Anastasopoulos and Mannering, 2009) have indicated that an increase in HGVs results in a decreased frequency of overtaking vehicles and lane changing behaviour leading to a reduction in the number of accidents. De Palma, Kilani and Lindsey (2008) also found that HGVs inflict more congestion than accident costs in comparison to light goods vehicles. In another study in Australia (Mitchell, Driscoll and Healey, 2004), rigid trucks and prime movers were found to be more likely to be involved in non-vehicular collisions on a curve in the roadway on national and state highways as well as other rural roads with high speed appearing to be a factor in many of the incidents for this category of vehicles. In addition to the above findings, accident statistics also show that large trucks present a serious safety problem, particularly with regard to the severity of accidents in which they are involved (Khorashadi et al., 2005; Evgenikos et al., 2016).

Speed limit *(Table 12 and Table 13):* The speed limit observed was found to have a negative effect on the number of *fatal and serious* accidents. The same observation was made for *slight* accidents. At less than 0.05 significance level there was enough evidence (p = 0.000) provided by the data to suggest that speed limit has an effect on *fatal and serious* accidents numbers and on *slight* accident numbers.

Aljanahi, Rhodes and Metcalfe (1999) in their study of the Tyne and Wear, UK region showed a stronger relationship existed between accidents and the traffic speed variability but the speed limit of the road was not considered. However, Ossiander and Cummings (2002) analysing the effects of increasing speed limits on a rural freeway in the USA found speed variance not to have been affected by the increase in speed limit. Even though the fatal crash rate increased due to the increase in speed limit, the total crash rate showed little change. The generally low numbers of recorded fatal accidents could be attributed to this. Results from the study by Taylor, Baruya and Kennedy (2002) showed the frequency of all injury accidents rose quickly with the mean speed of traffic. Since only 60km/h roads were considered in the Taylor, Baruya and Kennedy (2002) study, comparison with other classes of roads having different speed limits cannot be made. In another study by Taylor, Lynam and Baruya (2000) the frequency of road traffic accidents was found to increase with traffic speed and higher speeds showed a quick increase in accident frequency. Even though greater speed variation along a road may be related to an increase in road accidents this has not been proved experimentally due to lack of data and the possible difficulties in the analysis involved (Thomas et al., 2012).

Homogeneous Segment (HSegLength) *(Table 12 and Table 13):* The homogeneous segment length of road was found to have a positive effect on the number of *fatal and serious* accidents. The same observation was made for *slight* accidents. At less than 0.05 significance level there was enough evidence (p = 0.000) provided by the data to suggest that the homogenous segment length has an effect on *fatal and serious* accidents numbers and on *slight* accident numbers.

AADT *(Table 12 and Table 13):* The logarithm of the AADT was found to have a positive influence on *fatal and serious* accident numbers and on *slight* accident numbers. At less than 0.05 significance level there was evidence (p = 0.000) suggested by the data that AADT has an influence on the number of *fatal and serious* accidents and also on the number of *slight* accidents.

A study by Hau (1992) showed that traffic density determines speed and not vice-versa. Ceder and Livneh (1978) also provided some understanding into the relationship between accident and average daily traffic on interurban road sections by fitting a power function model. The total accident density was found to increase with an increase in average daily traffic (ADT) which is consistent with the findings obtained in this study. Another study by Martin (2002) found crash incidence rates increased steadily as traffic increased on 2 and 3 lane motorways with traffic levels of about 3000 vehicles/hour. The work of Golob and Recker (2001) revealed that the severity of accidents was influenced more by the volume of traffic than by speed.

Radius *(Table 12 and Table 13):* The logarithm of the radius of curvature was found to have a negative influence on both *fatal and serious* accidents and *slight* accidents. This finding is undeviating from other studies (Milton and Mannering, 1998; Haynes et al., 2007). At less than 0.05 significance level there was evidence provided by the data that the radius has an effect on the number of *fatal and serious* accidents (p = 0.000) and *slight* accidents (p = 0.000). Milton and Mannering (1998) found an increase in horizontal curve radius to decrease the number of accidents which is in agreement with results from this research. Haynes et al. (2007) also found that collision numbers were negatively related to road curvature with the cumulative angle being most strongly related to fatal road crashes. In the work of Berhanu (2004) a decrease in road curvature showed decreases in accidents. A possible explanation can be attributed to the urban environment used where traffic speed is likely to be low and allows for more driver reaction time to reduce speed on curved sections of the road and thereby have better control of the vehicle. The study by Berhanu (2004) also recognised that drivers tend to speed on straight sections of a road than within bends but there is still the possibility of accident risk being increased at areas of increased road curvature.

Multicollinearity *(Table 16 and Table 17):* Multicollinearity arises when high correlations exist within predictor variables resulting in unstable and unreliable estimates of regression coefficients (Allison, 2012). The highest correlation of parameter estimates was noted to be -0.317 (Table 16) occurring between the logarithm of the Radius (LogRadius) and homogeneous segment length (HSegLength) for *fatal and serious* accidents. For slight accidents the highest correlation between variables was -0.59 (Table 17) occurring between the logarithm of the annual average daily traffic (LogAADT) and the number of lanes (Lanes). The highest correlation figure obtained was -0.59 suggesting that multicollinearity is not an issue for the predictor variables. The correlation of parameter estimates are given in Table 16 and Table 17 for *fatal and serious* accidents and *slight* accidents.

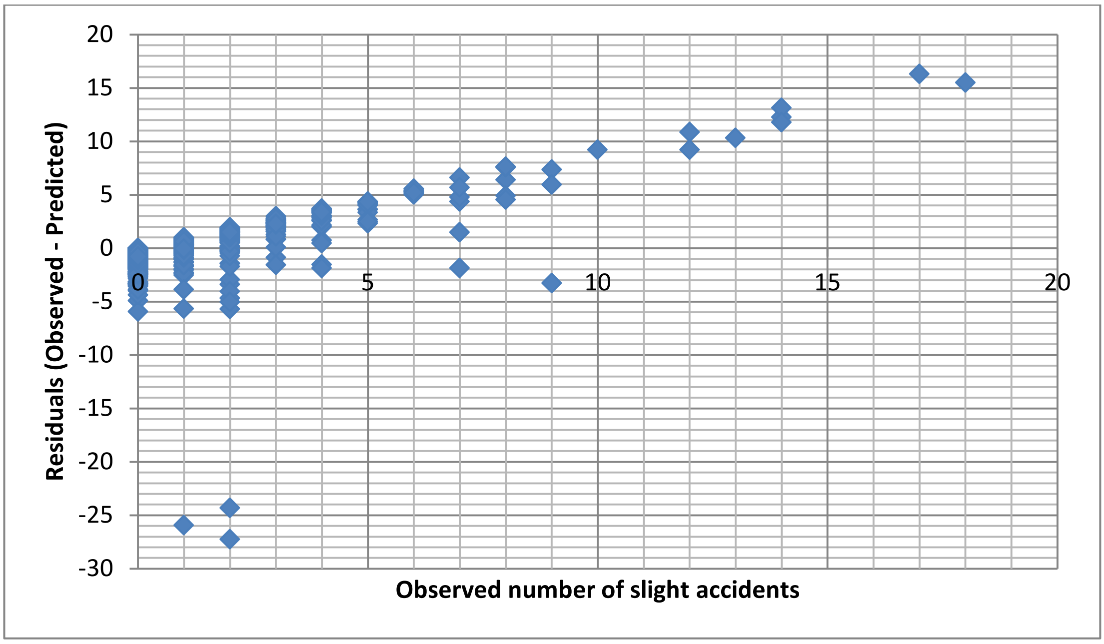
	Intercept	Jtns=0	Jtns=1	Dxn=1	Dxn=2	Dxn=3	Dxn =4	SpeedLimit	HSegLength	LogAADT	LogRadius
Intercept	1.000										
Jtns=0	0.146	1.000									
Jtns=1	.	.									
Dxn=1	-0.175	-0.046	.	1.000							
Dxn =2	-0.162	-0.027		0.693	1.000						
Dxn =3	-0.049	0.031.	.	0.499	0.462	1.000					
Dxn=4	.	.	.	.	.	.	.				
SpeedLimit	-0.124	-0.257	.	-0.121	-0.132	0.030	.	1.000			
HSegLength	-0.003	0.221	.	-0.059	-0.063	-0.052	.	-0.244	1.000		
LogAADT	-0.940	-0.130		0.065	0.068	-0.072	.	-0.087	0.053	1.000	
LogRadius	-0.139	-0.155	.	0.059	0.042	0.070	.	-0.096	**-0.317**	-0.035	1.000

Table 16 Correlation of parameter estimates for fatal and serious accidents on A-roads

	Intercept	Jtns=0	Jtns=1	Presof CR=0	Presof CR=1	Dxn=1	Dxn=2	Dxn=3	Dxn =4	Jtns SUM	Lanes	Percent HGV	Speed Limit	HSeg Length	LogAADT	LogRadius
Intercept	1.000															
Jtns=0	-0.044	1.000	.													
Jtns=1	.	.	.													
PresofCR=0	-0.373	0.046	.	1.000												
PresofCR=1	.	.	.	.	.											
Dxn=1	-0.067	-0.044	.	-0.102	.	1.000										
Dxn =2	-0.065	-0.026	.	-0.135	.	0.752	1.00									
Dxn =3	-0.153	0.040	.	0.108	.	0.485	0.461	1.000								
Dxn=4	.	.	.	.	.	.	.	.	.							
Jtns	-0.184	0.781	.	0.099	.	-0.042	-0.022	0.033	.	1.000						
Lanes	0.336	-0.022	.	-0.099	.	0.306	0.309	-0.124	.	-0.056	1.000					
PercentHGV	0.082	0.000	.	-0.062	.	-0.244	-0.264	0.031	.	0.083	-0.286	1.000				
SpeedLimit	-0.101	-0.022	.	-0.155	.	-0.008	0.010	-0.024	.	0.068	0.090	-0.458	1.000			
HSegLength	0.023	-0.026	.	-0.037	.	-0.050	-0.062	-0.082	.	-0.186	0.004	-0.105	-0.156	1.000		
LogAADT	-0.898	-0.080	.	0.242	.	-0.099	-0.096	0.055	.	0.023	**-0.599**	0.058	-0.088	0.036	1.000	
LogRadius	-0.118	-0.121	.	0.019	.	0.091	0.081	0.140	.	-0.007	-0.011	-0.035	-0.024	-0.298	-0.027	1.000

Table 17 Correlation of parameter estimates for slight accidents on A-roads

Figure 16 shows the relationship between the observed and the residual (observed - predicted) values of accidents on A-roads tested and validated using the Negative Binomial model developed. The pattern shown in Figure 16 for both slight accidents and fatal and serious accidents, indicate residuals increase with increase in observed accident numbers.



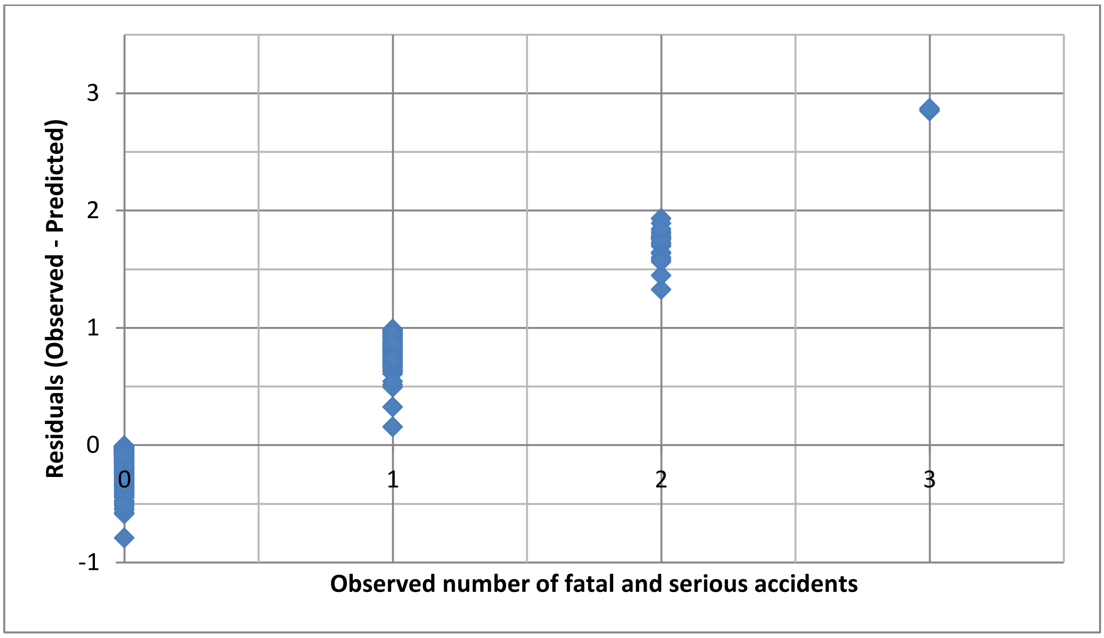


Figure 16 Relationship between the observed and residual values (observed-predicted) of accidents from Negative Binomial model

5.3 Empirical Bayes model results

Count models have the tendency to ignore spatial correlations which may influence results from the model estimates and usually tend to be corrected by using spatial models such as Bayesian hierarchical models which are supposed to control for spatial correlations. Even though Full Bayesian hierarchical models are able to handle problems associated with spatial correlation, for fatal and serious accidents this is not always the case. Wang, Quddus and Ison (2011) noted the pattern of residuals from *fatal and serious* accidents from count models to be similar to that obtained from a Bayesian spatial model. This may be attributed to the higher regression to mean effects associated with *fatal and serious* accidents as compared to *slight* accidents since the occurrence of *fatal and serious* accidents are rare. Full Bayesian hierarchical models are also disadvantaged in that they may not be easily transferable to other datasets with correlation results sometimes being difficult to interpret (Lord and Mannering, 2010). However, quantifying road safety using the Empirical Bayes (EB) method has evolved over the years and it has been described as the preferred method to measuring the expected number of accidents (Elvik, 2008). This is because EB methods were developed for the purpose of controlling for regression-to-the-mean effects in before and after studies that assess the impact of road safety measures as well as in identifying hazardous road locations (Elvik, 2008).

The Empirical Bayes method works on the basis that the best estimate of safety is obtained by combining two sources of information; (i) the accident record for a given study unit (driver, intersection, road section, etc.), and (ii) an accident prediction model showing how various factors affect accident occurrence. Complexities are involved in developing good accident prediction models and the Empirical Bayes method is no exception. Discussions in Chapter 3 about the limitations associated with models such as the negative binomial model led to the improvement of the model by the Empirical Bayes approach.

Roads data available for use was split into two portions with approximately 75% of the data used to develop the negative binomial model with the remaining 25% used to validate it. The negative binomial model developed was further developed into an empirical Bayes model. This section discusses results obtained from testing the model on the remaining 25% of the roads data. The length of roads used in validating the Negative Binomial Empirical Bayes accident model was 185,040 metres. The roads comprised 1771 homogeneous segments, 159 *fatal and serious* accidents and 960 *slight* accidents. There were 1872 junctions and 11120m of cycle routes along the roads investigated. A summary of results obtained for slight accidents and fatal and serious accidents combined from observed data, negative binomial model and empirical Bayes model are given in Table 18.

Road number	Observed		Negative Binomial Predicted		Empirical Bayes Predicted		Road number	Observed		Negative Binomial Predicted		Empirical Bayes Predicted	
	Slight	*Fatal and serious*	*Slight*	*Fatal and serious*	*Slight*	*Fatal and serious*		*Slight*	*Fatal and serious*	*Slight*	*Fatal and serious*	*Slight*	*Fatal and serious*
A6211SN	12	4	23	4	20	4	A6211NS	32	6	37	12	34	11
A6130SN	101	8	40	7	80	7	A6130NS	22	6	43	4	29	5
A6117SN	53	14	48	11	52	12	A6117NS	18	3	43	7	34	7
A6030SN	72	7	46	13	58	13	A6030NS	31	1	38	8	37	7
A6005WE	116	18	262	37	154	32	A6005EW	71	17	270	24	116	22
A6002SN	68	15	71	19	68	18	A6002NS	31	14	89	13	60	13
A47WE	183	19	57	13	102	13	A47EW	150	27	99	35	130	34

Table 18 Results of accident predictions

A paired t-test was carried out for the observed accidents, negative binomial model and empirical Bayes model. The t-test was carried out in order to determine if there is any statistical significance between observed accidents and either the negative binomial model or empirical Bayes model. A t-test was also carried out to determine any statistical significance between the negative binomial model and the empirical Bayes model results. The test was carried out separately for slight accidents and for fatal and serious accidents combined and a 5 percent significance level was used. The paired t-test compared

- observed slight accidents to negative binomial slight accidents
- observed slight accidents to empirical Bayes slight accidents
- negative binomial slight accidents to empirical Bayes slight accidents
- observed fatal and serious accidents to negative binomial fatal and serious accidents
- observed fatal and serious accidents to empirical Bayes fatal and serious accidents
- negative binomial fatal and serious accidents to empirical Bayes fatal and serious accidents

The null hypothesis H_o tested were

1. There is no difference between observed accidents and the accidents generated from the negative binomial model
2. There is no difference between the observed accidents and the accidents generated from the empirical Bayes model
3. There is no difference between the accidents generated from the negative binomial model and the empirical Bayes model

The research (alternative) hypothesis H_A tested were

1. There is a difference between observed accidents and the accidents generated from the negative binomial model
2. There is a difference between the observed accidents and the accidents generated from the empirical Bayes model

3. There is a difference between the accidents generated from the negative binomial model and the empirical Bayes model

Results obtained from the t-test paired sample statistics are given in Table 19 with results of the t-tests given in Table 20. Abbreviations used were Obsvd for Observed, NB for Negative Binomial and EB for Empirical Bayes.

Severity of accident	Models compared	Mean	N	Std. Deviation	Std. Error Mean
Slight	Obsvd	0.54	1770	1.497	0.036
	NB	0.66	1770	1.416	0.034
	Obsvd	0.54	1770	1.497	0.036
	EB	0.55	1770	0.994	0.024
	NB	0.66	1770	1.416	0.034
	EB	0.55	1770	0.994	0.024
Fatal and Serious	*Obsvd*	0.09	1770	0.339	0.008
	NB	0.12	1770	0.108	0.003
	Obsvd	0.09	1770	0.339	0.008
	EB	0.11	1770	0.117	0.003
	NB	0.12	1770	0.108	0.003
	EB	0.11	1770	0.117	0.003

Table 19 Results of t-test paired sample statistics for accidents

Severity of accident	Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
Slight	Obsvd - NB	-0.115	1.798	0.043	-0.199	-0.032	-2.703	1769	0.007
	Obsvd - EB	-0.008	0.696	0.017	-0.041	0.024	-0.505	1769	0.614
	NB – EB	0.107	1.333	0.032	0.045	0.169	3.381	1769	0.001
Fatal and Serious	Obsvd - NB	-0.028	0.332	0.008	-0.043	-0.012	-3.534	1769	0.000
	Obsvd - EB	-0.023	0.279	0.007	-0.036	-0.010	-3.398	1769	0.001
	NB – EB	0.005	0.061	0.001	0.002	0.008	3.652	1769	0.000

Table 20 Results of t-test for accidents

Analysing Table 19 and Table 20, it can be seen that there is a significant difference in observed (Obsvd) slight accidents (M= 0.54, SD= 1.497) and slight accidents obtained from the Negative Binomial (NB) model (M=0.66, SD=1.416) condition; t (1769) = -2.703, p=0.007. For observed (Obsvd) fatal and serious accidents (M= 0.09, SD= 0.339) and fatal and serious accidents obtained from the Negative Binomial (NB) model (M=0.12, SD=0.108) condition; t (1769) = -3.534, p=0.000 a significant difference was obtained. The same pattern of significance was observed between the Negative Binomial model and Empirical Bayes model for both slight accidents and fatal and serious accidents combined.

For Negative Binomial (NB) fatal and serious accidents (M= 0.12, SD= 0.108) and fatal and serious accidents obtained from the Empirical Bayes (EB) model (M=0.11, SD=0.117) condition; t (1769) = 3.652, p=0.000 a significant difference was obtained. For Negative Binomial (NB) slight accidents (M= 0.66, SD= 1.416) and slight accidents obtained from the Empirical Bayes (EB) model (M=0.55, SD=0.994) condition; t (1769) = 3.381, p=0.001 a significant difference was obtained.

For Observed slight accidents (M= 0.54, SD= 1.497) and the Empirical Bayes model slight accidents (M= 0.55, SD= 0.994), there was no evidence of a difference, condition; t (1769)= -0.505, p=0.614. However, for fatal and serious accidents, it is noted that there is a significant difference in observed accidents (M= 0.09, SD= 0.339) and the Empirical Bayes model (M= 0.11, SD= 0.117) condition; t (1769) = -3.398, p=0.001.

Paired t-test values indicating significant difference between the samples tested was obtained in the case of slight accidents for the observed and negative binomial, negative binomial and empirical Bayes, and for fatal and serious accidents between observed and negative binomial, observed and empirical Bayes and negative binomial and empirical Bayes. These t-tests indicating significant difference had p-values less than 0.05 and the 95 percent confidence interval obtained for these samples given in Table 20 are noted not to include zero. For t-test carried out for slight accidents between observed and empirical Bayes samples, where no significant different was found with p-value greater than 0.05 (p=0.614) the 95 percent confidence interval (-0.041 to 0.024) was found to include zero further indicating no significant difference.

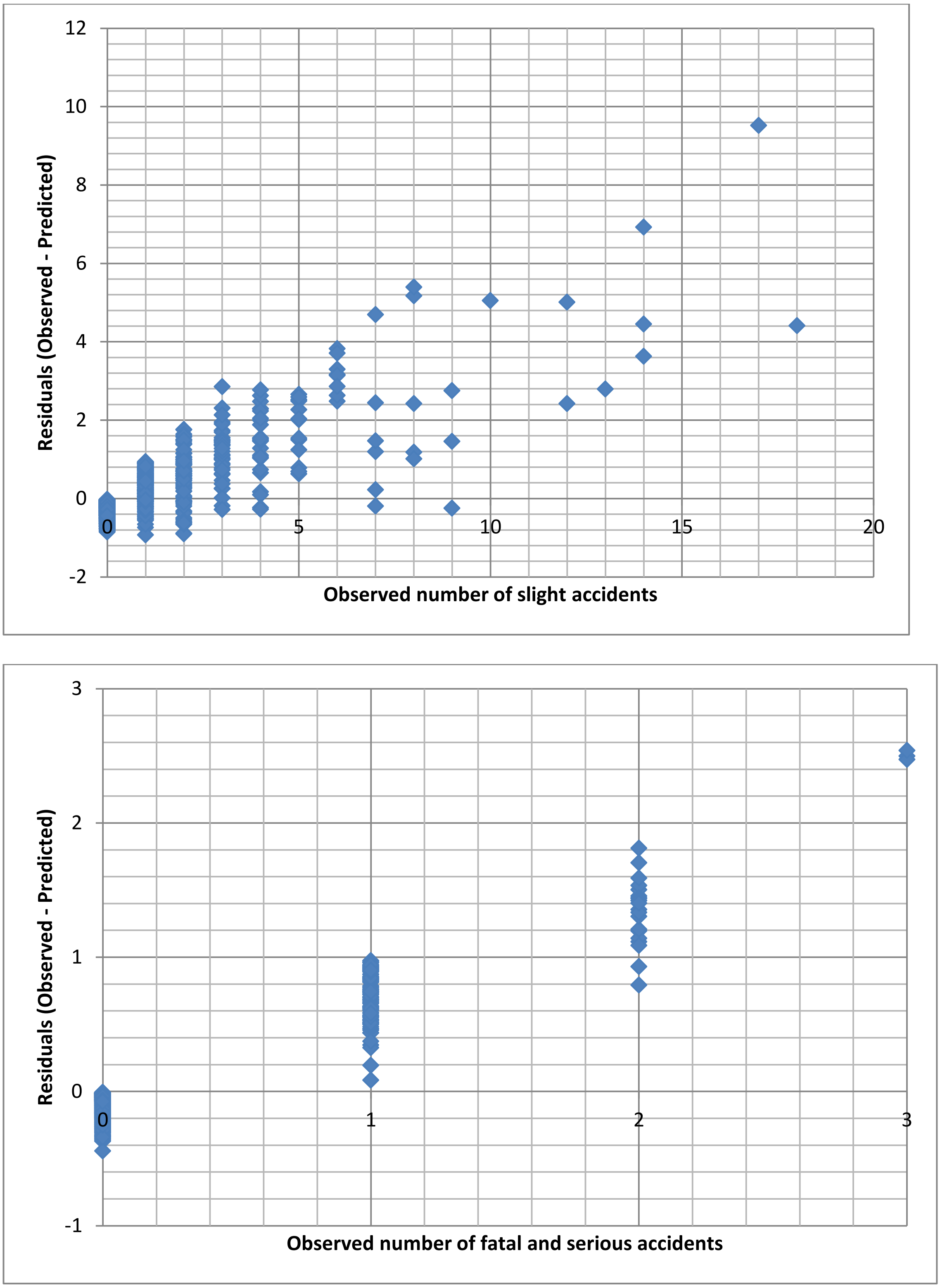


Figure 17 Relationship between the observed and residual values (observed-predicted) of accidents from Empirical Bayes model

In Figure 17, the relationship between the observed and the residual (observed - predicted) values of accidents on A-roads tested and validated using the Empirical Bayes method shows a pattern to be compared with Figure 16. For fatal and serious accidents, the pattern of residuals observed in Figure 17 is similar to the pattern of residuals in Figure 16. The pattern shows residuals increasing with observed numbers. A possible explanation for this can be attributed to the effect of high regression-to-mean in fatal and serious accidents due to the rarity of occurrence of this type of accident. For slight accidents, the pattern shown in Figure 17 is different from that shown in Figure 16 showing a refinement in the model obtained from Empirical Bayes.

5.4 Summary

Roads used for the accident prediction model was 789,560 metres in total. 604,520 metres was used in developing the model with the remaining 185,040 metres used for validating the model. A negative binomial model was developed and improved by means of an empirical Bayes method. The parameters initially used in developing the models were junctions presents (options being Yes or No), Presence of a Cycle route (options being Yes or No), direction of road (e.g. North-South), presence of pedestrian crossing (options being Yes or No), the sum of junctions within a homogeneous length of road, the number of lanes, percentage of heavy goods vehicles, speed limit, homogeneous segment length of road, logarithm of the Annual Average Daily Traffic (AADT), logarithm of the radius of the road, sum of cycle route, average speed and percentage slope. After running the initial models for *fatal and serious* accidents combined and *slight* accidents separately, model variables found not to be statistically insignificant were removed and the model was rerun. For the *fatal and serious* accidents mode, parameters found to be statistically significant are the presence of junctions, direction of travel along a road, speed limit, homogeneous segment length of the road, logarithm of the Annual Average Daily Traffic and the logarithm of the radius. For the *slight* accidents model, variables found to be statistically significant are the presence of junctions, presence of cycle route, direction of travel along a road, sum of junctions within a homogeneous length of road, number of lanes per driving direction, percentage heavy goods vehicles (HGV), speed limit, homogeneous segment length of road, logarithm of the Annual Average Daily Traffic (AADT) and the logarithm of the radius. The negative binomial model developed was tested on 185,040 metres of roads independent of what was used in developing the model. The negative binomial model was further developed into an empirical Bayes model to enhance the predictive power as well as help deal with the effects of regression-to-the-mean. The residual plots (observed – predicted) against observed accidents for the empirical Bayes model fatal and serious accidents combined was found not be very different from that obtained from the negative binomial model. However, the residual plots (observed –predicted) against observed accidents for the empirical Bayes model slight accidents was found to be different from that obtained from the negative binomial model showing an improvement in the empirical Bayes model predictions. When a paired t-test was carried out between data samples, apart from observed slight accidents and empirical Bayes slight accidents which showed no statistical difference, all other paired t-test sample data showed statistical difference at the 5 percent significance level with 95 percent confidence interval limit values confirming the significance of values obtained.

6 Results (II): Optimisation model

6.1 Introduction

The facility location model in this research is done to obtain the optimum number and location for speed control devices (speed cameras and vehicle activated signs (VAS)). Consideration has been given to recommendations made by the Home Office, Highways Agency and from other available guidelines regarding the installation of speed cameras and vehicle activated signs. These recommendations and guidelines have been used as a starting guide and incorporated into the models as appropriate.

The question to be answered is 'where should a speed control device be placed so that the resulting benefits are reduction in road traffic accidents through vehicle speed reduction?'. To answer this question, results from the practical application of the objectives and constraints defined in Chapter 4 have been applied to some A-roads within the Nottinghamshire and Leicestershire areas of United Kingdom. A-roads used to validate the accident prediction models in Chapter 5 have been used in this chapter to test the optimisation model. Appendix D provides GIS plots of the variation of accidents along the roads used. A list of roads used in this chapter is given in Table 21.

Road	Start location (approximate)	End location (approximate)
A6211	A60 Mansfield road/A6211 Thackerays Lane junction	A6211 Colwick Loop road/A612 Colwick Loop road junction
A6002	A6007 Ilkeston Road/A6002 Coventry Lane junction	A6002 Sandhurst road/Hucknall Lane junction
A6130	A6130 Castle Boulevard/Sherwin Road junction	A6130 Gregory Boulevard/Mansfield Road junction
A6117	A60 Leeming Lane South/A6117 Old Mill Lane junction	A617 Sherwood Way East/A6117 junction
A6030	A6 London road/Stoughton road junction	A563 Troon Way/A6030 Victoria Road East junction
A6005	Hills road/Draycott road	Castle Boulevard/Wilford Street
A47	Normandy Way/Leicester road junction	A47/A43 junction

Table 21 Roads used in optimisation

It should be mentioned that the optimisation model was developed on the basis of roads having accidents of a certain magnitude and severity. It is worth stating that the optimisation model does not take into account roads used for construction purposes with the need to have speed cameras placed on them to reduce vehicle speeds for the safety of road workers.

6.2 Pattern Search Results

Pattern search optimisation normally depends on the smoothness of the cost function whereas genetic algorithms on the other hand may not converge with a smooth cost function (Wetter and Wright, 2003). Both pattern search and genetic algorithms can be used on dis-continuous and un-differentiable functions (Mathworks, 2015). The pattern search algorithm evaluates a series of points approaching an optimal point. During each step, the algorithm searches a set of points referred to as a mesh around the present point. The present point acts as the calculated point from the previous step of the algorithm. If a point in the mesh is found by the pattern search which improves the objective function at the present point the new point is used as the present point for the following step of the algorithm. In pattern search the search seeks to obtain a better point than the present one. (A better point implies one with lower objective function value).

The optimisation model developed for the A-roads made use of two main optimisation techniques of which Pattern Search was one. Fourteen different segments of A-roads (in Nottinghamshire and Leicestershire) were used in the *pattern search* optimisation models. Roads used have been split into south north (SN), north south (NS), east west (EW) and west east (WE) directions. The optimisation was carried out using the road centreline as the reference point.

For all roads used in the optimisation, the fitness plots at each generation are given in Figure 18 to Figure 22. Computation times obtained from pattern search varied from a minimum of 13 minutes to a maximum of 595 minutes. Fitness plots obtained from the *pattern search* optimisation produced results indicating more progress in lowering the fitness value. The plots indicate the objective function value of the best point at each iteration with the objective function values typically improving quickly during the early stages of iteration and then it begins to level off as the optimal value is approached. A better point which is one with lower objective function value is always searched for and desired. In Figure 18, the A6002SN shows a gradual stepwise reduction in the function value to the end of iteration with a final function value of 1.72 obtained around iteration 72. A gradual progress in the refinement of the fitness function is observed for the A6002SN. For the A6002NS in Figure 18, a similar observation was made however it is noted that the function value started at a much lower value of 0.08 at the start of iteration in comparison with a function value of 3500 at start of iteration for the A6002SN.

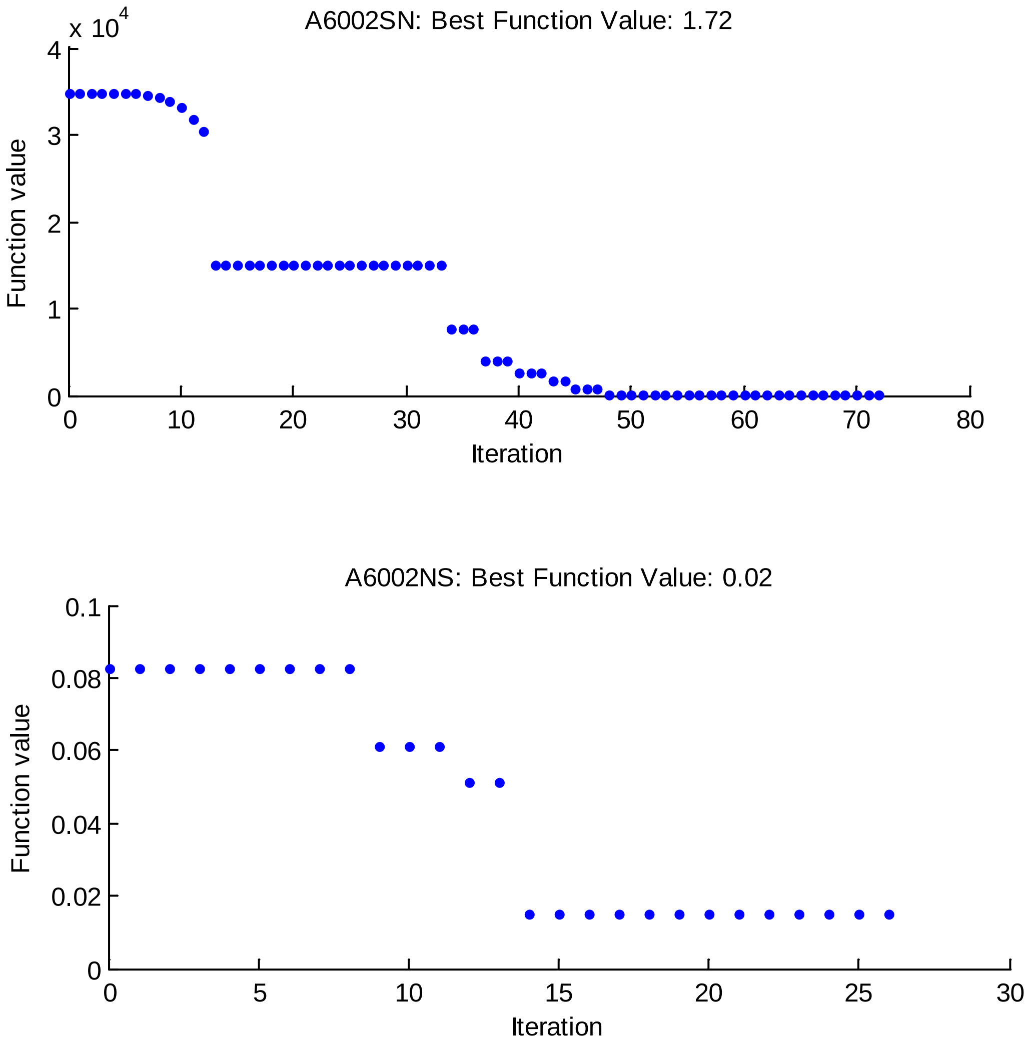


Figure 18 Plot of best fitness values for A6002

The A6130SN in Figure 19 shows a plot of the function value with the iteration. The fitness value for the A6130SN is observed to be lowered step by step from start of iteration to a point where the fitness value remains the same to the end with a value of 2.76. There was however no optimal point located for the A6130NS. For the A47WE in Figure 19, pattern search produced a drop in function value after the first twelve iterations after which minor drops in fitness value was observed before terminating in a final function value of 2.67.

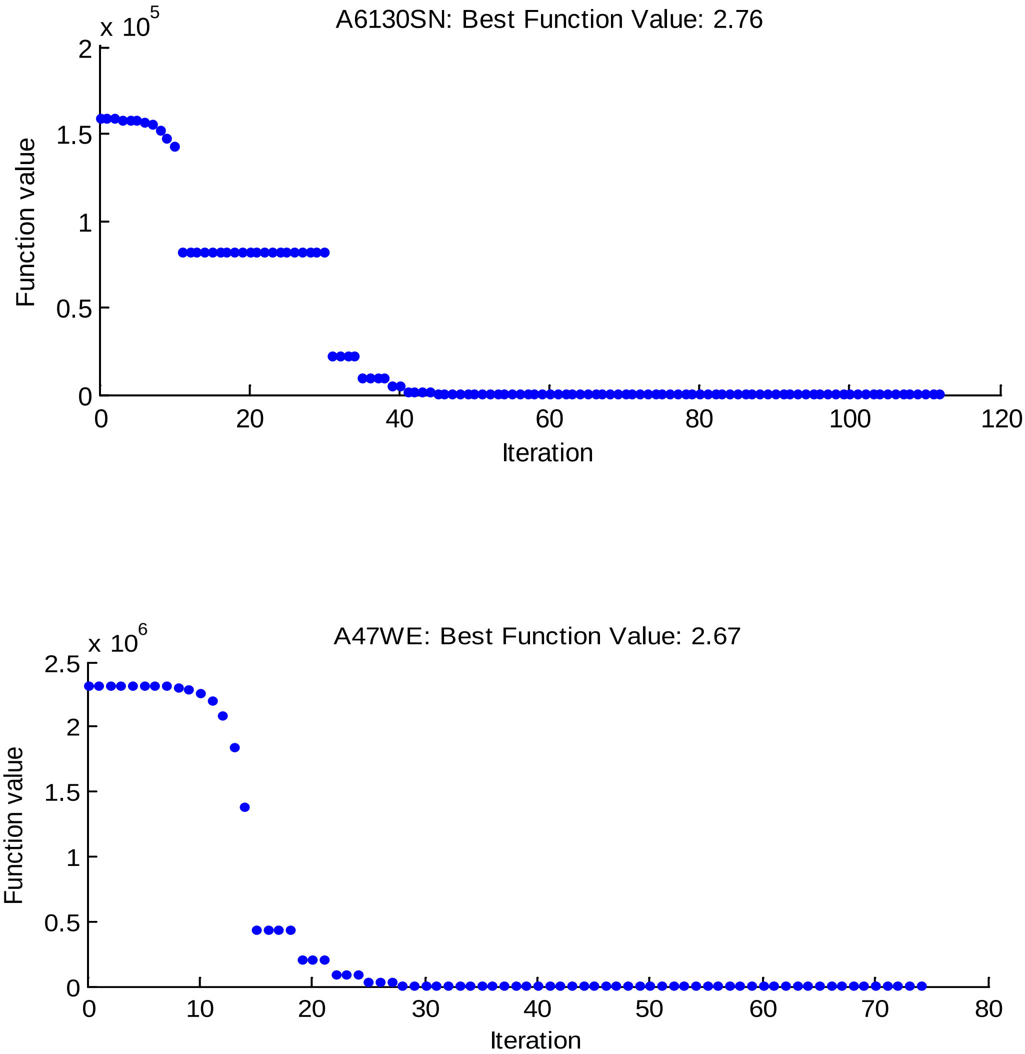


Figure 19 Plot of best fitness value for A6130SN and A47WE

The A6117SN and A6117NS in Figure 20 shows the refinement in function value from start of iteration to end. Both figures show the function value reaching a final reduced value from iteration 11. The profile of both plots are similar however the

A6117SN finishes off with a final function value of 7407.44 while the A6117NS finishes with a function value of 15120.8

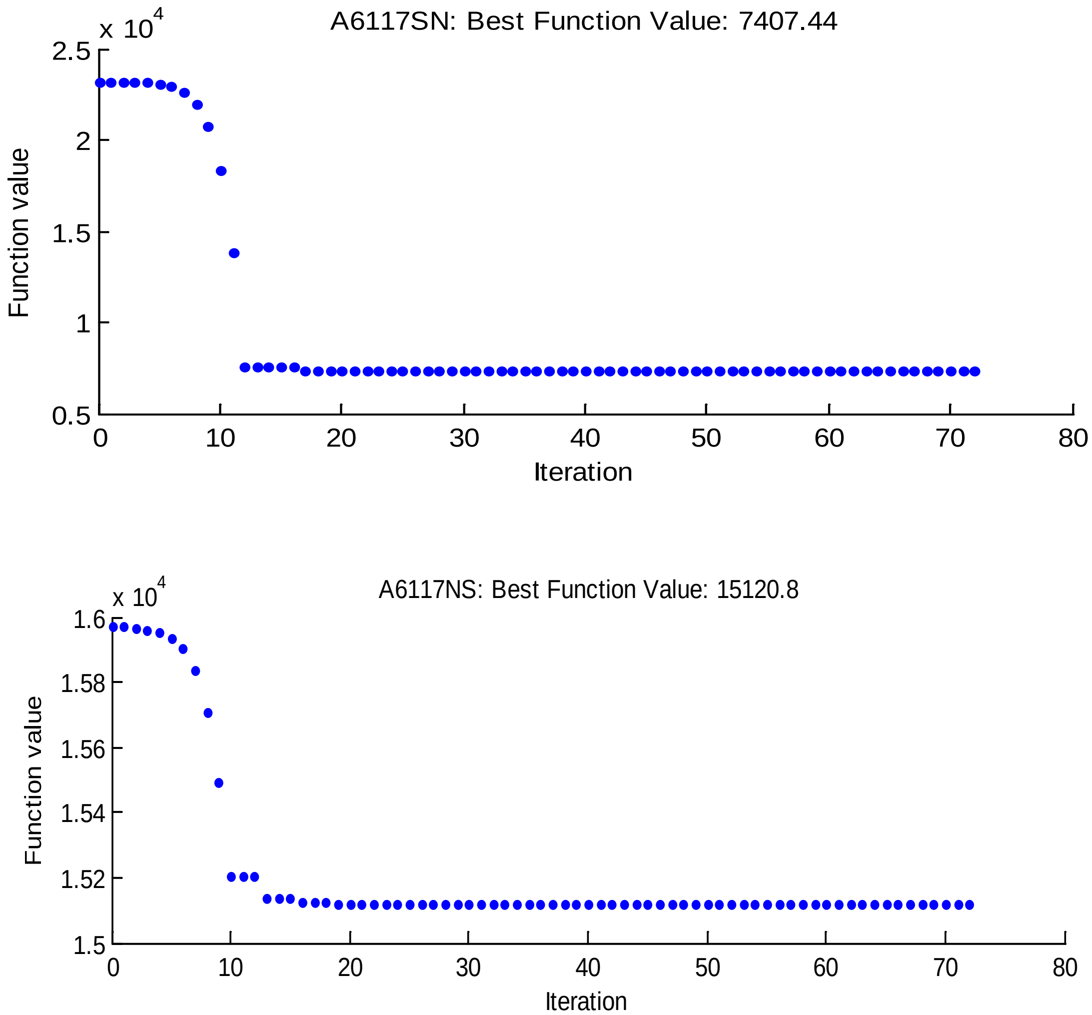
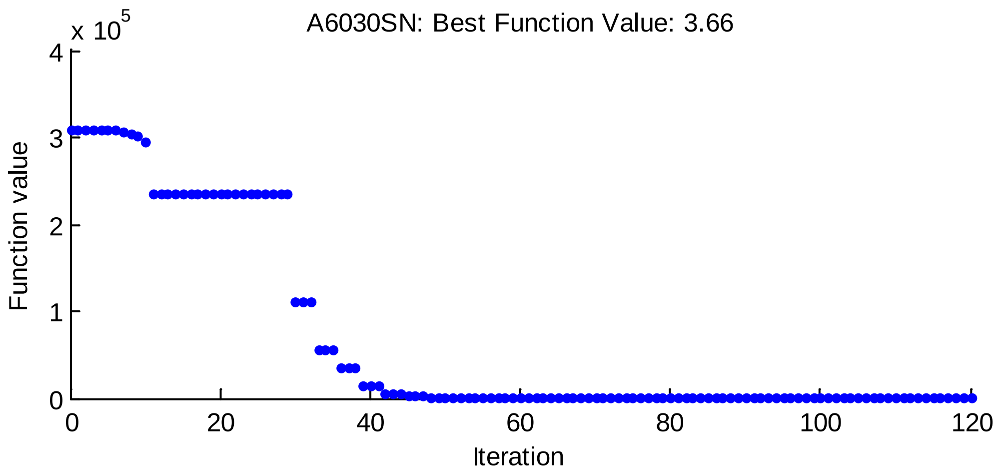


Figure 20 Plot of best fitness value for A6117

The A6030SN and A6030NS in Figure 21 shows the function value plot against the number of iterations. The A6030SN shows a stepwise reduction in the function value in comparison to the plot shown for the A6030NS. The A6030SN has a final function value of 3.66 with the A6030NS having a final function value of 0.05



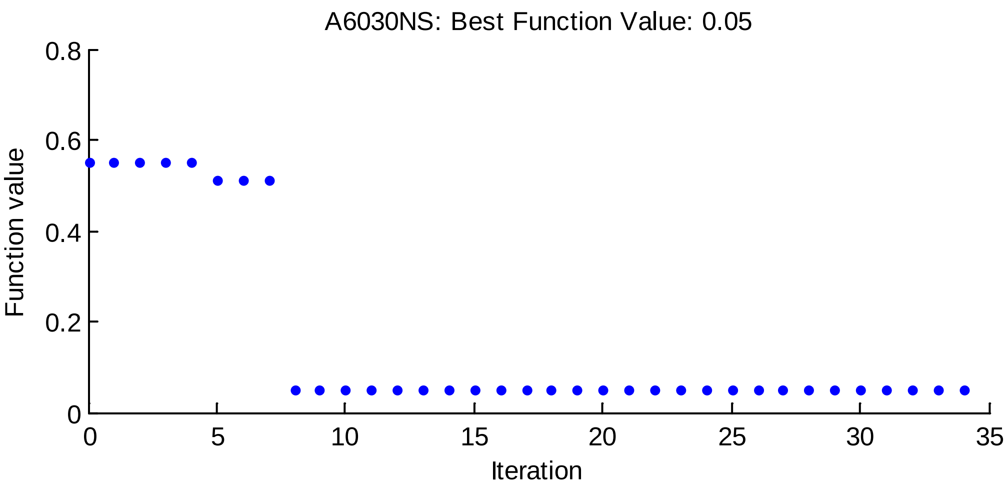
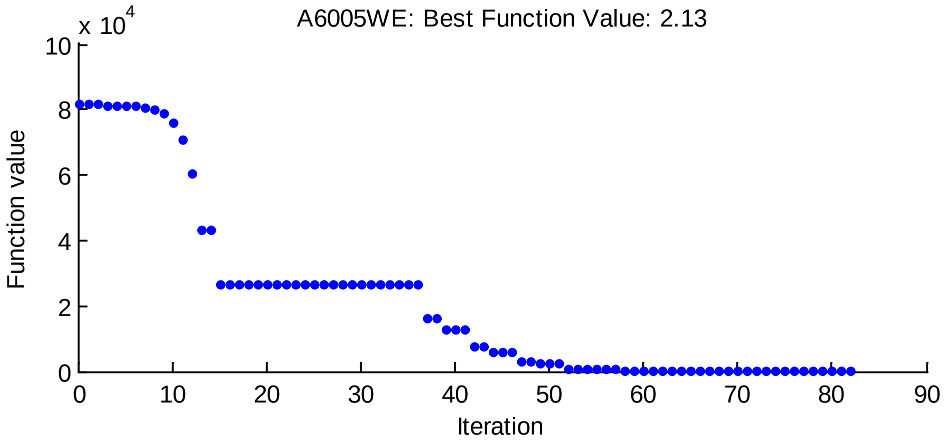


Figure 21 Plot of best fitness value for A6030

The A6005WE and A6005EW shows the function value plot with the A6005WE having a final function value of 2.13. The A6005EW on the other hand produces a much lower function value of 1.01.



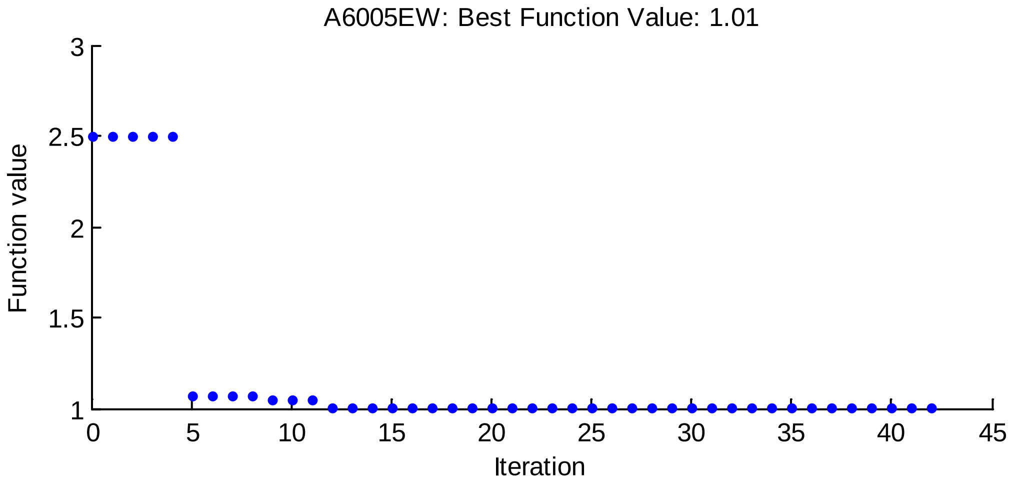
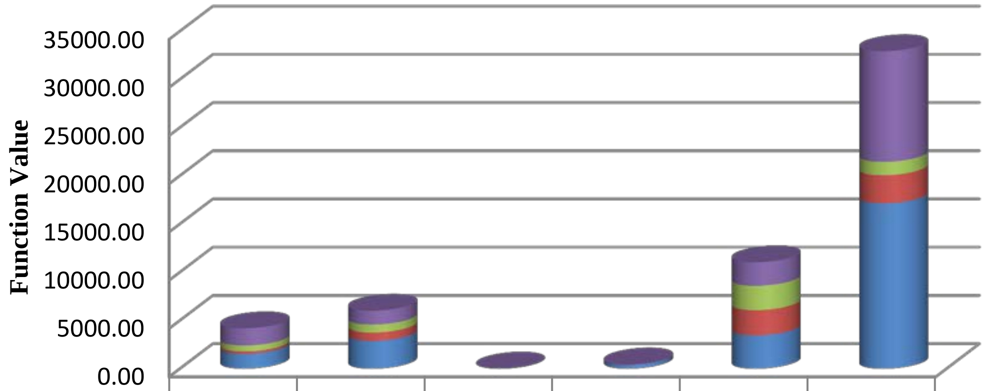


Figure 22 Plot of best fitness value for A6005

6.3 Genetic Algorithm Results

Genetic Algorithm was used as the other optimisation technique. The same roads used for the pattern search were used for the genetic algorithm.

In order to get results closer to the actual minimum value, a number of operations can be applied to a genetic algorithm run. In this research, values were set for the population size, elite count, crossover fraction and number of generations. The elite counts are individuals in the present generation who have the best fitness value and they normally proceed to the next generation. In this research an elite count of 2 was used such that there should be two individuals good enough to proceed to the next generation for optimisation. The value of the elite count was chosen based on recommendations from other research (Mathworks, 2015). An increase in the size of the population also allows the genetic algorithm to explore more points to achieve a better outcome. As a guide the population size is normally set to be at least the value of the number of variables so that individuals in each population can spread to the space being explored. Based on these suggestions a population size of 20 was used. During crossover, children are generated as a result of blending a vector from a pair of parents. The cross over fraction used was 0.8. This value was chosen because studies carried out indicate that very good results are obtained when the crossover fraction is set from 0.4 to 0.8 (Mathworks, 2015). Also, crossover fraction values of 0.6, 0.7, 0.8 and 0.9 were used to run all road sections used in the optimisation and a crossover fraction of 0.8 was found to improve the fitness value in most cases. Figure 23 shows the fitness values obtained for the various crossover fractions (CrOvFr) of the roads.



	A6002	A6130	A6117	A47	A6030	A6005
0.9 CrOvFr	1774.55	1436.99	0.64	37.29	2436.83	11485.47
0.8 CrOvFr	613.49	830.43	0.52	5.86	2526.40	1389.14
0.7 CrOvFr	249.92	861.32	0.19	7.63	2616.83	2918.53
0.6 CrOvFr	1535.46	2877.42	3.71	380.77	3447.80	17124.71

Figure 23 Plot of function values with crossover fractions

Mutation however applies arbitrary modifications to individuals in the present generation to form a child. The adaptive feasible mutation was used. Taking into consideration the constraints applied to the models, this type of mutation was considered suitable. With the inclusion of constraints, this type of mutation randomly generates directions that are adaptive with respect to the last successful or unsuccessful generations and choose a direction and step length that satisfies bounds and linear constraints (Mathworks, 2015). Mutation without crossover and vice versa is not an effective option for optimising functions in genetic algorithm. A balance in the choice of mutation and crossover values is required to obtain good results.

The best fitness values given in the figures indicate the fitness of the best individual obtained so far till the current iteration. The mean fitness is the average of the fitness values throughout the entire population. For each generation, the population gets refined and new average population fitness values are obtained. The best fitness tends to get better as the iterations progress. This normally happens quickly initially as the individuals are away from the optimum solution point and then slows down as the algorithm identifies better solutions that are difficult to improve upon.

Figure 24 shows the fitness function plot for the A6002SN obtained from genetic algorithm. There is a gentle drop in fitness value occurring from iteration 1 to iteration 5 with the plot showing that the fitness value does not improve any further after iteration 5. The A6130SN in Figure 24 displays the fitness value remaining constant for a very short period of iteration starting from the beginning to iteration 1 and dropping very slightly at iteration 1. After iteration 1 the model does not provide any improvement in the fitness value up to the final point of iteration.

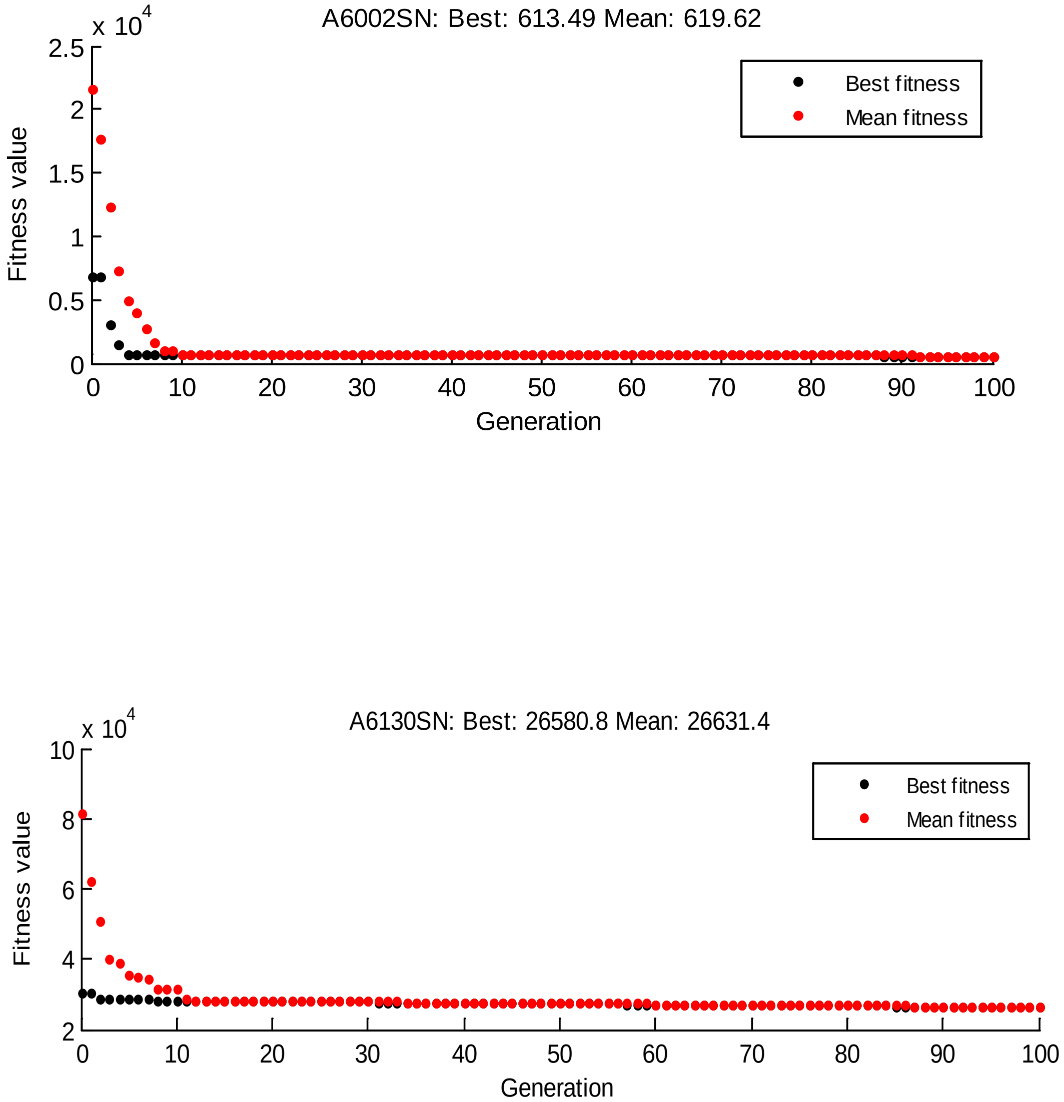
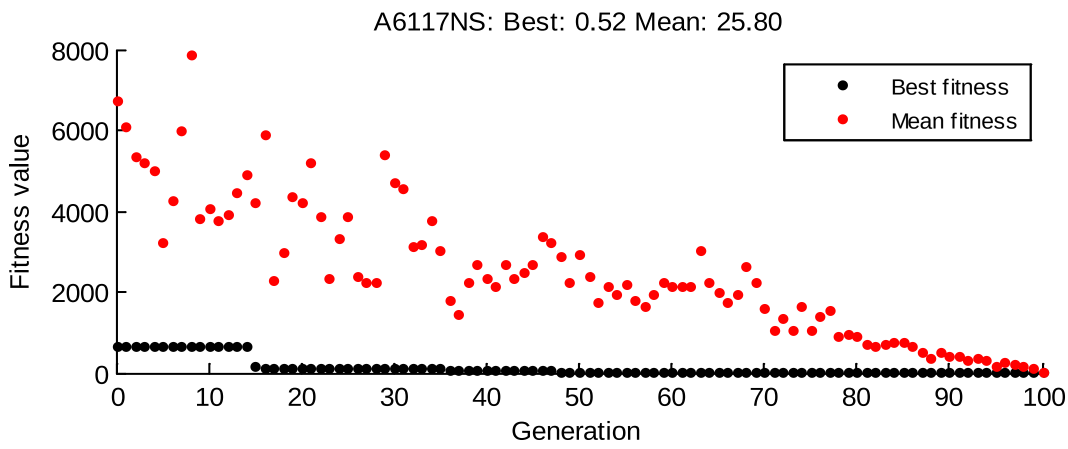


Figure 24 Results from Genetic Algorithm Optimisation

In Figure 25 the A6117NS shows the fitness value remaining constant from iteration 1 to iteration 14. There is a reduction in fitness value from iteration 14 to iteration 16 and the fitness value remains constant from iteration 16 to the final iteration point indicating there is no further improvement in the fitness value.



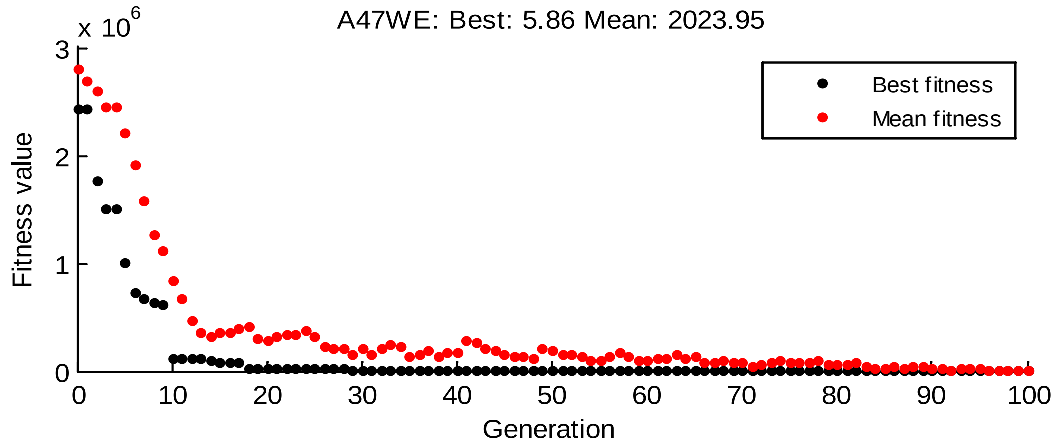
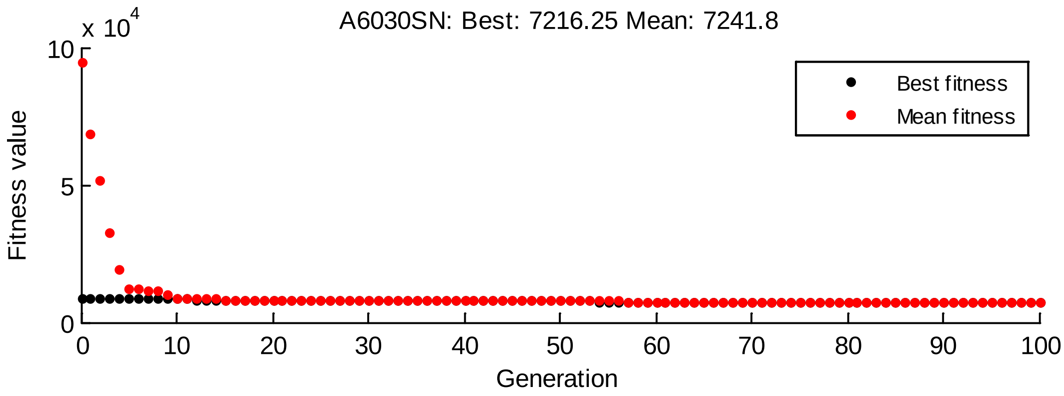


Figure 25 Results from Genetic Algorithm Optimisation

In Figure 25 the A47WE shows a steep drop in the fitness value from iteration 1 to iteration 10. From iteration 10 to iteration 15 there is very little refinement in the fitness value with the fitness value remaining constant from iteration 16 to the end of iteration of the model.

The A6030SN in Figure 26 shows no improvement in the fitness value from the beginning to the end of iteration. A similar pattern of no improvement in the fitness value from start of iteration to the end of iteration can be noted in the A6005WE plot in Figure 26.



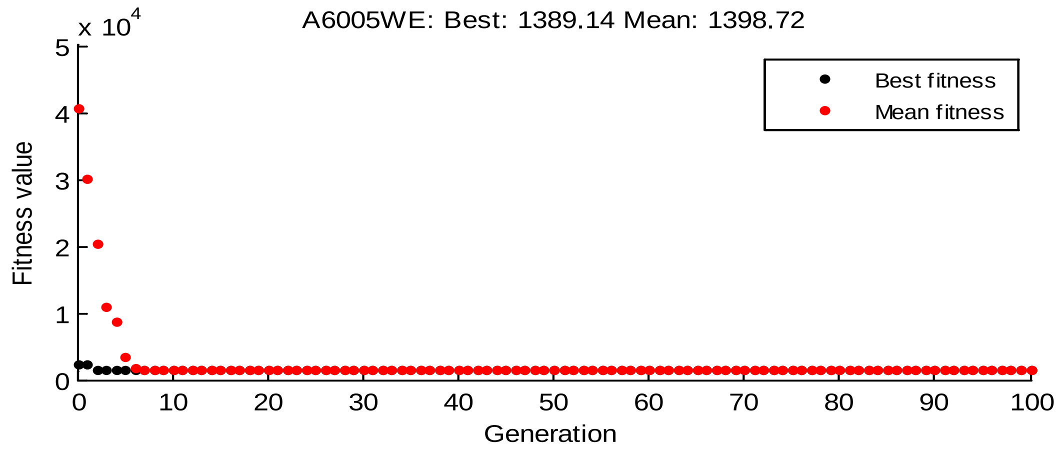


Figure 26 Results from Genetic Algorithm Optimisation

A paired t-test was computed for the fitness values obtained from genetic algorithm and pattern search for the same road samples and the values obtained are given in Table 22 and Table 23.

Models compared	Mean	N	Std. Deviation	Std. Error Mean
Genetic Algorithm	6323.56	120	9500.76	867.30
Pattern Search	1251.24	120	3710.05	338.68

Table 22 Results of t-test paired sample statistics for optimisation function values

Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Genetic Algorithm – Pattern Search	5072.32	10950.62	999.65	3092.91	7051.72	5.07	119	0.000

Table 23 Results of t-test for optimisation function values

Analysing Table 22 and Table 23, it can be seen that there is a significant difference in the function values for genetic algorithm (M= 6323.56, SD= 9500.76) and the function values obtained from pattern search (M=1251.24, SD=3710.05) condition; t (119) =5.07, p=0.000.

The minimum computation time obtained for genetic algorithm was 3 minutes with a maximum computation time being 22 minutes. Figure 27 to Figure 32 show the plotted easting and northing coordinates of optimised locations obtained from pattern search and genetic algorithm. Appendix E and Appendix F provides information on the easting and northing coordinate locations produced from pattern search and genetic algorithm for roads found to have optimum location points. The function values are provided in addition to the optimum location coordinate points.

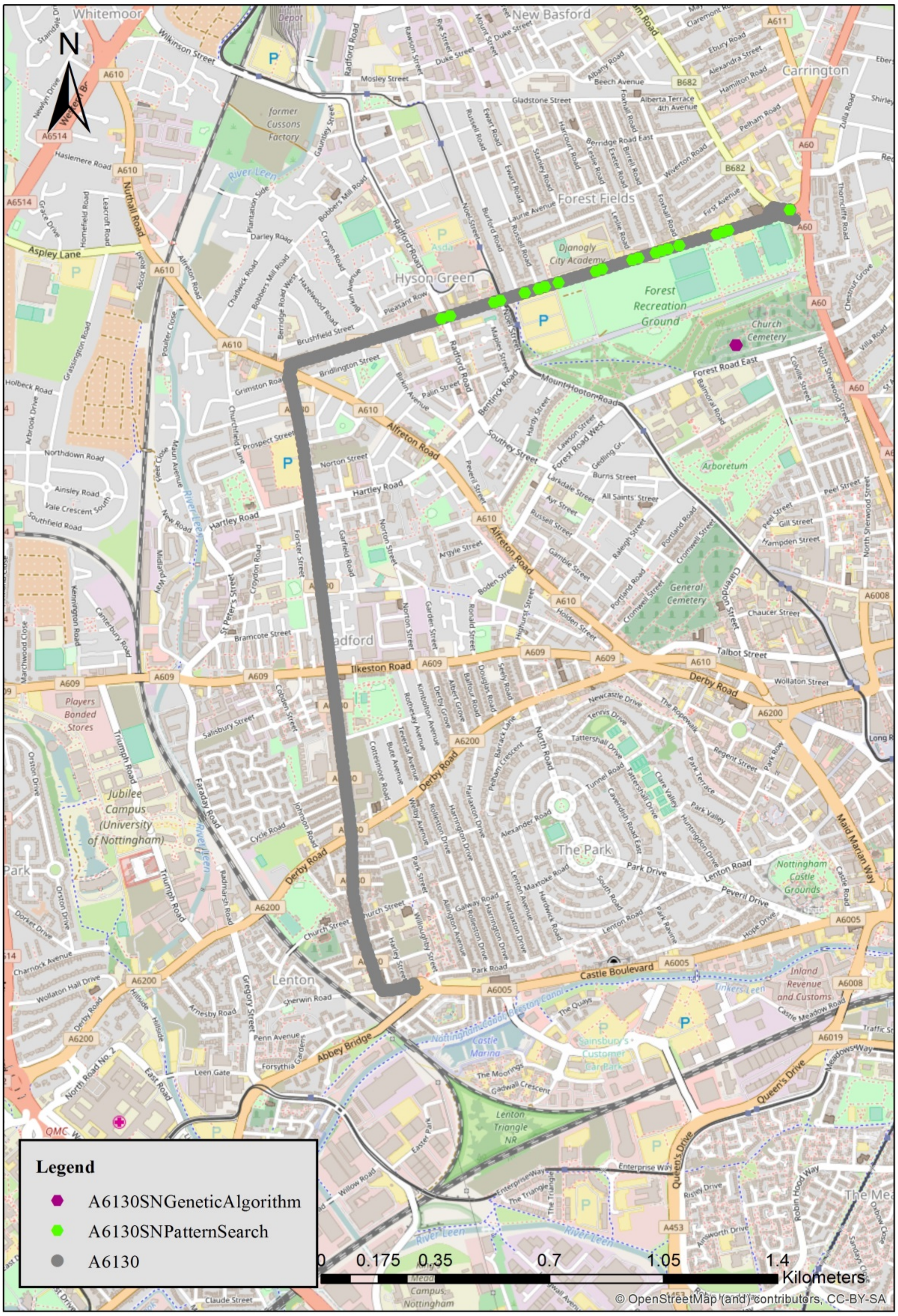


Figure 27 Genetic Algorithm and Pattern Search plotted results for A6130

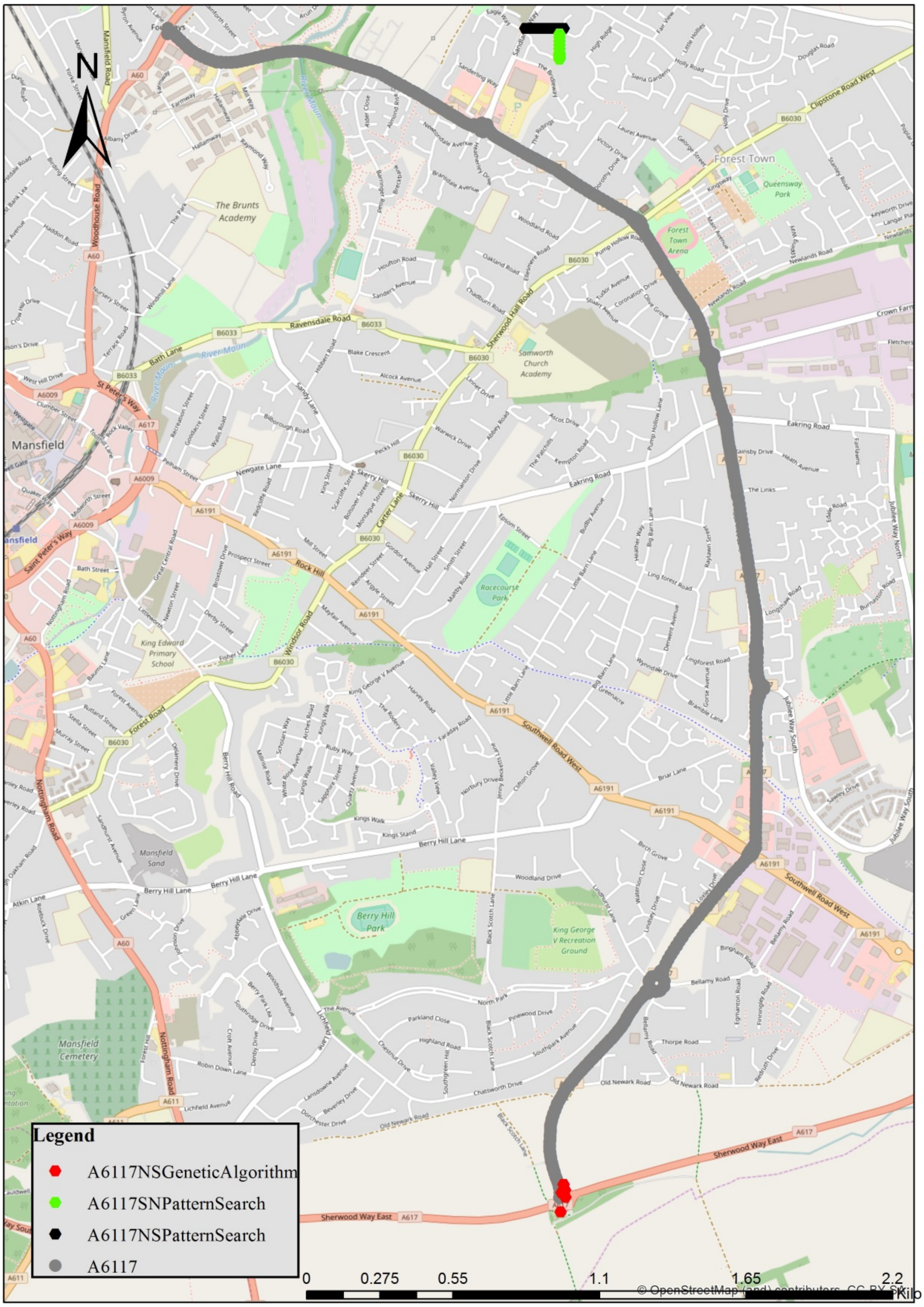


Figure 28 Genetic Algorithm and Pattern Search plotted results for A6117

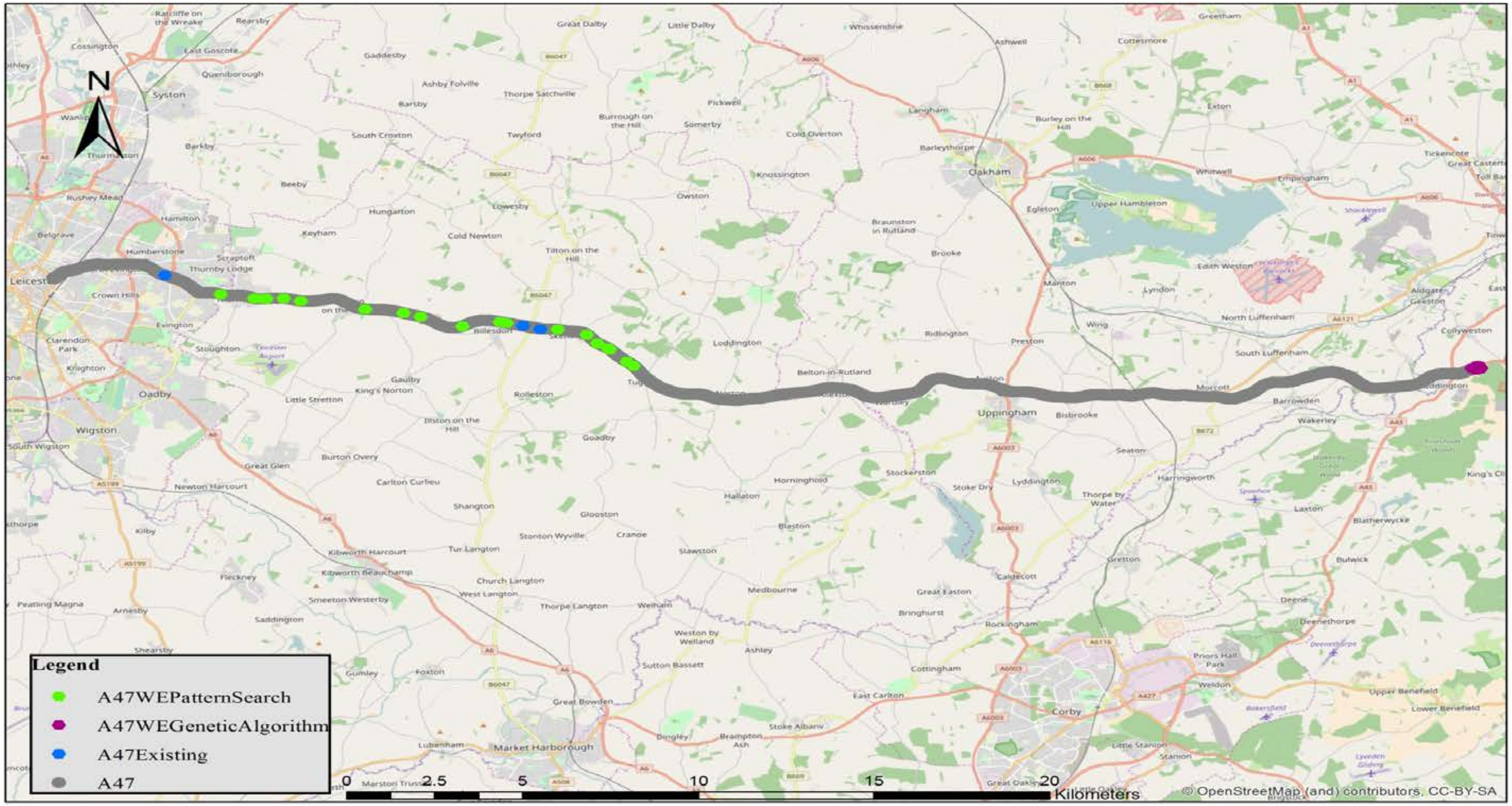


Figure 29 Genetic Algorithm, Pattern Search and Existing plotted results for A47

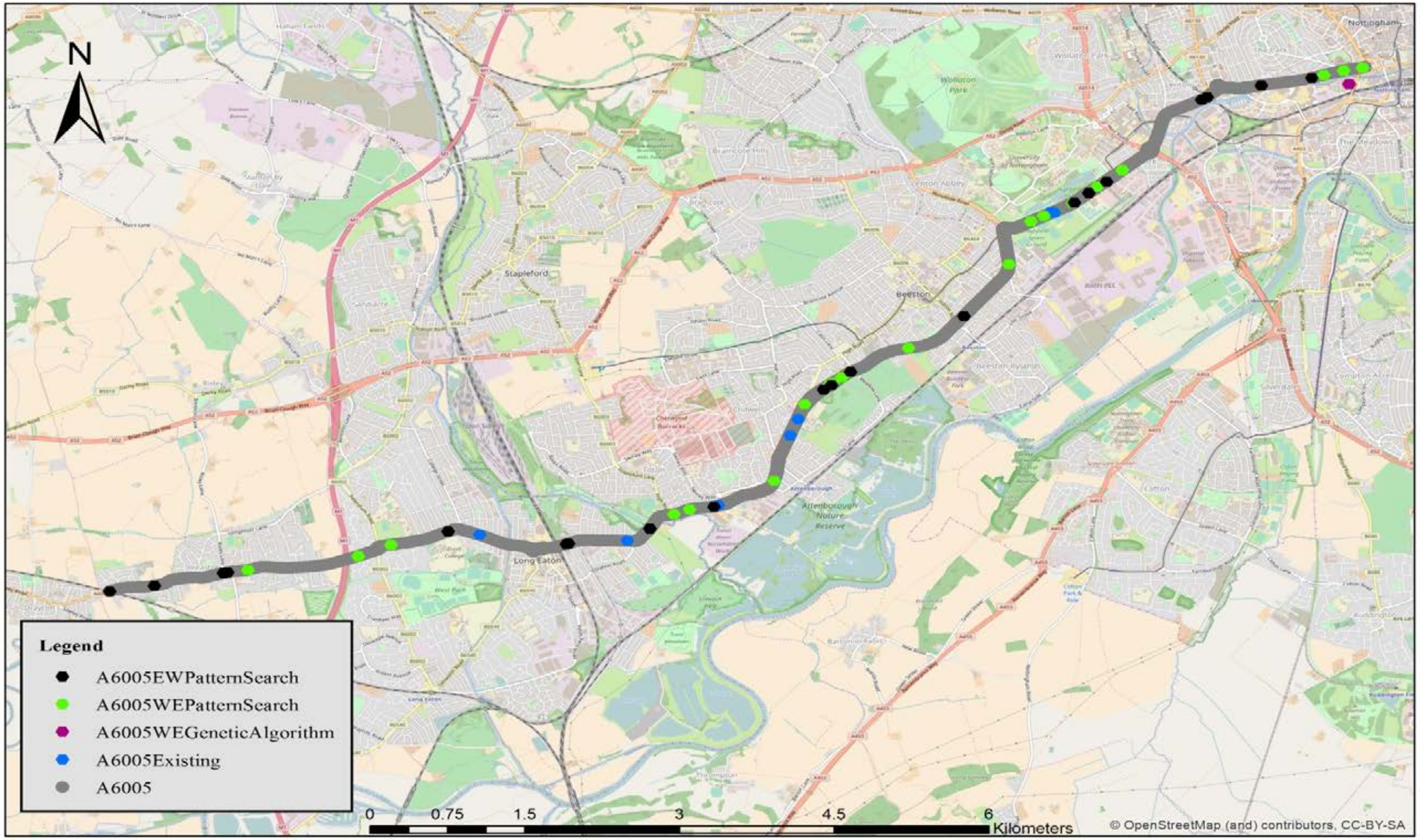


Figure 30 Genetic Algorithm, Pattern Search and Existing plotted results for A6005

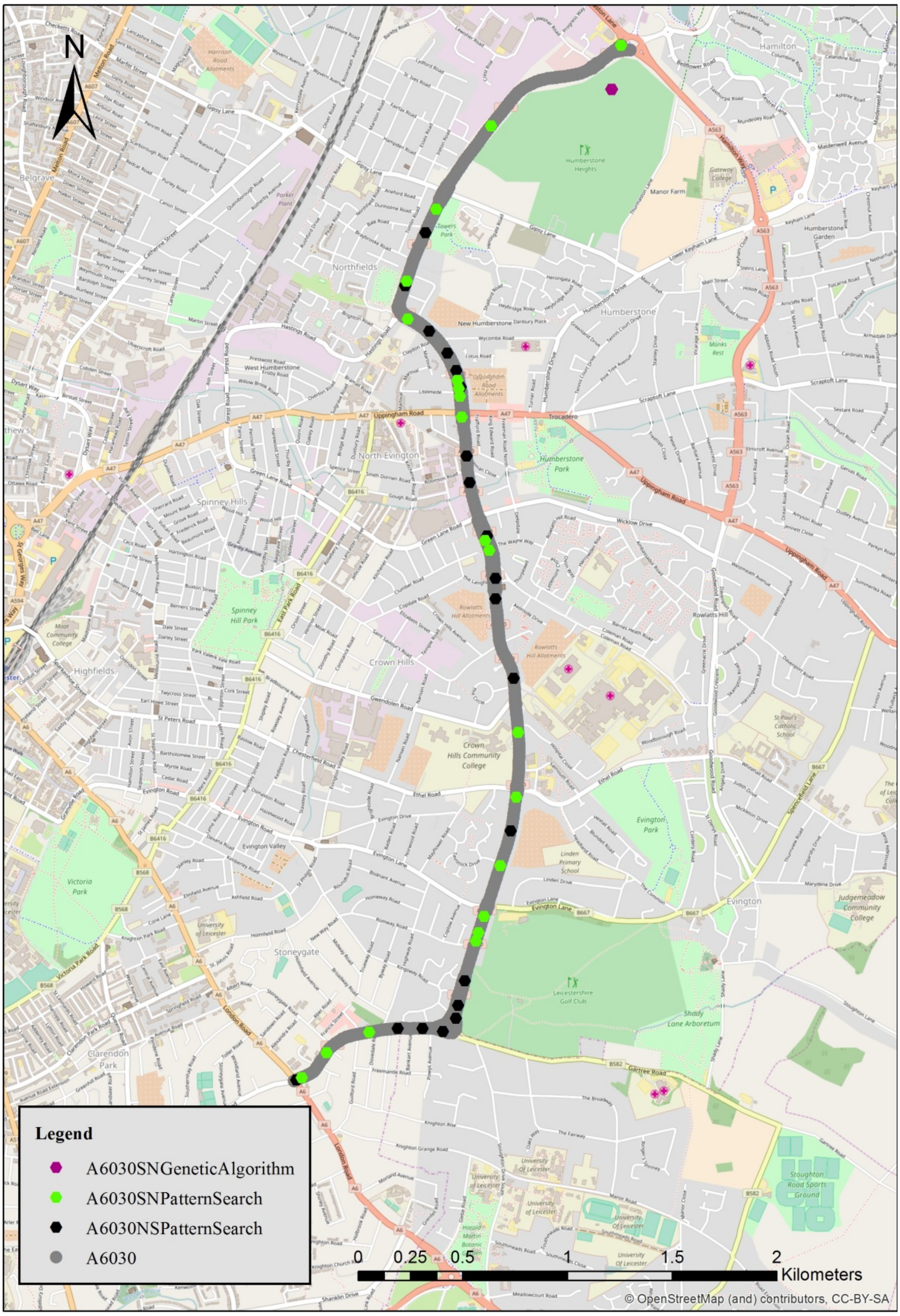


Figure 31 Genetic Algorithm and Pattern Search plotted results for A6030

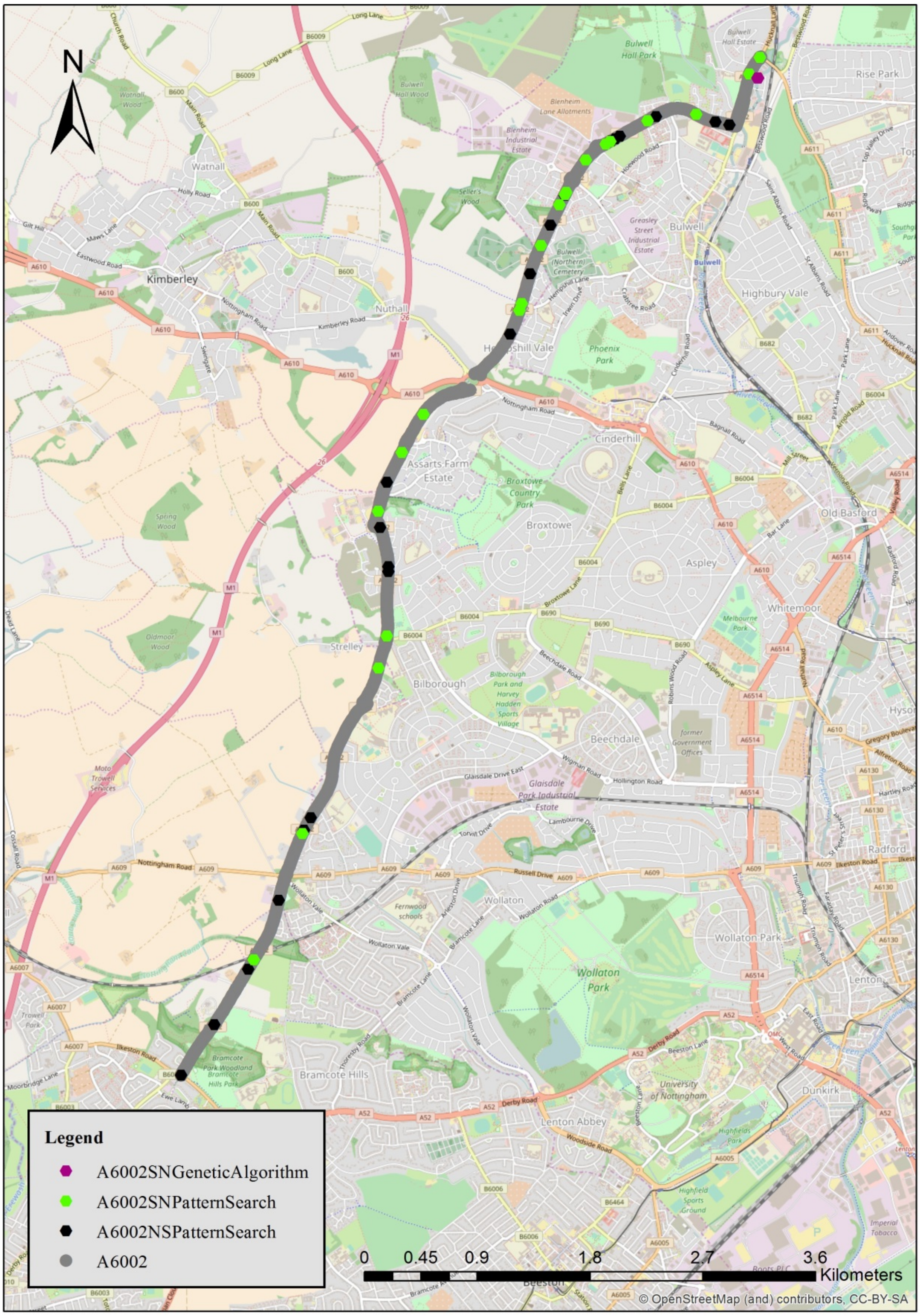


Figure 32 Genetic Algorithm and Pattern Search plotted results for A6002

6.4 Model validation

Some validation methods used for evolutionary algorithms include objective function evaluation, construct validity, statistical models, blinding techniques to eliminate bias, convergence and design validation (Langerman and Ehlers, 2006). Some of these validation methods have been applied to the optimisation in this thesis and are presented in this section.

Design validation using statistical methods is being adopted as the method of validation for the optimisation techniques used. In validating the optimisation model there was the need to reconsider variables that will alter the number of road traffic accidents as well as the position of a speed control device. With the accident prediction model being the main determinant in predicting the number of accidents, some variables in the accident prediction model noted to have significance on the occurrence of an accident were altered. The main variables altered were the number of lanes, radius, speed limit, Annual Average Daily Traffic (AADT) and percentage heavy goods vehicles (HGV). An improvement in road infrastructure will result in a change in these variables. An increase in speed limit and the introduction of an additional lane will result in a change in the radius of the road. The introduction of an additional lane will have arisen from an increase in AADT with a possible increase in percentage HGV. Altering these variables will change and shift accident numbers and severities from one location along a road network to another and these variables are also noted in literature to have a significant association with road accidents (McDonald, 2004; Bhatnagar, 1994; Golob, Recker and Alvarez, 2004; Taylor, Baruya and Kennedy, 2002).

Variation in parameters applied include a 10mph increase in speed limit, the addition of one lane, a 2.4 percent increase in AADT, a 2 percent increase in percentage HGV and a 41% percent increase in radius. The speed limit of a road has a specified radius with which a road must comply with (Highways Agency, 2002), the percentage increase in radius resulting from an increase from one speed limit to another was noted and averaged and the value applied in this research. The percentage increase in figures for the AADT and HGVs were chosen based on the 2014 road traffic estimates (Department for Transport, 2015). The adjusted parameter values were used with all other parameters remaining the same to remodel the accident prediction. It was observed that accident numbers for *fatal and serious* accidents combined and for *slight* accidents changed. The predicted *fatal and serious* accidents was found to reduce compared with the validated *fatal and serious* accidents with the predicted *slight* accidents slightly increasing in comparison with the validated *slight* accidents. The effect of changes in accident numbers was analysed using a paired t-test statistical analysis. The purpose of the t-test was to determine if there is any statistical significance between the predicted empirical Bayes fatal and serious accidents and the validated empirical Bayes fatal and serious accidents.

The first null hypothesis H_o being that there is no difference between predicted empirical Bayes fatal and serious accidents and validated empirical Bayes fatal and serious accidents. The first research (alternative) hypothesis H_A tested was that there is a difference between predicted empirical Bayes fatal and serious accidents and validated empirical fatal and serious accidents.

The second null hypothesis H_o being that there is no difference between predicted empirical Bayes slight accidents and validated empirical Bayes slight accidents. The second research (alternative) hypothesis H_A tested was that there is a difference between predicted empirical Bayes slight accidents and validated empirical Bayes slight accidents.

Results from the t-test analysis are given in Table 24 to Table 27.

Models compared	Mean	N	Std. Deviation	Std. Error Mean
Fatal and Serious Empirical Bayes (Predicted)	0.112	1770	0.117	0.003
Fatal and Serious Empirical Bayes (Validated)	0.094	170	0.099	0.002

Table 24 Results of t-test paired sample statistics for predicted and validated Empirical Bayes fatal and serious accidents

Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Predicted – Validated (Fatal and Serious Empirical Bayes)	0.02	0.02	0.00	0.017	0.019	45.48	1769	0.000

Table 25 Results of t-test paired sample statistics for predicted and validated Empirical Bayes fatal and serious accidents

Analysing Table 24 and Table 25, it can be seen that there is a difference in the predicted Empirical Bayes fatal and serious accidents (M= 0.112, SD= 0.117) and the validated Empirical Bayes fatal and serious accidents (M=0.094, SD=0.099) condition; t (1769) =45.48, p=0.000.

Models compared	Mean	N	Std. Deviation	Std. Error Mean
Slight Empirical Bayes (Predicted)	0.550	1770	0.994	0.024
Slight Empirical Bayes (Validated)	0.552	1770	0.996	0.024

Table 26 Results of t-test paired sample statistics for predicted and validated Empirical Bayes slight accidents

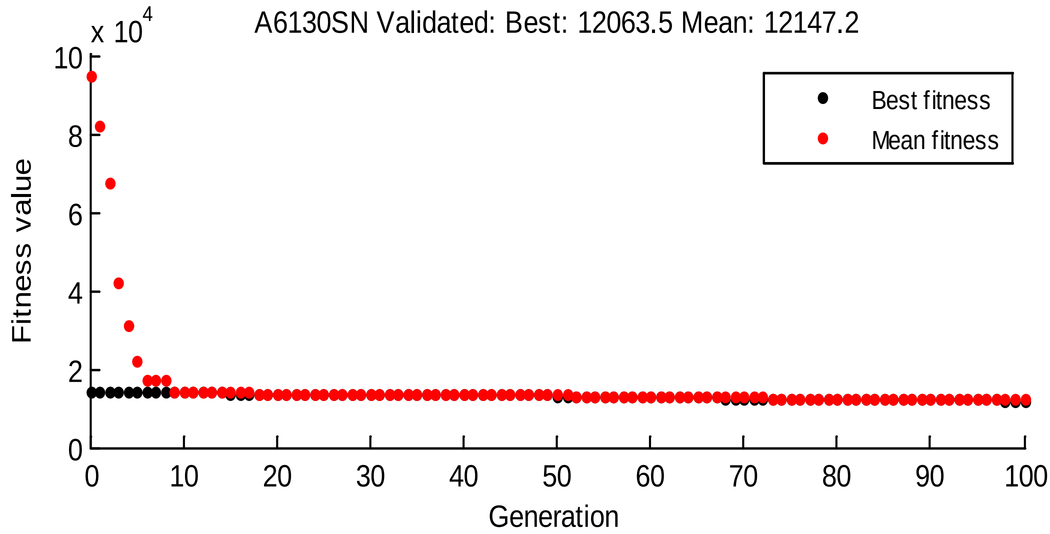
Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Predicted – Validated (Slight Empirical Bayes)	0.00	0.002	0.00	-0.002	-0.001	-30.34	1769	0.000

Table 27 Results of t-test paired sample statistics for predicted and validated Empirical Bayes slight accidents

Analysing Table 26 and Table 27, it can be seen that there is a significant difference in the predicted Empirical Bayes slight accidents (M= 0.550, SD= 0.994) and the validated Empirical Bayes slight accidents (M=0.552, SD=0.996) condition; t (1769) =-30.34, p=0.000.

Results obtained from a re-run of the optimisation using genetic algorithm and pattern search are given in Figure 33 to Figure 36.

.



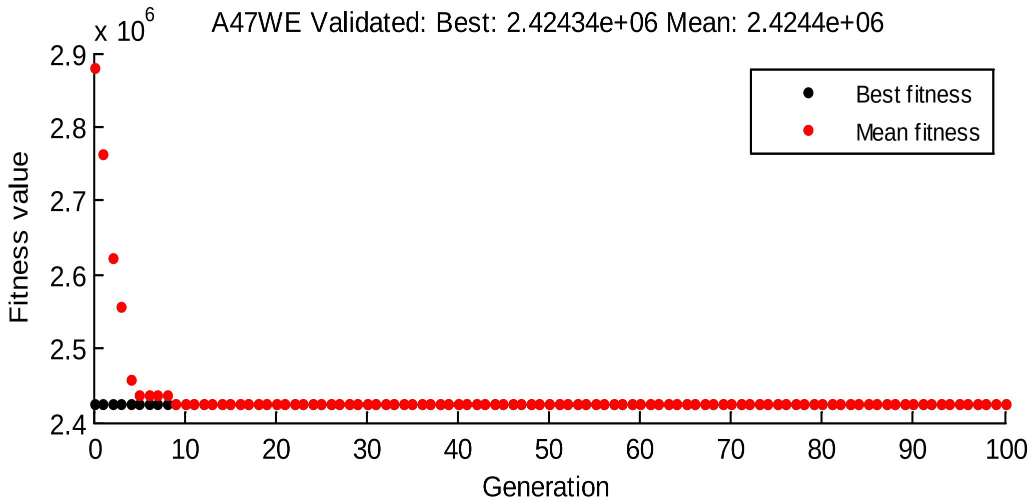
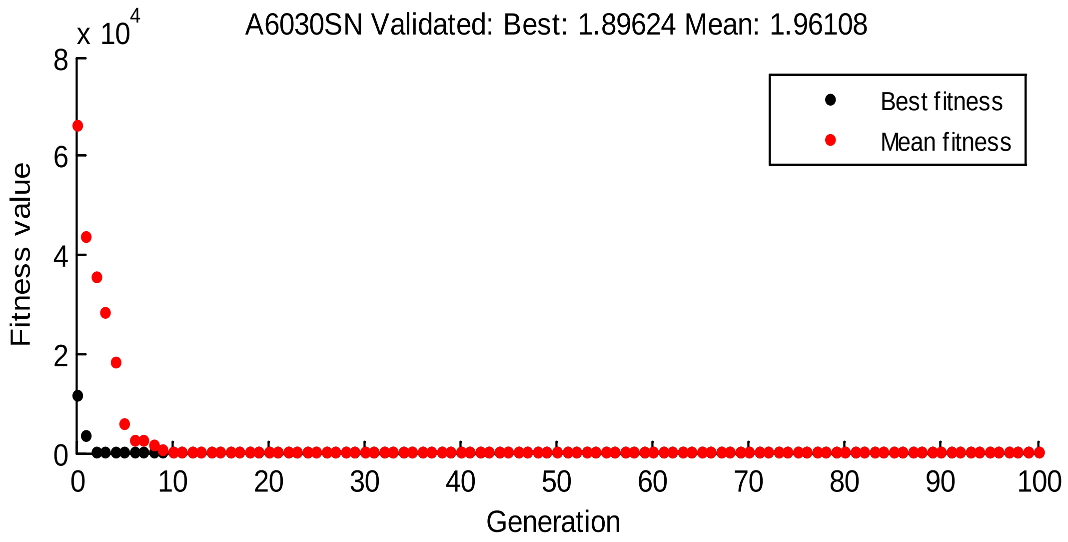


Figure 33 Validated results for Genetic Algorithm Optimisation



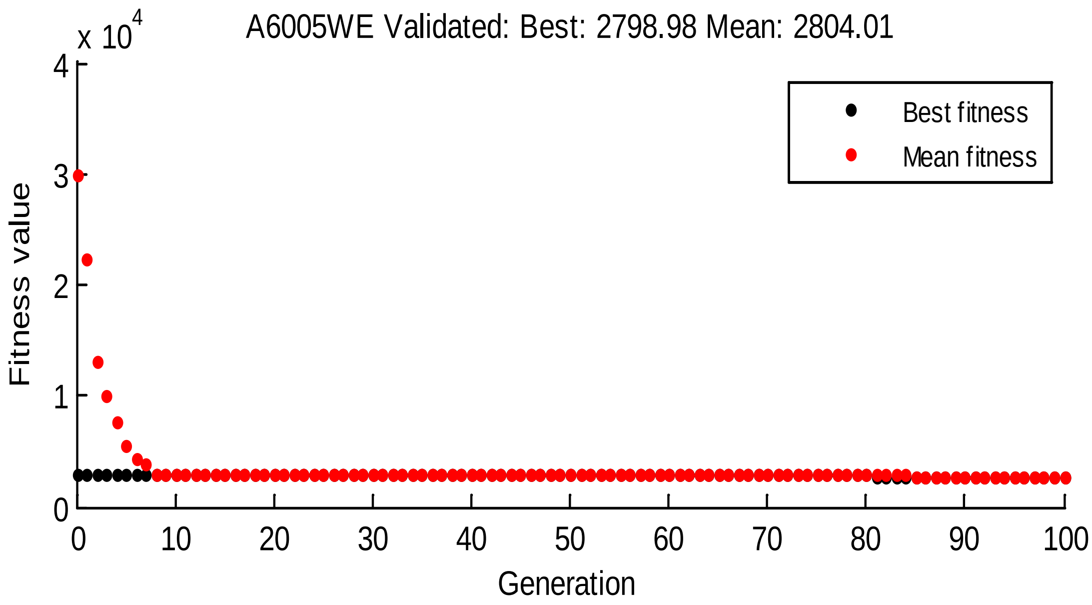


Figure 34 Validated results for Genetic Algorithm Optimisation

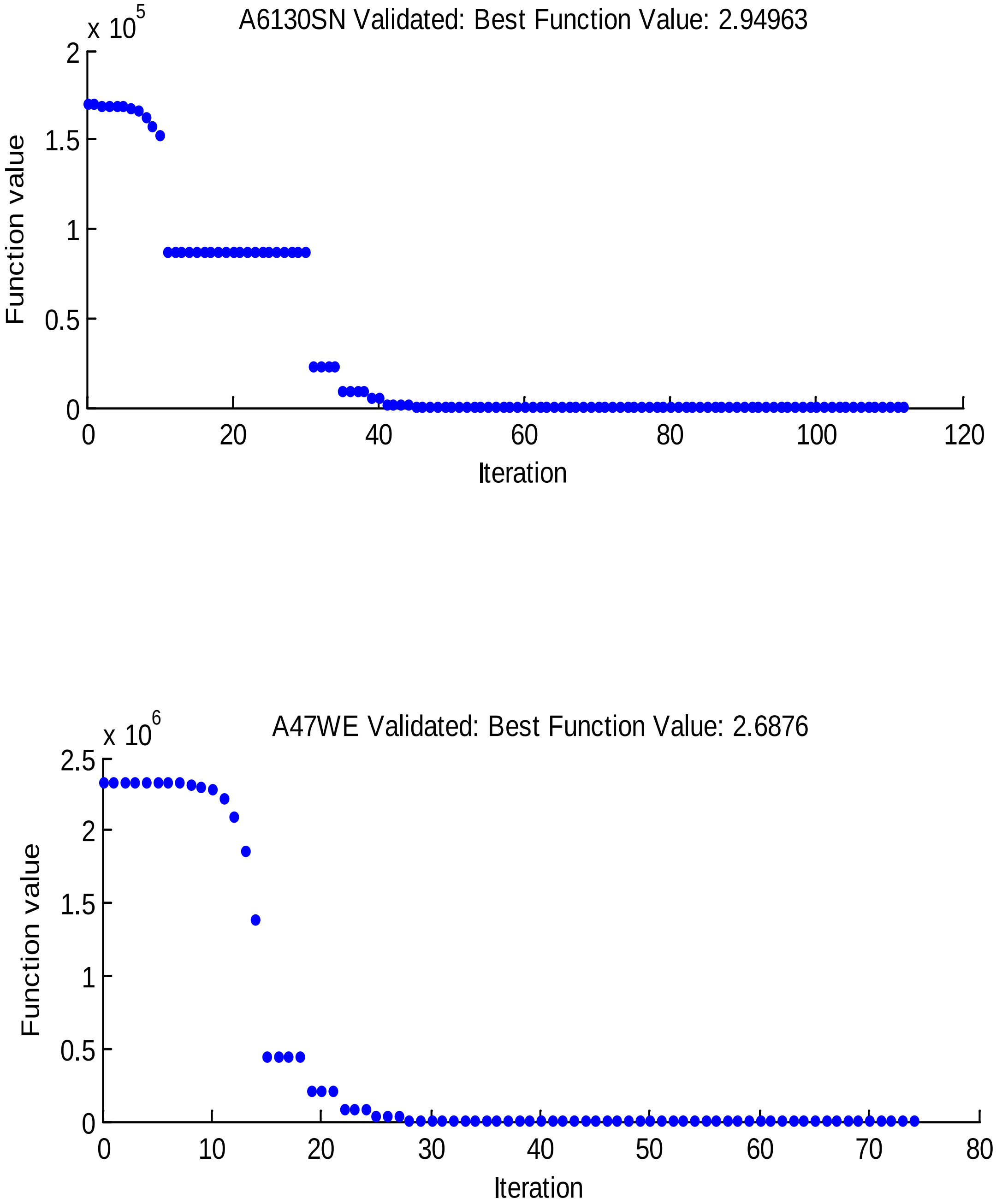
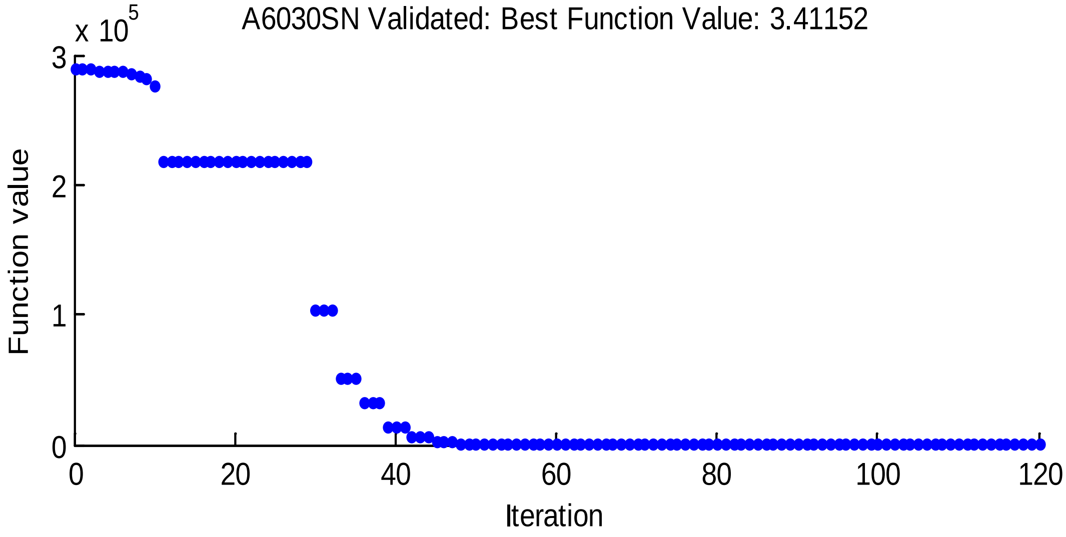


Figure 35 Validated results for Pattern Search Optimisation



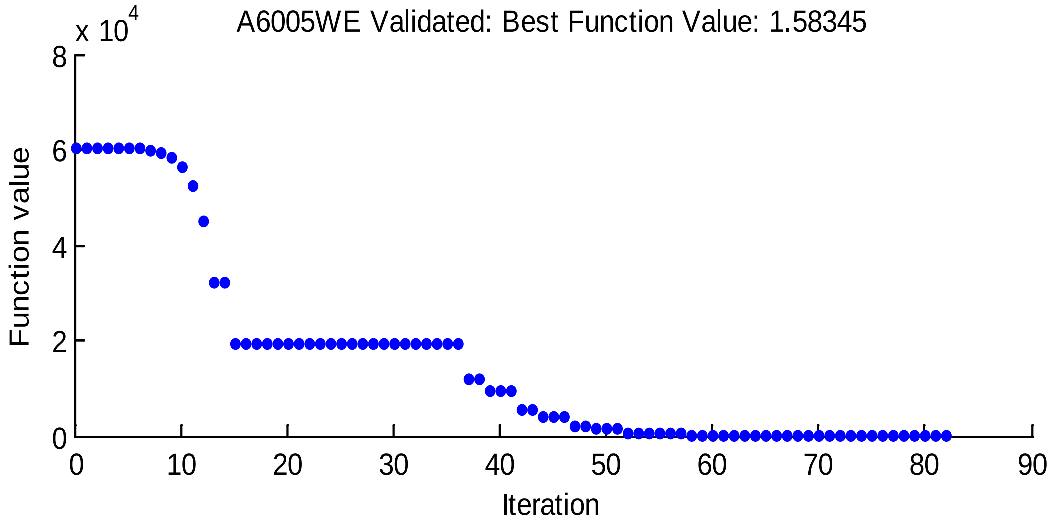


Figure 36 Validated results for Pattern Search Optimisation

A paired t-test was performed for the function values obtained from genetic algorithm and pattern search for the initial optimised locations and validated locations. This test was conducted in order to establish if there is any statistical significance between the paired function values. The compared t-tests were between the function values for

- Validated genetic algorithm and validated pattern search
- Validated genetic algorithm and the initial optimised genetic algorithm

- Validated pattern search and the initial optimised pattern search

The null hypothesis H_o were

1. There is no difference between the validated genetic algorithm function values and the pattern search function values.
2. There is no difference between the validated genetic algorithm function values and the initial optimised genetic algorithm function values.
3. There is no difference between the validated pattern search function values and the initial optimised pattern search function values.

The research (alternative) hypothesis H_A were

1. There is a difference between the validated genetic algorithm function values and the pattern search function values.
2. There is a difference between the validated genetic algorithm function values and the initial optimised genetic algorithm function values.
3. There is a difference between the validated pattern search function values and the initial optimised pattern search function values.

Results obtained for these hypothesis tests are given in Table 28 to Table 33. The values for paired t-test computed for the fitness values obtained from validated genetic algorithm and validated pattern search for the same road samples are given in Table 28 and Table 29.

Models compared	Mean	N	Std. Deviation	Std. Error Mean
Genetic Algorithm (Validated)	607649.34	80	1055520.75	118010.81
Pattern Search (Validated)	2.06	80	1.17	0.13

Table 28 Results of t-test paired sample statistics for validated optimisation function values

Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Genetic Algorithm (Validated) – Pattern Search (Validated)	607647	1055521	118010	372752	842541	5.15	79	0.000

Table 29 Results of t-test for validated optimisation function values

Analysing Table 28 and Table 29, it can be seen that there is a significant difference in the function values for genetic algorithm (M=607649.34, SD=1055520.75) and the function values obtained from pattern search (M=2.06, SD=1.17) condition; t (79) =5.15, p=0.000.

Computed paired t-test for the fitness values obtained from validated genetic algorithm and initial optimised genetic algorithm for the same road samples are given in Table 30 and Table 31.

Models compared	Mean	N	Std. Deviation	Std. Error Mean
Genetic Algorithm Validated	607649.34	80	1055520.75	118010.81
Genetic Algorithm Initial Optimised	9323.98	80	10419.83	1164.97

Table 30 Results of t-test paired sample statistics for validated and initial optimised genetic algorithm function values

Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Validated – Initial Optimised (Genetic Algorithm)	598325	1059808	118490	362476	834174	5.05	79	0.000

Table 31 Results of t-test for validated optimisation function values

Analysing Table 30 and Table 31, it can be seen that there is a significant difference in the function values for validated genetic algorithm (M=607649.34, SD=1055520.75) and the function values obtained from the initial optimised genetic algorithm (M=9323.98, SD=10419.83) condition; t (79)=5.05, p=0.000.

Computed paired t-test for the fitness values obtained from validated pattern search and initial optimised pattern search for the same road samples are given in Table 32 and Table 33.

Models compared	Mean	N	Std. Deviation	Std. Error Mean
Pattern Search Validated	2.06	80	1.17	0.13
Pattern Search Initial Optimised	2.80	80	1.00	0.11

Table 32 Results of t-test paired sample statistics for validated and initial optimised pattern search function values

Models compared	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
				Lower	Upper			
Pattern Search Validated – Pattern Search Initial Optimised	-0.73	0.67	0.07	-0.88	-0.59	-9.86	79	0.000

Table 33 Results of t-test for validated and initial optimised pattern search function values

Analysing Table 32 and Table 33, it can be seen that there is a significant difference in the function values for validated pattern search (M=2.06, SD=1.17) and the function values obtained from the initial optimised pattern search (M=2.80, SD=1.00) condition; t (79) =-9.86, p=0.000.

6.5 Limitations of model

The effect of driver behaviour has not been taken into account in the model since the focus of the research is aimed at using road feature characteristics. Also this model does not take into consideration the starting speed as well as changes in speed of vehicles along the link of road being investigated. The model assumes traffic to be free flowing and thus any effects of congestion are excluded.

Another point worth discussing is the dependence of the density of the number of speed control devices proposed with its effectiveness. Even though the effectiveness of speed cameras at reducing vehicle speed has been proved in various studies, it is noted from some of these studies that additional input (density) of speed cameras did not convert into comparable reductions in fatal injuries with preventive effect found to remain stable at about 21% (Elvik, 2011; Carnis and Blais, 2013). Also, even though speed cameras do have longer term effects in reducing vehicle speeds it is not the same at all locations. Considering the punitive measures associated with this type of device it is not anticipated that drivers will get complacent. For vehicle activated signs, despite the fact that they do not have punitive measures it is proposed that they be placed closer to locations where it is intended to reduce vehicle speeds in order to obtain maximum effect (Santiago-Chaparro, Chitturi and Noyce, 2012) rather than increasing the density when it is not desired. It is also worth stating that the density of speed control devices has cost implications. The model developed in this research was such that the best ‘x’ locations can be identified from the model so the density of these speed control devices can be effectively managed.

Spatial variations in speed and accidents which deals with the effect of traffic diversion or spill over from nearby roads as a result of improvement in the link being considered are not included in the model. Other spatial variations which include changes in speed limits on nearby roads that will divert traffic away from the road being investigated have not been factored into the model. This is because data on such a phenomenon is not readily available to be evaluated. Jones, Sauerzapf and Haynes (2008) however mention there is little evidence to suggest that crash migration occurs in the vicinity of cameras.

7 Discussion of findings

7.1 Introduction

Road traffic accidents continue to be a public health problem and indications are that this is likely to increase if no action is taken to curb the situation. Speeding continues to be a contributory factor in most road accidents. In the UK, vehicle speeds are managed through the use of speed cameras and vehicle activated signs in addition to other engineering measures. These are measures that have been proved to be effective at controlling and managing vehicle speeds.

The primary aim of this research was to develop an optimisation model to assist decision makers in determining the optimum location to mount a speed control device. This research developed an optimisation model using genetic algorithm and pattern search to optimally locate a speed camera or vehicle activated sign. This research started by carrying out a literature review of factors identified to contribute to road traffic accidents, accident prediction models and optimisation techniques. Some of the advantages and disadvantages of the accident prediction models were detailed out as well as a justification for choosing the one used. In the review of optimisation techniques, the applications of genetic algorithms and pattern search was discussed as well as the advantages, disadvantages and justification for adopting these techniques. Once accident causative factors had been identified, it became necessary to establish how data was to be obtained and used. An appropriate accident prediction model was required to be subsequently used in the optimisation model. With the absence of a suitable accident prediction model, there was the need to develop one. The accident prediction model developed included parameters identified in literature to contribute to road traffic accidents. It is worth reiterating that human factors are excluded from the model. The accident prediction models developed in Chapter 5 involved the accident frequency at a particular level of severity (*slight* or *fatal and serious* combined) to various road characteristics. A negative binomial model was first developed and this was improved upon using the empirical Bayes approach. The negative binomial empirical Bayes accident prediction models developed were for A-roads in the study area of Nottinghamshire and Leicestershire in the United Kingdom. Nottinghamshire and Leicestershire were chosen as the study area for this research mainly because local councils in these regions have historically cooperated with Loughborough University on projects. In addition to this, these areas have speed cameras and vehicle activated signs installed along roads which are of interest in this research. The accident prediction model was tested on an independent set of A-roads to validate the model. After developing the accident prediction model, an optimisation model was then developed using pattern search and genetic algorithm. The optimisation model was to help determine the optimum location to place a speed control device such as a speed camera subject to a given objective and constraints.

7.2 Accident prediction model

Approximately seven hundred and ninety (790km) kilometres of A-roads within Nottinghamshire and Leicestershire were randomly selected and used ensuring that they had the required data necessary for further analysis. In order to ascertain the validity of the accident prediction models developed, approximately 75 percent of the roads was used in developing the model with the remaining 25 percent used in validating and testing the model. All roads were split into homogeneous segments where a homogeneous segment of road is the length of road within which all risk indicating variables remain constant allowing for a uniform risk along that segment of road. Once a variable changes, the road moves into the next homogeneous segment.

The 75 percent of roads was made up of 604,523 metres (605km) of roads which comprised 4177 homogeneous segments, 350 fatal and serious accidents, 1,646 slight accidents, 1,580 junctions and 18,240m length of cycle route. The 25 percent of roads was made up of 185,000 metres (185km) of roads which comprised 1,771 homogeneous segments, 159 fatal and serious accidents, 960 slight accidents, 1,872 junctions and 11,120m of cycle route.

The predictor variables used in the negative binomial accident prediction model were the road direction, number of lanes, slope, radius, Annual Average Daily Traffic (AADT), Heavy Goods Vehicle (HGV), speed limit, homogeneous segment length, average speed, presence of cycle lane, length of cycle lane, presence of a junction or not, number of junctions and number of pedestrian crossings. After the initial output of the slight accident prediction model and fatal and serious accidents combined prediction model, it was noted that some of the variables used were not statistically significant. Variables found to be statistically significant ($p < 0.05$) in the slight accidents negative binomial model were presence of junctions (Jtn), direction of travel along the road (Dxn), the summation of junctions (JtnsSUM), number of lanes, percentage Heavy Good Vehicles (HGV), speed limit, homogeneous segment length of road (HSegLength), the logarithm of the Annual Average Daily Traffic (LogAADT) and logarithm of the radius (LogRadius). Variables found to be statistically insignificant ($p > 0.05$) were excluded from the model and rerun. For the fatal and serious accidents negative binomial model, variables found to be statistically significant were the presence of junctions (Jtn), direction of travel along the road (Dxn), speed limit, homogeneous segment length of road (HSegLength), the logarithm of the Annual Average Daily Traffic (LogAADT) and logarithm of the radius (LogRadius). Variables found to be statistically insignificant were also removed from the model and rerun.

The categorical variables in the rerun fatal and serious accidents model were the presence or absence of junctions (Jtn) along a homogeneous segment of road and the direction of travel. It was noted that 74.5 percent of the homogeneous segments of roads had no junctions. Despite the fact that a lower proportion of homogeneous segments of roads had junctions, there was evidence that the presence of junctions has an influence on fatal and serious accidents and this is consistent with previous findings. A negative relationship existed between the presence of junctions and fatal and serious accidents. Almost equal proportions of roads used were in the south-north (37.8%) and north-south (37.2%) directions with the east-west (12.3%) and west-east (12.8%) directions also sharing a relatively equal proportion of roads in those directions of travel. The direction of travel was found to have an effect on the number of fatal and serious accidents.

In the rerun slight accidents model, categorical values observed to be statistically significant were the presence or absence of junctions (Jtn), presence or absence of a cycle route (PresofCR) and direction of travel along a homogeneous segment of road. For the presence or absence of cycle route, a great proportion of homogeneous segments of roads had no cycle route with 3.9 percent of homogeneous segments of roads studied having cycle routes. However, it was found that at significance level of

0.05, the presence or absence of a cycle route did have an effect on the number of slight accidents.

7.2.1 Fatal and serious accidents

In the case of continuous variables for fatal and serious accidents, the speed limit, homogeneous segment length, logarithm of the Annual Average Daily Traffic and the logarithm of the radius of the road were represented as statistically significant in the rerun model. The speed limit of the road which was found to have a statistically significant effect on the number of fatal and serious accidents had a negative relationship with fatal and serious accidents. Some studies (Taylor, Lynam and Baruya, 2000; Ossiander and Cummings, 2002) have however indicated that higher speeds signify an increase in accidents which is not consistent with the findings obtained in this research and this can be explained to be the result of the small number of observations for fatal and serious accidents. The homogeneous segment of road which represents a length of road with all risk indicating variables remaining the same showed a negative relationship with the number of fatal and serious accidents. A less than 5 percent significance level was applied to the model and it was observed that there was sufficient evidence to suggest that the length of homogeneous segment of road has an influence on the number of fatal and serious accidents. The logarithm of the annual average daily traffic (AADT) which is sometimes referred to as the traffic flow was noted to have a positive effect on the number of fatal and serious accidents. This means with an increase in traffic volume, road accidents are likely to increase. This finding is similar to that obtained in studies by Ceder and Livneh (1978), Ceder (1982), Golob and Recker (2001), and Martin (2002). The logarithm of the radius was found to have a negative influence on the number of fatal and serious accidents implying that an increase in the radius of curvature results in a decrease in the number of accidents. There was also sufficient evidence at 95 percent significance level to suggest that the radius has an effect on the number of fatal and serious accidents. An increase in road curvature is expected to provide adequate comfort to drivers when negotiating curves and thus reduce the incidence of an accident in comparison to shorter and tighter curves which are a discomfort to drivers and at high speeds vehicles tend to lose control resulting in road accidents. Findings observed in this research were shown by Milton and Mannering (1998) and Haynes et al., (2007). A study by Berhanu (2004) however proved the opposite with a decrease in the radius of curvature showing a decrease in the number of accidents. The most generally used and important parameter in deriving operating speeds along horizontal curves has been identified as the curve radius (Bennett, 1994; Abdul-Mawjoud and Sofia, 2008). Curve radii smaller than 400 metres cause road accidents with curve radii less than 600 metres shown to be over represented in roads accidents (Choueiri and Lamm, 1987; Johnston, 1982). McDonald (2004) noted a 34 percent increase in the frequency of accidents per sharp curve per kilometre for single vehicle accidents.

The highest correlation of parameter estimates within the fatal and serious accidents model was found to occur between the logarithm of the radius and the length of homogeneous segment of road. A correlation value of -0.317 was obtained suggesting that multicollinearity is not an issue for the remaining explanatory variables.

7.2.2 Slight accidents

For the slight accidents model, the continuous variables represented were the sum of junctions within a homogeneous segment of road, number of lanes, percentage heavy good vehicles, speed limit, length of homogeneous segment of road, logarithm of the annual average daily traffic and logarithm of the radius. The summation of junctions within a homogeneous segment of roads studied indicates a positive relationship with the number of slight accidents. This implies that an increase in the number of junctions will result in an increase in the number of junctions. Even though Haynes et al. (2008) found that the number of junctions per kilometre had a negative association with crash rates, this finding is attributed to the small numbers of data used distributed over a large number of roads. Yagar (1984) also showed that junctions were statistically significant on speeds and accidents. The effect of the number of lanes on road traffic accidents has been adequately studied in literature with consistent results noted amongst researchers. Wang, Quddus and Ison (2011a), Kononov, Bailey and Allery (2008); Noland and Oh (2004); Garber (2000); Abdel-Aty and Radwan (2000) and Milton and Mannering (1998) were all able to show that an increase in the number of lanes results in an increase in road traffic accidents. This research was able to also show that an increase in the number lanes results in increased slight accidents with evidence also provided by the data that at a significance level less than 0.05 the number of lanes has an effect on slight accident numbers.

Heavy goods vehicles are noted to be over represented in severe road accidents largely attributed to the high masses of these vehicles resulting in severe outcome for other road users who get involved in the accident (Evgenikos et al., 2016; Khorashadi et al., 2005). This observation was however not consistent with findings from the current research since the effect of road accidents on fatal and serious accidents was found to be statistically insignificant. For slight accident numbers, the percentage of heavy good vehicles was found to have a negative effect meaning that an increase in heavy goods vehicle results in a decrease in slight accidents. There was also enough evidence provided by the data at the 0.05 significance level to show that the percentage of heavy goods vehicles has an influence on slight accident numbers. This outcome has been discussed in other studies (Anastasopoulos and Mannering, 2009; Shankar, Milton and Mannering, 1997; Miaou, 1994) noting that an increase in heavy goods vehicles results in a decrease in the prevalence of overtaking vehicles and lane changing behaviour resulting in fewer accidents.

The speed limit of the road was observed to have a statistically significant result on the number of slight accidents with a negative effect also observed. Johansson (1996) showed that reduced speed limits can reduce accident numbers that involve minor injuries and vehicle damage. Aljanahi et al., (1999) also arrived at a similar result to Johansson (1996) by finding that the number of accidents would reduce if there is a lowering of speed limit. Ossiander and Cummings (2002) were also able to show that an increase in speed limit leads to a high fatality rate. Even though these studies show results contrary to what was obtained in the current research, it is worth stating that the negative relationship observed between the speed limit and the number of slight accidents was not huge (-0.022).

The length of homogeneous segment of road was found to have a positive effect on slight accident number showing that an increase in homogeneous segment of road will results in an increase in slight accidents. The model data also showed strong evidence that the length of homogeneous segment of road has an effect on the number of slight accidents. The use of homogeneous road segments in road accident modelling has been used in studies by Milton and Mannering (1998), Abdel-Aty and Radwan (2000) and Li et al., (2007). Li et al., (2007), discussed the need to capture the risk levels of different road segments by differentiating the risks into separate directions of the road.

In order to reduce the large variations existing within values for the Annual Average Daily Traffic (AADT), a logarithm of the variable was used. The logarithm of the AADT was noted to have a positive effect on the number of slight accidents. There was evidence suggested by the data that at a 0.05 significance level the logarithm of the AADT has an effect on the number of slight accidents.

A logarithm of the radius was also used to decrease the extent of variations occurring in the values for the radius. The logarithm of the radius (LogRadius) was found to have a negative effect on slight accident numbers with evidence suggesting that at a 0.05 significance level, the logarithm of the radius has an effect on the number of slight accidents.

7.2.3 Other statistical tests

An investigation into the correlation coefficients between independent variables was carried out with a value of -0.59 obtained as the highest value of correlation occurring between the logarithm of the annual average daily traffic and the number of lanes. The value obtained for the correlation suggests that multicollinearity is not an issue in the model for slight accidents. Residual plots usually evaluate if the observed error or residuals (observed – predicted) is uniform with stochastic error. It is possible to find out if residuals are predictable with random errors in which case the residual plots offers an opportunity to enhance the model if residuals indicate model to be symmetrically inaccurate. For the fatal and serious accidents model, there was a pattern of increasing residuals as the observed number of accidents increased. For the slight accidents model a similar outcome was observed with about 3 points indicating the likelihood of the presence of outliers. The Goodness of fit for both the fatal and serious accidents model (1.116) and the slight accidents (2.385) models had Value/df values for the Pearson Chi-Square to be greater than 0.05 which indicate a good fit of the model to the data. Parameters which were statistically insignificant during the initial development of the models were removed to allow for better accuracy of the models.

To refine and improve upon the results obtained from the Negative Binomial model, an empirical Bayes method was used. The empirical Bayes approach helps control for regression-to-the-mean effects. The observed, negative binomial predicted and empirical Bayes accidents predicted for fatal and serious accidents combined and slight accidents was compared. A paired t-test method was used by comparing the observed data to each of the accident prediction models and also comparing the two accident prediction models. Observed slight accidents were compared to slight accidents from the negative binomial model; observed slight accidents were compared to empirical Bayes slight accidents and negative binomial slight accidents were compared with empirical Bayes slight accidents. The same pairs of comparison were carried out for fatal and serious accidents combined. The null hypothesis was that there was no difference between the two pairs of results being compared and the alternative hypothesis being that there is a difference between the two pairs of results being compared. All paired results using the t-test showed a significant difference for both slight accidents only and fatal and serious accidents combined with the exception of observed slight accidents (M= 0.54, SD= 1.497) and the empirical Bayes slight accidents (M= 0.55, SD= 0.994) which showed no evidence of a difference under the condition $t\ (1769) = -0.505$, $p=0.614$. The 95 percent confidence interval (-0.041 to 0.024) for the t-test carried out between observed slight accidents and the empirical Bayes slight accidents was noted to contain the number zero implying no significant difference.

A residual plot for fatal and serious accidents generated by the empirical Bayes method showed residuals (observed – predicted) increasing with increasing fatal and serious accident numbers with a very little improvement in the residuals. For the fatal and serious accidents empirical Bayes model, residuals varied from -0.5 to 2.7 (see Figure 17) which was found to be slightly lower than the residuals obtained from the fatal and serious accidents negative binomial model which ranged from -0.8 to 2.9. There was a refinement in the residual plots obtained for the slight accidents from the empirical Bayes method in comparison to that obtained for slight accidents from the negative binomial model. Residuals were found to range from -1 to 9 (see Figure 17) for the slight accidents empirical Bayes model which are much lower as compared to residual values ranging from -28 to 17 obtained for the slight accidents negative Binomial model (see Figure 16). In addition to this, the increasing pattern in the residuals is observed to have been removed to a large degree. This indicates the extent to which the model has dealt with the effect of regression-to-the-mean. The empirical Bayes model for the slight accidents can thus be said to be better than the slight accidents negative Binomial model with respect to the model fit for slight accidents.

7.3 Optimisation model

Optimisation problems apart from having a number of objectives, some of which may be conflicting are also normally constrained by some limitations. In Chapter 6 Genetic Algorithms and Pattern Search optimisation techniques were used to optimise the location of speed control devices ie. speed cameras with the main objective being to minimise the costs associated with fatal and serious accidents and slight accident numbers. Three main objectives were derived with the first being to minimise the total set-up cost of a speed control device. The second objective was to minimise the maintenance cost of a speed control device. The final objective was to minimise the total lost cost associated with an accident at the required level of severity. The optimal number of speed control devices in the form of speed cameras in this research was obtained by using the first and third objective which minimises total set-up cost as well as minimise the costs associated with the number of accidents at the required level of severity. The task of minimising the costs associated with the number of accidents at the required level of severity directly affects and includes the total set-up cost. This is because the available amount of money will affect the possible number of speed control devices which can be deployed. The optimisation was set up to provide information for twenty locations so depending on set-up cost constraints, the first ‘x’ numbers of locations can be used. The set of possible coordinate points given as (x,y) was used to denote the chromosome in the genetic algorithm and the same (x,y) search locations were used for the pattern search.

Rules and guidance for the National Safety Camera Programme for England and Wales as recommendations provided by the Department for Transport (2006) on speed were used in addition to road design standards as specified in the Design Manual for Roads and Bridges (Highways Agency, 2002). This approach was adopted because it represents a more practical and considered way to address the optimisation problem. Also these are guidance notes that have been approved by industry and this research was able to combine the relevant standards together to help determine the optimum location to place a speed control device. The examples used in this research show that the method can be adopted and used by planners, engineers and decision makers to optimally locate speed cameras or vehicle activated signs for speed reduction. The number of accidents at specified levels of severity was required to be obtained over a given length of road varying from 1000m to 3000m. This length of road was chosen because it was found that at lower lengths of road the number of fatal and serious accidents combined or slight accidents required was not achievable. Additionally, the guidelines for mounting speed cameras was used as a guide (Parliamentary Office of Science and Technology, 2004) as well as research revealing the distance over which these speed control devices are effective Li, Graham & Majumdar (2013), Høye (2014), Høye (2015), Høye (2015a). Other criteria in the form of radius, gradient, number of lanes and speed limit had to be satisfied in addition to the number of fatal and serious accidents combined or slight accidents. The number of fatal and serious accidents combined was to be greater than or equal to 3 with the number of slight accidents being greater than or equal to 15.

7.3.1 Pattern Search

Results of proposed coordinate locations obtained on plotted maps for pattern search were found to be mostly located referenced to the road centreline as was expected. The A6002SN shows the fitness function remained stable from iteration 0 to approximately iteration 7 after which there was a drop in the function value up to iteration 12. From iteration 12, the fitness function value dropped from a value of 30000 to about 15000 after which this value remained stable from iteration 12 to iteration 32. After iteration 32 the function value gradually dropped over short iteration periods of about 2 iteration intervals. The gradual drop occurred over approximately three instances of 2 iteration interval periods. At about iteration 45, the fitness function value remained stable and there was no further improvement in the function value. The A6002NS shows the fitness function starting at a low value of 0.08 and remaining stable for the first 8 generations. After that, the fitness value drops to 0.06 from iteration 8 to 11 with another drop occurring from iteration 11 to 13 at a function value of 0.05. This is followed by a final drop in fitness value to 0.02 and this value remained stable from iteration 13 to the final iteration point of 27.

The A6130SN presents the function value from iteration 0 to approximately iteration 5 remaining stable with a reduction in fitness value occurring between iteration 5 and iteration 10. After iteration 10, the function value drops to about half its value at iteration 10 maintaining this function value from iteration 11 to iteration 30. At iteration 30 the function value experiences another drop to about half its value at iteration 11 to a value of approximately 75000. The function value remained constant over a relatively short period of iteration from iteration 30 to iteration 34. A slight drop in function value occurred after iteration 34 with the function value remaining relatively stable over a short iteration period from iteration 35 to iteration 38. At iteration 39 another very small drop in function value occurred continuing with no reduction in function value to iteration 40 and no further refinement in function value being observed after iteration 40. The A47WE shows a relatively stable function value from the start of iteration to iteration 9 after which there is a steep drop in the function value from iteration 10 to iteration 14. From iteration 14, the function value remains constant up to iteration 18 where a drop in function value occurs. From iteration 19 to iteration 21 the function value remains constant and another slight drop in function value occurs at iteration 22. The function value remains constant over a very short period from iteration 22 to iteration 24 after which the model experiences no further improvement in the function value.

The A6117SN shows the fitness value remaining constant from the start of iteration to an iteration point of 5 after which there is a drop in fitness value from 22500 to 7400 between iteration 5 and iteration 11. From iteration 11 to the end of iteration, the function value remained stable at a value of approximately 7400. The A6117NS reveals the function value drop gradually from the start of iteration to iteration 5. From iteration 5, a steep drop in function value occurs up to iteration 10. The drop remains stable over a relatively short period from iteration 10 to about iteration 12. A slight drop in function value occurs again at iteration 12 and remains constant over another relatively short period from iteration 12 to iteration 14. From iteration 15, the function value stays constant with the model not providing any further improvement in the function value.

The A6030SN displays the function value remaining constant from the beginning of iteration to approximately iteration 8 after which the function value experiences a slow reduction from iteration 8 to iteration 10. From iteration 10, a steep drop in function value occurs and this was followed by a constant function value from iteration 10 to iteration 30. From iteration 30 to iteration 41, a drop in function value occurs four times on average at 3 iteration intervals terminating in a constant function value from iteration 41 to the end of iteration. The A6030NS displays the function value starting off at a value of 0.55 occurring at the start of iteration to iteration 4. At iteration 4, there is a slight drop in function value to 0.5 occurring between iteration 4 and iteration 7.5. The function value drops from 0.5 to 0.05 at iteration 7.5 and this function value remains constant to the end of iteration.

In the A6005WE, the plot shows the function value staying constant from the start of iteration to iteration 9. The function value experiences a drop in value from iteration 9 to iteration 14. The function value remains constant from iteration 15 to iteration 35 after which a stepwise drop in function value occurs. The short iteration stepwise drop in function value terminates at around iteration 50 where the model fails to produce any further improvement in the function value to the end of iteration. The A6005EW shows the function value starting at a figure of 2.5 from the beginning of iteration to iteration 5. From iteration 5, there is a drop in the fitness value to a value of 1.05 up to iteration 11. After iteration 11 there is another slight drop in the fitness value to 1.008 after which it remains stable and unchanged to the end of iteration. The A6005WE shows a better gradual refinement in the function value from start of iteration to completion in comparison to the A6005EW which even though produces a very low final fitness value than the A6005WE does not gradually refine the fitness value.

Computation times obtained from pattern search varied from a minimum of 13 minutes to a maximum of 595 minutes. The large variation in computation time was observed in the different lengths of road sections used. The shorter road segments run for a shorter period of time with the longer road segments taking more computation time. The A47 investigated is the third longest road used in this research with a total length for both directions of traffic being 87920m so the high computation time can be accounted for by this. Generally, pattern search plots produced were found to start from a high function value and progress towards a refined lowered function value. The initial optimised function values from pattern search are compared with the validated function values from pattern search using a t-test. The null hypothesis tested was that there is no difference between the initial optimised pattern search function values and the validated pattern search function values. The alternative hypothesis was that there is a difference between the initial optimised pattern search function values and the validated pattern search function values. Results using the t-test showed a significant difference in the function values for the validated pattern search (M=2.06, SD=1.17) and the initial optimised pattern search function values (M=2.80, SD=1.00) under the condition t(79)= -9.86, p=0.000. The 95 percent confidence interval was from -0.88 to -0.59.

7.3.2 Genetic Algorithm

The genetic algorithm optimisation was carried out on the same roads used for the pattern search. However for roads such as the A6002NS, A6030NS, A6005EW and A6117SN no optimal solutions were obtained. Results of proposed locations identified and plotted on maps for genetic algorithm were found not to be reference to the road centreline as proposed. Results obtained from genetic algorithm were also found to be inconsistent in certain instances with some proposed locations for the speed control devices placed more than 100m away perpendicularly from the road centreline. Consistent with other studies (Whitley et al, 1998; Wetter and Wright, 2003; Basak et. al, 2013) carried out, a crossover fraction of 0.8 was found to be able to better lower the function value. The fitness value for the A6002SN was found to drop to its final value over just five generations after which there was no refinement in the fitness value. Other roads such as the A6130SN also showed very little refinement in the fitness value from start of generation to the end. The A6117NS started with a fitness value which progressed constantly over fifteen generations after which the fitness value dropped and remained steady for another approximately 30 generations. After this there was a very slight drop in the fitness value after which there was no further refinement in fitness value over the rest of the generations. The

A47WE showed a steep stepwise refinement in the fitness value from start to generation 10. From generation 10, there was a very gradual drop in the fitness value up to generation 20 after which no further improvement in the fitness value was obtained. The A6030SN and the A6005WE show a small refinement in the fitness value from the beginning of the generation to the end of the generation. The results obtained for genetic algorithm can be attributed to the large search space and many generations for optimisation.

The initial optimised function values from genetic algorithm are compared with the validated function values from genetic algorithm using a t-test. The null hypothesis tested was that there is no difference between the initial optimised genetic algorithm function values and the validated genetic algorithm function values. The alternative hypothesis was that there is a difference between the initial optimised genetic algorithm function values and the validated genetic algorithm function values. Results using the t-test showed a significant difference in the function values for the validated genetic algorithm ($M=607649.34$, $SD=1055520.75$) and the initial optimised genetic algorithm function values ($M=9323.98$, $SD=10419.83$) under the condition $t(79)=5.05$, $p=0.000$. The 95 percent confidence interval was from 362476 to 834174. The computation time obtained for genetic algorithm varied from 3 minutes to 22 minutes. This short computation time is partially attributed to the function values being unrefined for a greater part and in most cases throughout the iteration period.

7.3.3 Comparing Pattern Search and Genetic Algorithm

Pattern Search and Genetic Algorithms were the two main types of optimisation techniques used in this study and it is necessary that the results obtained from these methods are compared and discussed for any similarities or differences. Generally it was observed that *pattern search* produced lower fitness function values than *genetic algorithm* with refinement in fitness function value being observed in later generation stages for *pattern search* than in *genetic algorithm.*

The validated function values from genetic algorithm are compared with the validated function values from pattern search using a t-test. The null hypothesis tested was that there is no difference between the validated genetic algorithm function values and the validated pattern search function values. The alternative hypothesis was that there is a difference between the validated genetic algorithm function values and the validated pattern search function values. Results using the t-test showed a significant difference in the function values for the validated genetic algorithm (M=607649.34, SD=1055520.75) and the validated pattern search function values (M=2.06, SD=1.17) under the condition $t(79) = 5.15$, $p=0.000$. The 95 percent confidence interval was from 372752 to 842541. The computation time obtained for genetic algorithm varied from 3 minutes to 22 minutes with that for pattern search varying from 13 minutes to 595 minutes. This short computation time observed in genetic algorithm can be explained to be the result of non-refinement in the fitness values from generation to generation. Pattern search did perform better than genetic algorithm in producing better and lower fitness values and this is also reflected in the computation times achieved.

It can be commented that *pattern search* optimisation made more progress in refining the fitness function in comparison with the *genetic algorithm* optimisation for most roads considered in this research. Lower fitness values were thus obtained from *pattern search* in comparison with *genetic algorithm* indicating better output and performance of *pattern search*.

Six main sets of results are discussed in this section from six different roads. The easting and northing values have been plotted and shown in Figure 27 to Figure 32. The first two (Figure 27 and Figure 28) will discuss results looking at two roads modelled using both genetic algorithm and pattern search highlighting some of the issues identified with the results. The next two (Figure 29 and Figure 30) results discuss roads modelled using genetic algorithm and pattern search with existing speed camera locations also shown. The last two (Figure 31 and Figure 32) will discuss roads modelled using pattern search and genetic algorithms.

In Figure 27 and Figure 28, the A6130 shows some easting and northing plotted values from the *genetic algorithm* optimisation run being far away from the expected road centreline reference point. Distance from the centreline of the A6130 road to the west of the identified *genetic algorithm* location was in excess of 300m. This shows some of the inconsistencies identified in the *genetic algorithm* optimisation. Given that genetic algorithms operates on the basis of natural selection, identifying the best individuals with better fitness value to proceed to the next generation is often a challenge to genetic algorithm and individuals in the current population with low fitness are often chosen as elite to proceed to the next generation. There is no guarantee that the genetic algorithm will find the global optimum location. Genetic algorithm however has the advantage of being able to solve problems where the objective function is discontinuous, non-differentiable, stochastic or highly nonlinear since the method does not depend on the error surface of the function. It is also able to solve problems and give multiple solutions. The A6117 in Figure 28 produced pattern search optimisation location points outside the expected location points. However, despite the fact that genetic algorithm has been found not to perform well in this research, the optimised genetic algorithm locations of A6117 in Figure 28 are observed to be referenced to the road centreline.

In Figure 29 and Figure 30, plotted easting and northing results obtained from *pattern search* for the A47 and A6005 were found to be similar to those obtained from *genetic algorithm*. In Figure 29 and Figure 30 pattern search, genetic algorithm and existing easting and northing locations for speed cameras are plotted. For the A47 where *existing* speed camera locations were identified, it was observed that the location was not close to that obtained from the *genetic algorithm* with a distance in excess of 20 kilometres apart between an existing speed camera location and a genetic algorithm obtained location. The optimised pattern search location was closer to the existing speed camera location as compared to the location obtained for genetic algorithm. A distance of approximately 470m and 540m at two different pattern search locations were observed from an existing location.

For the A6005 in Figure 30, some *existing* camera locations were found to be close to results obtained from *pattern search*. The minimum distance apart between a *pattern search* location and an *existing* camera location was found to be approximately 20m at two separate locations with other greater distances apart also observed. The least distance of separation found between a *genetic algorithm* and *pattern search* location along the A6005 was approximately 160m. The similarities in the locations established by pattern search indicate that the variables used in the optimisation in this research are somewhat consistent with what other decision makers may have obtained without the use of an optimisation technique. This finding is positive and suggests the possibility for the use of optimisation techniques in settings like this in the future. It is however suggested that this thesis be viewed as a good starting point to what could potentially become a very valuable tool to decision makers in deciding on the optimum location to place a speed camera or vehicle activated sign to satisfy a set of objectives and constraints.

In Figure 31 and Figure 32 results obtained for the A6030 and A6002 show the locations obtained for *pattern search* optimisation in comparison with *genetic algorithm* optimisation results. It can be observed that the *pattern search* results are referenced to the road centreline as is expected however for *genetic algorithm* it can be observed that results are not strictly referenced to the road centreline with some results found to be outside the confines of the road.

It was also observed that results obtained for the *pattern search* optimisation were found to be consistent in terms of function values obtained after subsequent runs for all the roads considered but this was not the case for *genetic algorithm*. This can be attributed to the characteristic features of pattern search which searches through a pattern under a given set of conditions and constraints. Genetic algorithm works using population and elite counts so if the population is not good enough to meet the requirements desired, the next generation of parents will not be 'healthy' enough to produce good 'offspring'. Despite the use of large search space and many generations by genetic algorithm, the search space used in this research has been found to be not large enough for genetic algorithm to use in searching for optimum solutions. In addition to this, checking for convergence in genetic algorithm is difficult so a specified number of generations was used in running it. Pattern search is however able to obtain better execution for a relatively smaller search space. Diversity in population which is the average distance between individuals affects the performance of *genetic algorithm* (Mathworks, 2015). The larger the average distance between individuals the higher the diversity and vice versa. However if the diversity is too high or low the performance of the *genetic algorithm* is affected. Other parameters affecting diversity are the initial range of the population and the amount of mutation. Even though mutation in addition to lower and upper bounds was applied to the genetic algorithm results obtained did not show much diversity.

There are some distinctive characteristics when comparing the approach adopted in this thesis with existing methods of speed camera or vehicle activated sign location; this research clearly considers one of the main aims of installing a speed camera or vehicle activated sign which is to reduce the number of accidents as the main objective function. Secondly recommendations from government guidelines are that speed cameras should only be considered where other engineering measures have failed. In addressing this, constraints set to the objective function were such that a given speed limit of road had to meet the criteria for variables such as the radius, number of lanes etc. This approach was considered more practical and realistic to the decision maker than existing methods in use whereby a speed camera can be mounted where improvements in engineering measures may be the lacking action required to improve safety and not the introduction of a speed camera. The procedure adopted in this research can be modified to suit other problems. This is because this research optimised locations based on the assumption that engineering measures are adequate and roads meet the desired standards as set out in the respective design guidelines (Highways Agency, 2002). In the case where a decision maker decides to place a speed camera or vehicle activated sign along a road which does not meet all the design requirements, the model can be altered to accommodate the objectives and constraints to be proposed.

This research can be described as *work in progress*; however it has demonstrated the huge potential evolutionary algorithms can play in managing vehicle speeds through the use of speed cameras or vehicle activated signs. The use of optimisation techniques in other fields of engineering and outside engineering is quite evident. In the field of transport safety this has been absent for some time. It is hoped that this research will generate more interest into the field of transport safety where there is high potential for benefits.

Improvement in vehicle technology in recent times means the presence of speed adaptation devices in vehicles may affect how people respond to speed cameras or vehicle activated signs. Intelligent Speed Adaptation (ISA) mainly considers two areas; the use of an advisory system where the driver receives a warning or through an intervention system where there is an automatic control of the driving systems of the vehicle aimed at lowering the vehicle speed (Jamson, Chorlton & Carsten, 2012;

Chorlton & Conner, 2012; Chorlton et al., 2012; Lai & Carsten, 2012). The former approach is currently what vehicle activated signs and other variable message signs provide on roads however as to whether these are always adhered to remains to be answered. The latter approach is more likely to be associated with in-vehicle technology. New generation of vehicles are likely to have this technology depending on the take-up rates of these new technologies (Fildes et al, 2013) and it is notable that there are still pools of "old generation" cars on the roads that will not have this technology and until this hurdle is overcome, vehicle speed management in the form of out of vehicle technology still needs to be used.

7.4 Methodological reflections

Limitations are always identified in studies carried out and this research is no exception. The limitations associated with each study or work carried out has been discussed in the respective chapters. The rest of this section makes mention of some other findings associated with the research.

7.4.1 Validity of findings

Improving road safety is not only a subject of national interest but cuts across nations. The Empirical Bayes accident prediction model developed in this research was tested against an independent sample of roads with further statistical analysis carried out to reveal the validity of the results. The use of pattern search and genetic algorithm also enabled two optimisation techniques to be compared and validated. This revealed some similarities and differences between the techniques used with statistical analysis also carried out to validate the results. Even though various accident prediction models have been developed over the years no literature has been found which tries to translate the derivation of the accident prediction model into use through the application of an optimisation technique as undertaken in this research. Considering that this is the first research of its kind in terms of this application it is anticipated that interest in this subject matter will continue to grow, rather than decline. This research presents a valid insight and use of a technique that has great potential to influence the decision making process of planners, engineers, decision makers etc. in identifying an optimum location to place a speed camera or vehicle activated sign based on a set of objectives and constraints. The original contribution of this research to the literature will raise some interesting debate among researchers in the near if not immediate future with more literature expected to be generated in this area for knowledge transfer.

7.4.2 Reliability of findings

The choice of accident prediction model developed and used in this research was also subjective. It became almost impossible to identify an accident prediction model previously developed in the UK and containing the prediction variables desired for the purposes of this research. Developing an accident prediction model requires careful attention into how the variables are identified and derived. Variables used in the accident prediction model were carefully computed, checked and used. The reliability of the variables were deemed fit for the model.

The selection of the optimisation techniques used was influenced by a literature review on the subject matter into other research areas of engineering given that none was found for the area of speed camera or vehicle activated sign usage. Optimisation techniques used in areas such as fire station location and ambulance location were identified as close enough to be emulated in this research. It is noted however that results obtained from the two methods used are quite similar in some instances with differences also noted. It is expected that other techniques will be investigated in future to validate against methods already used.

7.4.3 Policy implications

Every year about 1.2 million people die from road traffic accidents and as many as 50 million are injured on a world-wide level. Speed has been identified in studies (Perez at al., 2007; World Health Organisation , 2013) as a contributor to road traffic accidents. The World Health Organisation (2013) mentions speeding as a major road safety problem in all countries and mechanisms aimed at lowering speed can result in remarkable reductions in road traffic injuries. Excessive speed is also mentioned as a worldwide problem impacting the whole road network (motorways, highways, rural and urban roads). It is considered important for engineers, planners and decision makers to contribute to vehicle speed reduction to help reduce road traffic accidents. The output from this research has provided an increased understanding into the contribution of algorithms in improving road safety through speed reduction. It has been demonstrated that speed cameras and vehicle activated signs can be optimally located along roads to attain vehicle speed reduction. There are notable benefits to be derived from reducing vehicle speeds with its consequent effect on road safety being enormous. This is because in addition to the suffering caused to families, road traffic accidents can cause poverty to society in the form of medical care and rehabilitation costs as well as funeral expenses and the loss of the bread winner (World Health Organisation, 2009; Gabbe et al., 2014). A 1km/h reduction in vehicle speed has been shown to lead to a 3 percent reduction in accident risk (Finch et al., 1994) so that if vehicle speeds can be effectively managed road traffic accidents can be reduced to the benefit of society.

One of either *exceeding the speed limit* or *travelling too fast for the conditions* was reported in 10 percent of all accidents and these types of accidents accounted for 23 percent of all fatalities in 2013 (Department for Transport, 2014). It is obvious that the effect of speed on road traffic accidents and its severities is significant and it is considered essential that vehicle speeds are reduced in order to improve road traffic accidents. Speed cameras and vehicle activated signs have been shown in various studies to be effective at reducing vehicle speed (Pilkington and Kinra, 2005; Champness, Sheehan and Folkman, 2005; Jone, Sauerzaf and Haynes, 2008).

Additional advantages can be derived from encouraging existing initiatives aimed at reducing vehicle speeds. An example is the use of vehicle activated signs to display real time information to drivers to warn of a hazard or relay information about vehicle speed to the driver. Informing drivers in advance of the warning to provide enough driver-reaction time will have a positive impact on road safety. It is therefore essential that these devices are optimally located. A similar measure can be used for speed cameras by providing real time information to drivers about the presence of a speed camera ahead and the need to check vehicle speed.

Apart from developing the optimisation model to optimally locate speed cameras or vehicle activated signs, this research also discussed the contribution of accident prediction models in the optimisation model to aid in improving road safety by planner, engineers and policy makers. Chapter 2 discussed in detail factors affecting road traffic accidents and Chapter 3 provided a literature on accident prediction models and their limitations. It is essential to state again that there is no exact accident prediction model and engineers, planners and decision makers need to weigh the pros and cons of any proposed accident prediction model against the outcome desired as well as any cost and benefit implications that may arise in using a particular chosen method.

8 Conclusions and further research

8.1 Conclusions

Road traffic accidents continues to be a public health problem and indications are that this is likely to increase if no action is taken to curb the situation (World Health Organisation, 2013). Speeding continues to be a contributory factor in most road accidents. In UK, vehicle speeds are managed through the use of speed cameras and vehicle activated signs in addition to other engineering measures. These measures have been proved to be successful at controlling vehicle speeds and reducing road traffic accidents. Regarding installations of speed control devices (speed cameras and vehicle activated signs), there are guidelines which provides some information about the circumstances and conditions under which these speed control devices should be used. The decision of where a speed control device should be eventually located is that of the responsible engineer or designated person to make and this can be subjective. It is one of the most important and difficult decisions encountered by planners, engineers and decision makers. This thesis aimed at contributing to speed reduction by developing a model to help decision makers determine the optimum location for a speed control device in order to minimise road traffic accidents based on a set of objectives and constraints. A summary of relevant results obtained from the research carried out are presented as follows;

8.1.1 Accident prediction model

- 790 kilometres of A-Roads in Nottinghamshire and Leicestershire was investigated.
- 75 percent of a random selection of these roads was used in developing the model with the remaining 25 percent used in validating the model.
- The 75 percent of roads comprised 4177 homogeneous segments, 350 fatal and serious accidents, 1,646 slight accidents, 1,580 junctions and 18,240m length of cycle route.
- The 25 percent of roads was made up of 185,000 metres (185km) of roads comprising 1,771 homogeneous segments, 159 fatal and serious accidents, 960 slight accidents, 1,872 junctions and 11,120m of cycle route.

- The key parameters used in developing the model were related to road geometry and other traffic characteristics with no human factors included in the accident prediction model.
- A Negative Binomial model was initially developed for fatal and serious accidents combined and slight accidents separately.
- The Negative Binomial model was refined to an Empirical Bayes Negative Binomial regression model in order to take account of the regression to the mean effect.
- The predictor variables used in the negative binomial accident prediction model were the road direction, number of lanes, slope, radius, Annual Average Daily Traffic (AADT), Heavy Goods Vehicle (HGV), speed limit, homogeneous segment length, average speed, presence of cycle lane, length of cycle lane, presence of a junction or not, number of junctions and number of pedestrian crossings.
- Variables found to be statistically significant ($p < 0.05$) in the slight accidents negative binomial model were the presence of junctions (Jtn), direction of travel along the road (Dxn), the summation of junctions (JtnsSUM), number of lanes, percentage Heavy Goods Vehicles (HGV), speed limit, homogeneous segment length of road (HSegLength), logarithm of the Annual Average Daily Traffic (LogAADT) and logarithm of the radius (LogRadius) with variables found to be statistically insignificant ($p > 0.05$) excluded from the model and rerun.
- For the fatal and serious accidents negative binomial model, variables found to be statistically significant were the presence of junctions (Jtn), direction of travel along the road (Dxn), speed limit, homogeneous segment length of road (HSegLength), logarithm of the Annual Average Daily Traffic (LogAADT) and logarithm of the radius (LogRadius). Variables found to be statistically insignificant ($p > 0.05$) were excluded from the model and rerun.
- Categorical variables used in the rerun fatal and serious accidents model were the presence or absence of junctions (Jtn) along a homogeneous segment of road and the direction of travel.

- For the continuous variables for fatal and serious accidents, the speed limit, homogeneous segment length, logarithm of the Annual Average Daily Traffic and logarithm of the radius of the road were represented as statistically significant in the rerun model.
- The rerun slight accidents model had categorical values observed to be statistically significant being the presence or absence of junctions (Jtn), presence or absence of a cycle route (PresofCR) and direction of travel along a homogeneous segment of road.
- The slight accidents model had continuous variables represented in the model to be the sum of junctions within a homogeneous segment of road, number of lanes, percentage Heavy Goods Vehicles (HGV), speed limit, length of homogeneous segment of road, logarithm of the Annual Average Daily Traffic and logarithm of the radius.
- A paired t-test was used to compare the negative binomial slight accidents with the empirical Bayes slight accidents and there was a significant difference. For Negative Binomial (NB) slight accidents (M=0.66, SD=1.416) and slight accidents obtained from the Empirical Bayes (EB) model (M=0.55, SD=0.994) condition; t (1769) =3.381, p=0.001 a significant difference was obtained.
- A paired t-test was used to compare the negative binomial fatal and serious accidents with the empirical Bayes fatal and serious accidents and results showed a significant difference. For Negative Binomial (NB) fatal and serious accidents (M=0.12, SD=0.108) and fatal and serious accidents obtained from the Empirical Bayes (EB) model (M=0.11, SD=0.117) condition; t (1769) = 3.652, p=0.000 a significant difference was obtained.

8.1.2 Optimisation model

- Fourteen segments of roads were optimised.
- Genetic Algorithm and Pattern search optimisation techniques were used.
- For both genetic algorithm and pattern search, no optimal solution was obtained for the A6211.
- Results of proposed locations identified and plotted on maps for genetic algorithm were found not to be referenced to the road centreline as should be

expected. Results obtained from genetic algorithm were also found to be inconsistent in certain instances with some proposed locations for the speed control devices placed more than 100m away perpendicularly from the road centreline.

- Proposed coordinate location points obtained on plotted maps for pattern search were found to be mostly located referenced to the road centreline as was expected.
- Results obtained for the pattern search optimisation were found to be consistent in terms of function values obtained after subsequent runs for all the roads considered but this was not the case for genetic algorithm.
- The pattern search optimisation was noted to be able to lower the function value much better than the genetic algorithm optimisation.
- The A47 and A6005 had existing speed camera locations along and the positions were found to be closer to results obtained from pattern search than for genetic algorithm.
- Computation time obtained for genetic algorithm varied from 3 minutes to 22 minutes with that for pattern search varying from 13 minutes to 595 minutes.
- The A47 investigated is the third longest road used in this research with a total length for both directions of traffic being 87920m and this had the highest computation time of 595 minutes.
- T-test analysis show a significant difference in the function values for the validated pattern search (M=2.06, SD=1.17) and the initial optimised pattern search function values (M=2.80, SD=1.00) under the condition $t\ (79) = -9.86$, $p=0.000$
- Results from t-test showed a significant difference in the function values for the validated genetic algorithm (M=607649.34, SD=1055520.75) and the initial optimised genetic algorithm function values (M=9323.98, SD=10419.83) under the condition $t(79)=5.05$, $p=0.000$.
- T-test results show a significant difference in the function values for the validated genetic algorithm (M=607649.34, SD=1055520.75) and the validated pattern search function values (M=2.06, SD=1.17) under the condition $t\ (79)=5.15$, $p=0.000$.

- In this research pattern search optimisation performed better than the genetic algorithm optimisation.

8.2 Contribution to knowledge

No clear approach to identifying an appropriate location to place a speed control device using optimisation techniques was available at the start of this research. Even though there is some information available for other areas of study such as in fire station location there was nothing available in the area of road safety for speed management using speed cameras or vehicle activated signs. The work presented in this thesis is an original and useful contribution to vehicle speed management for the installation of speed cameras or vehicle activated signs. The key contributions to knowledge are as follows;

- The development of a new Empirical Bayes Negative Binomial regression Model for A-roads in the United Kingdom incorporating road geometry and traffic characteristics to predict road traffic accidents at specified injury severity levels. The use of the Empirical Bayes Negative Binomial regression Model in the prediction of road traffic accidents having predictor variables relevant to that desired for this research was absent with the development of the new model proving useful for subsequent work carried out in this thesis.
- The use of genetic algorithms and pattern search optimisation techniques in this thesis is considered unique since no literature has been found on the application of these optimisation techniques in the area of transport safety.
- The use of Geographical Information Systems (GIS) in this thesis enables decision makers to visually assess proposed speed control device locations and make informed decisions about the suitability of the location.
- The developed model can be used for road safety management. The use of Geographical Information Systems means road accidents can be better managed for both non-accident and existing accident sites. Existing non-accident prone sites with no known accident records can be checked for future accidents and such sites can have speed control devices proposed offering better speed management to improve road safety. For existing sites, the model can be used to better plan and implement the locations for mounting speed control devices. Considering that homogeneous segments of

road were used in the accident prediction model, accidents along each homogeneous segment can be calculated and plotted to better understand the characteristics of the types of accidents and locations at which the accidents are occurring. The Geographical Information Systems interface will allow accident sites to be geo-referenced and better studied.

- Road safety management can be incorporated into the wider area of infrastructure management. Considering that road designs are expected to be optimal both in terms of monetary value as well as in the purpose for which it is expected to serve, designers can use accident prediction models to identify unsafe locations in proposed designs and amend the design before the road gets built in order to save both time and money. With the use of accident prediction models, various alternative designs can be assessed and discounted or accepted leading to the optimum design being chosen. On the other hand accident prediction models can also be incorporated into Road Safety Audit measures to determine accident prone areas in order to recommend corrective measures.

8.3 Further Work

A number of areas have been identified for future work to be carried out on and these are discussed below.

- Pattern search and genetic algorithms are the main types of optimisation techniques used in this research. Even though pattern search produced more consistent and practical solutions to the optimisation problem, genetic algorithms did not in some instances. It is proposed that other optimisation techniques are investigated to identify any similarities or differences in results obtained.
- An accident prediction model that better predicts fatal and serious injury accidents will aid in providing a more concise estimation of these types of accidents. It is worth stating that accidents occurring at higher levels of severity can sometimes be caused under rare circumstances or conditions such as inclement weather and lack of knowledge about the road terrain and environment which are described as confounding factors and are mostly difficult to control.

- The accident prediction model used in the optimisation model was developed using A-roads in Nottinghamshire and Leicestershire, UK. Motorways and B-roads and other classes of roads in UK and other countries can be developed and compared.
- Decision makers are normally attracted to the aesthetics of the decision making process instead of the technicalities. The comprehension, development and application of the methodology used in developing the optimisation model will require a lot of software coding effort for engineers, planners, decision makers etc. which may not appeal to such professionals. The development of software materials in the form of a Graphical User interfaced decision making process will make the decision making process more flexible and user friendly. This proposal can be considered in another research.

References

Aarts, L. & Van Schagen, I. 2006, "Driving speed and the risk of road crashes: A review", *Accident Analysis and Prevention,* vol. 38, no. 2, pp. 215.

AASHTO. 2004, "American Association of State Highway and Transportation Officials" in *A Policy on Geometric Design of Highways and Street*.

AASHTO. 2001, in *A Policy on Geometric Design of Highways and Streets*, 4th edn, American Association of State Highways and Transportation Officials.

Abdel-Aty, M. & Radwan, A.E. 2000, "Modeling traffic accident occurrence and involvement", *Accident Analysis and Prevention,* vol. 32, no. 5, pp. 633-642.

Abdel-Aty, M. 2003, "Analysis of driver injury severity levels at multiple locations using ordered probit models", *Journal of Safety Research,* vol. 34, no. 5, pp. 597-603.

Abdul-Mawjoud, A.A. & Sofia, G.G. 2008, "Development of models for predicting speed on horizontal curves for two lane rural highways", *The Arabian Journal for Science and Engineering,* vol. 33, no. 2B.

Ackora-Prah, J., Gyamerah, S. A., Andam, P. S. & Gyamfi, D. 2014, "Pattern Search for Portfolio Selection", *Applied Mathematical Sciences,* vol. 8, no. 143 pp. 7137-7147.

Ackroyd, S. & Hughes, J.A. 1992, *Data collection in context,* Longman, London.

Aguero-Valverde, J. & Jovanis, P.P. 2006, "Spatial analysis of fatal and injury crashes in Pennsylvania", *Accident Analysis and Prevention,* vol. 38, no. 3, pp. 618-625.

Al-Harbi, M., Yassin, M., Bin Shams, F. & Bin Shams, M. 2012, "Stochastic modeling of the impact of meteorological conditions on road traffic accidents", *Stochastic Environmental Research and Risk Assessment,* vol. 26, no. 5, pp. 739-750.

Al Hajri, M. F., El-Naggar, K. M., AlRashidi, M. R. & Al-Othman, A. K. 2012, "Optimal extraction of solar cell parameters using pattern search", *Renewable Energy,* vol. 44, pp. 238-245.

Aljanahi, A.A.M., Rhodes, A.H. & Metcalfe, A.V. 1999, "Speed, speed limits and road traffic accidents under free flow conditions", *Accident Analysis and Prevention,* vol. 31, no. 1, pp. 161-168.

Allison, P. 2012, 10th September 2012-last update*, When Can You Safely Ignore Multicollinearity?* [Homepage of Statistical Horizons], [Online]. Available: http://www.statisticalhorizons.com/multicollinearity [2014, September/20th].

Allison, B., Owen, A., Rothwell, A., O'Sullivan, T., Saunders, C., Rice, J. & De, M.U. 1996, *Research skills for students*, Kogan Page, London.

Allsop, R. 2010, *The Effectiveness of speed cameras: A review of evidence*, RAC Foundation, UK.

Ameen, J.R.M. & Naji, J.A. 2001, "Causal models for road accident fatalities in Yemen", *Accident Analysis and Prevention*, vol. 33, no. 4, pp. 547-561.

Anastasopoulos, P.C. & Mannering, F.L. 2009, "A note on modeling vehicle accident frequencies with random-parameters count models", *Accident Analysis and Prevention,* vol. 41, no. 1, pp. 153-159.

Anderson, D.R. 2010, *Statistics for business and economics,* 2nd ed. edn, South-Western Cengage Learning, Andover.

Andersson, A.K. & Chapman, L. 2011, "The impact of climate change on winter road maintenance and traffic accidents in West Midlands, UK", *Accident Analysis and Prevention,* vol. 43, no. 1, pp. 284-289.

Andrey, J. & Yagar, S. 1993, "A temporal analysis of rain-related crash risk", *Accident Analysis and Prevention*, vol. 25, no. 4, pp. 465-472.

Armstrong, M.P. 2000, "Geography and Computational Science", *Annals of the Association of American Geographers,* vol. 90, no. 1, pp. 146-156.

Arostegui, M.A., Kadipasaoglu, S.N. & Khumawala, B.M. 2006, "An empirical comparison of Tabu Search, Simulated Annealing, and Genetic Algorithms for facilities location problems", *International Journal of Production Economics,* vol. 103, no. 2, pp. 742-754.

Ashton, S.J. & Mackay, G.M. 1979, "Some characteristics of the population who suffer trauma as pedestrians when hit by cars and some resulting implications", *Proceedings of the 4th IRCOBI Conference on Biomechanics of Trauma*.

Badri, M.A. 1999, "Combining the analytic hierarchy process and goal programming for global facility location-allocation problem", *International Journal of Production Economics*, vol. 62, no. 3, pp. 237-248.

Badri, M.A., Mortagy, A.K. & Alsayed, C.A. 1998, "A multi-objective model for locating fire stations", *European Journal of Operational Research,* vol. 110, no. 2, pp. 243-260.

Barker, J. 1997, *Trials of rural road safety engineering measures*. *TRL 202,* Transport Research Laboratory, UK.

Basak, R., Sanyal, A., Nath, S.K & Goswami, R. 2013, "Comparative view of Genetic Algorithm and Pattern Search for Global Optimization", *International Journal of Engineering and Science*, vol. 3, Issue 4, pp. 09-12.

Bellman, R.E. & Zadeh, L.A. 1970, "Decision- making in a fuzzy environment", *Management Science*, vol. 17, no. 4, pp. 141-164.

Belmont, D.M. 1953, "Effect of average speed and volume on motor vehicle accidents on two-lane tangents", *Highway Research Board Proceedings*, vol. 32, pp. 383-395.

Ben-Bassat, T. & Shinar, D. 2011, "Effect of shoulder width, guardrail and roadway geometry on driver perception and behavior", *Accident Analysis and Prevention*, vol. 43, no. 6, pp. 2142-2152.

Bennett, C.R. 1994, *A speed prediction model for rural two-lane highways*, University of Auckland NewZealand.

Berhanu, G. 2004, "Models relating traffic safety with road environment and traffic flows on arterial roads in Addis Ababa", *Accident Analysis and Prevention*, vol. 36, no. 5, pp. 697-704.

Bhati, A.S. 2005, "Modeling count outcomes with spatial structures: An information-theoretic approach", Discussion Paper, Justice Policy Center, The Urban Institute, Washington DC.

Bhatnagar, Y.S. 1994, "Observations on the use of chevron alignment markers", *Proceedings of 17th Australian Road Research Bard Ltd (ARRB) Conference*, pp. 65.

Bjørnskau, T. & Elvik, R. 1992, "Can road traffic law enforcement permanently reduce the number of accidents?", *Accident Analysis and Prevention*, vol. 24, no. 5, pp. 507-520.

Bolstad, W.M. 2007, *Introduction to Bayesian Statistics*, 2nd edn, John Wiley, Hoboken, N.J.

Brijs, T., Karlis, D. & Wets, G. 2008, "Studying the effect of weather conditions on daily crash counts using a discrete time-series model", *Accident Analysis and Prevention*, vol. 40, no. 3, pp. 1180-1190.

Brill, E.D. 1979, "The Use of Optimization Models in Public- Sector Planning", *Management Science*, vol. 25, no. 5, pp. 413-422.

Brodsky H and Hakkert A.S. 1988. "Risk of a road accident in rainy weather", *Accident Analysis and Prevention*, vol. 20, no. 3, pp. 161-176.

Brotcorne, L., Laporte, G. & Semet, F. 2003, "Ambulance location and relocation models", *European Journal of Operational Research*, vol. 147, no. 3, pp. 451-463.

Brüde, U. & Larsson, J. 1988, "The use of prediction models for eliminating effects due to regression-to-the- mean in road accident data", *Accident Analysis and Prevention*, vol. 20, no. 4, pp. 299-310.

Burbridge, A., Eveleigh, M. & Van Eysden, P. 2010, "Queensland experience with vehicle activated signs", *2010 Australasian Road Safety Research, Policing and Education Conference* Canberra, Australian Capital Territory, 31 August – 3 September.

Buzeman, D.G., Viano, D.C. & Lövsund, P. 1998, "Car occupant safety in frontal crashes: A parameter study of vehicle mass, impact speed, and inherent vehicle protection", *Accident Analysis and Prevention,* vol. 30, no. 6, pp. 713-722.

Cafiso, S., Di Graziano, A., Di Silvestro, G., La Cava, G. & Persaud, B. 2010, "Development of comprehensive accident models for two- lane rural highways using exposure, geometry, consistency and context variables", *Accident Analysis and Prevention,* vol. 42, no. 4, pp. 1072-1079.

Cameron, M.H., Vulcan, A.P., Finch, C.F. & Newstead, S.V. 1994, "Mandatory bicycle helmet use following a decade of helmet promotion in Victoria, Australia—An evaluation", *Accident Analysis and Prevention,* vol. 26, no. 3, pp. 325-337.

Carlin, B.P. & Louis, T.A. 2000, *Bayes and Empirical Bayes Methods for Data Analysis,* 2nd edn, Chapman & Hall/CRC, London.

Carnis, L. & Blais, E. 2013, "An assessment of the safety effects of the French speed camera program", *Accident; Analysis and Prevention,* vol. 51, pp. 301-309.

Casella, G. 1992, "Illustrating empirical Bayes methods", *Chemometrics and Intelligent Laboratory Systems,* vol. 16, pp. 107-125.

Ceder, A. & Livneh, M. 1978, "Further evaluation of the relationships between road accidents and average daily traffic", *Accident Analysis and Prevention,* vol. 10, no. 2, pp. 95-109.

Ceder, A. 1982, "Relationships between road accidents and hourly traffic flow—II: Probabilistic approach", *Accident Analysis and Prevention,* vol. 14, no. 1, pp. 35-44.

Ceder, A. & Livneh, M. 1982, "Relationships between road accidents and hourly traffic flow—I: Analyses and interpretation", *Accident Analysis and Prevention,* vol. 14, no. 1, pp. 19-34.

Ceder, A. & Livneh, M. 1978, "Further evaluation of the relationships between road accidents and average daily traffic", *Accident Analysis and Prevention,* vol. 10, no. 2, pp. 95-109.

Champness, P., Sheehan, M. & Folkman, L. 2005, "Time and distance halo effects of an overtly deployed mobile speed camera", *Road Safety Research, Policing and Education Conference* Wellington, New Zealand, 14-16 November.

Chang, L. & Wang, H. 2006, "Analysis of traffic injury severity: An application of non- parametric classification tree techniques", *Accident Analysis and Prevention,* vol. 38, no. 5, pp. 1019-1027.

Charlton, S.G. & dePont, J.J. 2007, *Curve speed management. Research Report 323*, Land Transport NewZealand, NewZealand.

Chau, K. 2004, "Two- stage dynamic model on allocation of construction facilities with genetic algorithm", *Automation in Construction,* vol. 13, no. 4, pp. 481-490.

Chen, C. & Ting, C. 2008, "Combining Lagrangian heuristic and Ant Colony System to solve the Single Source Capacitated Facility Location Problem", *Transportation Research Part E,* vol. 44, no. 6, pp. 1099-1122.

Chen, G., Meckle, W. & Wilson, J. 2002, "Speed and safety effect of photo radar enforcement on a highway corridor in British Columbia", *Accident Analysis and Prevention,* vol. 34, no. 2, pp. 129-138.

Chen, G. & Warburton, R.N. 2006, "Do speed cameras produce net benefits? Evidence from British Columbia, Canada", *Journal of Policy Analysis and Management,* vol. 25, no. 3, pp. 661-678.

Chen, S.J. & Hwang, C.L. 1992, *Fuzzy multiple attribute decision making: Methods and Applications,* Springer-Verlag, New York.

Cheung, B.K., Langevin, A. & Villeneuve, B. 2001, "High performing evolutionary techniques for solving complex location problems in industrial system design", *Journal of Intelligent Manufacturing,* vol. 12, no. 5, pp. 455-466.

Chin, H.C. & Quddus, M.A. 2003, "Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections", *Accident Analysis and Prevention,* vol. 35, no. 2, pp. 253-259.

Chorlton, K. & Conner, M. 2012, "Can enforced behavior change attitudes: Exploring the influence of Intelligent Speed Adaptation", *Accident Analysis and Prevention,* vol. 48, pp. 49-56.

Chorlton, K., Hess, S., Jamson, S. & Wardman, M. 2012, "Deal or no deal: Can incentives encourage widespread adoption of intelligent speed adaptation devices?", *Accident Analysis and Prevention,* vol. 48, pp. 73-82.

Choueiri, E.M. & Lamm, R. 1987, *Rural roads speed inconsistencies design methods: Part I, Operating speeds and accident rates on two lane rural highway curved sections*, Department of Civil and Environmental Engineering, Clarkson University, New York.

Clarke, D.D., Ward, P., Bartle, C. & Truman, W. 2010, "Killer crashes: Fatal road traffic accidents in the UK", *Accident Analysis and Prevention,* vol. 42, no. 2, pp. 764-770.

Clarke, D.D., Forsyth, R.. & Wright, R. 1998, "Behavioural factors in accidents at road junctions: The use of a genetic algorithm to extract descriptive rules from police case files", *Accident Analysis and Prevention,* vol. 30, no. 2, pp. 223-234.

Cochran, W.G. 1977, *Sampling techniques,* 3rd edn, Wiley, New York ; London.

Coley, D.A. 1999, *An Introduction to Genetic Algorithms for Scientists and Engineers ISBN 981 02 3602 6,* World Scientific Publishing.

Comte, S., Wardman, M. & Whelan, G. 2000, "Drivers' acceptance of automatic speed limiters: implications for policy and implementation", *Transport Policy,* vol. 7, no. 4, pp. 259-267.

Congdon, P. 2001, *Bayesian Statistical Modelling,* John Wiley and Sons, Chichester, UK.

Corbett, C. & Simon, F. 1999, *The effects of speed cameras: how drivers respond, DTLR Report 11*, Department for Transport, London.

Corbett, C. 1995, "Road traffic offending and the introduction of speed cameras in England: The first self- report survey", *Accident Analysis and Prevention,* vol. 27, no. 3, pp. 345-354.

Corbett, C., Delmonte, E., Quimby, A., and Grayson, G. 2008. "Does The Threat of Disqualification Deter Drivers from Speeding?" Road Safety Research Report No.96. London, Department for Transport. Available: http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme2/threat.pdf. [2015, May/4th]

Corbett, C. & Grayson, G. 2010, "Speed Limit Enforcement as Perceived by Offenders: Implications for Roads Policing", *Policing: A Journal of Policy and Practice,* vol. 4, no. 4, pp. 364-372.

Council, F. M., and J. R. Stewart. 2000. *Safety Effects of the Conversion of Rural Two-Lane to Four-Lane Roadways Based on Cross-Sectional Models.* FHWA-990327. FHWA, U.S. Department of Transportation.

Cowen D. 1988,"GIS versus CAD versus DBMS: what are the differences". *Photogrammetric Engineering and Remote Sensing*, vol.54, pp. 1551-1555.

Crashmap , *Crashmap* [Homepage of Crashmap], [Online]. Available: crashmap.co.uk [2014, July/18th].

Crombie, H. 2002, *The impact of transport and road traffic speed on health*. Available: http://www.nice.org.uk/niceMedia/documents/road_traffic_%20issues%20_270302.pdf [2011, October/20].

Cycle-routes UK. http://www.cycle-route.com/routes/Leicestershire-Routes-87.html. [2016, January/10].

Daniels, S., Vanrie, J., Dreesen, A. & Brijs, T. 2010, "Additional road markings as an indication of speed limits: Results of a field experiment and a driving simulator study", *Accident Analysis and Prevention,* vol. 42, no. 3, pp. 953-960.

Davis, G.A. & Swenson, T. 2006, "Collective responsibility for freeway rear-ending accidents?(Report)", *Accident Analysis and Prevention,* vol. 38, no. 4, pp. 728.

De Brabander, B., Nuyts, E. & Vereeck, L. 2005, "Road safety effects of roundabouts in Flanders", *Journal of Safety Research,* vol. 36, no. 3, pp. 289-296.

De Brabander, B. & Vereeck, L. 2007, "Safety effects of roundabouts in Flanders: Signal type, speed limits and vulnerable road users", *Accident Analysis and Prevention,* vol. 39, no. 3, pp. 591-599.

de Palma, A., Kilani, M. & Lindsey, R. 2008, "The merits of separating cars and trucks", *Journal of Urban Economics,* vol. 64, no. 2, pp. 340-361.

Deb, K. 2001, *Multi-objective optimisation using evolutionary algorithms,* Chichester: John Wiley and Sons.

Department for Transport. 2016, *Flow weighted vehicle speeds (CGN02):* Department for Transport, London. https://www.gov.uk/government/statistical-data-sets/cgn02-flow-weighted-vehicle-speeds.

Department for Transport. 2015, *Road Traffic Estimates: Great Britain 2014*, Department for Transport, London.

Department for Transport. 2014, *Reported Road Casualties Great Britain; Annual Report*, Department for Transport, London.

Department for Transport. 2013, *Traffic Signs Manual Chapter 4 Warning Signs*, Department for Transport; London, UK.

Department for Transport.2013a, *Traffic counts,* http://www.dft.gov.uk/traffic-counts/download.php [2013, July/5th].

Department for Transport. 2012, *Guidance on Road Classification and the Primary Route Network*, Department for Transport, UK.

Department for Transport. 2012a, *Reported Road Casualties Great Britain: 2011 Annual Report*, Department for Transport, UK.

Department for Transport. 2011, *Reported Road Casualties Great Britain: 2010 Annual Report*, Department for Transport, UK.

Department for Transport. 2011a, *Free Flow Vehicle Speeds in Great Britain 2010 Statistical Release.* Department for Transport, London.

Department for Transport. 2010, *General overview and trends in reported casualties. Reported Road Casualties in Great Britain: 2010*, UK.

Department for Transport. 2010a, *Reported Road Casualties Great Britain: 2009 Annual Report*, Department for Transport, UK.

Department for Transport. 2009, *Reported Road Casualties Great Britain: 2008 Annual Report*, Department for Transport, UK.

Department for Transport. 2008, *Reported Road Casualties Great Britain: 2007 Annual Report*, Department for Transport, UK.

Department for Transport. 2007, *DfT Circular 01/2007. Use of speed and red-light cameras for traffic enforcement: Guidance on deployment, visibility and signing*, Department for Transport, London, UK.

Department for Transport. 2006, *Handbook of Rules and Guidance for the National Safety Camera Programme for England and Wales for 2006/07*, Safety Camera Programme, Department for Transport, UK.

Department for Transport. 2004, *SATAS 20, Instructions for the Completion of Road Accident Reports*, Department for Transport, UK.

Department for Transport. 2003, *Vehicle Activated Signs. Traffic Advisory Leaflet 1/03*, Department for Transport, London.

DETR, 2001. Road Safety Good Practice Guide. Department of the Environment Transport and the Regions. The Stationary Office, London.

Deublein, M., Adey, B.T., Schubert, M., Köhler, J. & Faber, M.H. 2013, "Prediction of road accidents: A Bayesian hierarchical approach", *Accident Analysis and Prevention*, vol. 51, pp. 274-291.

Digimap-EDINA. https://digimap.edina.ac.uk/. [2012, January/15].

Dobson, A.J. & Barnett, A.G. 2008, *An introduction to Generalized linear Models*, 3rd edn, CRC Press, Taylor and Francis Group, Boca Raton, USA.

Donnell, E.T., Ni, Y., Adolini, M. & Elefteriadou, L. 2001, *Speed Prediction Models for Trucks on Two-Lane Rural Highways*, Transportation Research Board, Washington DC.

Dunteman, G.H. & Ho, M.R. 2006, *An Introduction to generalized linear models*, SAGE, London.

Easa, S.M. & Hassan, Y. 2000, "Development of transitioned vertical curve II Sight distance", *Transportation Research Part A*, vol. 34, no. 7, pp. 565-584.

Edwards, J.B. 2002, "Motorway speeds in wet weather: the comparative influence of porous and conventional asphalt surfacings", *Journal of Transport Geography,* vol. 10, no. 4, pp. 303-311.

Edwards, J.B. 1999, "Speed adjustment of motorway commuter traffic to inclement weather", *Transportation Research Part F: Psychology and Behaviour,* vol. 2, no. 1, pp. 1-14.

Edwards, J.B. 1998, "The Relationship Between Road Accident Severity and Recorded Weather", *Journal of Safety Research,* vol. 29, no. 4, pp. 249-262.

Edwards, J.B. 1996, "Weather-related road accidents in England and Wales: a spatial analysis", *Journal of Transport Geography,* vol. 4, no. 3, pp. 201-212.

Eisenberg, D. & Warner, K.E. 2005, "Effects of snowfalls on motor vehicle collisions, injuries, and fatalities.(Author Abstract)", *The American Journal of Public Health,* vol. 95, no. 1, pp. 120.

El-Basyouny, K. & Sayed, T. 2009, "Accident prediction models with random corridor parameters", *Accident Analysis and Prevention,* vol. 41, no. 5, pp. 1118-1123.

El-Basyouny, K. & Sayed, T. 2009, "Collision prediction models using multivariate Poisson-lognormal regression", *Accident Analysis and Prevention,* vol. 41, no. 4, pp. 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,* vol. 40, no. 3, pp. 1033-1054.

Elvik, R. 2013, "A re-parameterisation of the Power Model of the relationship between the speed of traffic and the number of accidents and accident victims.(Report)", *Accident Analysis and Prevention,* vol. 50, pp. 854.

Elvik, R. 2011, "Developing an accident modification function for speed enforcement", *Safety Science,* vol. 49, no. 6, pp. 920-925.

Elvik, R. 2009, *The Power Model of the relationship between speed and road safety: Update and new analysis. TOI Report 1034/2009*, The Institute of Transport Economics, Norway.

Elvik, R., Høye, A., Vaa, T., and Sorenson, M. 2009. *The handbook of road safety measures.* Bingley: Emerald Publishing.

Elvik, R. 2008, "The predictive validity of empirical Bayes estimates of road safety", *Accident Analysis and Prevention,* vol. 40, no. 6, pp. 1964-1969.

Elvik, R. 2006, "Laws of accident causation", *Accident Analysis and Prevention,* vol. 38, no. 4, pp. 742-747.

Elvik, R. 1997, "Effects on accidents of automatic speed enforcement in Norway", *Transportation Research Record 1595, TRB,* National Research Council, Washington, D.C, pp. 14-19.

Elvik, R., Christensen, P. & Amundsen, A. 2004, *Speed and road accidents: An evaluation of the Power Model.TOI Report 740/2004,* The Institute of Transport Economics, Norway.

Elvik, R. 2002, "The importance of confounding in observational before-and- after studies of road safety measures", *Accident Analysis and Prevention,* vol. 34, no. 5, pp. 631-635.

Erden, T. & Coskun, M.Z. 2010, "Multi-criteria site selection for fire services: the interaction with analytic hierarchy process and geographic information systems", *Natural Hazards and Earth System Sciences,* vol. 10, no. 10, pp. 2127-2134.

EU Directive 96/79/EC, Frontal Crash Protection, 1996.

EuroNCAP, 2004. European ne car assessment programme. Assessment protocol and biomechanical limits. *Version 4.1.*

Evans, L. & Wasielewski, P. 1982, "Do accident- involved drivers exhibit riskier everyday driving behavior?", *Accident Analysis and Prevention,* vol. 14, no. 1, pp. 57-64.

Evgenikos, P.,Yannis, G., Folla, K., Bauer, R., Machata, K., Brandstaetter, C., Thomas, P. & Kirk, A. 2016, "Characteristics and causes of heavy goods vehicles and buses accidents in Europe", *Proceedings of 6th European Transport Research Conference,* 18-21 April, Warsaw, Poland.

Fallowfield, L. 1995, "Questionnaire Design", *Archives of diseases in childhood,* vol. 72, no. 1, pp. 76-79.

Fambro, D.B., Fitzpatrick, K. & Russell, C.W. 2007, "Operating speed on crest vertical curves with limited stopping sight distance", *Transportation Research Record,* vol. 1701/2000, no. 0361-1981, pp. 25-31.

Fildes, B., Keall, M., Thomas, P., Parkkari, K., Pennisi, L & Tingvall, C. 2013. “Evaluation of the benefits of vehicle safety technology”, *Accident Analysis and Prevention,* vol. 55, pp. 274-281.

Finch, D.J., Kompfner, P., Lockwood, C.R. & Maycock, G. 1994, *Speed, speed limits and accidents,* Transport Research Laboratory, Crowthorne, UK.

Finn, P. & Bragg, B.W.E. 1986, "Perception of the risk of an accident by young and older drivers", *Accident Analysis and Prevention,* vol. 18, no. 4, pp. 289-298.

Fitzpatrick, K., Anderson, I.B., Bauer, K.M., Collins, J.M., Elefteriadou, L., Green, P., Harwood, D.W., Irizarry, N., Koppa, R., Krammes, R.A., Parma, K.D., Passetti, K., Poggioli, B., Tsimhoni, O. & Wooldridge, D. 2000, *Evaluation of Design Consistency methods for Two Lane Rural Highways. Report No. FHWA-RD-99-173*, Federal Highways Administration, USA.

Fitzpatrick, K., Elefteriadou, L., Harwood, D.W., Collins, J.M., McFadden, J., Anderson, I.B., Krammes, R.A., Irizzary, N., Parma, K.D., Bauer, K.M. & Passetti, K. 2000a, *Speed Prediction For Two-Lane Rural Highways.Report No. FHWA – RD -99 171*, Federal Highways Administration, USA.

Fleming, P. & Zalzala, A. 1997, "Special issue on genetic algorithms in engineering systems", *Proceedings Of The Institution Of Mechanical Engineers Part I-Journal Of Sy; Proc.Inst.Mech.Eng.Part I-J Syst Control Eng.*, vol. 211, no. 5, pp. I-II.

Forester, J. (2001). The bikeway controversy. Transportation Quarterly, 55.

Formula1. 2011, 21st January 2011-last update, *Thrilling 2010 season boosts Formula One TV audiences*. Available: http://www.formula1.com/news/headlines/2011/1/11660.html. [2011, November/15th].

Forward, S.E. 2009, “An assessment of what motivates road violations”, Transportation Research Part F, vol.12 pp.225-234.

Frampton, R.J., Welsh, R.H. and Thomas, P.D. 2002. "Belted Driver Protection in Frontal Impact - What Has Been Achieved and Where do Future Priorities Lie?". Proceedings 46th AAAM Conf, Tempe, Arizona, U.S.A.

Frampton, R., Page, M. & Thomas, P. 2006, "Factors related to fatal injury in frontal crashes involving European cars", *Proceedings of 50th Conference of the Association for the Advancement of Automotive Medicine*, 16-18 October, Chicago.

Fridstrøm, L., Ingebrigtsen, S., Ifver, J., Kulmala, R. & Thomsen, L.K. 1995, "Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts", *Accident Analysis and Prevention*, vol. 27, no. 1, pp. 1-20.

Gabany, S.G., Plummer, P. & Grigg, P. 1997, "Why drivers speed: The speeding perception inventory", *Journal of Safety Research*, vol. 28, no. 1, pp. 29-35.

Gabbe, B. L., Sleney, J. S., Gosling, C. M., Wilson, K., Sutherland, A., Hart, M., Watterson, D. & Christie, N. 2014, "Financial and employment impacts of serious injury: A qualitative study", *Injury; International Journal of the Care of the Injured*, vol. 45, pp. 1445-1451.

Garber, N. J. 2000. *The Effect of Speed, Flow, and Geometric Characteristics on Crash Rates for Different Types of Virginia Highways.* Virginia Transportation Research Council, Charlottesville.

Geroliminis, N., Kepaptsoglou, K. & Karlaftis, M.G. 2011, "A hybrid hypercube - Genetic algorithm approach for deploying many emergency response mobile units in an urban network", *European Journal of Operational Research,* vol. 210, no. 2, pp. 287-300.

Gloag, D. 1993, "Europe needs more road safety", *British medical journal,* vol. 306, no. 6871, pp. 165.

Godley, S.T., Triggs, T.J. & Fildes, B.N. 2004, "Perceptual lane width, wide perceptual road centre markings and driving speeds", *Ergonomics,* vol. 47, no. 3, pp. 237-256.

Goldenbeld, C. & Schagen, I.V. 2005, "The effects of speed enforcement with mobile radar on speed speed and accidents. An evaluation study on rural roads in the Dutch province Friesland", *Accident Analysis and Prevention,* vol. 37, pp. 1135-1144.

Golias, J.C. 1992, "Establishing relationships between accidents and flows at urba priority road junctions", *Accident Analysis and Prevention,* vol. 24 no. 6, pp. 689-694.

Golob, T.F. & Recker, W.W. 2001, *Relationships Among Urban Freeway Accidents, Traffic Flow, Weather and Lighting Conditions,* Institute of Transportation Studies, University of California, California, Berkeley.

Golob, T.F., Recker, W.W. & Alvarez, V.M. 2004, "Freeway safety as a function of traffic flow", *Accident Analysis and Prevention,* vol. 36, no. 6, pp. 933-946.

Gombard, J.M. & Louah, J. 1986, *Free speed as a function of road geometrical characteristics*.

Gomes, S. V. 2013, "The influence of the infrastructure characteristics in urban road accidents occurrence", *Accident Analysis and Prevention,* vol. 60, pp. 289-297.

Gomes, S. V. & Cardoso, J.L. 2012, "Safety effects of low-cost engineering measures. An observational study in a Portuguese multilane road", *Accident Analysis and Prevention,* vol. 48, pp. 346-352.

Graham, D.J. & Glaister, S. 2003, "Spatial variation in road pedestrian casualties: The role of urban scale, density and land- use mix", *Urban Studies,* vol. 40, no. 8, pp. 1591-1607.

Greene, J., & d'Oliveira, M. 1982, *Learning to use statistical tests in psychology,* 2nd edn., Open University Press, Buckingham.

Greibe, P. 2003, "Accident prediction models for urban roads", *Accident Analysis and Prevention,* vol. 35, no. 2, pp. 273-285.

Guého, L., Granié, M. & Abric, J. 2014, "French validation of a new version of the Driver Behavior Questionnaire (DBQ) for drivers of all ages and level of experiences", *Accident Analysis and Prevention,* vol. 63, pp. 41-48.

Hakkert, A.S., Gitelman, V., Cohen, A., Doveh, E. & Umansky, T. 2001, "The evaluation of effects on driver behavior and accidents of concentrated general enforcement on interurban roads in Israel", *Accident Analysis and Prevention,* vol. 33, no. 1, pp. 43-63.

Hale, T., & Moberg, C. 2003, "Location Science Research: A Review", *Annals of Operations Research,* vol. 123, no. 1, pp. 21-35.

Harwood, D.W., Mason, J.M. & Brydia, R.E. 1999, "Design policies for sight distance at stop-controlled intersections based on gap acceptance", *Transportation Research Part A*, vol. 33, no. 3, pp. 199-216.

Hau, T.D. 1992, *Economic fundamentals of road pricing: a diagrammatic analysis*, The World Bank.

Hauer, E. 2004, "Statistical road safety modelling", *Transportation Research Record*, vol. 1897, no. 1, pp. 81-87.

Hauer, E. 1997, *Observational before-after studies in road safety,* 1st edn, Elsevier Science Ltd., Oxford, UK.

Hauer, E. 1971, "Accidents, overtaking and speed control", *Accident Analysis and Prevention,* vol. 3, no. 1, pp. 1-13.

Haupt, R.L. & Haupt, S.E. 2004, *Practical genetic algorithms*, 2nd edn, John Wiley, Hoboken, N.J.

Haynes, R., Lake, I.R., Kingham, S., Sabel, C.E., Pearce, J. & Barnett, R. 2008, "The influence of road curvature on fatal crashes in New Zealand", *Accident Analysis and Prevention,* vol. 40, no. 3, pp. 843-850.

Haynes, R., Jones, A., Kennedy, V., Harvey, I. & Jewell, T. 2007, "District variations in road curvature in England and Wales and their association with road-traffic crashes", *Environment and Planning A,* vol. 39, no. 5, pp. 1222-1237.

Haywood, J.C. 1980, *Highway alignment and superelevation: Some design-speed misconceptions*, TRL, Crowthorne, Berkshire.

He, J., King, M., Watson, B., Rakotonirainy, A. & Fleiter, J. 2013, "Speed enforcement in China: National, provincial and city initiatives and their success", *Accident Analysis and Prevention,* vol. 50, pp. 282-288.

Heitzenger, C. 2002, *Simulation and Inverse Modeling of Semiconductor Manufacturing Processes*, Vienna University of Technology, Austria.

Henkel, R.E. 1976, *Tests of significance*, Sage Publications, Beverly Hills ; London.

Herrin, G.D & Neuhardt, J.B. 1974, "An Empirical Model for Automobile Driver Horizontal Curve Negotiation", *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 16, no. 2, pp. 129-133.

Highways Agency 2002, February 2002-last update, *Design Manual for roads and bridges, Volume 6, Road Geometry, Section 1- Links TD9/93* [Homepage of Department for Transport, UK], [Online]. Available: http://www.dft.gov.uk/ha/standards/dmrb/vol6/section1.htm [2013, June/4th].

Hirst, W.M., Mountain, L.J. & Maher, M.J. 2005, "Are speed enforcement cameras more effective than other speed management measures? An evaluation of the relationship between speed and accident reductions", *Accident Analysis and Prevention*, vol. 37, no. 4, pp. 731-741.

Hirst, W.M., Mountain, L.J., Maher, M.J., 2004a. “Sources of error in road safety scheme evaluation: a method to deal with outdated accident prediction models”. *Accident Analysis and Prevention*. Vol. 36, pp. 717–727.

Hirst, W.M., Mountain, L.J., Maher, M.J., 2004b. “Sources of error in road safety scheme evaluation: a quantified comparison of current methods”. *Accident Analysis and Prevention*. vol. 36, pp. 705–715.

Hiselius, L.W. 2004, "Estimating the relationship between accident frequency and homogeneous and inhomogeneous traffic flows", *Accident Analysis and Prevention*, vol. 36, no. 6, pp. 985-992.

Holland, C., Geraghty, J. & Shah, K. 2010, "Differential moderating effect of locus of control on effect of driving experience in young male and female drivers", *Personality and Individual Differences*, vol. 48, no. 7, pp. 821-826.

Hooke, R. & Jeeves, T. A. 1961, "’Direct Search’ Solution of Numerical and Statistical Problems", *Journal of the Association of Computing Machinery*, vol. 8, pp. 212-229.

Horswill, M. & McKenna, F. 1999, "The effects of perceived control on risk-taking", *Journal of Applied Social Psychology*, vol. 29, no. 2, pp. 377-391.

Horvath, C., Lewis I. & Watson, B. 2012. “The beliefs which motivate young male and female drivers to speed: A comparison of low and high intenders”, *Accident Analysis and Prevention*, vol.45, pp.334-341.

Høye, A. 2015, "Safety effects of fixed speed cameras – An Empirical Bayes evaluation", *Accident Analysis and Prevention*, vol. 82, pp. 263-269.

Høye, A. 2015a, "Safety effects of section control – An Empirical Bayes evaluation", *Accident Analysis and Prevention*, vol. 82, pp. 263-269.

Indriasari, V., Mahmud, A.R., Ahmad, N. & Shariff, A.R. 2010, "Maximal service area problem for optimal siting of emergency facilities", *International Journal of Geographical Information Science,* vol. 24, no. 2, pp. 213-230.

Ivan, J.N., Wang, C. & Bernardo, N.R. 2000, "Explaining two- lane highway crash rates using land use and hourly exposure", *Accident Analysis and Prevention,* vol. 32, no. 6, pp. 787-795.

Ivey, D.L., Griffin, L.I., Newton, T.M., Lytton, R.L. & Hankins, K.C. 1981, "Predicting wet weather accidents", *Accident Analysis and Prevention,* vol. 13, no. 2, pp. 83-99.

Jamson, S., Chorlton, K. & Carsten, O. 2012, "Could Intelligent Speed Adaptation make overtaking unsafe", *Accident Analysis and Prevention,* vol. 48, pp. 29-36.

Jiménez-Mejìas, E., Prieto, C., Martinez-Ruiz, V., Del Castillo, J., Lardelli-Claret, P. & Jimenez-Moleon, J. 2014, "Gender-related differences in distances travelled, driving behaviour and traffic accidents among university students", *Transportation Research Part F-Traffic Psychology And Behaviour,* vol. 27, pp. 81-89.

Johansson, P. 1996, "Speed limitation and motorway casualties: A time series count data regression approach", *Accident Analysis and Prevention,* vol. 28, no. 1, pp. 73-87.

Johnston, I.R. 1982, "Modifying driver behaviour on rural road curves: A review of recent research", *Proceedings of 11th Australian Road Research Board (ARRB) Conference* Australian Road Research Board,, pp. 115.

Jonah, B.A. 1997, "Sensation seeking and risky driving: a review and synthesis of the literature", *Accident Analysis and Prevention,* vol. 29, no. 5, pp. 651-665.

Jonah, B.A. 1986, "Accident risk and risk- taking behaviour among young drivers", *Accident Analysis and Prevention,* vol. 18, no. 4, pp. 255-271.

Jones, A.P. & Jorgensen, S.H. 2003, "The use of multilevel models for the prediction of road accident outcomes", *Accident Analysis and Prevention,* vol. 35, no. 1, pp. 59-69.

Jones, A.P., Sauerzapf, V. & Haynes, R. 2008, "The effects of mobile speed camera introduction on road traffic crashes and casualties in a rural county of England", *Journal of Safety Research,* vol. 39, no. 1, pp. 101-110.

Jung, S., Qin, X. & Noyce, D.A. 2010, "Rainfall effect on single-vehicle crash severities using polychotomous response models", *Accident Analysis and Prevention,* vol. 42, no. 1, pp. 213-224.

Kadiyali, L.R., Viswanathan, E., Jain, P.K. & Gupta, R.K. 1981, *Effect of Curvature and Sight Distance on the Free Speed of Vehicles on Curves on Two-Lane Roads,* Central Road Research Institute, New Delhi, India.

Kane, J.N. 1981, *A record of first happenings, discoveries and inventions in American history,* 4th edn, Hw Wilson Co.

Kanellaidis, G. 1995, "Factors affecting drivers' choice of speed on roadway curves", *Journal of Safety Research,* vol. 26, no. 1, pp. 49-56.

Kanellaidis, G., Golias, J. & Zarifopoulos, K. 1995, "A survey of drivers' attitudes toward speed limit violations", *Journal of Safety Research,* vol. 26, no. 1, pp. 31-40.

Karlaftis, M.G. & Golias, I. 2002, "Effects of road geometry and traffic volumes on rural roadway accident rates", *Accident Analysis and Prevention,* vol. 34, no. 3, pp. 357-365.

Keall, M.D., Povey, L.J. & Frith, W.J. 2002, "Further results from a trial comparing a hidden speed camera programme with visible camera operation", *Accident Analysis and Prevention,* vol. 34, no. 6, pp. 773-777.

Keall, M.D., Povey, L.J. & Frith, W.J. 2001, "The relative effectiveness of a hidden versus a visible speed camera programme", *Accident Analysis and Prevention,* vol. 33, no. 2, pp. 277-284.

Keay, K. & Simmonds, I. 2006, "Road accidents and rainfall in a large Australian city", *Accident Analysis and Prevention,* vol. 38, no. 3, pp. 445-454.

Khorashadi, A., Niemeier, D., Shankar, V. & Mannering, F. 2005, "Differences in rural and urban driver-injury severities in accidents involving large- trucks: An exploratory analysis", *Accident Analysis and Prevention,* vol. 37, no. 5, pp. 910-921.

Kim, J., Kim, S., Ulfarsson, G.F. & Porrello, L.A. 2007, "Bicyclist injury severities in bicycle–motor vehicle accidents", *Accident Analysis and Prevention,* vol. 39, no. 2, pp. 238-251.

Kleisen, L.M.B. 2013. “A positive view on road safety: Can ‘car karma’ contribute to safe driving style?”, *Accident Analysis and Prevention,* vol. 50, pp. 705-712.

Koetse, M.J. & Rietveld, P. 2009, "The impact of climate change and weather on transport: An overview of empirical findings", *Transportation Research Part D,* vol. 14, no. 3, pp. 205-221.

Kononov, J., Bailey, B. and Allery, B., 2008. Relationships between safety and both congestion and number of lanes on urban freeways. *Transportation Research Record:Journal of the Transportation Research Board,* No.2083, pp. 26-39.

Kopits, E. & Cropper, M. 2003, *Traffic fatalities and economic growth. Washington, DC, The World Bank.*

Koustanaï, A., Boloix, E., Van Elslande, P. & Bastien, C. 2008. “Statistical analysis of “looked-but-failed-to-see” accidents: Highlighting the involvement of two distinct mechanisms”, *Accident Analysis and Prevention,* vol. 40, pp. 461-469.

Krammes, R.A., Brackett, R.Q., Shafer, M.A., Ottesen, J.L., Anderson, I.B., Fink, K.L., Collins, K.M., Pendleton, O.J. & Messer, C.J. 1995, *Horizontal Alignment Design Consistency for Rural Two-Lane Highways.FHWA-RD-94-034*, Federal Highways Administration, USA.

Lai, F. & Carsten, O. 2012, "What benefit does Intelligent Speed Adaptation deliver: A close examination of its effect on vehicle speeds", *Accident Analysis and Prevention*, vol. 48, pp. 4-9.

Lamm, R. & Choueiri, E.M. 1987, *Recommendations for Evaluating Horizontal Design Consistency Based on Investigations in the State of New York. Transportation Research Board Record No. 1122*, Transportation Research Board, Washington DC.

Lamm, R., Psarianos, B. & Mailaender, T. 1999, *Highway design and traffic safety engineering handbook*, McGraw-Hill, New York; London.

Langerman, J.J. & Ehlers, E.M. 2006, *The validation of evolutionary algorithms* [Homepage of Rand Afrikaan University], [Online]. Available: http://www.uj.ac.za/EN/Faculties/science/departments/csweb/about/staff/ProfEhlers/Documents/Thevalidationofevolutionaryalgorithms.pdf [2015, May/8th].

Lave, C.A. 1985, "Speeding, Coordination and the 55MPH Limit", *The American Economic Review*, vol. 75, no. 5, pp. 1159-1164.

Lee, P.M. 2004, *Bayesian Statistics: An introduction*, 3rd edn, Arnold, London.

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*, vol. 34, no. 2, pp. 149-161.

Leicestershire County Council. 2011, *Local Transport Planning in Leicestershire 2011 -2026. Leicestershire Local Transport Plan 3*, Leicestershire County Council, Leicestershire, UK.

Leong, H.J. 1968, "The distribution and trend of free speeds on two-way rural highways in New South Wales", *Transportation Research Record*, pp. 791-814.

Levine, N., Kim, K.E. & Nitz, L.H. 1995 a, "Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns", *Accident Analysis and Prevention*, vol. 27, no. 5, pp. 663-674.

Levine, N., Kim, K.E. & Nitz, L.H. 1995 b, "Spatial analysis of Honolulu motor vehicle crashes: II. Zonal generators", *Accident Analysis and Prevention*, vol. 27, no. 5, pp. 675-685.

Lewis, R. M. & Torczon, V. 1996, *Rank ordering and positive bases in pattern search algorithms*, NASA Contract No. NAS1-19480, Institute for Computer Applications in Science and Engineering, NASA Langley Research Centre, Hampton, Virginia.

Li, H.J., Graham, D.J. & Majumdar, A. 2013, "The impacts of speed cameras on road accidents: An application of propensity score matching methods", *Accident Analysis and Prevention,* vol. 60, pp. 148-157.

Li, L.H., Zhu, L. & Sui, D.Z. 2007, "A GIS- based Bayesian approach for analyzing spatial-temporal patterns of intra-city motor vehicle crashes", *Journal of Transport Geography,* vol. 15, no. 4, pp. 274-285.

Li, W., Carriquiry, A., Pawlovich, M. & Welch, T. 2008, "The choice of statistical models in road safety countermeasure effectiveness studies in Iowa", *Accident Analysis and Prevention,* vol. 40, no. 4, pp. 1531-1542.

Li, X. & Yeh, A. 2005, "Integration of genetic algorithms and GIS for optimal location search", *International Journal of Geographical Information Science,* vol. 19, no. 5, pp. 581-601.

Li, X., He, J. & Liu, X. 2009, "Intelligent GIS for solving high-dimensional site selection problems using ant colony optimization techniques", *International Journal of Geographical Information Science,* vol. 23, no. 4, pp. 399-416.

Lim, S. & Kuby, M. 2010, "Heuristic algorithms for siting alternative- fuel stations using the Flow- Refueling Location Model", *European Journal of Operational Research,* vol. 204, no. 1, pp. 51-61.

Lord, D. 2006, "Modeling motor vehicle crashes using Poisson-gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter", *Accident Analysis and Prevention,* vol. 38, no. 4, pp. 751-766.

Lord, D., Guikema, S.D. & Geedipally, S.R. 2008, "Application of the Conway-Maxwell- Poisson generalized linear model for analyzing motor vehicle crashes", *Accident Analysis and Prevention,* vol. 40, no. 3, pp. 1123-1134.

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,* vol. 44, no. 5, pp. 291-305.

Lord, D., Washington, S. & Ivan, J.N. 2007, "Further notes on the application of zero- inflated models in highway safety", *Accident Analysis and Prevention,* vol. 39, no. 1, pp. 53-57.

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,* vol. 37, no. 1, pp. 35-46.

Lord, D., Manar, A. & Vizioli, A. 2005, "Modeling crash-flow- density and crash-flow-V/C ratio relationships for rural and urban freeway segments", *Accident Analysis and Prevention,* vol. 37, no. 1, pp. 185-199.

Loughborough University, *Data Protection Policy* [Homepage of Loughborough University], [Online]. Available: http://www.lboro.ac.uk/admin/ar/policy/dpact/ludpp/ [2012, April/12th].

Loughborough University, *Guidance on making a submission to the ethical advisory committee* [Homepage of Loughborough University], [Online]. Available: http://www.lboro.ac.uk/admin/committees/ethical/gn/sub-guide1.htm [2012, April/4th].

Lu Xiang-lin, L.X. & Hou Yun-xian, H.Y. 2009, *Ant Colony Optimization for Facility Location for Large-Scale Emergencies*.

Ma, J., Kockelman, K.M. & Damien, P. 2008, "A multivariate Poisson- lognormal regression model for prediction of crash counts by severity, using Bayesian methods", *Accident Analysis and Prevention*, vol. 40, no. 3, pp. 964-975.

Macnab, Y. 2004, "Bayesian spatial and ecological models for small- area accident and injury analysis", *Accident Analysis and Prevention*, vol. 36, no. 6, pp. 1019-1028.

Malczewski, J. 2006, "GIS-based multicriteria decision analysis: a survey of the literature", *International Journal of Geographical Information Science*, vol. 20, no. 7, pp. 703-726.

Martin, J. 2002, "Relationship between crash rate and hourly traffic flow on interurban motorways", *Accident Analysis and Prevention*, vol. 34, no. 5, pp. 619-629.

Mathworks, 2015, *MATLAB Global Optimization Toolbox User's Guide*, Mathworks, Massachusetts, USA.

McCullagh, P & Nelder, J.A. 1990, *Generalized Linear Models, 2nd Edition.* Chapman and Hall Publishers, London.

McDonald, N. 2004, "Look and Learn. Capitalising on individual responsibility in speed management", *Proceedings of the 2004 Australian Institute of Traffic Planning and Management (AITPM) National Conference* Australian Institute of Traffic Planning and Management, Adelaide, pp. 71.

McKenna, F.R. 2005, *Behavioural research in road safety: Why do drivers break the speed limit*, Department for Transport, UK.

McLean, J. 1981, "Driver speed behaviour and rural road alignment design", *Traffic Engineering and Control*, vol. 22, no. 4, pp. 208-211.

McLean, J.R. 1979, *An alternative to the design speed concept for low-speed alignment design*, Australian Road Research Board, Vermont, South Australia.

Miaou, S. 1994, "The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions", *Accident Analysis and Prevention,* vol. 26, no. 4, pp. 471-482.

Miettinen, K.M. 1999, *Nonlinear multiobjective optimisation,* Boston. Kluwer Academic Publishers.

Miloua, R., Kebbab, Z.,Chiker, F., Sahraoui, K., Khadraoui, M & Benramdane, N. 2011, "Determination of layer thickness and optical constants of thin films by using a modified pattern search method", *Optical Letters,* vol. 37, pp. 449-451

Miltner, E. & Salwender, H.-J. 1995, "Influencing factors on the injury severity of restrained front seat occupants in car-to-car head-on collisions", *Accident Analysis and Prevention,* vol. 27, no. 2, pp. 143-150.

Milton, J.C., Shankar, V.N. & Mannering, F.L. 2008, "Highway accident severities and the mixed logit model: An exploratory empirical analysis", *Accident Analysis and Prevention,* vol. 40, no. 1, pp. 260-266.

Milton, J. & Mannering, F. 1998, "The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies", *Transportation,* vol. 25, no. 4, pp. 395-413.

Mitchell, C.G.B. 2012. “Speed and safety: Evidence from Published Data”, *Royal Automobile Club Foundation for Motoring, London.*

Mitchell, C.G.B. 2013. “The licensing and safety of older drivers in Britain”, *Accident Analysis and Prevention,* vol. 50, pp. 732-741.

Mitchell, R., Driscoll, T. & Healey, S. 2004, "Work- related road fatalities in Australia", *Accident Analysis and Prevention,* vol. 36, no. 5, pp. 851-860.

Morris, A.P., Hancox, G., Martin, O., Bell, D., Johansson, C., Rosander, P., Scholliers, J. & Silla, A. 2013, "Critical accident scenarios for cyclists and how they can be addressed through ITS solutions", *Proceedings, International Cycling Safety Conference. 20-21 November, Helmond, The Netherlands*.

Morrison, D.S., Petticrew, M. & Thomson, H. 2003, "What are the most effective ways of improving population health through transport interventions? Evidence from systematic reviews. (Public Health Policy and Practice)", *Journal of Epidemiology & Community Health,* vol. 57, no. 5, pp. 327.

Morton, V. & Torgerson, D. J. 2003, "Effect of regression to the mean on decision making in health care", *BMJ: British Medical Journal,* vol. 326, no. 7398, pp. 1083-1084.

Moták, L., Gabaude, C., Bougeant, J. & Huet, N. 2014, "Comparison of driving avoidance and self-regulatory patterns in younger and older drivers", *Transportation Research Part F-Traffic Psychology And Behaviour,* vol. 26, pp. 18-27.

Mountain, L.J., Hirst, W.M. & Maher, M.J. 2005, "Are speed enforcement cameras more effective than other speed management measures?: The impact of speed management schemes on 30 mph roads", *Accident Analysis and Prevention,* vol. 37, no. 4, pp. 742-754.

Mountain, L.J., Hirst, W.M. & Maher, M.J. 2004, "Costing lives or saving lives? A detailed evaluation of the impact of speed cameras on safety", *Traffic Engineering and Control,* vol. 45, pp. 280-287.

Mountain, L., Fawaz, B. & Jarrett, D. 1996, "Accident prediction models for roads with minor junctions", *Accident Analysis and Prevention,* vol. 28, no. 6, pp. 695-707.

Murray, C.J.L. & Lopez, A.D. 1996, *The global burden of disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020,* 1st edn, Harvard School of Public Health, Boston, Massachusetts.

Murray, C.J.L., Lopez, A.D., Mathers, C.D. & Stein, C. 2001, *The Global Burden of Disease 2000 project: aims, methods and data sources*. World Health Organisation.

Myers, R.H., Montgomery, D.C. & Vining G.G 2001, *Generalized linear models: with applications in Engineering and the Sciences,* Wiley, Chichester, UK.

Navin, F., Zein, S. & Felipe, E. 2000, "Road safety engineering: an effective tool in the fight against whiplash injuries", *Accident Analysis and Prevention,* vol. 32, no. 2, pp. 271-275.

Newstead, S. 2009, *Evaluation of the crash effects of the Queensland mobile speed camera program in the years 2006-2007*, Monash University Accident Research Centre, Victoria, Australia.

Noland, R.B. & Oh, L. 2004, "The effect of infrastructure and demographic change on traffic-related fatalities and crashes: A case study of Illinois county-level data", *Accident Analysis and Prevention,* vol. 36, no. 4, pp. 525-532.

Noland, R.B. & Quddus, M.A. 2004, "A spatially disaggregate analysis of road casualties in England", *Accident Analysis and Prevention,* vol. 36, no. 6, pp. 973-984.

Noland, R. & Quddus, M.A. 2005, "Congestion and safety: A spatial analysis of London", *Transportation Research Part A-Policy And Practice,* vol. 39, no. 7-9, pp. 737-754.

Noland, R.B. 2003, "Traffic fatalities and injuries: the effect of changes in infrastructure and other trends", *Accident Analysis and Prevention,* vol. 35, no. 4, pp. 599-611.

Nottingham City Council , *Nottinghamshire Insight Mapping* [Homepage of Nottingham city council], [Online]. Available: http://info.nottinghamcity.gov.uk/insightmapping/# [2013, June/20th].

Nottingham City Council. 2011, *Nottingham Local Transport Plan Strategy 2011 – 2026*, Nottingham City Council, Nottingham, UK.

Nottinghamshire County Council. 2010, *'Safer Roads' – The Casualty Report 2009*, Highway Safety Group, Nottinghamshire County Council, Nottinghamshire, UK.

Novoa, A.M., Pérez, K., Santamariña-Rubio, E., Marí-Dell'Olmo, M. & Tobías, A. 2010, "Effectiveness of speed enforcement through fixed speed cameras: a time series study", *Injury prevention : journal of the International Society for Child and Adolescent Injury Prevention*, vol. 16, no. 1, pp. 12.

O'Flaherty, C.A.O. & et.al. 1997, *Transport Planning and Traffic Engineering*.

Oppenheim, A.N. 2000, *Questionnaire design, interviewing and attitude measurement*, New ed., Pinter, London.

Ossiander, E.M. & Cummings, P. 2002, "Freeway speed limits and traffic fatalities in Washington State", *Accident Analysis and Prevention*, vol. 34, no. 1, pp. 13-18.

Özkan, T. & Lajunen, T. 2005, "A new addition to DBQ: Positive Driver Behaviours Scale", *Transportation Research Part F-Traffic Psychology And Behaviour*, vol. 8, no. 4-5, pp. 355-368.

Park, E.S., Park, J. & Lomax, T.J. 2010, "A fully Bayesian multivariate approach to before-after safety evaluation", *Accident Analysis and Prevention*, vol. 42, no. 4, pp. 1118-1127.

Parliamentary Office of Science and Technology. 2004. "*Speed Cameras*" POSTnote 04/218. London, UK

Peden, M., McGee, K. & Sharma, G. 2002, , *The Injury Chartbook: A graphical overview of the global burden of injuries*. Available: http://whqlibdoc.who.int/publications/924156220X.pdf [2011, October/20].

Peden, M.et.al. 2004, *World report on road traffic injury prevention*. World Health Organisation, Geneva.

Peng R, Xu R-H and Qin J. 2008. "Bi-level Simulated Annealing Algorithm for Facility Location Problem". *International Conference on Information Management, Innovation Management and Industrial Engineering*. Dec. 2008, vol. 3, pp. 17-22.

Pérez, K., Marl-Dell'Olmo, M., Tobias, A. & Borrell, C. 2007, "Reducing road traffic injuries: Effectiveness of speed cameras in an urban setting", *American Journal of Public Health*, vol. 97, no. 9, pp. 1632-1637.

Pérez, I. 2006, "Safety impact of engineering treatments on undivided rural roads", *Accident Analysis and Prevention*, vol. 38, no. 1, pp. 192-200.

Persaud, B., Lan, B., Lyon, C. & Bhim, R. 2010, "Comparison of empirical Bayes and full Bayes approaches for before-after road safety evaluations", *Accident Analysis and Prevention*, vol. 42, pp. 38-43.

Persaud, B. & Lyon, C. 2007, "Empirical Bayes before-after studies: lessons learned from two decades of experience and future directions", *Accident Analysis and Prevention*, vol. 39, pp. 546-555.

Phillips, R.O., Bjørnskau, R.H., Hagman, R. & Sagberg, F. 2011, "Reduction in car-bicycle conflict at a road-cycle path intersection: Evidence of road user adaptation?", *Transportation Research Part F*, vol. 14, pp. 87-95.

Pilkington, P. & Kinra, S. 2005, "Effectiveness Of Speed Cameras In Preventing Road Traffic Collisions And Related Casualties: Systematic Review", *BMJ: British Medical Journal*, vol. 330, no. 7487, pp. 331-334.

PR Newswire. 2005. *World's first car celebrates 120th birthday*. *New York*. Available: http://www.prnewswire.com/news-releases/worlds-first-car-celebrates-120th-birthday-53519292.html [2011, October/21st].

Pucher, J. (2001). Cycling safety on bikeways versus roads. Transportation Quarterly, 55.

Quddus, M.A. 2008, "Modelling area- wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data", *Accident Analysis and Prevention*, vol. 40, no. 4, pp. 1486-1497.

Quddus, M.A. 2008, "Time series count data models: An empirical application to traffic accidents", *Accident Analysis and Prevention*, vol. 40, no. 5, pp. 1732-1741.

Rakotonirainy, A., Steinhardt, D., Delhomme, P., Darvell, M. & Schramm, A. 2012. “Older drivers’ crashes in Queensland, Australia”, *Accident Analysis and Prevention*, vol. 48, pp. 423-429.

Reeves, C.R. 1993, *Modern heuristic techniques for combinatorial problems*, Blackwell Scientific Publications, Oxford.

Reid, L.D. 1983, "A survey of recent driver steering behavior models suited to accident studies", *Accident Analysis and Prevention*, vol. 15, no. 1, pp. 23-40.

Reinfurt, D.W., Zegeer, C.V., Shelton, B.J. & Neuman, T.R. 1992, *Analysis of vehicle operations on horizontal curves (Transportation Research Record No. 1318)*, Transportation Research Board, Washington DC.

Retting, R.A., Kyrychenko, S.Y. & McCartt, A.T. 2008, "Evaluation of automated speed enforcement on Loop 101 freeway in Scottsdale, Arizona", *Accident Analysis and Prevention*, vol. 40, no. 4, pp. 1506-1512.

Retting, R., Farmer, C. & Mccartt, A. 2008, "Evaluation of Automated Speed Enforcement in Montgomery County, Maryland", *Traffic Injury Prevention*, vol. 9, no. 5, pp. 440-445.

Ritchie, M.L., McCoy, W.K. & Welde, W.L. 1968, "A study of the relation between forward velocity and lateral acceleration in curves during normal driving", *Human factors*, vol. 10, no. 3, pp. 255.

RoadPeace. 2012, , *August National Road Safety Month: RoadPeace calls for road crime to be treated as real crime*. Available: http://roadpeace.org/resources/PR_010812_Road_Crime_is_Real_Crime.pdf [2011, October 22].

Roberts, I. 2005, "Death On The Road To International Development". *British Medical Journal*, vol. 7497, pp330-972.

Robson, C., 2011, *Real world research: a resource for users of social research methods in applied settings*, 3rd edn, Wiley, Chichester.

Rosén, E., Stigson, H. & Sander, U. 2011, "Literature review of pedestrian fatality risk as a function of car impact speed", *Accident Analysis and Prevention*, vol. 43, no. 1, pp. 25-33.

Rosén E and Sander U. 2009, Pedestrian fatality risk as a function of car impact speed. *Accident Analysis and Prevention*. Vol. 41, no. 3, pp. 536-542.

ROSPA 2011, December 2011-last update, *Speed cameras* [Homepage of The Royal Society for the Prevention of Accidents], [Online]. Available: http://www.rospa.com/rospaweb/docs/advice-services/road-safety/drivers/speed-camera-factsheet-1211.pdf [2014, December/10th].

Russell, S.J. 2010, *Artificial intelligence : a modern approach*, 3rd ed., International edn, Pearson, Boston, Mass.] ; London.

Safe Motoring. 2013, 19th August 2013-last update, *Understanding speed limits* [Homepage of Safe Motoring], [Online]. Available: http://www.safermotoring.co.uk/understandingspeedlimits.html [2014, September/20th].

Sakawa, M. 1993, *Fuzzy sets and interactive multiobjective optimization*, Plenum Press, New York.

Sakawa, M., Kato, K., Sunada, H. & Shibano, T. 1997, "Fuzzy programming for multiobjective 0- 1 programming problems through revised genetic algorithms", *European Journal of Operational Research; Eur.J.Oper.Res.,* vol. 97, no. 1, pp. 149-158.

Salhi, S. & Gamal, M.D.H. 2003, "A Genetic Algorithm Based Approach for the Uncapacitated Continuous Location–Allocation Problem", *Annals of Operations Research,* vol. 123, no. 1, pp. 203-222.

Santiago-Chaparro, K.R., Chitturi, M., Bill, A. & Noyce, D.A. 2012, "Spatial effectiveness of speed feedback signs", *Transportation Research Record, Journal of the Transportation Research Board,* vol. 2281, pp. 8-15.

Saris, W.E. & Gallhofer, I.N. 2007, *Design, evaluation, and analysis of questionnaires for survey research,* Wiley ; John Wiley distributor, Hoboken, N.J. : Chichester.

Sasaki, S., Comber, A.J., Suzuki, H. & Brunsdon, C. 2010, "Using genetic algorithms to optimise current and future health planning - the example of ambulance locations.(Research Report)", *International Journal of Health Geographics,* vol. 9, pp. 4.

Scheaffer, R.L., Mendenhall, W. & Ott, L. 2005, *Elementary survey sampling,* 6th edn, Brooks-Cole, Belmont, Calif. ; London.

Schelling, T.C. 1978, *Micromotives and Macrobahaviour, New York,* W.W.Norton & Company, Inc. Publishers.

Schepers, J.P., Kroeze, P.A., Sweers, W. & Wüst, J.C. 2011, "Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections", *Accident Analysis and Prevention,* vol. 43, pp. 853-861.

Shankar, V.N., Albin, R.B., Milton, J.C. & Mannering, F.L. 1998, "Evaluation of median crossover likelihoods with clustered accident counts: an empirical inquiry using the random effect negative binomial model", *Transport. Res. Record,* vol. 1635, pp. 44-48.

Shankar, V., Milton, J. & Mannering, F. 1997, "Modeling accident frequencies as zero-altered probability processes: An empirical inquiry", *Accident Analysis and Prevention,* vol. 29, no. 6, pp. 829-837.

Shankar, V., Mannering, F. & Barfield, W. 1996, "Statistical analysis of accident severity on rural freeways", *Accident Analysis and Prevention,* vol. 28, no. 3, pp. 391-401.

Shefer, D. 1994, "Congestion, air pollution, and road fatalities in urban areas", *Accident Analysis and Prevention,* vol. 26, no. 4, pp. 501-509.

Shin, K., Washington, S.P. & van Schalkwyk, I. 2009, "Evaluation of the Scottsdale Loop 101 automated speed enforcement demonstration program", *Accident Analysis and Prevention,* vol. 41, no. 3, pp. 393-403.

Siddiqui, M.Z., Everett, J.W. & Vieux, B.E. 1996, "Landfill siting using geographic information systems: a demonstration", *Journal of Environmental Engineering,* vol. 122, no. 6, pp. 515.

Silverman, D. 2006, *Interpreting qualitative data: methods for analyzing talk, text and interaction,* 3rd edn, SAGE, London.

Smeed, R.J. 1972, "The usefulness of formulae in traffic engineering and road safety", *Accident Analysis and Prevention,* vol. 4, no. 4, pp. 303-312.

Smeed, R.J. 1949, "Some Statistical Aspects of Road Safety Research", *Journal of the Royal Statistical Society. Series A (General),* vol. 112, no. 1, pp. 1-34.

Song, J.J., Ghosh, A., Miaou, S. & Mallick, B. 2006, "Bayesian multivariate spatial models for roadway traffic crash mapping", *Journal of Multivariate Analysis,* vol. 97, no. 1, pp. 246-273.

Speed camera UK. 2013, Feb 2013-last update, *UK speed camera locations*. Available: http://www.speedcameramap.co.uk/ [2015, May/10th].

Stewart, J.A. 1994, "The Poisson-lognormal model for bibliometric/scientometric distributions", *Information Processing & Management,* vol. 30, no. 2, pp. 239-251.

Summala, H. 1996, "Accident risk and driver behaviour", *Safety Science,* vol. 22, no. 1, pp. 103-117.

Tarriére, C. 1995, "Children are not miniature adults", *Proceedings of the 1995 International IRCOBI Conference on the Biomechanics of Impact, Brunnen, Switzerland* , pp. 15.

Tay, R. 2010, "Speed Cameras Improving Safety or Raising Revenue?", *Journal of Transport Economics and Policy;* vol. 44, pp. 247-257.

Taylor, M.C., Baruya, A. & Kennedy, J.V. 2002, *The relationship between speed and accidents on rural single-carriageway roads.TRL511*, Transport Research Laboratory, UK.

Taylor, M.C., Lynam, D.A. & Baruya, A. 2000, *The effects of drivers' speed on the frequency of road accidents.TRL421*, Transport Research Laboratory, Crowthorne, UK.

The Economist. 2009, *Boost for Formula One* [Homepage of British Broadcasting Corporartion], [Online]. Available: http://www.economist.com/node/13350762 [2013, October/18th].

The World Bank. 2011, *Ensuring the decade is action. UN decade of action for road safety 2011-2020*. Available: http://siteresources.worldbank.org/EXTTOPGLOROASAF/Resources/2582212-1265307800361/decade_of_action_2011.pdf [2012, June/10].

Thomas, P., Welsh, R., Quddus, M., Frampton, R & Lenard, J. 2012, *The relationship between speed and casualties*, Draft Final Report. DfT Project Ref: PPRO 04/06/39, Department for Transport, UK.

Train, K.E., 2003. Discrete choice methods with simulation. Cambridge University Press.

Tranter, P. & Warn, J. 2008, "Relationships between interest in motor racing and driver attitudes and behaviour amongst mature drivers: An Australian case study", *Accident Analysis and Prevention*, vol. 40, no. 5, pp. 1683-1689.

Trimpop, R & Kirkcaldy, B. 1997. “Personality predictors of driving accidents”, *Personality and Individual Differences*, vol. 23 Issue 1, pp. 147-152.

Tzeng, G.H. & Chen, Y.W. 1999, "The optimal location of airport fire stations: A fuzzy multi-objective programming and revised genetic algorithm approach", *Transportation Planning and Technology*, vol. 23, no. 1, pp. 37-55.

UK Motorists, *Speed limits in the UK* [Homepage of UK Motorists], [Online]. Available: http://www.ukmotorists.com/speed_history.asp [2014, September/20th].

United Nations. 2012. *One year on, global commitment to the Decade of Action for Road Safety 2011-2020 intensifies*. Available: htto://www.who.int/roadsafety/en/ [2012, June/06].

Usman, T., Fu, L. & Miranda-Moreno, L. 2012, "A disaggregate model for quantifying the safety effects of winter road maintenance activities at an operational level", *Accident Analysis and Prevention*, vol. 48, pp. 368-378.

Usman, T., Fu, L. & Miranda-Moreno, L. 2010, "Quantifying safety benefit of winter road maintenance: Accident frequency modeling", *Accident Analysis and Prevention*, vol. 42, no. 6, pp. 1878-1887.

Vorko-Jović, A., Kern, J. & Biloglav, Z. 2006, "Risk factors in urban road traffic accidents", *Accident Analysis and Prevention*, vol. 37, pp. 93-98.

Vose, M.D. & Wright, A.H. 1998, "The Simple Genetic Algorithm and the Walsh Transform: Part II, The Inverse", *Evolutionary computation*, vol. 6, no. 3, pp. 275-289.

Wahlgren, O. 1967, *The Dependence of Vehicle Speeds on Different Factors – Particularly Road Geometry on Two-Lane Highways in Finland.Scientific Researches No. 22*, Finland Institute of Technology, Helsinki.

Walter, L. & Broughton, J. 2011, "Effectiveness of speed indicator devices: An observational study in South London", *Accident Analysis and Prevention*, vol. 43, no. 4, pp. 1355-1358.

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,* vol. 43, no. 6, pp. 1979-1990.

Wang, C., Quddus, M.A. & Ison, S.G. 2011a, "A spatio-temporal analysis of the impact of congestion on traffic safety on major roads in the UK", *Transportmetrica A: Transport Science,* vol. 9, no. 2, pp. 124-148.

Wang, C., Quddus, M.A. & Ison, S.G. 2009, "Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England", *Accident Analysis and Prevention,* vol. 41, no. 4, pp. 798-808.

Ward, H. 2003, *The history and development of speed camera use in Great Britain*, Centre for Transport Studies, University College London, London.

Warner, H.W., Özkan, T. & Lajunen, T. 2010, "Can the traffic locus of control (T-LOC) scale be successfully used to predict Swedish drivers' speeding behaviour?", *Accident Analysis and Prevention,* vol. 42, no. 4, pp. 1113-1117.

Washington State Department of Transport 2010, *Design Manual M 22-01.07.*

Wasielewski, P. 1984, "Speed as a measure of driver risk: Observed speeds versus driver and vehicle characteristics", *Accident Analysis and Prevention,* vol. 16, no. 2, pp. 89-103.

Wei, L., Li, H., Liu, Q., Chen, J. & Cui, Y. 2011, "Study and implementation of fire sites planning based on GIS and AHP", *The 5th Conference on Performance-based Fire and Fire Protection Engineering* Procedia Engineering, , pp. 486.

Wetter, M. & Wright, J. 2003, "Comparison of a Generalized Pattern Search and Genetic Algorithm optimization method", *8th International IBPSA Conference,* vol. 3, pp. 1401-1408.

Whitley, L.D., Howe, A.E., Rana, S., Watson, J.P. & Barbulescu, L. 1998. Comparing Heuristic Search Methods and Genetic Algorithms for Warehouse Scheduling. *The Institute of Electrical and Electronics Engineers*.

Winnett, M.A. & Wheeler A.H. 2002, *Vehicle-activated signs - a large scale evaluation*.

Withers, S.D. 2002, "Quantitative methods: Bayesian inference, Bayesian thinking", *Progress in Human Geography,* vol. 26, no. 4, pp. 553-566.

World Health Organisation. 2013, *Global Status Report on Road Safety 2013. Supporting a decade of action*, World Health Organisation, Geneva, Switzerland.

World Health Organisation. 2009, *Global status report on road safety; Time for action*. Geneva, Switzerland.

World Health Organisation. 2005, 60/5 *Improving Global Road Safety*. Geneva, Switzerland.

World Health Organisation. 2001. *The World Health Report 2001. Mental health: new understanding, new hope*. Geneva, Switzerland.

World Health Organisation. 1974, *Prevention of road traffic accidents*. Geneva, Switzerland.

Xiang-lin Lu, X.L., Yun-xian Hou, Y.H. & Shen Qiang, S.Q. 2010, *A fuzzy queuing facility location model with ant colony optimization algorithm for large-scale emergencies*.

Xiao, N., Bennett, D.A. & Armstrong, M.P. 2007, "Interactive evolutionary approaches to multiobjective spatial decision making: A synthetic review", *Computers, Environment and Urban Systems,* vol. 31, no. 3, pp. 232-252.

Yagar, S. 1984, "Predicting speeds for rural 2-lane highways", *Transportation Research Part A: General*, vol. 18, no. 1, pp. 61-70.

Yagar, S. & Van Aerde, M. 1983, "Geometric and environmental effects on speeds of 2- lane highways", *Transportation Research Part A: General,* vol. 17, no. 4, pp. 315-325.

Yang, L., Jones, B.F. & Yang, S. 2007, "A fuzzy multi-objective programming for optimization of fire station locations through genetic algorithms", *European Journal of Operational Research,* vol. 181, no. 2, pp. 903-915.

Zalzala, A. M. & Fleming, P. J. 1997, *Genetic Algorithms in Engineering Systems,* Institution of Electrical Engineers.

Zeleny, M. 1982, *Multiple criteria decision making,* McGraw-Hill, New York.

Zheng, X. & Chen, Z. 2010, "Back-calculation of the strength and location of hazardous materials releases using pattern search method", *Journal of Hazardous Materials,* vol. 183, pp. 474-481

Zou, Y., Zhang, Y. & Lord, D. 2013, "Application of finite mixture of negative binomial regression models with varying weight parameters for vehicle crash data analysis", *Accident Analysis and Prevention,* vol. 50, pp. 1042-1051.

Appendices

Appendix A: Site visits risk assessment

	Special Considerations
Loughborough Design School **Research Site Visits risk assessment** Project: PhD Research Date: as required/to be confirmed Project manager: Dr. Richard Frampton and Dr. Andrew Morris Risk Assessment by: Agnes Wallace-Frimpong	m Ethical Committee Approval needed m CRB clearance needed m External staff / partners involved ⦿

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
1.	**Travel to and from sites by car (hire car, personal car etc)**					
	Driving	Discomfort/injury	2	4	8	• Driver must have full driving licence • Any points on the driver's licence must be made known to insurance company but to colleagues if necessary. • Driver must have enough rest prior to driving • Driver must try and make seat and posture comfortable prior to setting off. • Personal car must have insurance for business purposes

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
						• Ensure conditions for car use are complied with • Where possible, share driving task ie. take turns to drive • Ensure seat belts are worn at all times when driving • Non-driver must be encouraged to wear seat belt • Driving should be in accordance with the highway code rules
	Breakdown Road traffic Accident Break in to car	Stress, delay, anxiety Injury, death, anxiety Stress, delay, loss of valuable items	2 3 2	4 7 4	8 21 8	• Allow adequate time for journeys • Inform colleagues when setting off and let them know when you arrive at your destination • Organise journeys such that as much work as is possible can be carried out on successful journey days to make up for unsuccessful/unfruitful days • Ensure hired car has good breakdown assistance • In the case of personal cars, there should be adequate breakdown assistance • Ensure car is well locked, windows rolled up for all doors, keep items that can be burgled out of sight of passers-by, park car in a safe place where you can keep an eye on it from time to time.

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
	Break/Interruption in communication	Stress, delay, loss of valuable items	2	4	8	• All researchers must carry a fully charged mobile phone at all times with a spare one as well if possible. • In the case of pay-as-you-go phones, there must be enough credit available
2. Travel to and from sites by public transport (bus, train, taxi etc)						
	Breakdown Accident Personal safety	Stress, delay, anxiety Injury, death, stress, delay Stress, injury, death	2 3 2	4 7 4	8 21 8	• Allow enough time for journeys • Organise journeys such that as much work as is possible can be carried out on successful journey days to make up for unsuccessful/unfruitful days • Travel with another colleague if possible • Avoid lone working as much as possible: if this is unavoidable, arrangements must be made with a colleague in the Design school for site worker to call colleague from time to time eg. every 1 hour • If travel tickets are to be bought in advance do that through reputable companies • If mode of travel is by taxi, ensure a reliable company is used • Avoid late night travel/working • Inform colleagues about journey details • Ensure enough rest has been obtained before journey to avoid stress and lack of concentration
	Break/Interruption in	Stress, delay, loss of	2	4	8	• All researchers must carry a fully

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
	communication	valuable items				charged mobile phone at all times with a spare one as well if possible. • In the case of pay-as-you-go phones, there must be enough credit available
3. Traffic conditions						
	Accident Traffic Noise Objects being thrown by motorists at researchers, verbal abuse and aggressive behaviours by passing motorists	Injury, death Ability of researchers to communicate Stress, anxiety, injury	3	7	21	• Avoid stepping into live traffic • Ensure the right Personal Protective Equipment (PPE) is worn and it is comfortable • Walk facing traffic as much as is possible • Use walkways where this is available • Keep personal/survey items well secured to avoid it falling into carriageway • Acquaint selves with road layout and geometric constraints before commencing actual study. • Work away from live traffic as much as is possble • Try and get nearer to aid good communication and avoid shouting from a distance • Avoid confrontations with passing motorists • Keep survey items out of reach of passing motorists

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
4. Personal Safety and Personal Welfare						
	Accidents (slips, trips, falls, electrocution etc) Harassment Assault and attack by members of the public	Stress, anxiety, injury, death Stress, injury, anxiety Stress, injury, death	3	4	12	• Ethical committee approval obtained to collect data • Appropriate insurance obtained if possible • Documentation of any pre-arranged appointments with other parties not in LDS • Any necessary safety training/induction must be obtained. • Appropriate personal protective equipment (PPE) and clothing must be worn • If necessary, appropriate first aid training must be carried out and equipment made available • Before the visit, try to find out about the area to be visited by asking someone with local knowledge about the area • Do not create any suspicion to community members • Where possible explain in lay-man's language what the whole project is about when asked in order to allay any fears • If confrontation arises just walk away and leave vicinity as soon as possible • Report to Police in extreme

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
						circumstances • Ensure breaks are taken to have some food/refreshment to renew energy • Avoid lone working as much as possible: if this is unavoidable, arrangements must be made with a colleague in the Design school for site worker to call colleague from time to time eg. every 1 hour • Avoid working in areas with known bad reputation. Where this is unavoidable lone working should be omitted. • Work within sight of other colleagues (avoid working down alleyways out of sight from public view/other colleagues) • Consider the use of two-way radio if working some distance apart where two or more researchers are involved. • Where possible each researcher should have a key to the vehicle
	Break/Interruption in communication	Stress, delay, loss of valuable items	2	4	8	• All researchers must carry a fully charged mobile phone at all times with a spare one as well if possible. • In the case of pay-as-you-go phones, there must be enough credit available on phone • Make sure there is good mobile phone signalling at location before starting work and if not move to a good location

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
						to enable effective communication • A nominated colleague contactable by phone must be made aware of the destination of the researchers, how long they expect to be out for and an expected return time as well as phone numbers of all the researchers out on site and vice-versa • The nominated colleague must be kept informed by contacting initially before any work starts, when work is finished, if returning to the office or not, when the alternative location other than the office is reached safely and more importantly contacted regularly • Nominated colleague should keep written/documented record of all contact details this should include researcher/s involved, location, start time, end time, expected return time in addition to times regular calls are received • If nominated colleague does not receive call at agreed time for regular calls, colleague should try contacting researchers on site and if feasible make a trip to the site if not the Police must be contacted about situation.
5. Ground Conditions						
	Uneven ground	Slips, trips, falls and electrocution	2	5	10	• Researchers to wear comfortable footwear

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
	Soft ground Exposed cables Kerbing Signing					• Walk and avoid running • Stay clear of excavations • Stay clear of exposed cables • Take extra care should you encounter soft/wet ground • In areas where there is kerbing, care should be taken not to slip, fall or trip into live traffic and ground • Care should be taken not to walk into erected signs by the road
6.	**Environment**					
	Inclement weather (windy, rainy etc) Nuisance (noise, dust etc) Wind blown debris	Delay in carrying out works Books will get wet Inability to communicate well Debris getting into eye, harming vision etc.	3	4	12	• Try and check weather forecast to help in planning site visits • Use waterproof clipboards to reduce the likelihood of writing book getting wet • In cold weather, regular breaks back in the vehicle should be taken in order to keep warm. Wear appropriate clothing and have some hot drinks/beverage to drink to keep warm. • In hot weather, appropriate clothing must also be worn and cold drinks must be taken regularly to avoid dehydration. Sun creams should be used as and when required • Ensure papers to be used for recording information is well secured on a clip board and not blown away by windy weather • Use appropriate personal protective

Tick	Hazard	Potential Harm	Likelihoo	Severity	Risk	Control measures and Actions
						equipment (PPE) • Avoid shouting from a distance to communicate • Step out of the way of any flying debris
						•

Risk assessment system:

Likelihood of occurrence		**Severity (effect on safety)**		**Risk Priority Number (RPN) = Occurrence x Severity**
No likelihood	0	No effect on safety	0	Single mode risk = RPN of
Unlikely	1	Hardly noticeable effects	1	25 or more
Possible	2 - 3	Insignificant effects	2 - 3	
Probable	4 - 6	Moderate effects	4 - 6	Multiple mode risk= 3 x RPN
Likely	7 - 8	Severe effects	7 - 8	of 20 or more
Almost inevitable	9 - 10	Very high severity	9 - 10	

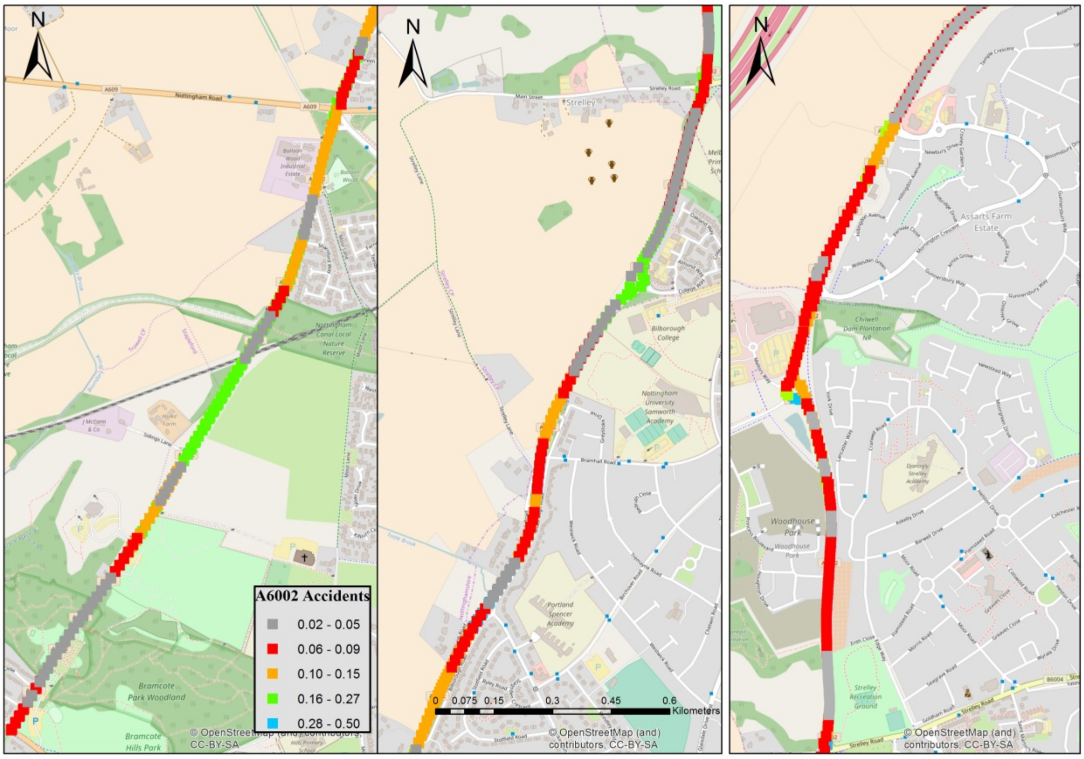
Appendix B: Matlab code used to run Genetic Algorithm for a typical road

NOT AVAILABLE

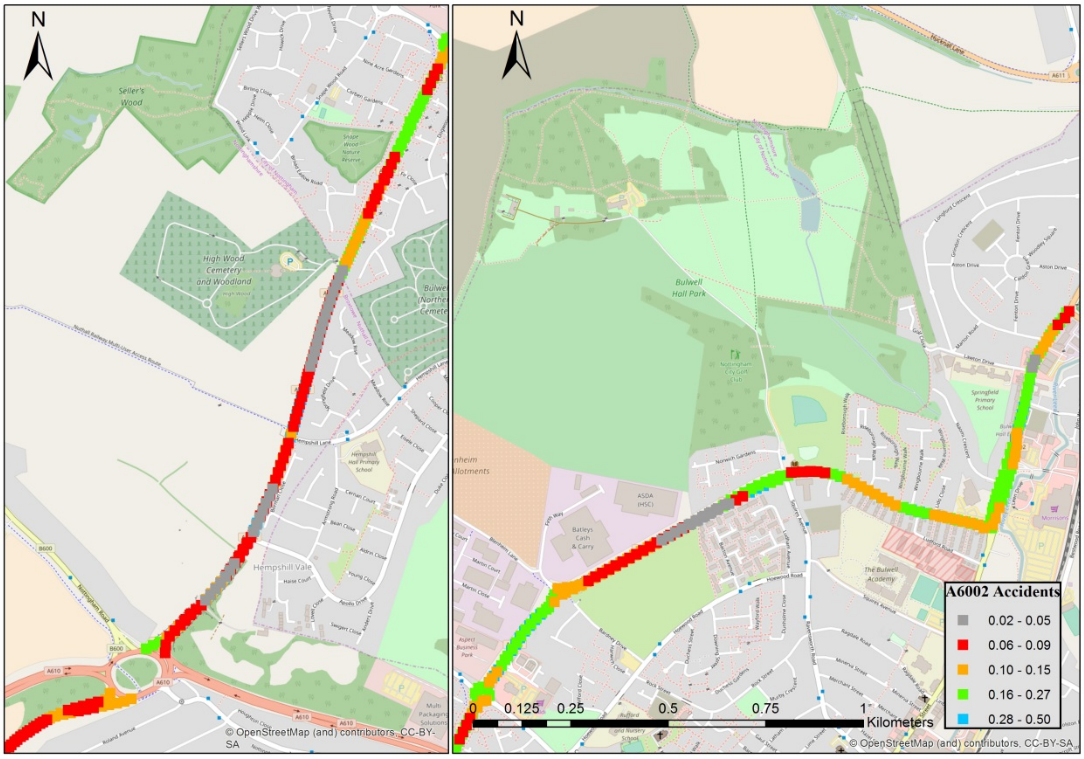
Appendix C: Matlab code used to run Pattern Search for a typical road

NOT AVAILABLE

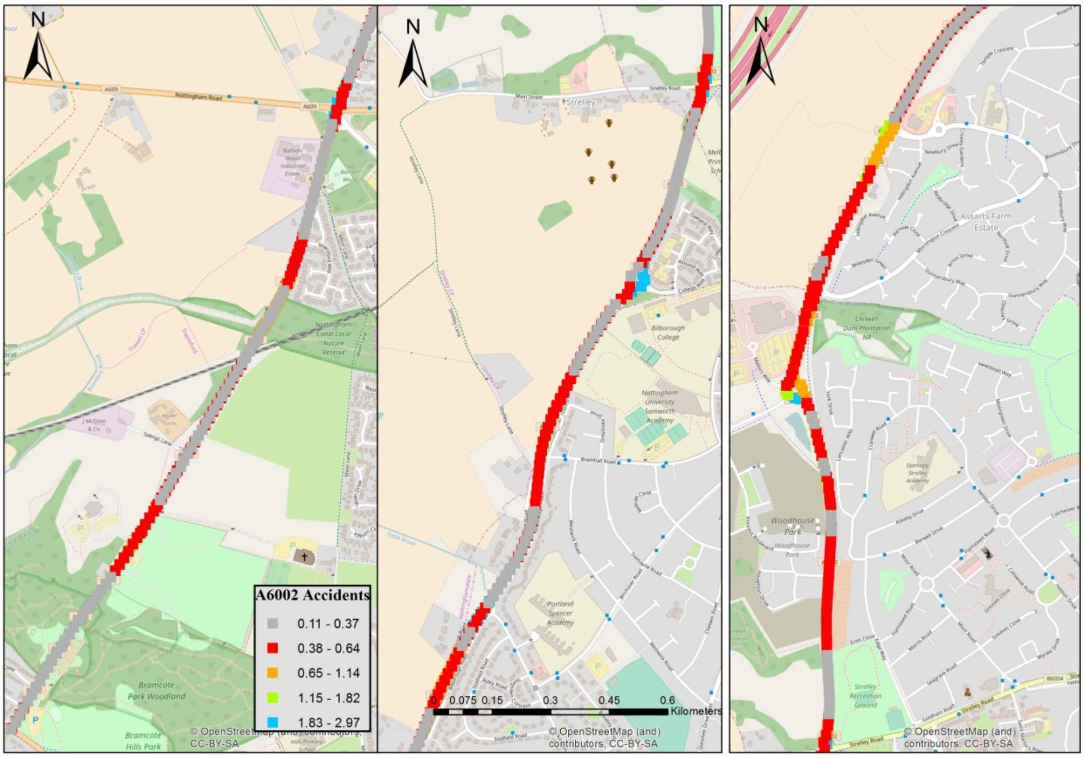
Appendix D: GIS plots of the variation of accidents along roads



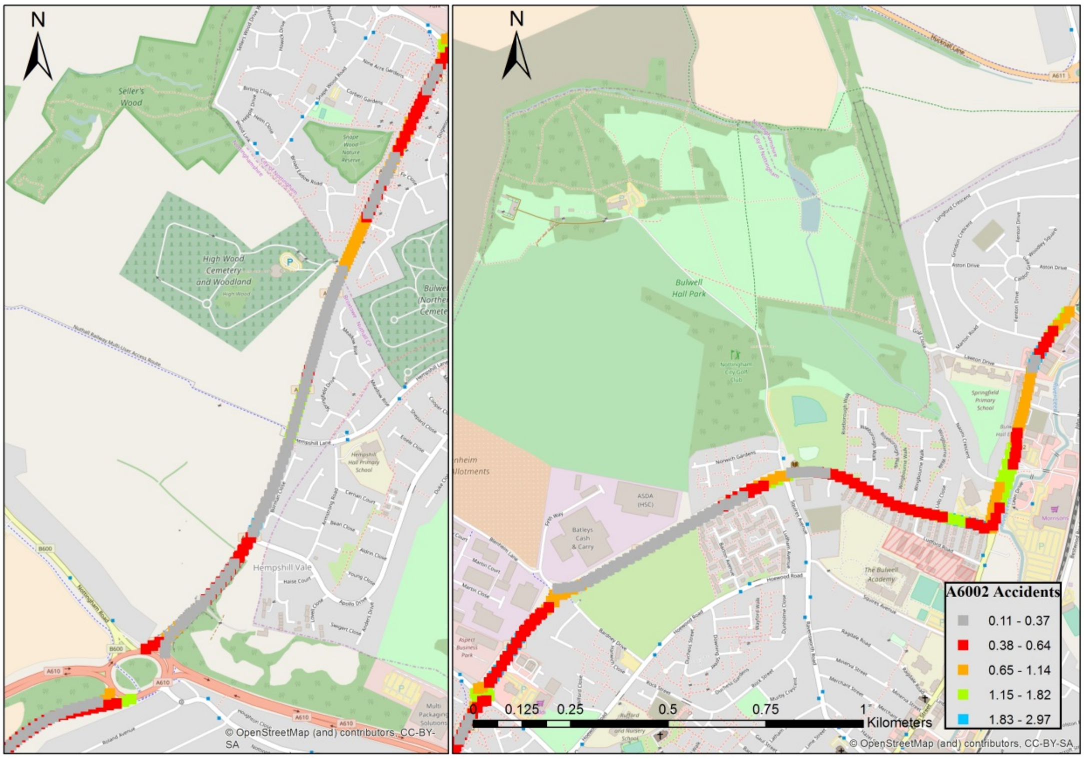
Appendix D1: A6002 Empirical Bayes Fatal and Serious Accidents (1 of 2)



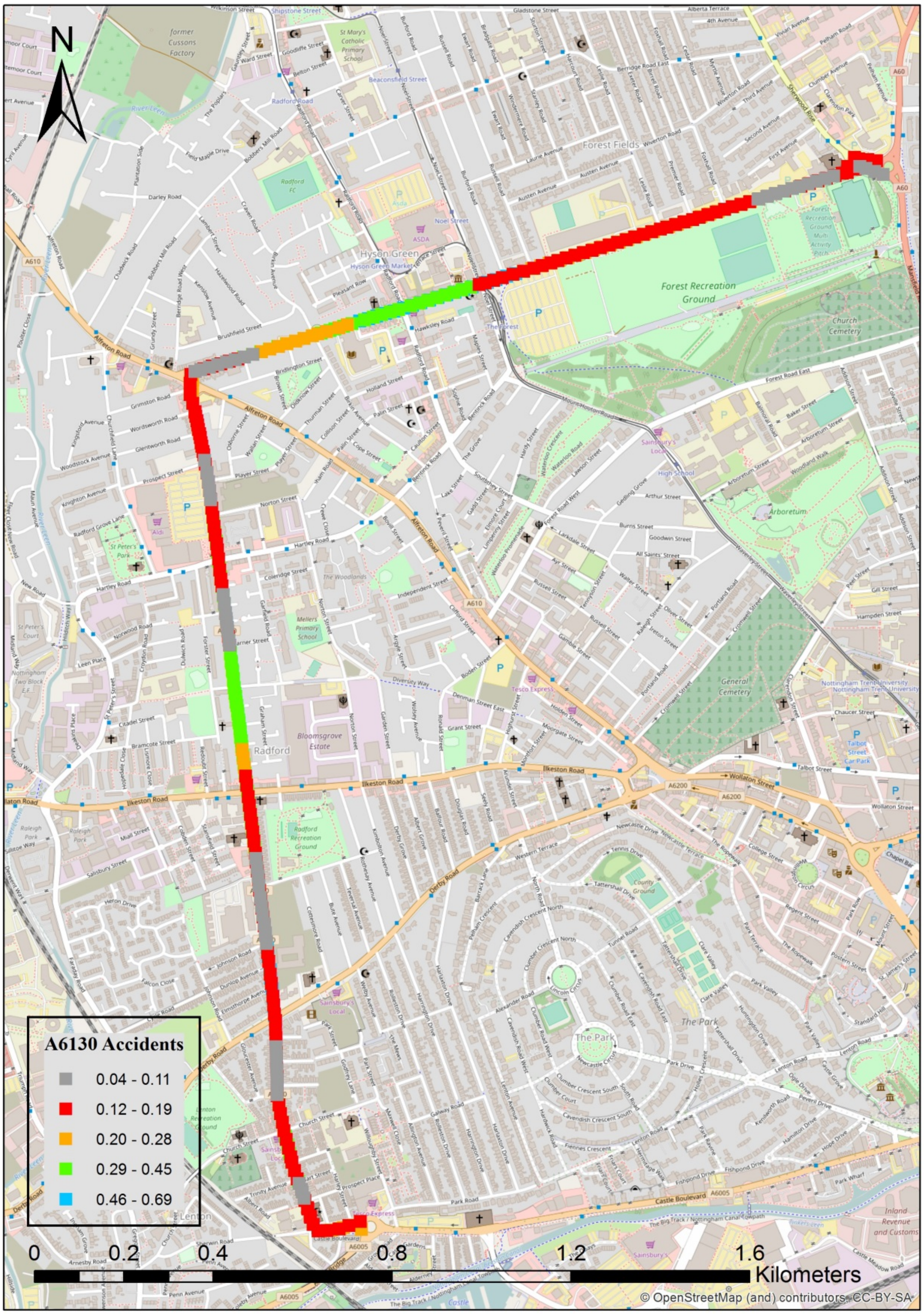
Appendix D2: A6002 Empirical Bayes Fatal and Serious Accidents (2 of 2)



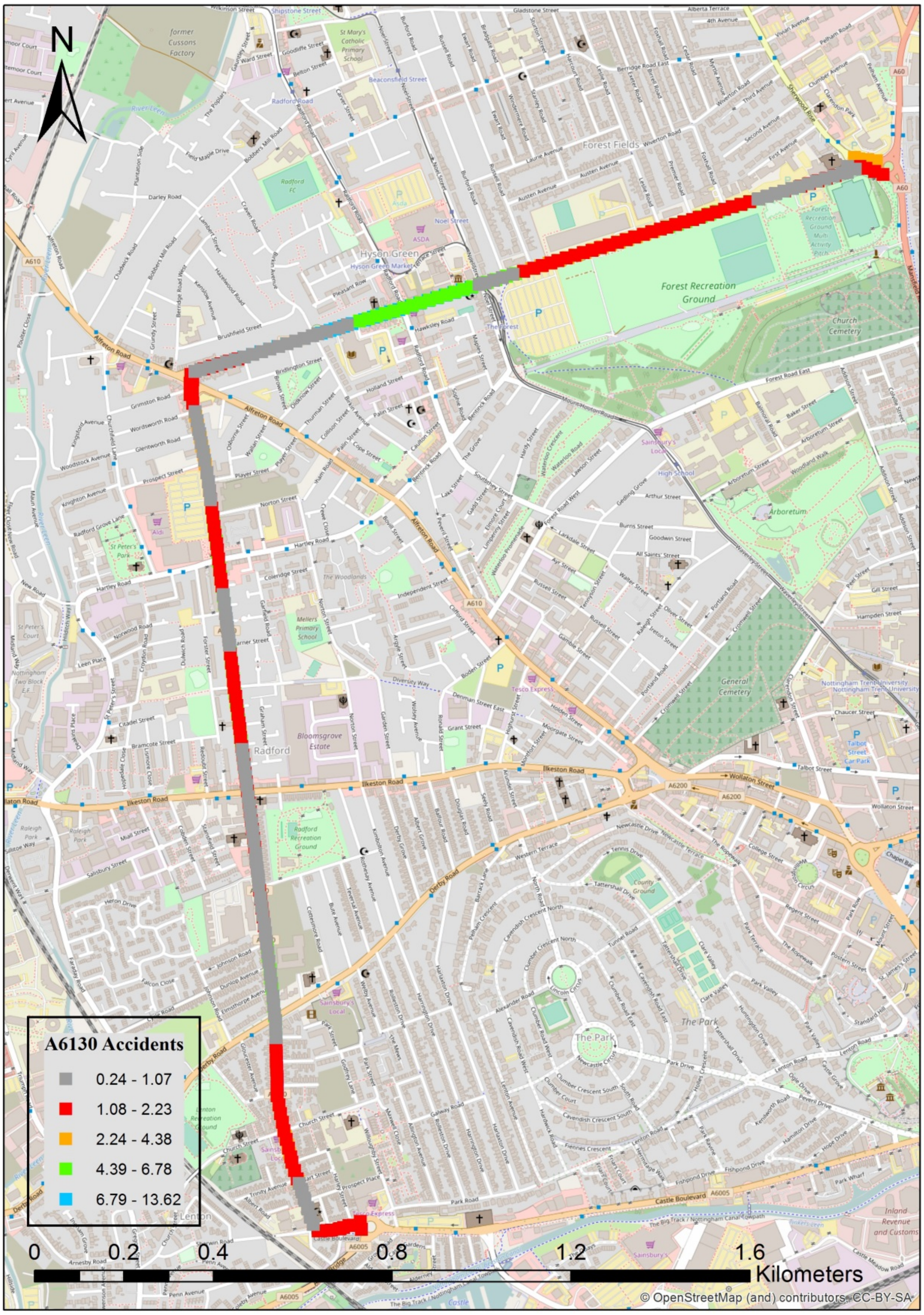
Appendix D1: A6002 Empirical Bayes Slight accidents (1 of 2)



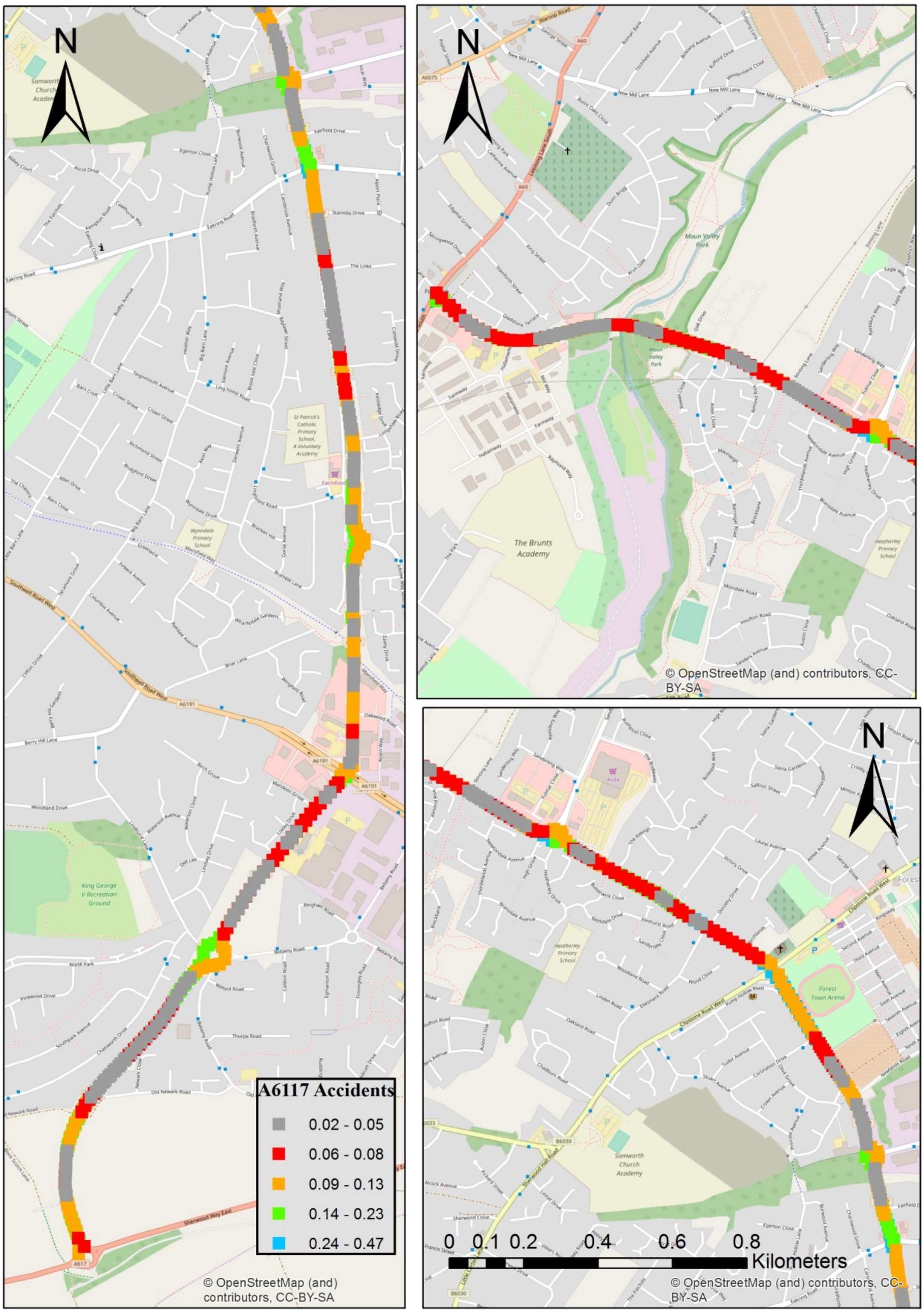
Appendix D1: A6002 Empirical Bayes Slight Accidents (2 of 2)



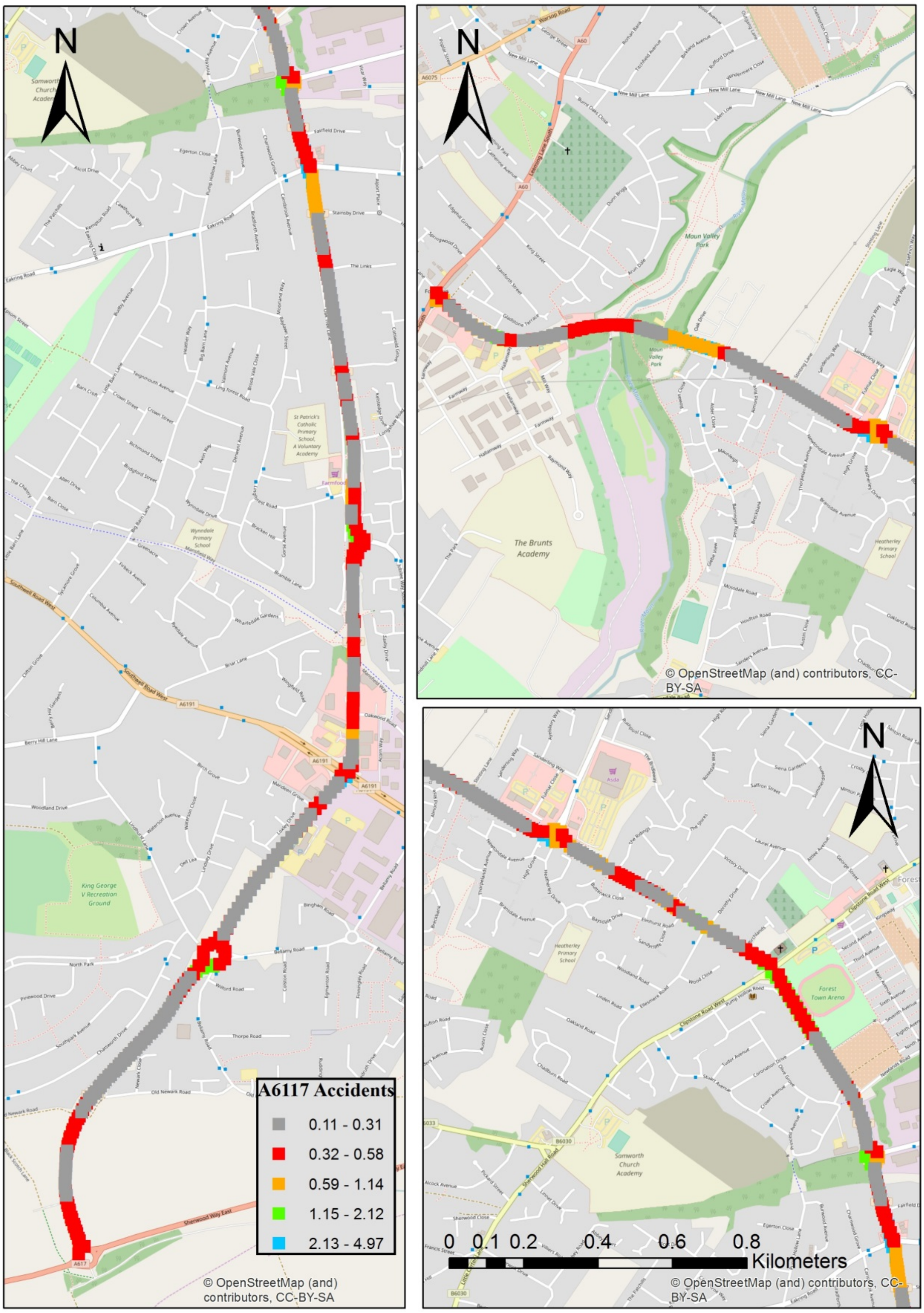
Appendix D2: A6130 Empirical Bayes Fatal and Serious Accidents



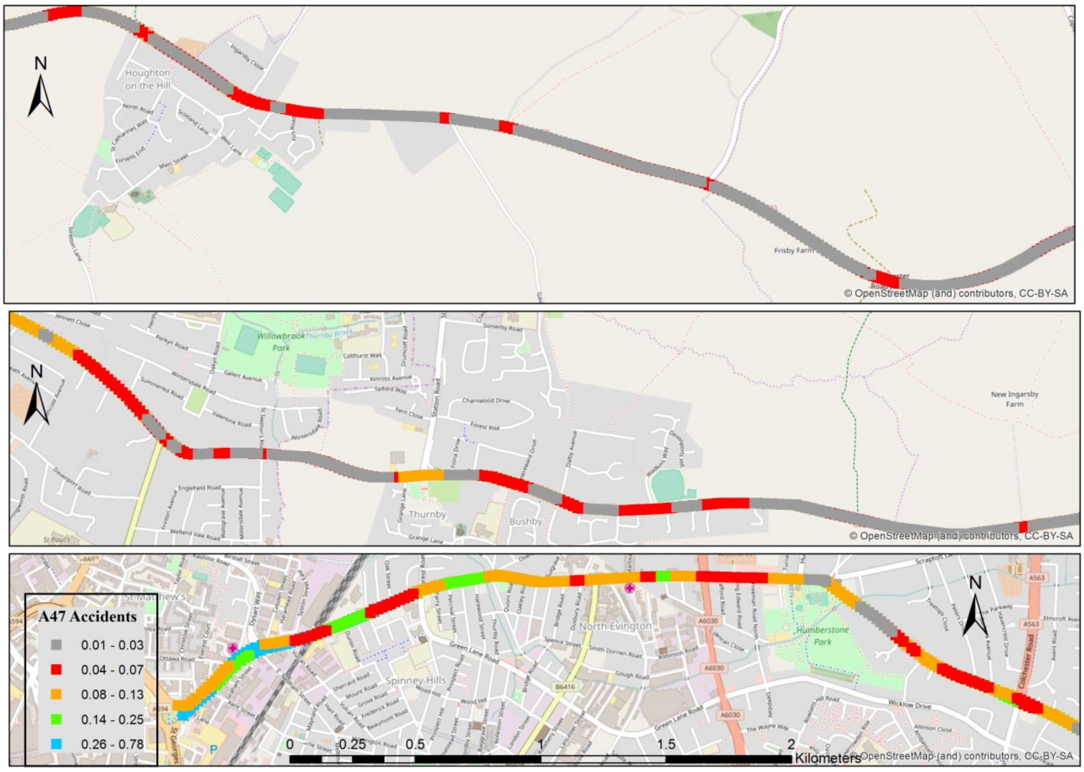
Appendix D2: A6130 Empirical Bayes Slight Accidents



Appendix D3: A6117 Empirical Bayes Fatal and Serious Accidents



Appendix D3: A6117 Empirical Bayes Slight Accidents

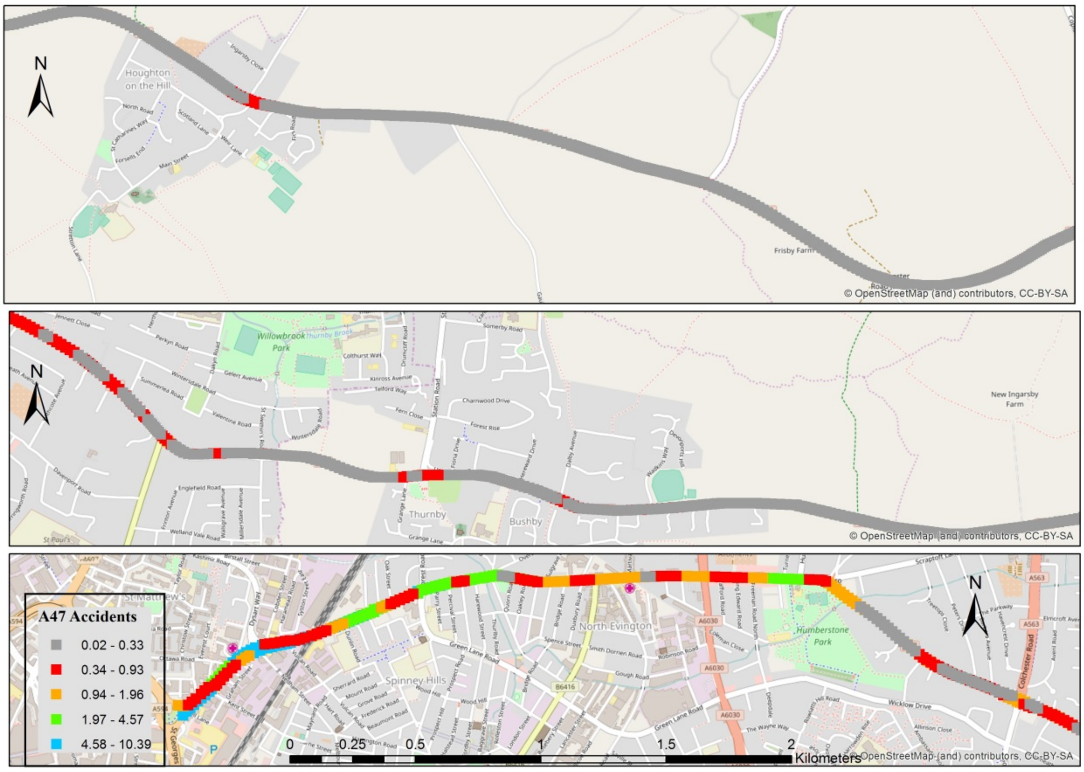


Appendix D4: A47 Empirical Bayes Fatal and Serious Accidents (1 of 4)

Appendix D4: A47 Empirical Bayes Fatal and Serious Accidents (2 of 4)

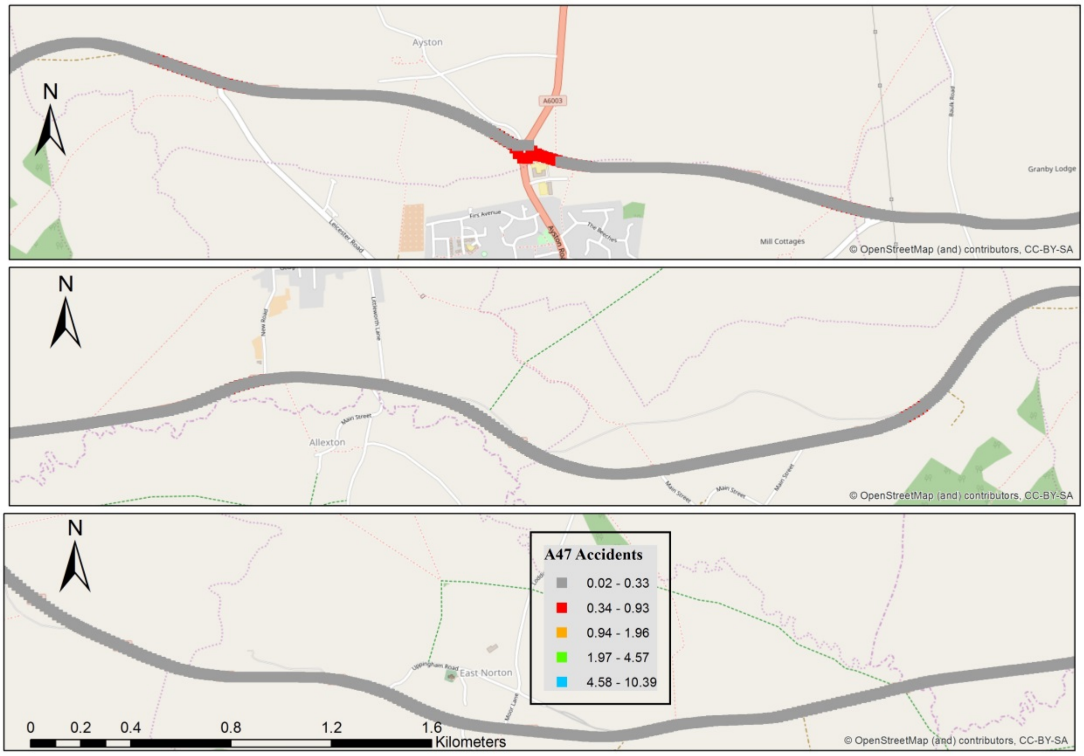
Appendix D4: A47 Empirical Bayes Fatal and Serious Accidents (3 of 4)

Appendix D4: A47 Empirical Bayes Fatal and Serious Accidents (4 of 4)



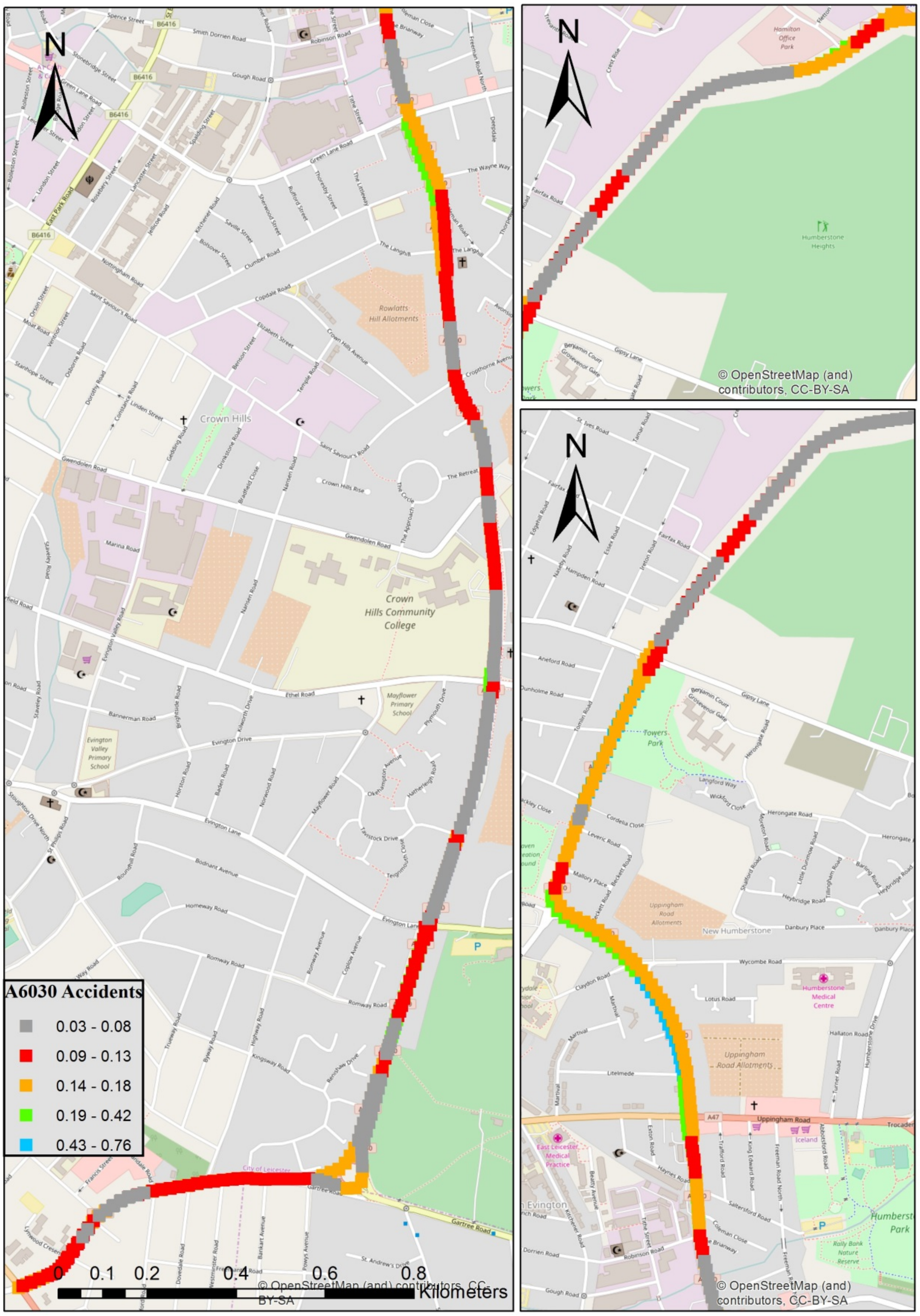
Appendix D4: A47 Empirical Bayes Slight Accidents (1 of 4)

Appendix D4: A47 Empirical Bayes Slight Accidents (2 of 4)

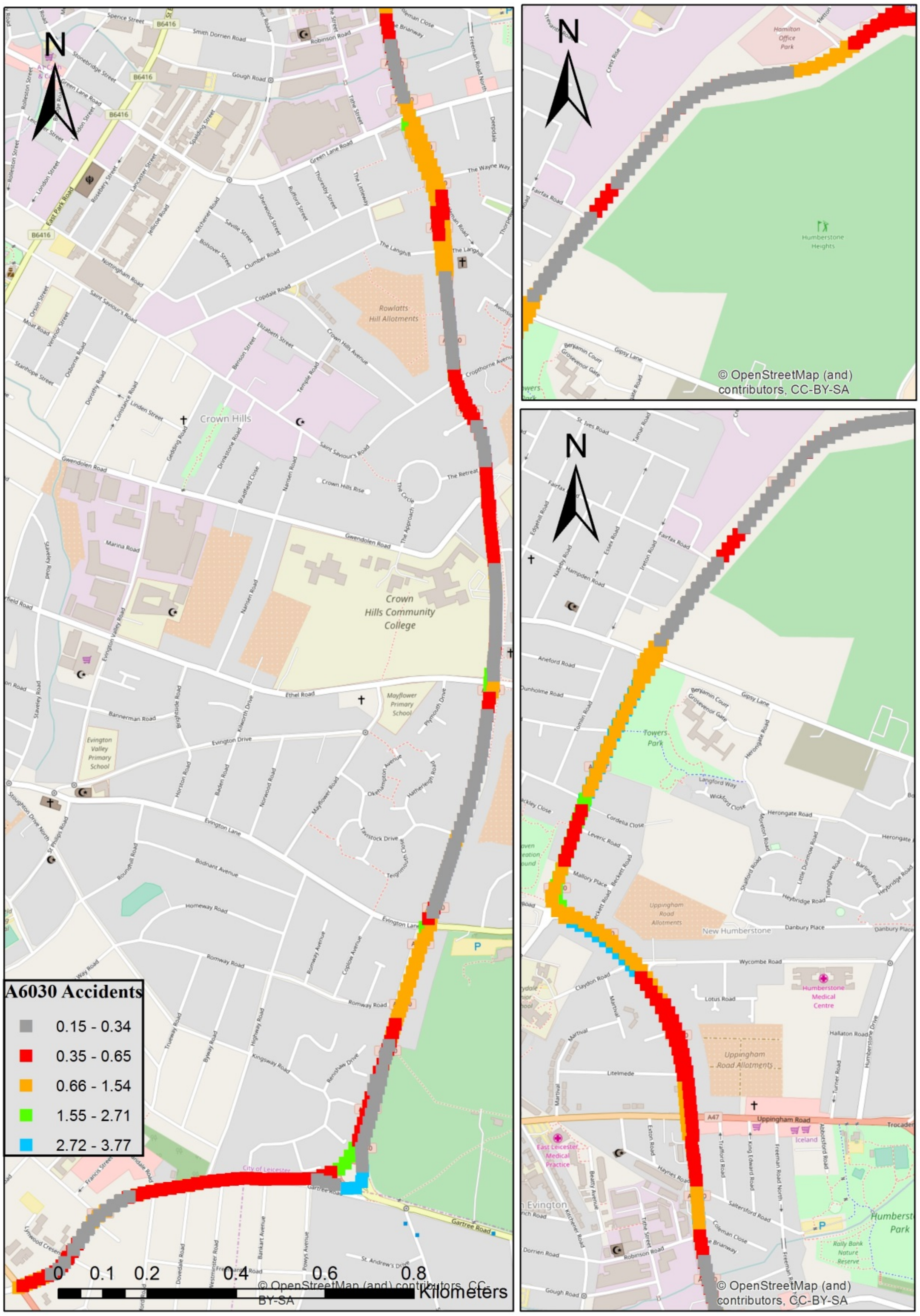


Appendix D4: A47 Empirical Bayes Slight Accidents (3 of 4)

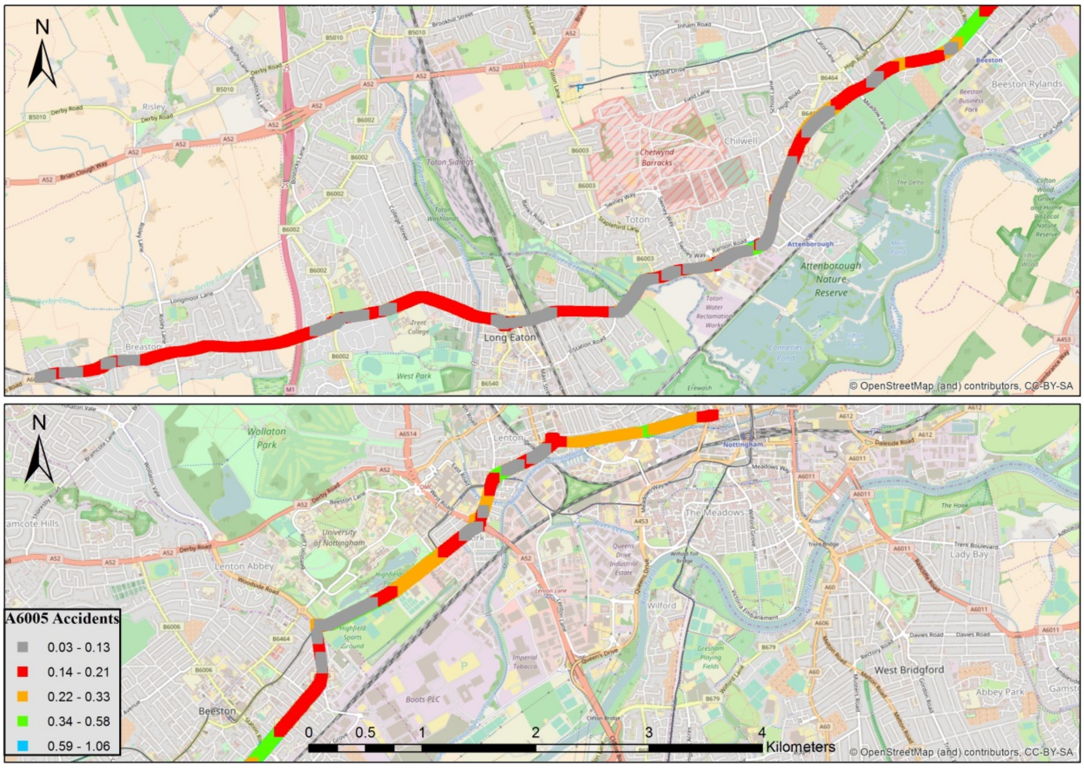
Appendix D4: A47 Empirical Bayes Slight Accidents (4 of 4)



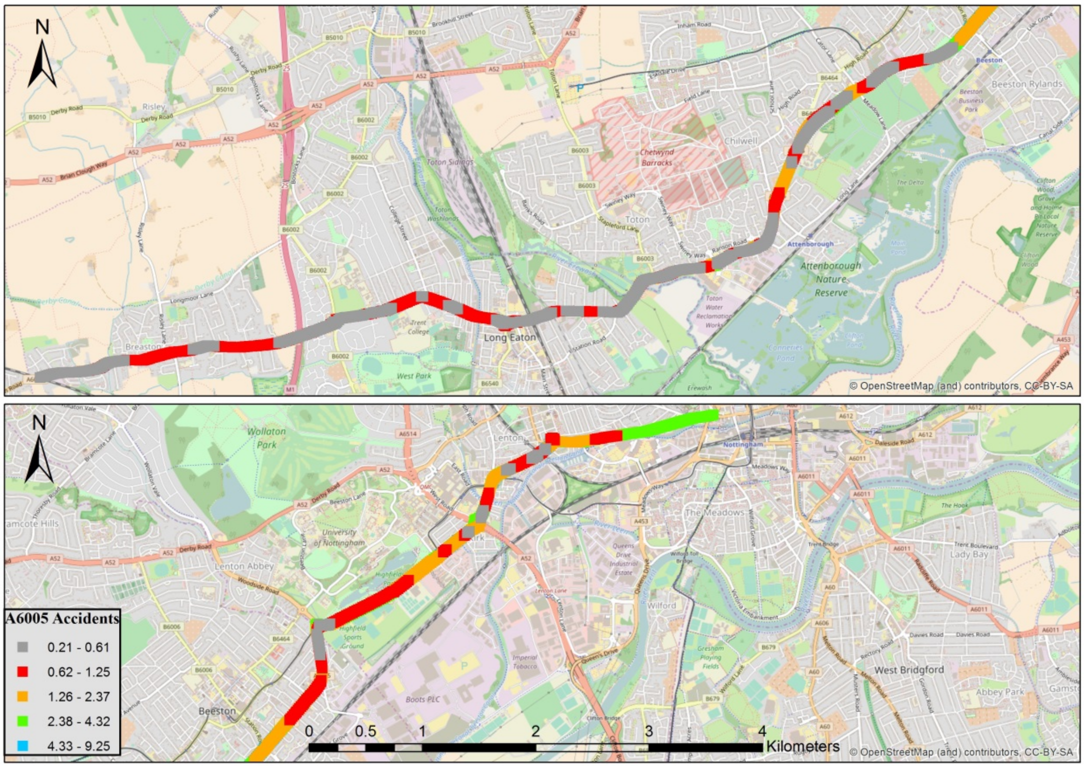
Appendix D5: A6030 Empirical Bayes Fatal and Serious Accidents



Appendix D5: A6030 Empirical Bayes Slight Accidents



Appendix D6: A6005 Empirical Bayes Fatal and Serious Accidents



Appendix D6: A6005 Empirical Bayes Slight Accidents

Appendix E: Easting, Northing and function value for optimised roads (Pattern Search)

easting (x)	northing(y)	model	road	fval	direction
449897.412	338418.382	PS	A6002	0.014884	NS
452683.41	344812.38	PS	A6002	0.555555	NS
453362.41	345889.38	PS	A6002	0.621422	NS
452958.41	345438.38	PS	A6002	0.676188	NS
451486.41	342788.88	PS	A6002	0.739838	NS
451539.41	343149.88	PS	A6002	0.817858	NS
452523.41	344330.38	PS	A6002	0.85473	NS
450433.41	339264.88	PS	A6002	0.856178	NS
451551.41	342475.38	PS	A6002	0.887755	NS
452843.41	345199.88	PS	A6002	0.951979	NS
454160.41	346022.88	PS	A6002	0.969422	NS
451551.41	342435.38	PS	A6002	1.040752	NS
453327.41	345869.88	PS	A6002	1.130645	NS
453687.41	346066.88	PS	A6002	1.167222	NS
450931.41	340472.38	PS	A6002	1.173683	NS
450677.41	339812.88	PS	A6002	1.183185	NS
454271.41	345998.38	PS	A6002	1.217347	NS
450162.41	338822.38	PS	A6002	1.248898	NS
450878.41	340371.88	PS	A6002	1.290745	NS
453397.41	345908.88	PS	A6002	1.299508	NS

easting (x)	northing(y)	model	road	fval	direction
460761.9324	302209.1	PS	A6030	0.048746	NS
461589.9	305000.6	PS	A6030	0.079797	NS
461575.9	305124.1	PS	A6030	0.089715	NS
461281.9	305920.1	PS	A6030	0.369902	NS
461712.9	304553.1	PS	A6030	0.464588	NS
461397.9	305708.1	PS	A6030	0.48055	NS
461482.9	305605.6	PS	A6030	0.976289	NS
461672.9	304750.6	PS	A6030	1.097991	NS
461713.9	304458.1	PS	A6030	1.455109	NS
461546.9	305446.6	PS	A6030	1.481592	NS
461566.9	302672.6	PS	A6030	1.618747	NS
461525.9	305523.6	PS	A6030	1.633629	NS
461364.9	302449.6	PS	A6030	1.654274	NS
461524.9	302498.6	PS	A6030	1.688849	NS
461378.9	306168.6	PS	A6030	1.836085	NS
461534.9	302557.1	PS	A6030	1.86623	NS
461462.9	302436.6	PS	A6030	1.896127	NS
461245.9	302450.1	PS	A6030	1.953092	NS
461800.9	304085.6	PS	A6030	1.970741	NS
461786.9	303374.1	PS	A6030	1.976654	NS

easting (x)	northing(y)	model	road	fval	direction
456164.9	339172.9	PS	A6005	0.117955	EW
450239.9	334071.9	PS	A6005	0.271531	EW
455590.9	339006.9	PS	A6005	0.411269	EW
456656.9	339258.9	PS	A6005	0.48402	EW
450860.9	334322.4	PS	A6005	0.597261	EW
453283.9	336518.9	PS	A6005	0.778418	EW
452184.9	335873.9	PS	A6005	0.813155	EW
454664.9	338064.9	PS	A6005	0.843865	EW
454495.9	337932.9	PS	A6005	0.851369	EW
449449.9	333896.9	PS	A6005	0.871803	EW
446104.9	333557.9	PS	A6005	0.898945	EW
454355.9	337823.9	PS	A6005	0.914115	EW
446143.9	333566.4	PS	A6005	0.988712	EW
445002.9541	333349.9	PS	A6005	1.008444	EW
451923.9	335669.4	PS	A6005	1.030584	EW
445434.9	333411.9	PS	A6005	1.038183	EW
449430.9	333890.9	PS	A6005	1.067121	EW
448282.9	334036.9	PS	A6005	1.086266	EW
451998.9	335723.9	PS	A6005	1.095317	EW
455639.9	339040.9	PS	A6005	1.183525	EW

easting (x)	northing(y)	model	road	fval	direction
454517	346537	PS	A6002	1.719498	SN
452962	345444.5	PS	A6002	1.720109	SN
451469	342918	PS	A6002	1.720358	SN
454432	346406.5	PS	A6002	1.720483	SN
452916	345363	PS	A6002	1.721843	SN
451472	341667.5	PS	A6002	1.722676	SN
451830	343691	PS	A6002	1.723643	SN
451659	343387	PS	A6002	1.728799	SN
453322	345867	PS	A6002	1.731853	SN
450477	339338.5	PS	A6002	1.733089	SN
450865	340345	PS	A6002	1.736561	SN
452970	345462.5	PS	A6002	1.738774	SN
453127	345720.5	PS	A6002	1.740382	SN
454008	346082.5	PS	A6002	1.741116	SN
451539	341920.5	PS	A6002	1.742731	SN
452615	344573	PS	A6002	1.753084	SN
452599	344520.5	PS	A6002	1.753418	SN
453620	346032.5	PS	A6002	1.75644	SN
452769	345036.5	PS	A6002	1.756638	SN
453287	345847.5	PS	A6002	1.758375	SN

easting (x)	northing(y)	model	road	fval	direction
456113	341301	PS	A6130	0	SN
456400	341389	PS	A6130	0	SN
456472	341411	PS	A6130	0	SN
456544	341434	PS	A6130	0	SN
456695	341480	PS	A6130	0	SN
456883.7	341542.5	PS	A6130	2.756458	SN
455995	341265	PS	A6130	1.26814	SN
456491	341417	PS	A6130	1.726304	SN
456289	341355	PS	A6130	2.300383	SN
456132	341307	PS	A6130	2.351602	SN
455842	341218	PS	A6130	2.606424	SN
456308	341361	PS	A6130	2.858565	SN
456070	341288.5	PS	A6130	3.056724	SN
455815	341209.5	PS	A6130	3.08855	SN
456174	341320	PS	A6130	3.29662	SN
456510	341423	PS	A6130	3.452593	SN
456419	341395	PS	A6130	3.461291	SN
455976	341259	PS	A6130	3.61266	SN
456659	341469	PS	A6130	3.627623	SN
456678	341475	PS	A6130	3.722367	SN

easting (x)	northing(y)	model	road	fval	direction
455742	362602.0949	PS	A6117	2771.213	SN
455742	362602.0949	PS	A6117	2812.49	SN
455742	362607.8228	PS	A6117	2831.823	SN
455742	362613.5506	PS	A6117	2892.433	SN
455742	362596.3671	PS	A6117	3037.052	SN
455742	362613.5506	PS	A6117	3093.442	SN
455742	362617.605	PS	A6117	3351.648	SN
455742	362612.9319	PS	A6117	3413.156	SN
455742	362608.2587	PS	A6117	3474.665	SN
455742	362603.5855	PS	A6117	3536.174	SN
455742	362598.9124	PS	A6117	3597.683	SN
455742	362594.2392	PS	A6117	3659.192	SN
455742	362612.9318	PS	A6117	4165.955	SN
455742	362634.0914	PS	A6117	4397.715	SN
455742	362647.7702	PS	A6117	4443.865	SN
455742	362661.449	PS	A6117	4490.016	SN
455742	362675.1279	PS	A6117	4536.166	SN
455742	362688.7	PS	A6117	4582.317	SN
455742	362634.0914	PS	A6117	5007.308	SN
455742	362688.6696	PS	A6117	7407.44	SN

easting (x)	northing(y)	model	road	fval	direction
455614.7297	362709	PS	A6117	1458.112	NS
455632.7643	362709	PS	A6117	1487.045	NS
455650.5287	362709	PS	A6117	1515.5	NS
455668.2931	362709	PS	A6117	1543.955	NS
455686.0574	362709	PS	A6117	1572.41	NS
455703.8218	362709	PS	A6117	1600.864	NS
455721.5861	362709	PS	A6117	1629.319	NS
455738.6581	362709	PS	A6117	1661.286	NS
455632.7644	362709	PS	A6117	1663.109	NS
455754.9494	362709	PS	A6117	1697.214	NS
455705.5968	362709	PS	A6117	12973.46	NS
455705.9287	362709	PS	A6117	13035.17	NS
455708.472	362709	PS	A6117	13096.6	NS
455711.8871	362709	PS	A6117	13157.53	NS
455716.6141	362709	PS	A6117	13217.71	NS
455722.2092	362709	PS	A6117	13277.11	NS
455728.6892	362709	PS	A6117	13335.7	NS
455735.1689	362709	PS	A6117	13394.3	NS
455741.6489	362709	PS	A6117	13452.89	NS
455757.4	362709	PS	A6117	15120.84	NS

easting (x)	northing(y)	model	road	fval	direction
464209	304196.3	PS	A47	2.663251	WE
469418	303487.8	PS	A47	2.810815	WE
475284	302127.8	PS	A47	2.849202	WE
475237	302164.8	PS	A47	2.876509	WE
474630	302673.8	PS	A47	2.895373	WE
469899	303351.8	PS	A47	2.906668	WE
475770	301658.8	PS	A47	2.938712	WE
465510	304026.3	PS	A47	2.945825	WE
465160	304037.3	PS	A47	3.013989	WE
465490	304025.3	PS	A47	3.049839	WE
472156	303154.3	PS	A47	3.054322	WE
471090	302994.8	PS	A47	3.13508	WE
465454	304023.3	PS	A47	3.139494	WE
466498	303943.3	PS	A47	3.144187	WE
473802	302872.8	PS	A47	3.160747	WE
472351	303110.8	PS	A47	3.179533	WE
474933	302355.8	PS	A47	3.261387	WE
466001	304042.8	PS	A47	3.266192	WE
468319	303630.8	PS	A47	3.286827	WE
475958	301509.8	PS	A47	3.291955	WE

easting (x)	northing(y)	model	road	fval	direction
462313	307044.2	PS	A6030	3.656524	SN
461111	302431.5	PS	A6030	3.680969	SN
461812	303531.5	PS	A6030	3.697823	SN
461552	305306	PS	A6030	3.757051	SN
461531	305477.5	PS	A6030	3.757434	SN
460789	302220.5	PS	A6030	3.768962	SN
461544	305404	PS	A6030	3.77131	SN
461620	302859	PS	A6030	3.772267	SN
461294	305765	PS	A6030	3.790873	SN
461665	304730	PS	A6030	3.791648	SN
461431	306277	PS	A6030	3.829171	SN
461737	303211	PS	A6030	3.879667	SN
461693	306667.5	PS	A6030	3.889836	SN
461633	302897	PS	A6030	3.89432	SN
460907	302336	PS	A6030	3.941818	SN
461289	305941	PS	A6030	3.94896	SN
461539	305438.5	PS	A6030	4.006766	SN
461658	302973	PS	A6030	4.01724	SN
461683	304683.5	PS	A6030	4.021997	SN
461822	303834	PS	A6030	4.042093	SN

easting (x)	northing(y)	model	road	fval	direction
457150	339378	PS	A6005	2.128415	WE
453721	337115	PS	A6005	2.129198	WE
454336	337822.5	PS	A6005	2.129573	WE
456772	339291	PS	A6005	2.131714	WE
454819	338195.5	PS	A6005	2.133537	WE
450623	334292.5	PS	A6005	2.134501	WE
453932	337608.5	PS	A6005	2.135092	WE
452100	335813.5	PS	A6005	2.135343	WE
447727	333884	PS	A6005	2.13595	WE
454578	338009.5	PS	A6005	2.136686	WE
447413	333754.5	PS	A6005	2.136703	WE
451740	335501	PS	A6005	2.137724	WE
451978	335725.5	PS	A6005	2.13791	WE
452745	336146.5	PS	A6005	2.13813	WE
446343	333595	PS	A6005	2.138836	WE
451442	334618.5	PS	A6005	2.140178	WE
454057	337671.5	PS	A6005	2.14216	WE
452012	335749	PS	A6005	2.143215	WE
456966	339346	PS	A6005	2.144631	WE
450474	334236	PS	A6005	2.14739	WE

Appendix F: Easting, Northing and function value for optimised roads (Genetic Algorithm)

easting (x)	northing(y)	model	road	fval	direction
454499.5553	346373.5154	GA	A6002	615.6459	SN
454499.3053	346373.0154	GA	A6002	617.604	SN
454500.5553	346371.0154	GA	A6002	624.5469	SN
454499.3053	346373.0154	GA	A6002	617.604	SN
454499.5553	346373.5154	GA	A6002	615.6459	SN
454499.0553	346370.9529	GA	A6002	625.3653	SN
454499.5553	346371.4529	GA	A6002	623.307	SN
454499.3053	346373.0154	GA	A6002	617.604	SN
454500.5553	346372.0154	GA	A6002	620.8299	SN
454499.5553	346372.0154	GA	A6002	621.2175	SN
454499.3053	346373.0154	GA	A6002	617.604	SN
454499.3053	346373.0154	GA	A6002	617.604	SN
454499.3053	346373.0154	GA	A6002	617.604	SN
454498.5553	346371.4529	GA	A6002	623.7155	SN
454499.5553	346373.0154	GA	A6002	617.5031	SN
454500.0553	346372.4529	GA	A6002	619.3953	SN
454499.0553	346371.4529	GA	A6002	623.5085	SN
454499.3053	346371.0154	GA	A6002	625.0321	SN
454500.3053	346374.0154	GA	A6002	613.4925	SN
454499.3053	346373.0154	GA	A6002	617.604	SN

easting (x)	northing(y)	model	road	fval	direction
456718.97	341128.7895	GA	A6130	26614.06	SN
456718.97	341128.7895	GA	A6130	26614.06	SN
456718.97	341128.7895	GA	A6130	26614.06	SN
456718.095	341127.727	GA	A6130	26692.41	SN
456719.345	341128.352	GA	A6130	26630.07	SN
456717.845	341128.0551	GA	A6130	26679.75	SN
456718.97	341128.7895	GA	A6130	26614.06	SN
456717.845	341128.0551	GA	A6130	26679.75	SN
456718.095	341128.7895	GA	A6130	26633.45	SN
456718.97	341128.7895	GA	A6130	26614.06	SN
456719.345	341128.7895	GA	A6130	26605.77	SN
456718.97	341128.7895	GA	A6130	26614.06	SN
456718.97	341128.7895	GA	A6130	26614.06	SN
456718.095	341128.352	GA	A6130	26657.73	SN
456718.345	341128.352	GA	A6130	26652.18	SN
456719.97	341126.852	GA	A6130	26699.7	SN
456718.97	341129.2895	GA	A6130	26586.3	SN
456720.47	341128.352	GA	A6130	26605.34	SN
456719.47	341128.352	GA	A6130	26627.32	SN
456719.22	341129.2895	GA	A6130	26580.77	SN

easting (x)	northing(y)	model	road	fval	direction
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455746.0609	358270.8065	GA	A6117	253.6931	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455750.1736	358343.2596	GA	A6117	25.09363	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358374.7042	GA	A6117	109.2205	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455765.5776	358326.3637	GA	A6117	64.70292	NS
455763.598	358353.4554	GA	A6117	41.56041	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455758.6744	358347.1652	GA	A6117	14.50744	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS
455757.2985	358343.2596	GA	A6117	0.519337	NS

easting (x)	northing(y)	model	road	fval	direction
499982.8064	301480.7554	GA	A47	5.85895	WE
499982.8064	301480.7554	GA	A47	5.85895	WE
499981.8114	301479.8992	GA	A47	75.00491	WE
499935.6724	301501.2781	GA	A47	2940.949	WE
499804.7587	301419.402	GA	A47	10790.92	WE
499973.5425	301489.0678	GA	A47	707.5133	WE
500054.3417	301417.2718	GA	A47	5487.494	WE
499983.601	301465.3722	GA	A47	887.1941	WE
500016.9746	301487.7437	GA	A47	2002.105	WE
500046.5508	301458.7594	GA	A47	3870.114	WE
499978.5631	301478.1644	GA	A47	283.6171	WE
499982.3261	301480.3017	GA	A47	38.13428	WE
499982.8064	301480.7554	GA	A47	5.85895	WE
499983.1395	301472.3101	GA	A47	488.7135	WE
499962.0997	301392.3591	GA	A47	5206.566	WE
499916.9307	301367.3347	GA	A47	7518.991	WE
499982.8064	301480.7554	GA	A47	5.85895	WE
499982.8064	301480.7554	GA	A47	5.85895	WE
499982.5231	301483.3758	GA	A47	146.5575	WE
499982.8064	301480.7554	GA	A47	5.85895	WE

easting (x)	northing(y)	model	road	fval	direction
462268.5802	306839.0645	GA	A6030	7219.869	SN
462268.5802	306839.0645	GA	A6030	7219.869	SN
462268.0802	306839.0645	GA	A6030	7223.53	SN
462269.5802	306838.0645	GA	A6030	7246.311	SN
462268.5802	306839.0645	GA	A6030	7219.869	SN
462268.5802	306838.0645	GA	A6030	7253.483	SN
462268.0802	306838.5645	GA	A6030	7240.328	SN
462268.5802	306838.0645	GA	A6030	7253.483	SN
462268.5802	306838.0645	GA	A6030	7253.483	SN
462269.8302	306837.5645	GA	A6030	7261.373	SN
462269.5802	306838.0645	GA	A6030	7246.311	SN
462268.5802	306839.0645	GA	A6030	7219.869	SN
462268.5802	306838.0645	GA	A6030	7253.483	SN
462268.5802	306838.5645	GA	A6030	7236.675	SN
462268.5802	306837.0645	GA	A6030	7287.104	SN
462269.5802	306838.0645	GA	A6030	7246.311	SN
462268.0802	306839.0645	GA	A6030	7223.53	SN
462268.5802	306837.5645	GA	A6030	7270.292	SN
462269.8302	306838.0645	GA	A6030	7244.543	SN
462269.0802	306839.0645	GA	A6030	7216.247	SN

easting (x)	northing(y)	model	road	fval	direction
457016.97	339188.9094	GA	A6005	1393.456	WE
457016.74	339188.3469	GA	A6005	1397.036	WE
457018.04	339188.3469	GA	A6005	1392.564	WE
457016.97	339188.9094	GA	A6005	1393.456	WE
457018.04	339187.3469	GA	A6005	1397.508	WE
457016.99	339187.9094	GA	A6005	1398.328	WE
457016.74	339187.3469	GA	A6005	1401.965	WE
457016.97	339188.3469	GA	A6005	1396.226	WE
457016.99	339187.4719	GA	A6005	1400.486	WE
457016.99	339187.9094	GA	A6005	1398.328	WE
457016.41	339186.2219	GA	A6005	1408.647	WE
457016.99	339188.3469	GA	A6005	1396.172	WE
457016.97	339188.9094	GA	A6005	1393.456	WE
457015.47	339187.2219	GA	A6005	1406.956	WE
457016.99	339187.4094	GA	A6005	1400.795	WE
457016.74	339186.1594	GA	A6005	1407.828	WE
457017.74	339187.3469	GA	A6005	1398.526	WE
457015.97	339187.2844	GA	A6005	1404.917	WE
457019.04	339188.3469	GA	A6005	1389.136	WE
457017.74	339187.3469	GA	A6005	1398.526	WE