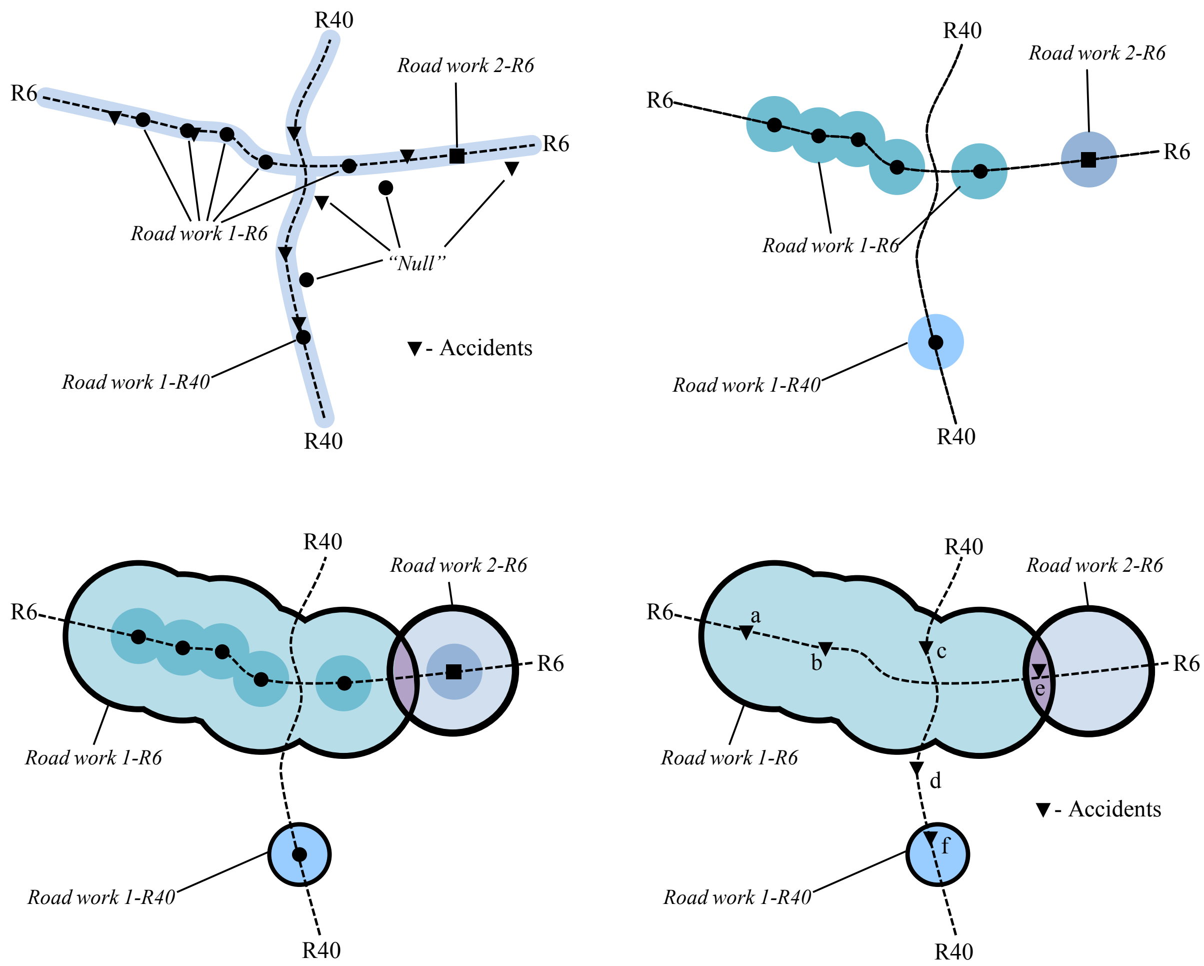
CHALMERS
UNIVERSITY OF TECHNOLOGY



Traffic Safety at Road Works

Usage of GIS as a tool to locate and quantify accidents at road works

Master's Thesis in the Master's Programme Infrastructure and Environmental Engineering

JOACHIM JENSEN
FREDRIK SVENSSON

Department of Civil and Environmental Engineering
Division of GeoEngineering
Road and traffic
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2015

Master's Thesis 2015:96

MASTER'S THESIS 2015:96

Traffic Safety at Road Works

Usage of GIS as a tool to locate and quantify accidents at road works

Master's Thesis in the Master's Programme Infrastructure and Environmental Engineering

JOACHIM JENSEN

FREDRIK SVENSSON

Department of Civil and Environmental Engineering

Division of GeoEngineering

Road and Traffic

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2015

Traffic Safety at Road Works
Usage of GIS as a tool to locate and quantify accidents at road works

Master's Thesis in the Master's Programme Infrastructure and Environmental Engineering

JOACHIM JENSEN

FREDRIK SVENSSON

Examensarbete 2015:96 / Institutionen för bygg- och miljöteknik,
Chalmers tekniska högskola 2015

Department of Civil and Environmental Engineering
Division of GeoEngineering
Road and traffic

Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: + 46 (0)31-772 1000

Cover:
The different steps done in ArcGIS in order to connect accident data together with road works. The whole process is described in detail in Chapter 7.

Reproservice. Göteborg, Sweden, 2015

Traffic Safety at Road Works

Usage of GIS as a tool to locate and quantify accidents at road works

Master's thesis in the Master's Programme Infrastructure and Environmental Engineering

JOACHIM JENSEN
FREDRIK SVENSSON
Department of Civil and Environmental Engineering
Division of GeoEngineering
Road and Traffic
Chalmers University of Technology

ABSTRACT

In 2008 and 2012, the Swedish Transport Administration made two studies whose purpose was to investigate road work related accidents. This was achieved by performing a free-text search in a traffic related accident database named STRADA, which both the police and health care reports into. Results from the studies showed that less than 1 % of all traffic related accidents is caused due to road works. However, it is believed that there exists a hidden figure and that the percentage should be higher, as the accident descriptions in STRADA is believed to be insufficient.

As the accidents and road works both are registered spatially and in time, a usage of geographical information systems (GIS) could be an approach in order to quantify this hidden figure, which is what this master's thesis aims to achieve. Therefore, by using a software named ArcGIS, a model was developed where accidents are linked to road work data on governmental roads. A filtration was made for those accidents which had occurred during a road work's active period as well as within its influence area. The results were then compared to the normal amount of accidents that occur on the same road segments, but when no road work was active on-site. Additionally, a free-text search was performed for the years 2009-2014, comparable to the ones in the earlier studies. This was done in order to validate the model later on.

It was found for 2009–2014, that 1.74 % more accidents occurs on road segments while a road work is active on-site, compared to normal conditions. Furthermore, about 8 % of all accidents that occurred on governmental roads were potentially caused by road works. Among the injured, a large proportion consists of young adults in the ages 25-34 for all severity degrees.

However, the results are inconclusive as it was seen that only 38 % of the confirmed road work related accidents were identified within the GIS model. This is believed to be due to limitations of the model, as well as uncertainties regarding the input data. Because of this, further investigations regarding road work related accidents are recommended, especially concerning the registration approach for STRADA as well as limiting the area that are being analysed.

Key words: *road works, accidents, traffic, GIS, ArcGIS, STRADA, FIFA*

Trafiksäkerhet vid vägarbeten

GIS som ett verktyg för att lokalisera och kvantifiera olyckor vid vägarbeten

Examensarbete inom masterprogrammet Infrastructure and Environmental Engineering

JOACHIM JENSEN
FREDRIK SVENSSON
Institutionen för bygg- och miljöteknik
Avdelningen för geologi och geoteknik
Väg och trafik
Chalmers tekniska högskola

SAMMANFATTNING

Trafikverket genomförde år 2008 och 2012 två utredningar vars syfte var att undersöka vägarbetsrelaterade olyckor. Dessa studier utfördes genom att fritextsöka i en trafikolycksdatabas kallad STRADA, i vilken både polis och sjukvård rapporterar till. Resultatet från dessa undersökningar visar att mindre än 1 % av alla trafikrelaterade olyckor sker på grund av vägarbeten. Det finns dock anledning att tro att det existerar ett mörkertal bland dessa och att procentsatsen bör vara högre, då olycksbeskrivningarna som återfinns i STRADA anses vara otillräckliga.

Då både vägarbetena och trafikolyckorna rapporteras in såväl rumsligt som i tiden, kan en användning av geografiska informationssystem (GIS) vara ett sätt att kvantifiera detta mörkertal, vilket också är vad detta examensarbete hoppas uppnå. Genom att använda en programvara, kallad ArcGIS, utvecklades en modell där olyckor länkades till vägarbetsdata på statliga vägar. En filtrering gjordes därefter för de olyckor som hade skett under ett vägarbetes aktiva period liksom innanför dess influensyta. Resultatet från modellen jämfördes senare mot det normala antal olyckor som sker på samma vägsegment, men då inget vägarbete är aktivt på plats. Utöver det gjordes en motsvarande fritextsökning för åren 2009-2014, med syfte att användas i en validering av modellen längre fram.

Det visade sig för perioden 2009–2014 att 1,74 % fler olyckor sker på vägsegment då ett vägarbete är aktivt på plats mot hur det är i normala fall. Därutöver var 8 % av alla olyckor som skedde på staliga vägar eventuellt orsakade av vägarbeten. En stor del av de som skadade sig bestod av unga vuxna i åldrarna 25-34 för alla skadegrader.

Resultatet är dock inte övertygande, då endast 38 % av de i fritextsökningen bekräftade vägarbetsolyckorna identifierades i GIS-modellen. Detta tros bero på begränsningar i modellen samt osäkerheter i indatan. På grund av detta vill vi rekommendera fortsatta studier gällande vägarbetsrelaterade olyckor, speciellt med hänseende till inrapporteringsmetoden för STRADA samt att analysområdet skall begränsas ytterligare.

Nyckelord: *vägarbeten, olyckor, trafik, GIS, ArcGIS, STRADA, FIFA*

Contents

Preface

In this master's thesis the object has been to evaluate if a usage of geographical information systems can help in locating and quantifying accidents that have been caused due to road works.

The project was carried out during the spring 2015 at the consultancy firm Reinertsen's head office in Sweden, as well as under the department of GeoEngineering at Chalmers University of Technology.

The work done this last half a year would not have been possible without help from some people, to which we would now like to extend our deepest gratitude to. First and foremost, we would like to thank Fredrik Johnson, head of section Civil Engineering at Reinertsen and also our primary supervisor. Your guidance and feedback during our weekly meetings have been plenty helpful when we needed a push to proceed forward with the project.

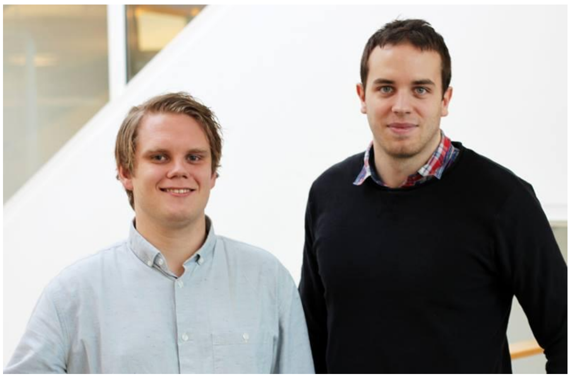
Following that, we would like to extend our regards to Gunnar Lannér at Chalmers, who were our secondary supervisor at the university. Thanks for the positive feedback and ideas for our thesis. We would also like to thank Anders Markstedt from Chalmers, for his guidance about the formalities regarding the examination process, as well as approving this master's thesis' subject.

We would also like to send out some "thank yous" to all those who we interviewed as part of our literature study. Those persons are in no particular order, Jimmie Sjöberg at the Swedish Transport Administration, Jonny Sandström at Reinertsen and Sebastian Hasselblom at WSP. Thanks as well to the Swedish Transport Administration and Agency for providing us with all the necessary data. Thanks as well to our opponents Hanna Porsgaard and Sofia Söderström, your feedback has been excellent during the thesis work.

Lastly, we want to thank each other for the pleasant days and occasional bickering, as well as our fellow thesis students at Reinertsen for the good laughs during all the *compulsary* coffee breaks.

Gothenburg, June 2015

Joachim Jensen and Fredrik Svensson



The two authors an early morning at Reinertsen's office in Gothenburg.
Joachim is the short one to the left and Fredrik is the stately one to the right.

Abbreviations

ADT	Annual Daily Traffic
BAS-P	Byggarbetsmiljösamordnare, planering (transl.)
BAS-U	Byggarbetsmiljösamordnare, utförande (transl.)
C	Capacity (vehicles/hour)
GIS	Geographical Information Systems
FIFA	Improved information on road works (transl.)
PHF	Peak Hour Factor
STRADA	Swedish Traffic Accident Data Acquisition
TA-plans	Traffic Arrangement plans

Translations

Arbetsmiljöverket	Swedish Work Environment Authority
Byggarbetsmiljösamordnare	Construction Work Environment Coordinator
Planering	Planning and design
Trafikverket	Swedish Transport Administration
Transportstyrelsen	Swedish Transport Agency
Utförande	Under development

“The best road to progress is freedom’s road.”

— John F. Kennedy, March 14 1961

1 Introduction

One hundred years have nearly passed since the introduction of the car as a mode of transportation for both commercial and private use (Foner & Garraty, 1991). Because of this, a need arose for a transportation infrastructure that was designed for the car in mind and the answer is the modern road network that today extends to even the farthest corners of our world (Joint Research Center, 2010). But in order for the roads to perform well for its intended users, the network continuously needs to be maintained and re-built.

In Sweden, the Swedish Transport Administration *Trafikverket*, has a vision for their infrastructural users; their trip should be secure, green and smooth to their destinations (Trafikverket, 2012). The safety aspect in the vision is something that has to be developed, as an average of 23 people were killed in a month due to traffic accidents during 2014, implying that the safety for the road users should be further improved (Transportstyrelsen, 2015a). The green and sustainable aspect in the Swedish Transport Administration's vision is steadily improving with new technology such as hybrid and electric cars and also less gas consuming asphalts. A vision of an efficient transportation on Swedish roads is essential since a lot of money and time is invested in both private and commercial transportation (Trafikverket, 2012).

In order to have an efficient transportation system within society it is imperative to maintain and modify the already existing road network, as well as constructing new roads. These road works need to be designed in a way that they do not have a too large of an impact on the flow of the traffic and at the same time not creating an unsafe environment for the road users or the road workers at the site. This has proven to be a challenging task not only for the construction companies that are building the road, but also for the consultants that design them.

1.1 Background

Several accidents occur each year during road works; the majority of the injured are users driving at the roads and not the employees working at it (Liljegren, 2008). The safety aspect at the road works does not only take into account the safety of the workers but also the safety of the road users which is why it is important to have a well-structured safety work at every road construction site (Sandström, 2015). The construction industry in Sweden has proven repeatedly to be a hazardous environment for the workers, which has led to tougher demands on safety at these road work sites. This is why every road construction, that will last a relatively long time, needs a Construction Work Environment Coordinator (Sandström, 2015). In relatively large road projects in Sweden there has to be persons responsible for the safety work, the two designated for this task are BAS-P and BAS-U. The first of the two is in charge of planning how to create a safe environment at the road work, not only for the workers at the site but also for the road users. The latter on the other hand is in charge of implementing those ideas at the road work, as well as improving those that are not sufficient enough (Sandström, 2015).

Despite that there are people in charge of handling the safety at road works accidents happen, many of which cannot be prevented due to factors that are hard to influence and foresee. These accidents are registered by both the police and hospitals and are entered into a database called the Swedish Traffic Accident Database Acquisition or

STRADA. This database is handled by the Swedish Transport Agency and can be accessed by persons working with statistics or traffic safety at governmental-, municipal offices or in the private sector. It is a belief that accidents happen more frequently at road works, this belief has led to several studies concerning how to make road works safer and reduce the amount of accidents happening at them. Accidents connected to road works have earlier been found in a study made by the Swedish Transport Administration where a free-text search was performed in STRADA. However, some of the descriptions are sparse concerning information on how the accident happened and as such may give an inaccurate representation on the amount of accidents that has occurred due to road works. As the accidents contain information about their location and the same information exists for road works, there could be an alternative way to connect them both spatially and in time.

Geographical information system (GIS) software products are mainly used in order to illustrate and analyse spatial data. They can be used solely as map programs showing roads, cities and forests, consequently not so different from a regular map. However, GIS can also be used as an analysing tool where roads can include information on how many lanes they have, how wide they are and what their speed limits are. This data can then for instance be used to perform analyses on how the severity of injuries in traffic are related to the road's speed limit or calculate a distribution route for products in order to make it as fuel efficient as possible (QGIS, 2015). Due to this, the usage of GIS may be a way to reach a more statistically certain representation on how more often an accident occurs because of a road work compared to when there is not any on-site.

1.2 Aim

This project aimed to examine if the occurrence of a road work resulted in a higher rate of accidents at any given road segment for the period of 2009-2014. This was done by the usage of the GIS program ArcGIS. The aim was to see if it was possible to quantify the number of accidents related to performed road works with ArcGIS and comparing it to a previously performed method. By doing this, the following research questions may then be answered:

- Is it possible by the usage of geographical information systems to provide more accurate statistics of road work related accidents?
- Is there an increase of accidents at a location when there is a road work active on-site?
- Which types of road work related accidents occur today and in what order of magnitude?
- How can the safety aspect be further integrated into the road construction phases?

1.3 Limitations

Accidents related to road works were only studied for roads owned by the Swedish government. This choice can be linked to the available data of earlier locations of road works, as the responsible authority on national traffic issues, the Swedish Transport Administration, has a large amount of this data compiled and archived. In contrast, information of road works performed on municipality owned roads are not located at a single source, making the compilation of this data too time-consuming for this project.

The type of road works examined was limited to those that were on a set location because of the difficulties of performing a GIS analysis for an event that has no fixed position. This is also tied to the fact that the data acquired from the Swedish Transport Administration did not contain as high level of detail for intermittent or moving road works.

Two studies were performed by the Swedish Transport Administration between the years 2003-2011 on traffic safety in relation to road works. This is why a study including the same time period is redundant, limiting this report's scope to the last full six years, or more accurately between 2009-2014. The start date of 2009 was chosen since it overlaps with the Swedish Transport Administration's results, making a comparable analysis possible, as well as showcasing if there is a reason to suspect if an estimated number of unknown cases were not shown in the earlier analyses.

1.4 Method

In order to get a wider insight in how the road construction process is performed, a literature study was carried out. Information were mainly sought from the two national transport authorities in Sweden, as they are responsible for maintaining and developing the national road network. Additionally, information was also gathered from the Swedish National Road Transport Research Institute, which specialises in research and development related to infrastructure and transport. Because of this, they have a large database containing previously performed studies that were of interest to this thesis.

To get a general overview of the subject interviews were performed in the beginning of the project, the first interview was with Jimmie Sjöberg at the Swedish Transport Administration in order to get an understanding of what kind of demands there are on traffic arrangement plans (TA-plans) and how they are designed when a road work is planned. A second interview was conducted at the consultancy firm WSP with the traffic planner Sebastian Hasselblom. This interview was done in order to give the authors more understanding on how the capacity is calculated on different road types and how road works affect the flow and speed of the overall traffic. A third interview, with Jonny Sandström at the consultancy firm REINERTSEN, was done in order to obtain which obligations and duties BAS-P has when designing and planning the safety work at a road works. Also a literature study was performed where reports concerning BAS-P and BAS-U together with the Work Environment Act were read and discussed.

The aim of this project was to investigate a new method of quantifying the number of accidents happening at road works. The traffic accident database STRADA was used in order to access information about where traffic accidents have happened the past six years, these accidents were then imported to a geographical information system software named ArcGIS. The police and hospitals are the ones registering the accidents into STRADA and the information is at times deficient, which is why several actions had to be done in ArcGIS in order to link if any accidents were connected to a nearby road work that was active at the time of the accident. The actions made in ArcGIS created a buffer area around every road work with a size dependent on the theoretical queue length that could form because of the road work. If an accident would have happened within this buffer area then it could potentially have happened because of the road work. This means that any accident within this buffer area was counted, even those happening several years later when the road work was finished. However, these buffer areas were constrained to only recognise those accidents that have happened at the same road and during the road work's active period.

These accidents, that were believed to have taken place at a road work, were then compared with accidents that happened within the same buffer zones and time of the year, however, at the years preceding or succeeding it. If for example a road work in January 2011 had any reported traffic accidents, these were compared with accidents that took place within the same buffer zones, however, with the difference that those accidents instead happened in January 2009, 2010, 2012, 2013 and 2014.

In order to find out whether or not the method of using ArcGIS was a reliable tool to find accidents happening at road works another method was performed. This method extracted those accidents that were reported to have occurred at a road work, the extraction was performed by doing a free-text search in STRADA where words related to road works were used. These accidents that were confirmed to have taken place at a road work were then compared to those accidents that were believed to have happened at a road work in ArcGIS. This was done in order to find out whether or not the method used in ArcGIS was able to find those accidents that were confirmed to have happened at a road work when the free-text search was performed.

In summary the subject of this master's thesis has been to extract the accident reports submitted into STRADA for the whole six-year period of 2009-2014 and to quantify those that occur at road works. This has been done by entering the data into ArcGIS, a GIS-software, and then cross-referencing the accidents to a nearby road work that has been active at the time that the accident happened.

1.5 Thesis outline

The thesis is structured so that it begins with three chapters that will delve into the background of the report's subject matter, followed by a chapter presenting previous comparable studies in the area. Lastly, a major part of the thesis is devoted to what the aims strives to achieve, namely providing an approach to assess accidents in GIS.

Beginning in Chapter 2, an explanation is given for how the planning process for road projects in Sweden is conducted. This is done in order to establish an understanding for the background factors that drive the development of new roads, as well as maintenance of existing ones.

Following that, Chapter 3 will provide an explanation for how road safety work is currently managed with the help of the BAS-P, BAS-U and TA-plans.

Chapter 4 then succeeds, which will go further into the current accident situation on Swedish roads as well as what type of accidents and injuries that can be expected on them.

Chapter 5 will present some of the results from two studies that were performed by the Swedish Transport Administration in 2008 and 2012. This is done so as to provide material that could be used in a later comparison to the results from this project's own GIS model.

In order to provide context for the later GIS modelling, Chapter 6 will introduce GIS as a method to investigate accident occurrences on Swedish roads. Some of the more relevant functions of the model will also be given an introduction.

At last, Chapter 7 will present the approach that was used in order to analyse the accident and road work data in both Excel and ArcGIS. This chapter is structured so that it can work as a manual for later studies in the subject, should a need arise.

The results from this analysis are presented in the following Chapter 8, of which a discussion regarding those same results will be held in Chapter 9.

Lastly, using both the results and accompanying discussion, some conclusions and suggestions for further studies will be given in Chapter 10. Additionally, in the back end of this thesis, some appendices are included in order to provide further information for the interested reader.

2 Planning Process for Road Projects

There is always going to be a need for new road constructions as well as maintaining the old ones in order to provide the basic services for the inhabitants within a society. It is not only the inhabitants that gain from a high-maintained infrastructure in society, also commercial companies that use the roads benefit from this. In Sweden there are two types of roads; the public – which are owned by the government and the municipalities – and the private roads. The public roads have a total length of 140 100 kilometres of which 70 % are owned by the government (Trafikverket, 2015).

This chapter is going to more thoroughly address the planning process in whole and how calculations are performed in order to determine a given road's capacity.

2.1 Overview of road planning phases

A governmental road construction project is often divided into three phases, which are *the planning phase*, *construction phase* and *management and maintenance phase* (Trafikverket, 2013). The planning phase is in itself then divided into four sections; *the strategic choice of measures*, *feasibility study*, *preliminary design phase* and *final design plan* (Svensk Byggtjänst, 2015). The following section will attempt to address this whole process and how the different parts of it relates to each other internally. See *Figure 2.1* for an illustration of the process.

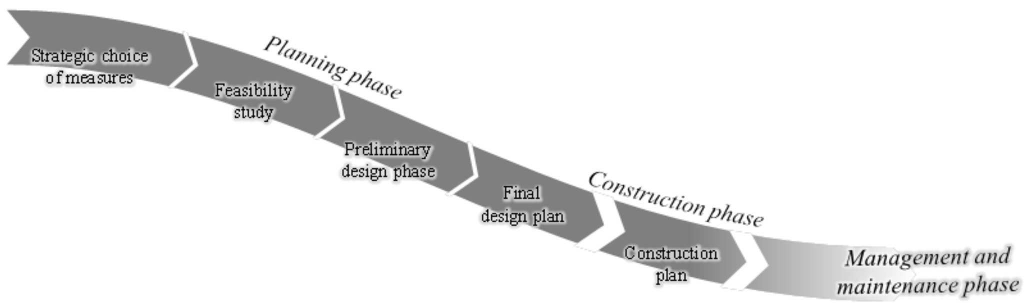


Figure 2.1 An overview of the three main road construction phases and the planning phase's four sections.

Strategic choice of measures: It is important that when the Swedish Transport Administration builds a new road that they think of the overall road network. This is where the strategic choice of measures is introduced - a phase where planned transport solutions are supposed to achieve a larger effect together with the rest of the transport system. This should be done by the help of concerned stakeholders that potentially could be affected by the decisions made by the Swedish Transport Administration. With more participants in the dialogue a broad knowledge base and experience is engaged within the process. The dialogue that is kept consists of four steps called *'Fyrstegsprincipen'* in Swedish. These four steps are *"Think differently"*, *"Optimise the use"*, *"Reconstruct"* and *"Construct new"* which in theory means that the entire road system should give a higher effect on the results rather than on the single transport solution (Trafikverket, 2014a).

The first step, *"Think differently"*, is where consideration on different measures are taken, these could for instance be regarding the needs on transportation and travels along with how the users chooses their way of travels. When *"Optimising the use",* in the second step, the focus lies on implementing measures that result in a more effective usage of the already existing infrastructure. The third step, *"Reconstruct"*, is to perform limited reconstructions only if there is a need for one. If the need cannot be fulfilled by implementing any of the previous three steps then a *new construction* should be built, this is however a last resort (Trafikverket, 2014b).

Feasibility study: The second step in the planning process in a road project is the feasibility study, also known as the preliminary study or pre-study. This pre-study evaluates and gives basic data in order to determine whether or not the project should move forward to the next phase. The pre-study also describes why the existing roads are not sufficient to today's standards; it could be that they are hazardous, not environmentally sustainable or that they do not have a high enough capacity to take care of the present traffic volume. The capacity of the given road type is evaluated in different flow-velocity diagrams. An example of these types of diagrams can be seen in *Figure 2.2* in the next section. The third point is used in order to find the capacity of a given road (Liikennevirasto, 2010).

The safety aspect during construction is often not considered in this phase but gets more and more important as the process advances. The pre-study is used to map the conditions required for a new road, to precise the need for a new road and to value the economic consequences if it is to be built. Certain allowable stretches – called road corridors – limiting the road's extent are considered where possible road alignments may lie within. The planning process looks similar whether it is a municipal or a government owned road network. (Liikennevirasto, 2010).

Preliminary design phase: Different solutions are analysed and studied in the preliminary design phase. Dependent on the shape of the road and which corridor is chosen the road will have different consequences on the environment and the safety of the road. These aspects are considered in this phase, however minor details such as noise barriers or lampposts are considered in a later stage. An environmental impact assessment needs to be done in order to determine what type of consequences the new road will have on the environment and the surrounding area. The safety aspect during this phase is primarily surrounding the road users and how to design the road in order to make it as safe as possible for them (Vägverket, 2007).

Final design plan: This is where the road is designed and a decision need to be made in order to find the most suitable road alignment that will fit the needs. This is also where, in a more detailed manner, investigations regarding soil layers and geo-technical properties are done in order to find out which challenges may arise during construction and usage. At the final design plan the proposal of the road in the horizontal-, vertical- and cross-sections are decided so that it will be safe, sustainable and effective (Liikennevirasto, 2010). This is where some larger infrastructural projects will have a construction work environment coordinator for the planning process (BAS-P). This person is supposed to identify risks that could appear at the site such as fast moving traffic or falling objects. BAS-P should also come up with solutions on how to avoid or remove these risks (Sandström, 2015). Chapter 3, Section 2 will go into the BAS-P's work more thoroughly.

Construction plan: During this phase all documents necessary for construction are drafted. A close cooperation between the consultancy firm and entrepreneur is needed in order for the construction to go as efficiently as possible, along with an open communication to easily clarify details and potential misunderstandings (Liikennevirasto, 2010). BAS-P is often not part of the construction phase, but is instead replaced by a BAS-U who is in charge of the work environment at the construction site. BAS-U is responsible for the implementation of the BAS-P's work and other work environment related matters (Sandström, 2015).

Management and maintenance: When the road is finished and open for traffic the last phase starts, the management and maintenance phase. It is during this phase that the road is properly managed with regular inspections and maintenance such as ploughing, patching and fixing barriers. BAS-P has a responsibility to set up documentations regarding those hazards that can be expected during the management and maintenance phase (Sandström, 2015).

2.2 Capacity calculation of roads

The capacity of a road tells how large flow it can handle before congestion starts to build-up and queues form. This capacity is difficult to calculate as it depends on several different factors. These factors can be visibility, speed, number of lanes, heavy traffic quota, et cetera (Trafikverket, 2014c). Because of the vast ranges of road types used as an input in this project a simplified version was used in order to calculate the capacity of the roads used. This simplified version uses the speed limit of the road, number of lanes, road type, percentage of daily traffic during the maximum peak hour and lastly the length of a mean vehicle travelling at the road (Hasselblom, 2015).

By using the number of lanes and speed together with what type of road that is being examined a relationship between the velocity and traffic flow can be reached. This relationship is illustrated in *Figure 2.2* and shows the relationship between the velocity and flow, where the y-axis is the travelling speed of the vehicles and the x-axis is the flow of the vehicles in vehicles/hour. The appearance of the diagram is depending on if it is a motorway, highway or a city bound road together with the number of lanes at that specific road segment. The speed is often known since it is a fixed value on specific road segment. The flow is however calculated by the use of the road's annual daily traffic (ADT) and the per cent of daily traffic during maximum peak hour. This per cent is depending on what type of road that is being studied, for instance, if it is a vacation road, an industrial road or a commuter's road (Johnson, 2015).

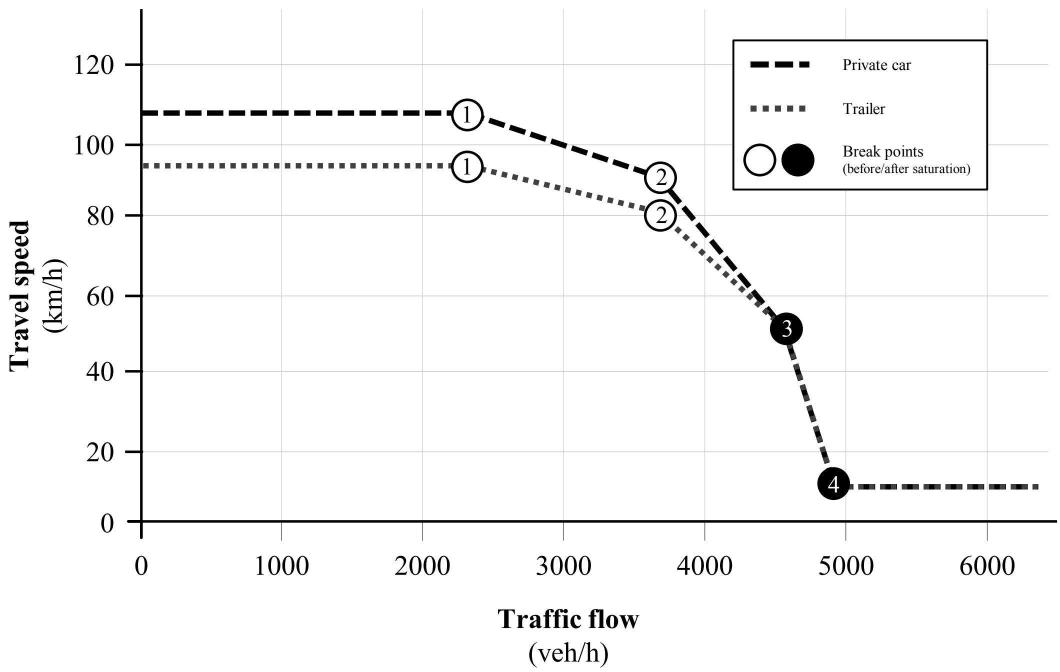


Figure 2.2 *Principle showing where the different breaking points occur for a specific traffic flow (numbered 1 to 4). Based on charts from* (Trafikverket, 2014c).

The travel speed and the traffic flow are then used as input into the diagram where two lines and four points can be seen. The two lines represent a private car and a trailer respectively. The four break points represent the different speeds at different flows. The speed before the first point represent the free-flow at that specific road segment, this means that the speed is presumed to be the same even if the flow is 1000 or 2000 vehicles per hour. The capacity of the road is reached at point three (in *Figure 2.2* about 4500 vehicles/hour), where the road experiences congestion (Trafikverket, 2014c).

Lastly, an explanation has to be given for the presumed speed of 10 km/h after point four. This value is explained by the Swedish Transport Administration to be a model value based on traffic statistics. At this point it is assumed that the speed will remain constant as the road is now so congested that no vehicles are able to overtake any other, resulting in a static relationship between them (Trafikverket, 2014c).

3 Safety Work at Road Projects

The construction industry is often labelled as one of the industries with the highest risk of work-related injuries among its employees (Arbetsmiljöverket, 2014). During 2013 the industry alone stood for about eleven per cent of all work-related accidents that led to an absence from work in Sweden. In absolute numbers it totalled at 3366 persons for the whole year, of which another eleven per cent could be related to infrastructural projects (Arbetsmiljöverket, 2013). Additionally, in a survey performed by the trade unions for construction workers and managers, about one fifth of all employees stated that they are worried at least once a week of hurting themselves at work (Fransson, 2008). Due to this, various legal and practical means have been developed in order to address these concerns, of which this chapter will present some of them, namely the *"Work Environment Act"*, *"Construction Work Environment Coordinator"* and *"Traffic Arrangement plans"*.

3.1 Work Environment Act

All employers who conduct business within Sweden are required to follow certain demands in order to ensure their employees' welfare. These demands are put forward by the Swedish Work Environment Authority as a way to prevent poor health and accidents while at work, as well as working for a satisfactory work environment. But to make certain that the employers actually work to guarantee their employees wellbeing, the same authority has the legal means with a law to enforce them to do so (Arbetsmiljöverket, 2015).

This law is called the Work Environment Act and it includes the businesses' work environment in relation to some different fields, such as; technical, physical, organisational, social and lastly the content of the work. The law is presented in several chapters, of which 6§ to 7§ in chapter three and 8§ in chapter four is specifically devoted to the work environment of the construction industry (Arbetsmiljöverket, 2015). An excerpt of 6§ in chapter three is presented, whereas the entirety of the relevant legal texts can be found in *Appendix A.*

> "*Chapter 3, section 6*
>
> The person who orders execution of building or construction work shall
>
> 1. during each phase of the planning and project, make sure that work environment viewpoints are considered when it is applicable to the building phase as well as future usage,
> 2. appoint a suitable building work environment coordinator for the planning and projecting of work with the information that is stated in Section 7 a, and
> 3. appoint a suitable building work environment coordinator for the execution of work with the information which is stated in Section 7 b and Section 7 f.
>
> The person who orders the execution of a building or construction work can appoint him/herself or someone else as building work environment coordinator. If someone else has been appointed the person who orders the execution of a building or construction is however not released from responsibility for such work information as is stated in the first paragraph of 2 or 3." (Arbetsmiljöverket, 2015)

Among other things, it states that those involved during the planning and design phase must ensure that the work environment during the latter stages of construction meets sufficient requirements throughout the whole process. Additionally, this includes the post-construction stage, when the considered users are driving on the finished road. (Arbetsmiljöverket, 2015).

However, in order to facilitate the safety work within an organisation it is of equal importance that its employees are involved in working to improve the work environment alongside the employers. Because of this, the employees have the legal right to nominate a person as their own safety representative. This person is then guaranteed by the Work Environment Act to be included in the organisation's dealings with said environment (Arbetsmiljöverket, 2015). Traditionally this safety representative is nominated by the employees' trade union.

3.2 Construction Work Environment Coordinator

As mentioned in the beginning of this chapter, the construction industry is considered to be one of the high-risk industries with work-related injuries. This has led to the development of an additional measure to guarantee that a safe work environment is ensured during the whole construction process. This position is called "*Byggarbetsmiljösamordnare*" (BAS) in Swedish, which literally translates to Construction Work Environment Coordinator (Sandström, 2015). However, due to the fact that the planning and design phase as opposed to the construction phase has very different considerations that need to be accounted for, the coordinator role has been divided into one for each phase. These two roles are called BAS-P and BAS-U respectively; the former handling the safety during the **p**lanning and design phase and the latter the construction **u**nder development. In Sweden it is required by law to have a BAS-P and a BAS-U in construction projects that are relatively large. How the two roles relate to the organisational structure of a project is shown in *Figure 3.1*.

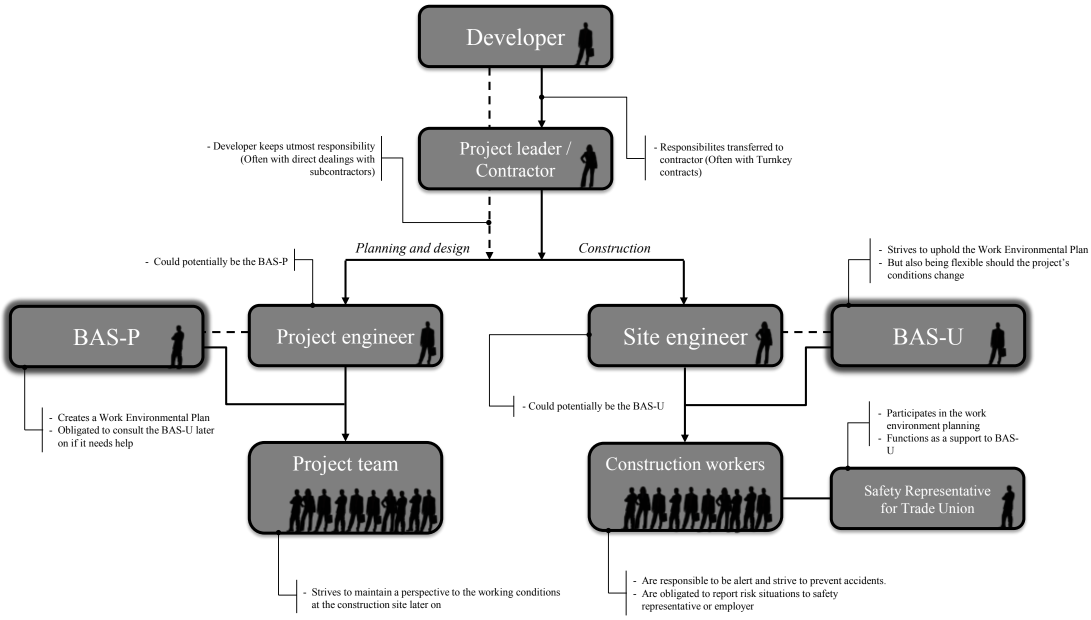


Figure 3.1 Organisation structure showing how the BAS-P and BAS-U relate to the overall construction hierarchy (Sandström, 2015).

The work of BAS-P is governed by the work environment act, as described in Chapter 3, Section 3.1. The BAS-P is part of the project because of hers or his understanding on work related risks and should, together with the engineers, identify those hazards that could arise during construction or by later usage. A list of risks should be established and together with a work environment plan – containing general safety procedures – be put up at the construction site. The most desirable scenario would be to eliminate the risks at an early planning phase, however, this is not at times possible because there will always exist dangerous construction steps. The best approach is instead to minimise those hazards that cannot be eliminated. Furthermore, BAS-P could potentially be the project engineer in a construction project but could just as well work under the project engineer within the project team. As such, the placement in the hierarchy is entirely dependent on what type of project that is being planned (Sandström, 2015).

When the planning phase of the project is done and it is time to begin construction, the BAS-P hands over the work environment responsibility to the BAS-U who is supposed to implement those preventative actions created by BAS-P. The type of risks that remain during management and maintenance of the road should then be compiled, which is part of the BAS-P's work (Sandström, 2015).

3.3 Traffic Arrangement Plans

By performing maintenance or construction work on a road, the traffic that usually flows on it will be affected in some way. This – combined with the fact that there are people working in an environment already stated to be associated with high-risk – increases the need to inform the road users of the upcoming road works as well as protecting the road workers. The easiest and preferable solution to these conflicting interests would be to close off the road during the allotted time period. However, as often is the case for larger roads, there might not be possible to re-direct its traffic onto a smaller, alternative route since a smaller road probably is not designed for the traffic volume in mind. Also, the possibility of an increase in travelling time affects the decision-making in such a way that it might result in a higher cost for the society.

By compromise, a plan for the traffic layout during the construction phase need to be proposed, as this makes it possible for the traffic to remain on the road while the work is performed, albeit in a slightly altered phase in order to accommodate the workers' safety as well. These plans are called traffic arrangement plans (TA-plans) and have to be submitted to the authority that is responsible for the road in question. Additionally, depending on which type of road that the road works is planned, the effect of it can be visualised to the public online on two accessible road maps called *"Trafiken.nu"* and *"Läget i trafiken"*. For roads owned by the Swedish government, the contractor is required to report the work in detail on a web-platform called "*Förbättrad Information För Arbete på väg"* (FIFA), an abbreviation that can roughly be translated to "*Improved information on road works"* which is going to be presented further in this chapter, Section 3.3.2.

The design of the TA-plans varies greatly from case to case, as the terms and environment surrounding the roads are not fixed between each project. This means that the regulations which govern these plans are not as strict and are more dependent on what demands the developer wants fulfilled for each plan, as well as what the road authority deems safe enough. However, there are some general guidelines of what the plans need to contain. For example, if the speed limit past a road work exceeds 70 km/h, a safety barrier is required in order to fully protect the workers inside the work zone. An aspect that also needs to be considered is whether or not the road work is planned to be set at a specific location or if it is moving either intermittently or continuously.

At the Swedish Transport Administration website an instructional document is given for the traffic engineers and road workers who are responsible for the layout of the TA-plans. The document is structured after a new principle, called the V3-principle, which the authority recommends its users to follow as a mean to ensure a safety device that protects the road users as well as the road workers (Trafikverket, 2014a).

3.3.1 The V3-principle – a way of ensuring safety at road works

The V3-principle consists of three stages to ensure that the safety zone around a road work is acceptable (Trafikverket, 2014a). The stages in order are: *Alert*, *Guide* and *Protect* and the name of the principle comes from the translation of the three stages' names in Swedish *"Varna"*, *"Vägleda"* and *"Värna"* of which an example is given in *Figure 3.2*. The three stages will be described in this section.

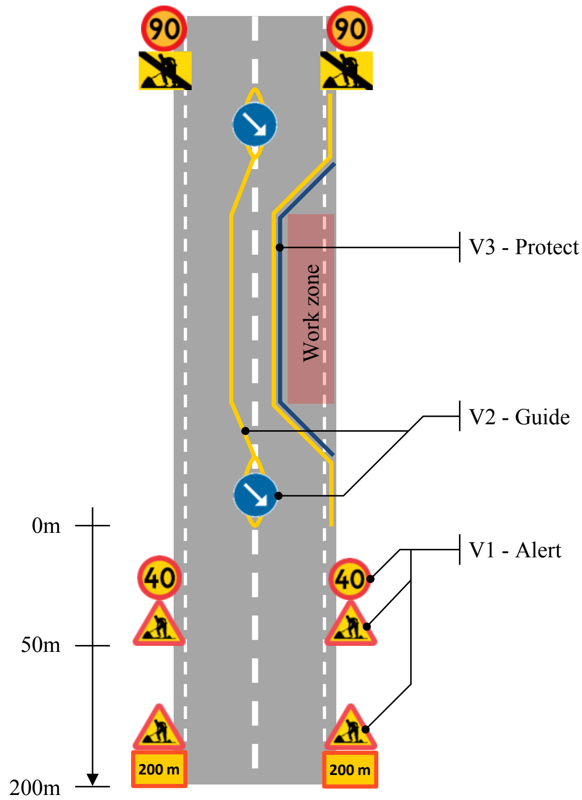


Figure 3.2 Overview of a TA-plan as well as giving a general idea of how the V3-principle works.

V1 – Alert: The first step is to alert the road users of the upcoming road work in order to prepare them for the manoeuvres they will have to perform later on. This can be done in various ways, but they all involve the usage of signs as a visual element. In this step the most widely used ones are temporary warning signs and informative signs, which requires that the drivers lower their speed as well as informing them of the upcoming situation, depending on how the road work is designed (Trafikverket, 2014a). Another more recent feature has been with the introduction of variable message-signs (VMS), which is a dynamic sign that can change its message depending on the situation by use of light emitting diodes. The difference when compared to the traditional usage of reflective signs is that it allows for flexibility in relation to the current activity of the work zone. However, due to VMS's usage of a more sophisticated technology it is also more costly. Additionally, it is not clear how much the road users distinguishes any differences between a VMS signalling a recommended speed as opposed to a static sign showing the speed limit (Friberg, 2007). The difference between the two is when there is a red border around the static sign, but it could be possible that the road users just interpret them both as the same, see *Figure 3.3*.

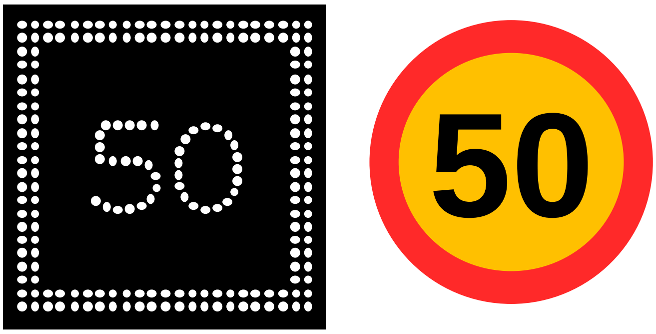


Figure 3.3 The VMS to the left shows the white border signalling a recommended speed, whereas the static sign always has a red border for its speed restriction.

V2 – Guide: After the road users have been alerted about the road work, it is advisable to guide them past it as safe as possible. The most preferable way to do this would be to re-direct the traffic using another route – this effectively eliminates any hazardous situation for the road users as well as minimises the work related risks for the road workers. But as this is not achievable at all times, the traffic will have to be guided safely past the work zone by other measures, namely either *calming* or *controlling* the traffic. In short, the two categories can be described as a way of influencing the drivers' speed and attention passively or actively (Friberg, 2007).

The traffic calming means could be summarised to include measures that indirectly affect the drivers' behaviour by decreasing the speed as they drive past the road work. By this definition the usage of signs, mobile speed bumps, road markings and the sighting of the police could be stated to calm the traffic, while still maintaining a steady flow (Trafikverket, 2014a). The disadvantage to this approach is that it does not create situations where both the road workers and road users are in total control. However, studies have on the other hand shown that it does manage to decrease the risk level considerably (Friberg, 2007).

The traffic controlling measures affect the drivers' behaviour directly by making them stop fully before gaining access past the road work. This can be done either with personnel such as a traffic guard or a pilot, or with the usage of signs and mobile traffic lights (Trafikverket, 2014a). This measure is most often preferred when the traffic is closed off on one side of the road, making passage limited in both directions. On the other hand it is often avoided on roads with a large amount of traffic, such as highways, as it disrupts the flow of the traffic too much and can in extension create congestions (Friberg, 2007).

V3 – Protect: The last step of the V3-principle is to enhance the safety of the road workers and road users by protecting them by various means. Barriers of different kinds are the most widely used for roads that will need to be able to maintain high speeds, as it both act as a buffer protection for the workers as well as guides the drivers. The kind of barrier used depends on what speed limit the passing road has; for lower speeds vertical reflective signs suffice; for higher speed a sturdy concrete wall might be needed. Another type of barrier is the usage of a TMA-vehicle (Truck Mounted Attenuator) placed in such a position that it stops a car that has lost control or that the driver have not realised that there is a road work. A TMA-vehicle is a type of collision protection that is, as the name suggests, mounted onto a truck's back end. The protection consists of a structure that is absorbing the kinetic energy of the colliding vehicle (Trafikverket, 2014a).

3.3.2 Reporting a planned road work

Due to the disturbances that a road work may generate in the traffic system, there is no possibility for an entrepreneur to start its construction plan without notification. For this purpose a reporting system has been put forward by the government and municipalities which own the majority of all roads in Sweden. This system is nearly consistent between the two, but there are however some differences, of which a presentation of both will follow.

Government owned roads: If a road work is to be performed on a road owned by the Swedish government, the approval of the planning and construction process has to be acquired from the Swedish Transport Administration. Because of the extensive national road network and the maintenance it often requires, the authority has created a web platform where contractors have to submit information about any planned project. If approved, a few parts of the information are sent out to various authorities who then can communicate to the public how the situation will be on the road network – facilitating both private and commercial users' need for an accessible transport system (Trafikverket, 2014b).

The web-platform is the aforementioned FIFA, which is the Swedish abbreviation for a translation of *Improved information for road works*. On the platform, persons tied to different companies primarily create accounts, but private persons are also allowed to do the same. The latter are given access in order to be able to report areas where there is a need for road works. Before the planned road work should be reported in, there are a number of aspects that need to be taken into account in order to begin the construction phase. The platform consists mainly of three parts that are relevant for the process; the creation of the application, a part for activating the project and lastly a possibility to search for older, approved plans in order to re-use their documentation (Trafikverket, 2014b).

Firstly, the submitter needs to give some general information as well as informing if the project can be limited to a single location, namely if it is a set, moving or intermittent work. Should it be the first of the three, the location will then need to be drawn on a map that uploads parts of the information from FIFA on a public map at two websites named "*Trafiken.nu*" and "*Läget i trafiken*", which several news media outlets and GPS-servers uses to inform the road users. Is it the latter two – moving or intermittent work – the traffic information will instead have to be phoned in to one of the Swedish Transport Administration's local offices, but some information about the upcoming project will have to be submitted into FIFA anyways. Among the other general information given, the submitter will need to describe its duration and if a risk assessment has been done, as well as providing its diary number for archive purposes (Trafikverket, 2014b).

If the work is to be on a set location, a drawing on a map needs to be provided as previously mentioned. Additionally for set road works, some other information is provided in this stage: the degree of influence on the road network, recommended speed, road number, effect on traffic lanes, time plan, et cetera. If the submitter is very detailed while filling in the information at this point, it will later on lead to more accurate estimations of how the planned road work will affect the road network, and in extension be more helpful for the road users (Trafikverket, 2014b).

Lastly, some further information will need to be provided in order to facilitate the Transport Administration's decision to greenlit the project. Among other things, the submitter has to upload their planned TA-plan and regulations that have been set beforehand in the process. Also, contact information to particular individuals of the project has to be supplied – these include the project engineer, main contractor, site engineer and the person responsible for the TA-plan's markings. Finally, when all this is done, the application is sent to the Transport Administration, which will then evaluate it and determine if it fulfils the requirements needed for it to be greenlit (Trafikverket, 2014b)

Municipal owned roads: When a contractor needs to perform road work on a road owned by a municipality, the process is near similar to the application process for the roads owned by the government. The difference comes to the fact that the Swedish municipal roads are not owned by only one stakeholder, as there are almost 290 municipalities in Sweden. Due to common infrastructural questions as well as other general ones regarding local government, these municipalities along with the Swedish counties have grouped together in an organisation called the Swedish Association of Local Authorities and Regions. This organisation has in turn put forward a guidebook that, according to national laws, provides the municipal traffic planners and contractors with basic information on how to perform a safe and secure road work (Sveriges Kommuner och Landsting, 2014).

As it is a guidebook it does not provide a binding framework that each party has to work towards, instead it gives support that both the municipality and private firms might need when conducting road works at these roads. In general, this support can be summarised to include what legal framework that exists, what responsibility each party has for the process as well as some examples of traffic arrangement plans that can be used. How the process is then conducted is up to the individual municipality, the guidebook is unspecified in that regard, as different conditions in different municipalities makes it difficult to create a uniform system on how to perform a road work (Sveriges Kommuner och Landsting, 2014). Lastly, due to the limitations of this thesis, no further focus will be given to municipal owned roads.

4 The Swedish Traffic Accident Database

Traffic related accidents happen every year, some of these accidents cause severe damages to the ones involved and some are even fatal. According to the Swedish Transport Agency 2976 persons were severely injured in the year of 2012 and 285 were killed in traffic-related accidents. These numbers can be compared to the statistics gathered in 1970 were 6614 persons were severely injured and 1307 were killed. The numbers of dead and severely injured have steadily been decreasing in numbers since the start with the compilation of statistics, but the vision in the future is to have zero deaths and severe injuries in traffic accidents. Statistics from recent years show that it is mainly elderly persons, 75 years or older, that die in traffic accidents. The percentage of fatalities is then steadily decreasing with younger age. It is mainly persons between the ages of 25-34 that have been severely injured in accidents the past years (Transportstyrelsen, 2013).

With the intention of decreasing the amount of accidents that occur on Swedish roads each year, the Swedish Transport Agency has developed a system for collecting accident data. This is also what this chapter will attempt to present as it will go into the background for the system, the way it works and how it can be used.

4.1 Vision Zero - a vision with no deaths or severe injuries

In October 1997, the Swedish parliament took a decision to vote yes on the Vision Zero (Sjöö & Ungerbäck, 2007). By taking this decision, Sweden changed their policy when looking at the future transportation system, a vision where road users should not need to worry about their life or their health when travelling at the government owned roads. The Vision Zero focuses on severe injuries and deaths that occurred on the roads – more focus than had been given earlier – as the idea is that no one should die or get severely injured in traffic related accidents. More effort was put into reducing these severe accidents and it was also decided that every stakeholder in the road community, whether they are part of the design or maintenance of the road or professional users such as truckers or taxi drivers, should work towards a safe environment on the road and prevent accidents.

The decision makers knew that if the Vision Zero would ever be a reality, the need of a proper database for collecting statistical data involving injuries and accidents that happen in traffic is essential. The old database that was being used at the time was the Swedish Transport Administration's accident database, "*Vägverkets OLYcksdatabas"*. This database only had accidents reported in by the police and in several cases did not account for accidents with only unprotected pedestrians involved since these accidents were rarely brought to the police's attention (Forward & Samuelsson, 2007).

4.2 Development of a new traffic accident database system

The new database system had already begun development when the decision to strive towards the Vision Zero was made. The task was charged to the Swedish Traffic Administration by the Swedish government in 1996. This new database, named the "*Swedish Traffic Accident Data Acquisition"* (STRADA), was going to not only take into account the accidents reported by the police, but also include reports from the national emergency hospitals (Forward & Samuelsson, 2007). The differences between the two authorities' reports are further presented in Chapter 4, Section 4.2.1and 4.2.2. By also having the emergency hospitals creating logs in STRADA more accidents could be reported and a more extensive injury log was thus created. Injuries could then be cross-referenced to what type of accidents that had happened which gave a wider understanding on how to design more safer roads, vehicles, et cetera.

In the decision made by the Swedish government in 1996, the following was stated:

> "The continued work should, in consideration of the legal matters in the data- and confidentiality legislation, clarify:
>
> - How the information from police and hospitals regarding traffic accidents could be developed further.
> - How collaboration with national and international statistics- and information systems would look like.
> - How people's integrity and confidentiality should be taken into consideration. How different user groups are supposed to gain access to the information that they are in need of.
> - How falling accidents among pedestrians should be included in the system
>
> The new information system that was going to be developed would be designed so that:
>
> - It supports the traffic safety work on a central, regional and local aspect.
> - It gives material that could make it easier to take the right decisions and measures in a traffic safety aspect.
> - It minimizes double work and costs in the public sector and at the same time give high quality and relevance." (SOU 2014:24, 2014)

4.2.1 STRADA – police

The police have an obligation to always report accidents to STRADA. Reports written by the police are important in order to know how and where the accident happened. As the police is not as properly educated to estimate how severe an injury is, when compared to medical personnel, their report should be regarded more critically, especially concerning the severity of the injuries (Forward & Samuelsson, 2007).

The form that the police fill in contains information of where and when the accident took place and what type of vehicles that were involved. An empty area is also included where the police are supposed to sketch a simple drawing on how the accident happened. The weather and road conditions should also be filled in together with a three-graded injury scale and whether or not someone is suspected to be driving under some form of influence (Forward & Samuelsson, 2007). Lastly, there is also a possibility for the police to fill in a space where they more thoroughly describe the accident situation; these text-strings are later used in the free-text search analysis performed in this master's thesis project. However, according to a study that reviewed STRADA as a system, these descriptions are from experience often deficient (Liljegren, 2008).

4.2.2 STRADA – health care

Differing from the police, the hospitals are not obligated to report any accidents into STRADA, however, they get financial compensation for every report they send in. This is due to the doctor-patient confidentiality where they need the patients consent in order to send in a report. When STRADA was introduced, 12 emergency hospitals started to report injuries caused by traffic accidents into this database. This was in 1999 and amounted to about 18 % of the emergency hospitals in Sweden. In 2003, the police's old database system OLY (*Vägverkets OLYcksdatabas*) was replaced by STRADA. At the same time, about 43 % of the emergency hospitals were reporting the traffic accidents into STRADA (VTI, 2013). Today, 98 % of the emergency hospitals are a part of STRADA, which means that only one emergency hospital is remaining (Transportstyrelsen, 2015b).

The form that the hospitals fill in is similar to the form that the police have but less demanding on how the accident actually happened. The hospitals are more thorough when describing injuries, such as if there were any fractures, lacerations or other damages as well as their locations on the body (Mattsson K, 2014).

4.2.3 Type of accidents in STRADA

Accident types are divided into twelve different categories in STRADA depending on what type of accident that have happened, for instance single-car accidents, rear-end collisions, intersection accidents, et cetera. These categories, which are listed in *Table 4.1*, are then divided in to sub-categories with a total of 80 posts. These sub-categories explain which type of transportation method and factors that were a part of the accident, such as bicycle/moped, motor vehicle, if there were any animals involved or from what type of road one of the drivers came from in accident at a junction. Those types of accidents that were of interest in this master's thesis project are shown in the same table, where they are denoted *"Used in analysis"* or *"Partially used in analysis"*.

Table 4.1 The 12 categories that traffic accidents can be classified as.

Index in STRADA	Definition of accident types	Comments
S0-S4	Single-car	Used in analysis
M0-M3	Head-on	Used in analysis
O0-O2	Overtaking	Used in analysis
U0-U2	Rear-end collision	Used in analysis
A0-A8	Turning	Used in analysis
K0-K7	Accident at a junction	Used in analysis
C0-C7	Bicycle/moped related	Not used in analysis
F0-F9	Pedestrian related	Not used in analysis
V0-V6	Other accidents	Partially used in analysis
W1-W5	Cloven-hoofed animal related	Not used in analysis
G0-G7	Pedestrian/bicycle/moped in an accident with no motor vehicles involved	Not used in analysis
J0-J9	Track-bound vehicle related	Not used in analysis

4.2.4 Extracting data from the web-platform

When the information has been gathered from both the police and hospitals, the Swedish Transport Agency compiles it and uploads it onto their own web-platform. This platform is then accessible for those who are registered users, which in most cases includes traffic planners at the municipality or consultancy firms, politicians and police officers (Transportstyrelsen, 2014). At the platform, it is then possible to extract the following information regarding an accident:

- Location given in GPS-coordinates
- Gender and age of the persons involved
- Severity of accident and injuries, using the ISS medical score
- Conditions of the road
- Description of the event that led to the accident
- Other specific information

As the data in STRADA is concerning personal information – which could in theory be linked to a specific person in an accident – some obligations have been set to the potential user of the system by the Swedish Traffic agency. In order to be able to extract any information at all the user needs to attend a course, provided by the agency, that presents how the system works and in what extent it can offer accident statistics. Also important, is the fact that the user needs to sign a confidentiality contract, which legally binds her/him not to disclose any personal information that could potentially reveal a registered person's identity (Transportstyrelsen, 2014).

4.3 Distribution of accident types at government owned roads

A preliminary analysis regarding what type of accidents that is generally occurring at governmental owned roads has also been performed. This was done in order to make a comparison later on against the results from the free-text search in STRADA and the ArcGIS model. To be able to make this comparison those accident types used as input into all the analyses had to be similar. This means that the accident types that were part of this analysis needed to be the same as those that were used in the other analyses such as the free-text search and the GIS model. As can be seen in *Table 4.1* this means that the vulnerable road users such as the pedestrians and the cyclists were not part of this report's later analysis. Those accidents gathered from STRADA had a total of about 90 000 different accidents and the distribution over the different accident types can be seen in *Figure 4.1*.

The data, acquired from STRADA, shows that *single-car* and *rear-end accidents* make up nearly 70 % of all accidents on roads owned by the Swedish government, when excluding accidents involving vulnerable road users.

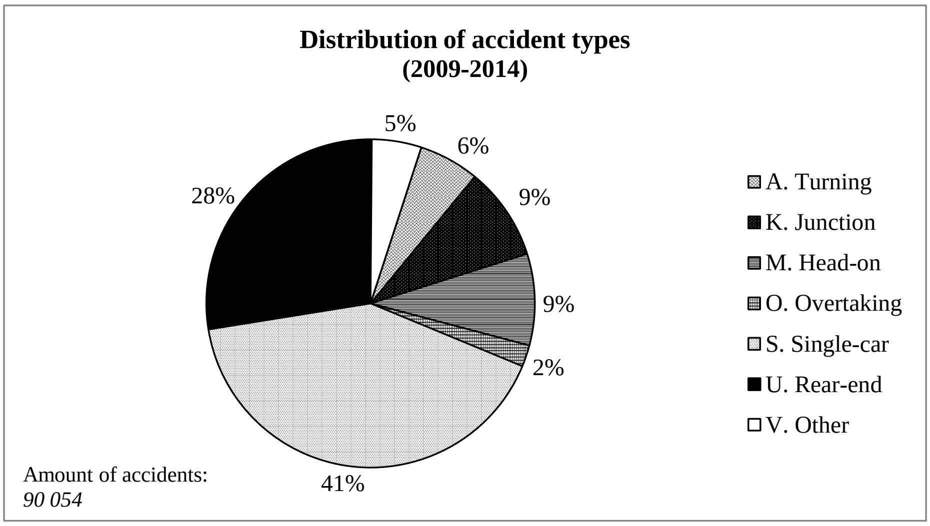


Figure 4.1 *Accident distribution on government owned roads for the years 2009-2014. Included types in the analysis are those listed from Table 4.1.*

4.4 Distribution of accidents among age groups at government owned roads

Another preliminary analysis was carried out in order to determine the distribution of accidents based on the severity of the injuries among the different age groups of the Swedish population. As stated above, this was done for a later comparison against the results acquired from the free-text search in STRADA and from the model in ArcGIS. The ranges of the different age groups are based on divisions used by the Swedish Traffic Administration (Liljegren, 2008).

Looking at *Figure 4.2* below, it can for instance be seen that young adults between the ages 25 to 35 years old are the well-represented for nearly all the severity grades.

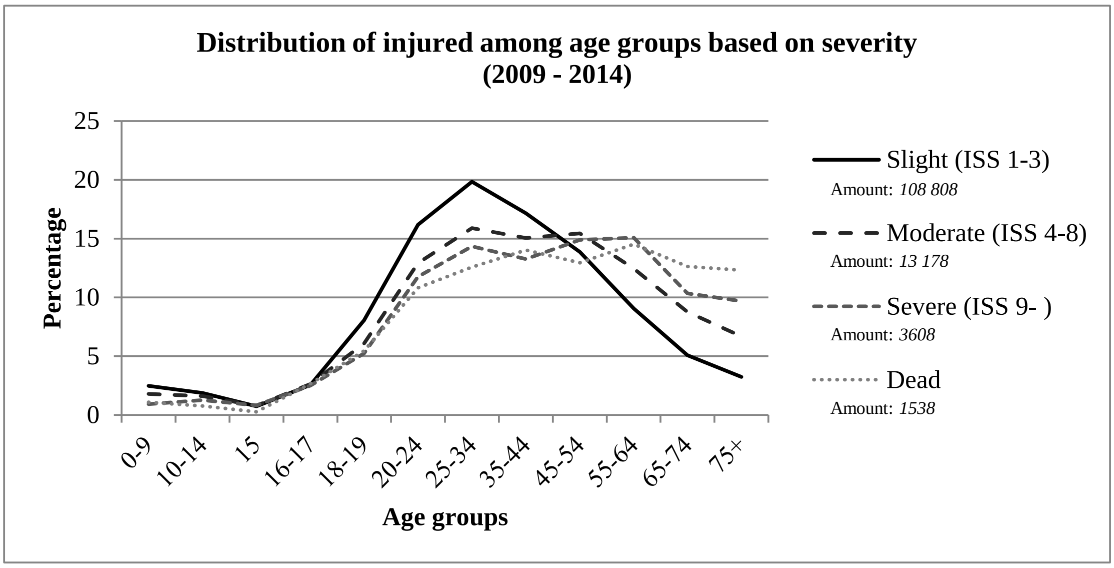


Figure 4.2 *Age group distribution on government owned roads for the years 2009-2014. The severity is based on the ISS medical score.*

5 Previous Studies of Accidents at Road Works

Two studies have previously been made by the Swedish Transport Administration on the risks at road works, one in 2008 and one in 2012. The study made in 2008 considered only the reports sent into STRADA by the police and was made in order to analyse how to design future road works in a safer way and reduce the number of accidents occurring at these. The report was written because there was a belief that a hidden number existed for how many accidents that occurred because of road works each year. This study showed that when making a free-text search with keywords[1] related to road works in STRADA, 1073 persons were found injured or killed in accidents in the years 2003-2007. These results are only based on the reports made by the police into STRADA, which is why it is believed that the number of road work related accidents should be even higher (Liljegren, 2008).

Following, the study made in 2012 was based on the reports written by the police together with the emergency hospitals between the years 2003-2011 and mainly focused on what type of injuries the persons had when they were in a rear-end collision, since this is one of the most common accidents at road work sites (Liljegren, 2012). The following sections present the results from these studies and what type of accidents that occurred at road works together with what type of injuries that often are inflicted to the persons involved.

5.1 Type of accidents identified by study made in 2008

As stated earlier, the study made in 2008 only took the police reports into account for the period 2003-2007. Rear-end collisions amounted to about 45 % of the total accidents between the years 2003-2007 (Liljegren, 2008). This percentage is produced when only taking those accident types into account listed in Chapter 4, section 4.2.3 in *Table 4.1*, since these can then be used in a comparison with this master's thesis project. Because of this, accidents representing vulnerable road users such as pedestrians and cyclists are not illustrated in the pie chart. However, these accidents represented about 15 % and were the third largest accident type between the years 2003-2007. The distribution of the different accident types is illustrated in *Figure 5.1* and represents accidents that have occurred at both municipal and government owned roads. Finally, in the study it was found that less than 1 % of all traffic related accidents are caused because of road works. The author of the study did however believe that there is a large hidden value of these accidents and that the percentage should be higher.

[1] Vägarb, gatuarb, nybygg, ombygg, reperation, reparation, Svevia, SKANSKA, PEAB, NCC, Vägverk, Trafikverket, Trafikkontor, plog, sand, salt, snöröj, slåtter, arbetsford, beläggningsarb, asfaltsa, vägskrap, hyvel, vägbygg, väghålln och gatukontor.

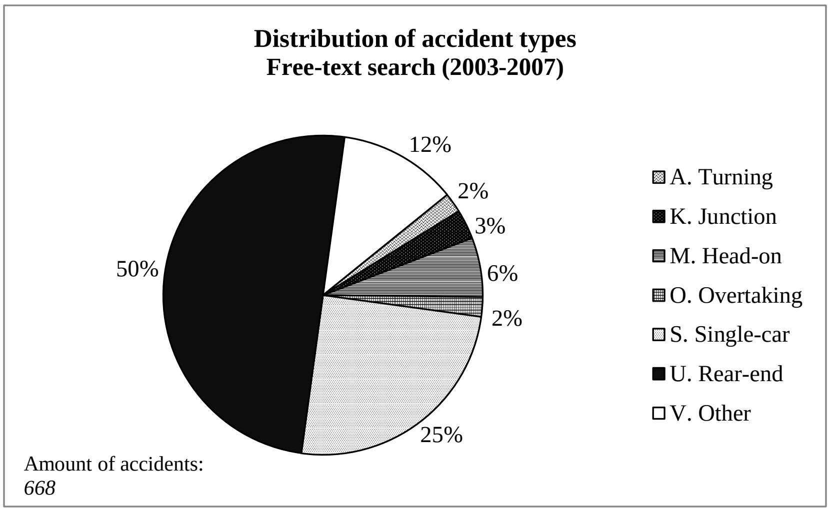


Figure 5.1 Accident distribution at road works on Swedish roads for the years 2003-2007, based on a free-text search in STRADA (Liljegren, 2008). *Included types in the analysis are those listed from Table 4.1.*

The rear-end collision accidents were found to mainly occur due to a sudden queue build-up or that a road user collides with a slow-moving or stationary road work vehicle. These collisions are more often between two road users rather than with a road work vehicle. In a rear-end collision the first road user slows down while the second is not paying enough attention and collides with the vehicle ahead. These types of accidents often occur far back in the queues, away from the road work. Sometimes they occur as far from the beginning of the queue that the alert of an upcoming road work has not yet been sighted by the driver. When this kind of accident occurs it is not obvious for the police to know why the accident happened, since there may be no road work in sight or that the queue may have dispersed by the time the police arrive to the scene (Liljegren, 2011).

Single car accidents are the second most common accident type at road works, these often have the most serious outcome. In the study made in 2008 it represented about 27 % of the total accidents happening at road works. The third largest accident is the *"Other"* accidents, these accidents were for some reason unidentified and are as such a source of error in the results. According to the author of the study made in 2008, a large part of these should be identified as the accident type rear-end collisions. This was noticed by the author when reading the accident descriptions more thoroughly.

Lastly, it was seen that pedestrian related accidents were a small contribution to the total number of accidents at road works. The reason why these accidents were not as represented in the study could be because they were rarely brought to the police's attention, as the injured were instead sent to the hospital immediately (Liljegren, 2008).

5.2 Type of accidents occurring identified by study made in 2012

As stated in the previous section, the study made in 2008 only took the police's reports into consideration; however, the study made in 2012 also included the hospitals' reports. The report was performed with data compiled between the years 2003-2011. The statistics between the two reports look a little different and the results for the report made in 2012 are illustrated in *Figure 5.2*. The study was taking both government and municipal owned roads into account, which is a difference to this master's thesis project where only the government owned roads were used.

In the study made in 2012 the rear-end collisions were 49 % of the total amount of accidents at road works, single car accidents were 29 % and the second highest were the accident with vulnerable road users such as cyclists, pedestrians, et cetera. As earlier stated, these types of accidents involving vulnerable road users are not directly comparable with this project, which is why these accidents are not included in the pie chart. The accidents named *"Other"* in *Figure 5.2* are accidents that for some reason could not be identified as any of the given categories. The emergency hospitals have just in the past couple of years reached 98 % coverage in Sweden, which means that many accidents were not reported at all into STRADA earlier (Liljegren, 2012).

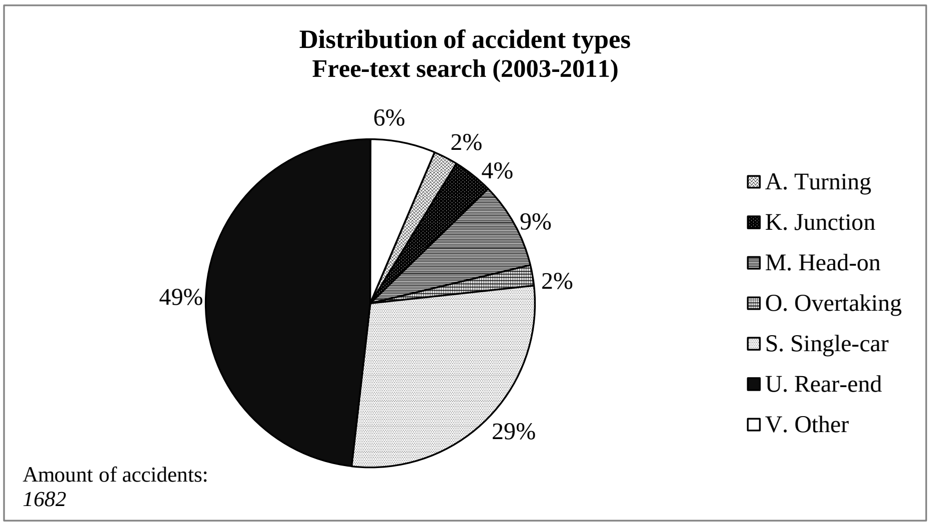


Figure 5.2 *Accident distribution at road works on Swedish roads for the years 2003-2011, based on a free-text search in STRADA* (Liljegren, 2012). *Included types in the analysis are those listed from Table 4.1.*

Many of the accidents that occur when a road work vehicle has been involved in an accident is during snow ploughing or when it is performing any other maintenance work. The road user does not expect it to move as slow as it is and because of that does not have the necessary time to brake resulting in a collision. Accidents also occur when road users are not aware that a road work vehicle is stationary which could end in a collision. This type of accident is often happening with TMA-vehicles where most of the collision energy is absorbed into the collision protection mounted onto the back of the vehicle. TMA-vehicles have proven to not only lessen the afflicted injuries to the road user but also provide the workers at the site with a safer environment (Liljegren, 2014).

5.3 Type of injuries occurring due to road works

The most common type of injuries at rear-end collisions are whiplash injuries. This type of injury is often classified as slight by the police even though the injury might cause problems for the remainder of a person's life (Liljegren, 2011).

According to a study made by Folksam – a Swedish insurance company – women more often develop whiplash injuries when they have been in an accident even though there were almost the same numbers of women and men involved in rear-end collisions in the years 2003-2009. This can be explained by that the car's safety system often is designed and tested for men (Liljegren, 2011). Whiplash injuries in traffic accidents are increasing fast; they represent over half of the disablement claims in Sweden that are caused by traffic accidents. About 25 000 persons in Sweden develop whiplash injuries each year in traffic related accidents of which about 1 400 of them are experiencing permanent injuries. Whiplash injuries are more severe if the speed is higher and are caused by the fast forward and backward movement of the head during a collision, which damages the soft parts of the neck. Symptoms do not always arise at the time of the accident, they could show up as late as six months after the incident (Söderström, 2004).

An injury that also is very common in traffic accidents is fractures. These injuries are often classified as severe by the police even though the recuperation period just lasts a couple of months while whiplash injuries can last a lifetime, which could be a reason to why they are not classified as severe by the hospitals (Söderström, 2004).

6 An Introduction to Geographical Information Systems

As most of the analysis will be undertaken in a computer software environment called Geographical Information System (GIS), there is a need to explain how this type of programming works, which is what the following chapter will attempt to do.

GIS is a way of showing and analysing geographically referenced information. More basically it can be compared to the traditional way of illustrating the physical world by maps. Whereas a traditional map only contains the information that can be seen on it in the form of symbols and representations of geographical features – a map in GIS contains concealed data that are linked to the specific elements shown on it. An example of this, which is related to this project, is the depiction of roads. On a regular map, roads only show up as lines connecting between different localities or objects. Different road types, speed limitations, road widths, et cetera could be depicted by usage of different colours or hatchings, but there is an extent for how much this can be done before it affects the map's readability. In a GIS-software on the other hand, the roads are still portrayed as lines and can be made distinct by the same usage of colouration and hatchings. However, the difference comes to how the software manages to store a larger quantity of information inside each of the road elements, which then can be displayed when needed. More importantly, the geographically referenced data inside each element can be combined with other features on the map – for example statistics of accidents – to provide a map showing how a certain type of road could be more prone to cause car collisions (QGIS, 2015).

The data displayed in GIS can be shown in two different ways, namely in a rasterised or vectorised form, the difference between the two is illustrated in *Figure 6.1*. Raster data is best described as a matrix containing a large quantity of blocks with colour information (also called pixels). An advantage with this kind of information is that it is easy to edit and display, but on the other hand it minimizes the potential to use the raster data in conjunction with other data for analysis. Meanwhile, vectorised data consists only of information about points and lines and how they might interconnect. Depending on how the points and lines do this, it allows for a variety of figures and shapes to be drawn that are at once also infinitely scalable. These features are called datasets and can be displayed as the already aforementioned points and lines, but also as polygons. An advantage with vector data is that it is built up mathematically and each feature is linked to a specific row of data in an attribute table. This allows for large quantities of information to be stored within each element and in extension makes it possible for the GIS software to perform calculations on it (QGIS, 2015). Vector data is saved in a data format named shape files (.shp) and is the chosen method for the analysed input data in this project.

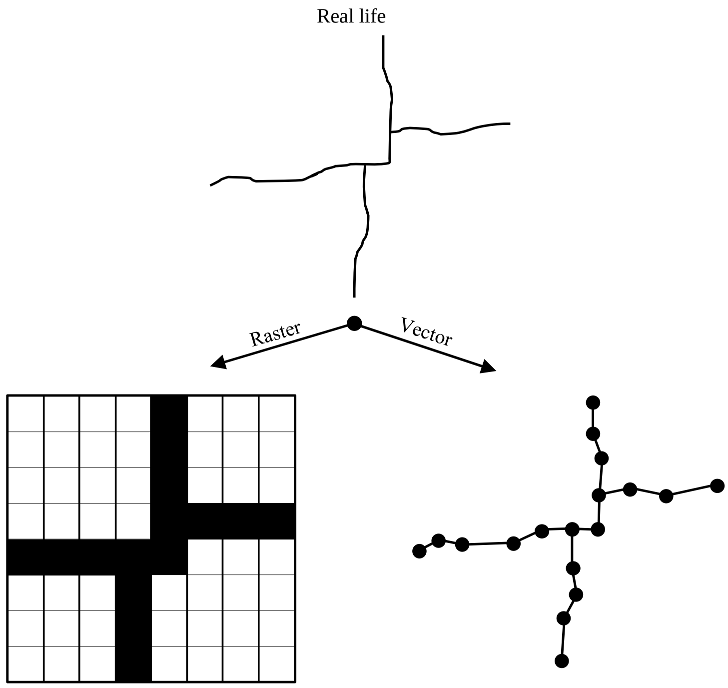


Figure 6.1 The two different representations of reality by usage of raster or vector data in a GIS-environment.

As mentioned earlier, the objects depicted on the map are shown as separate features. The way that they are displayed within the GIS software is often by the way of a *"layer tree"*, which has the extended function that it easily can toggle on and off their visibility. For vector data, this does not affect the ability to perform calculations with the features – it merely makes the map more readable. Lastly, the *"layer tree"* has the additional advantage that it provides a short-cut to extended functions within the GIS software.

For this project the GIS software ArcGIS has been used, this is a licensed program developed by an American company named Esri. By being licensed it is capable of handling large amounts of data as well as including several built-in functions. It is by the usage of these functions that the data of the different map features can be combined, of which a description now will follow for the more relevant ones related to this project. All of the following functions are only applicable for a vector dataset.

6.1 Dissolve and Merge

When working in ArcGIS with a map containing a large amount of different features, one way to optimise the computer's performance is to combine these into a smaller amount of groupings. This can be done in two ways depending on how the map elements are loaded into the program.

If the features are displayed in the same shape layer file, the available method is to use the *'Dissolve'*-function in ArcGIS. This allows the user to join together several features depending on a chosen factor. In *Figure 6.2* an example of this is given. Here a polygon dataset (containing eight smaller features) is made simpler by dissolving some of the smaller features together based on their common classification number. This reduces the amount from eight to three and thus the shape layer file is now made up of less data (Esri, 2013a).

Should the datasets instead belong to two different shape layer files, the other method will then be to use the *'Merge'*-function. Another example is given in the same *Figure 6.2*. The two different line layers are displayed by the overlapping dashed squares. By using the function these will then be merged together into a single file, which is then more easily displayed in ArcGIS (Esri, 2013b).

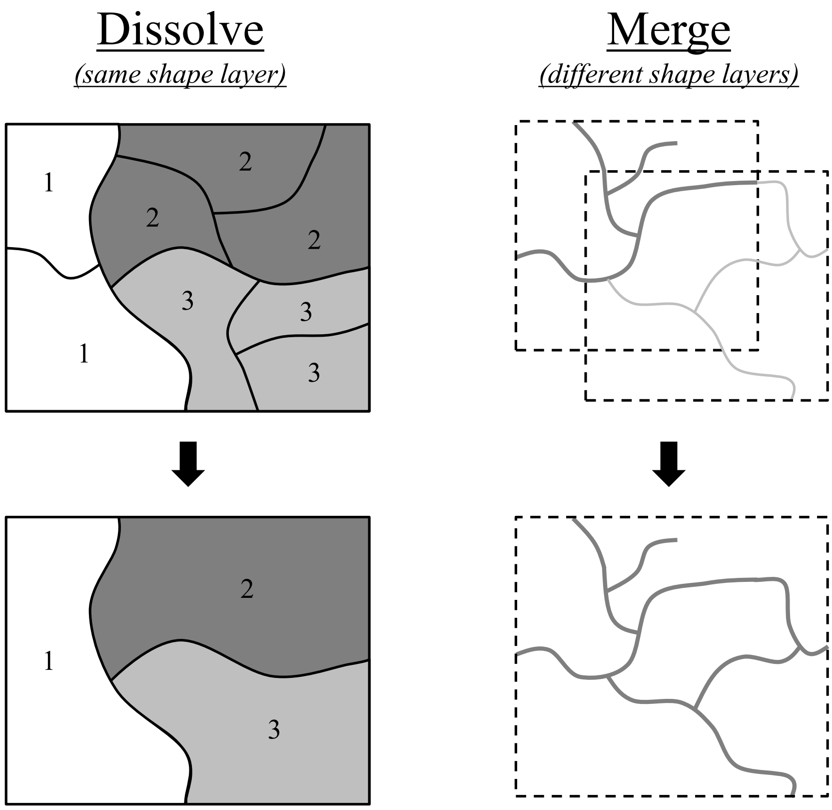
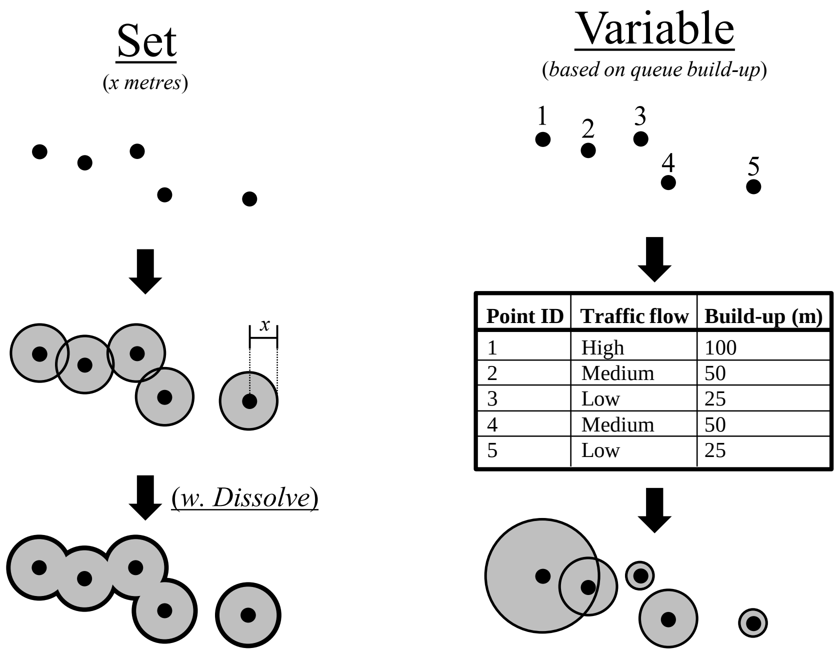


Figure 6.2 *The result of using the functions 'Dissolve' and 'Merge' in ArcGIS. 'Dissolve' is used when the dataset consists of only one shape layer file, whereas 'Merge' joins two different shape layer files into one.*

6.2 Buffer

In some instances there might be a need to decide whether or not an object is positioned within an acceptable distance to another object. This can most easily be decided by using a built-in ruler in ArcGIS, however, in most cases there is a need to perform the analysis multiple times and because of this the '*Buffer'*-function exists within the program. In *Figure 6.3* an example is given for how this function works for a point dataset, where it creates a set of polygons. By starting the function within ArcGIS, the user is given several options on how the program should perform the buffering. The buffer length could for example be set to a given value decided by the user as shown to the left in *Figure 6.3*, or change depending on a value linked to each specific feature within the dataset – as shown to the right. The '*Buffer'* tool can also perform a *'Dissolve'* simultaneously as it buffers, either based on a specific attribute within each feature or combine them all together. This last option often creates a new dataset that consists of a smaller amount of features compared to the original one from which it was created. Another important aspect is that the created polygons from the buffer also can inherit previous attributes linked to the original feature (Esri, 2013c).



Point ID	Traffic flow	Build-up (m)
1	High	100
2	Medium	50
3	Low	25
4	Medium	50
5	Low	25

Figure 6.3 *The result of using the two different 'Buffer' functions in ArcGIS. Using a set distance of x metres, a buffer area with the radius x metres is created. Using a value based on the 'Build-up' feature will instead give variable distance.*

6.3 Spatial Join

As a GIS environment is built up around its spatially referenced information, a need might arise to combine information between different datasets depending on how they are positioned in relation to each other. If the datasets consists of a common identifier, this can easily be performed by linking them through the *'Join'*-function, though this loses the spatial relationship between them. Another more powerful tool to use is *'Spatial Join'* which adds data from one dataset to another depending on what match option that is chosen (Esri, 2014).

The dataset that will have received data from another dataset is called the target feature, whereas the one delivering it is called the join feature. Looking at *Figure 6.4*, the point layer is the target feature in this case and the line layer is the latter one. For this particular example the match option is that the point layer should receive the name of the road it is closest to. ArcGIS then performs a vicinity analysis, searching for the closest road to each point and then generates a new layer which includes the road name in a new column for each point feature. Some other match options are also available depending on what type of analysis that needs to be performed, for example if a feature intersects another one; if it has one completely within it; if it is spatially identical to another one and so on (Esri, 2014).

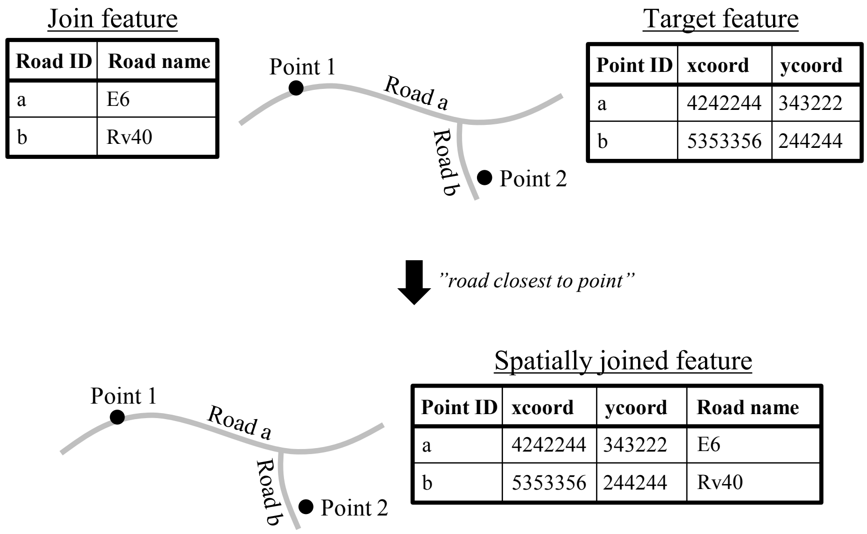


Figure 6.4 *The result of using the 'Spatial join' function in ArcGIS. The join feature (road names) is spatially joined to the closest join feature (point ID).*

Lastly, depending on the spatial relationships between two or more features the join operation can be performed in two different ways. Say that a point is located completely within two polygon features and the user wants to add the ID of each polygon to that point's attribute table. This relationship is called a *'One-to-many'*-relationship and in ArcGIS this generates two duplicate points at the same location, but having the two different IDs added to them each. The other spatial relationship is the *'One-to-one'*-relationship. In this case only one new point is generated acquiring one of the two IDs only, and in extension loses the relationship to the one polygon not chosen. The process by which a polygon is chosen is entirely dependent on where they are located in their own attribute table, e.g. the first found polygon that fulfils the match option is added to the new point's attribute table (Esri, 2014).

6.4 Selections

In ArcGIS different features can be chosen and highlighted by clicking on them with the built-in selection pointer. However, in most instances when an analysis has to be performed on only a limited amount of the displayed features the use of the two *‘Select by…*’-functions is more appropriate. See *Figure 6.5* where the difference between them is exemplified.

To the left, a selection is performed on the point dataset based on whether or not the points fall completely within the polygon that is chosen as the spatial constraint. This method is called *‘Select by location’* and as can be seen in the image, the two points located within the polygon is highlighted in grey making it possible to perform further analysis on them (Esri, 2012a).

The other instance is to use *‘Select by attributes*’, this is shown to the right in *Figure 6.5*. In this case the selection is performed based on something called a ‘*Structure Query Language’*-text string (SQL). This is a built-in programming language that makes it possible to select a feature based on a specific query. In the figure, the two points that fulfils the Boolean statement *“Year ≥ 2013”* is therefore chosen, while point 1 is ignored as it is linked to the year of 2012 (Esri, 2012b).

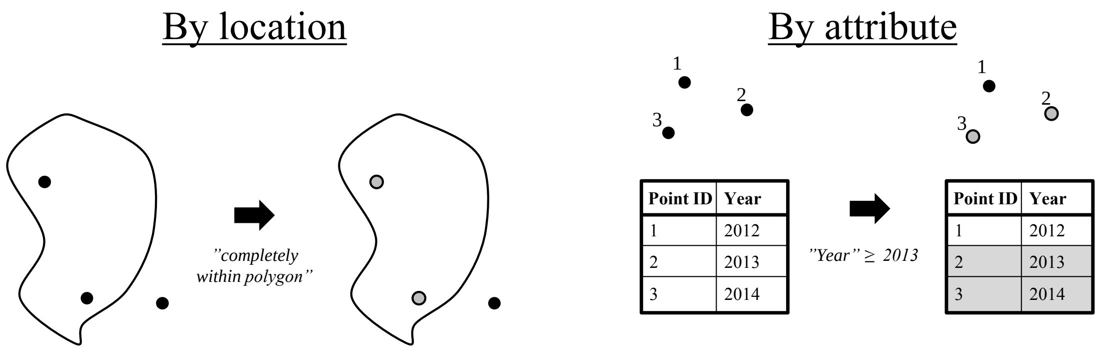


Figure 6.5 *The difference between the two ‘Select by…’-functions. ‘Select by location’ chooses the point located completely within the polygon, whereas ‘Select by attributes’ demonstrates how the points occurring during or after 2013 are selected.*

6.5 Model builder

When working in ArcGIS it is important to have a well-structured plan on when functions should be performed, in what order and with what type of input data. If the same functions are to be iterated on several input files there is a more efficient alternative than manually performing each function repeatedly. For this a built-in application called the *'Model builder'* exists. This application allows the user to create a large flow chart with functions that is to be performed for each of the input files and does not need any attention from the user each time a new function is to be initiated. How a typical model in *'Model builder'* is built up is illustrated in *Figure 6.6*. This particular model contains two input files, called '*Input 1'* and *'Input 2*', where two separate functions are performed – *'Buffer'* and *'Select by attributes'* – and they are later merged into one file. This output from the *'Merge'* function is then divided into two different files where one function is performed on each of them (Esri, 2012c).

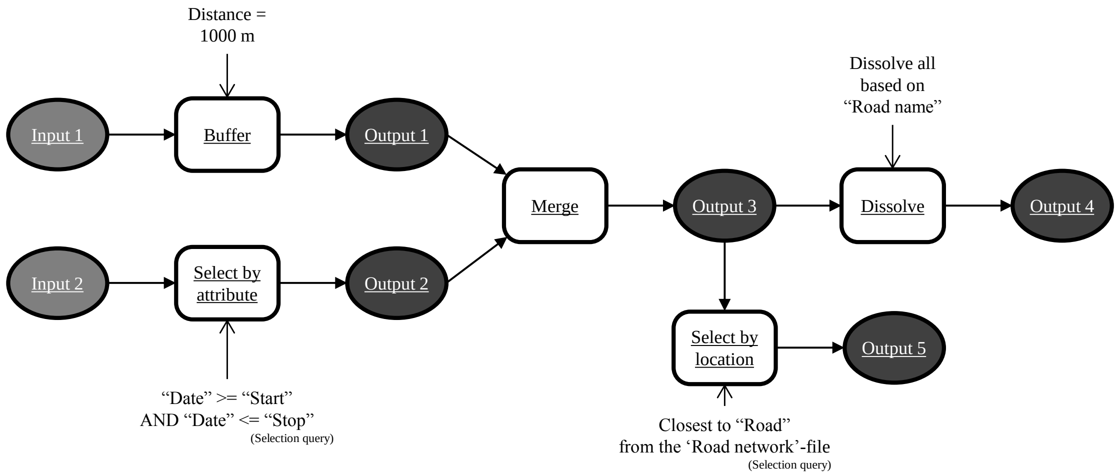


Figure 6.6 *Illustration of how a model in 'Model builder' can be built up. Some examples of user inputs are shown by the arrows going into the white function boxes. The lighter grey ovals are input data, whereas the darker ones are output data.*

If the model is very large and in the end it is found that there are a number of errors in the model, then the time it took for ArcGIS to run the specific model is lost. However, as long as a function is performed within *'Model builder'* and the input to it included no errors then that function does not have to be run again. Conversely, should a function in the model contain any errors then these will affect the consecutive functions and in the end give a wrong result, which means that a restart of the model is necessary. In *Figure 6.6*, if the settings in the *'Dissolve'* function is set up wrong and the output contains an error then there is no need to run the functions *'Buffer'*, *'Select by attributes'*, *'Merge'* or *'Select by location'* again since these have already been completed without any errors. When working outside of the *'Model builder'*, there could be an error in the settings of a function, which gives a wrong answer. There is however only a need to run that specific function again, as long as no actions have been done afterwards. Was there any error in the input then the previous function need to be re-run, just like in the *'Model builder'* (Esri, 2012c).

7 Model of Accidents Occurring due to Road Works

This chapter describes the actions that were made in both ArcGIS and Excel. The model that was performed is specific for this project's constraints but can also be used in similar projects if needed. As the model includes several operations that are complicated, this chapter has been structured so that each section begins with an overview of the tasks performed, followed by a deeper description of the necessary steps to achieve it.

For an overview of the whole process in ArcGIS, see *Appendix B.*

7.1 Pre-filtration of irrelevant accidents in Excel

Several hundred thousand accidents have been logged into STRADA the last six years, 2009-2014. This data has been assembled by the Swedish Transport Agency and was reported in by the police and emergency hospitals. The accident data has then been used for this analysis made in Excel and ArcGIS.

- Tasks:
  - Removal of irrelevant accidents
    - Those outside 2009-2014
    - Those with irrelevant accident types
    - Those without coordinates
  - Modification of the accidents' dates
  - Removal of accident descriptions

Removal of irrelevant accidents: Since ArcGIS, along with the computer used, both have their limitations, a series of optimisations had to be made. These optimisations removed the accidents that were considered not to have a large effect on the results. First of all, accidents that had happened before 2009 and after 2014 were removed as the project's limitations were set to just analyse these years.

The accidents that then were removed are listed as *'Not used in the analysis'* in *Table 4.1*, Chapter 4, Section 4.2.3. The more specific accident types such as V1, V2 and V4 are accidents dependent on motor vehicles in conflict with animals; track-bound vehicle in conflict with another vehicle; and lastly a bicycle or moped in conflict with another bicycle or moped. They were removed because these types of accidents often happen on municipal roads. As stated in Chapter 1, Section 1.4 it is only the road network owned by the Swedish government that was taken into consideration.

Lastly, accidents that were reported into STRADA without any coordinates for their specific locations were removed. These accidents were not of interest since they could not be used in the GIS-model in any case. After this initial removal of irrelevant accidents about 90 000 remained.

Modification of the accidents' dates: When the categories that did not fulfil the need were removed from the data the file needed to be adjusted in order to make it readable in ArcGIS. An earlier test run had to be made in order to find out how an interaction between the road works and the accidents from STRADA would work. This test run showed that the date was an essential part in the analysis and thus it needed to be addressed. Since the date in STRADA was compiled with the year, month and day in different columns the operations made in ArcGIS later would not work without making some adjustments.

Removal of accident descriptions: Additionally, more alterations had to be performed with some of the columns in the STRADA data. For instance, ArcGIS is not able to handle a too long text string, which meant that the description of the accidents had to be removed from the file. Several other columns in the STRADA data were removed, such as if a motorcyclist wore protective clothing, if the air bag had inflated during a collision or if the car had a safety seat for children, et cetera. This type of information would not be used in the analysis anyways and would as a consequence make the program slower. Should this information be needed in a later stage, it could be referenced back to their specific accident.

7.2 Manipulation of road work data

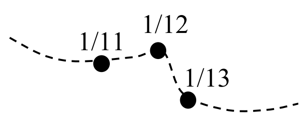
- Tasks
  - A comprehension of the road work data
  - Removal of irrelevant road works
    - Those with duplicates
    - Those without coordinates

A comprehension of the road work data: The data containing every road work was provided by the Swedish Transport Administration and contained data from the last six years. This was received as a semicolon-separated text-file, containing information about when a given road work started and ended, a description of it, its coordinates, et cetera. Each row in the text-file would represent one node in ArcGIS and every element between each semi-colon represents single features for that specific node, this is illustrated in *Figure 7.1*. These rows, in the text-file represented a single node in a given road work which together with the rest of the nodes in the same road work, build up a chain, which can be seen in *Figure 7.1*.

(Textfile)

ID; intID; start; stop; event; x-coord.; y-coord.
1; 11; 20120112; 20130201; renovation of railings; 6583791; 634214
1; 12; 20120112; 20130201; renovation of railings; 6583801; 634224
1; 13; 20120112; 20130201; renovation of railings; 6583811: 634204

(ArcGIS)



(Excel)

ID	intID	start	stop	event	x–coord.	y–coord.
1	11	20120112	20130201	renovation of railings	6583791	634214
1	12	20120112	20130201	renovation of railings	6583801	634224
1	13	20120112	20130201	renovation of railings	6583811	634204

Figure 7.1 The illustration consists for how the road work nodes were received and the following conversion from a semi-colon separated text-string to a table.

Removal of irrelevant road works: The text-file containing road work nodes was imported into Excel where it was discovered that it had 7.9 million rows and almost half of them were found to be duplicates. This amount of rows is far more than Excel can handle in a single sheet. To solve this, a Matlab program (a SQL programming environment) was written that removes the duplicates and writes a new text-file containing the IDs and coordinates for the road works, see *Appendix C* for how the program was set-up. When the new text file was created it only contained about 4.9 million points. Lastly, road works that did not have any coordinates were removed since they could not be used later in the spatial analysis. The text-string that described the road works was – just as with the accidents – too long and needed to be removed in order for it to be properly imported into ArcGIS. Again, should this information be needed it could always be referenced back.

7.3 Modelling in ArcGIS and Excel

The Swedish Transport Administration provided the road network that was used in the analysis since it contained thorough information on the amount of traffic, speed limits of the road segments, a more precise location, et cetera. This section describes how the model in ArcGIS was built up, which parameters that were taken into account and the results of each action.

7.3.1 Standardisation of different sources of input data

The whole process for this section is illustrated in *Appendix B, Section B.1*.

- Tasks
  - Standardisation of accident dates
  - Modification of the road network
  - Removal of deficient road work and accident data
  - Creation of the initial buffer zones (400 metres)
  - Adding road data to road works

Standardisation of accident dates: There are several aspects that needed to be considered when exporting from Excel to ArcGIS. The date of the accidents needed to be written with the entire date in one column in the format *yyyymmdd*, in order to be able to interact with the road works' date within ArcGIS. The two different columns for the road works' date would be the limiting values and mark when the road works started and ended, these are exemplified in *Figure 7.1*. These dates would also have the same format as the accident's dates. In theory, if the date of the accident would occur within the road work's active time period, along with the condition of it happening nearby and on the same road as a the road work, it could mean that the accident was caused by it. Below follows the selection criteria that the accident data will need to fulfil in the later stages of the analysis:

$$Date_{road\ work,start} \leq Date_{accident} \leq Date_{road\ work,stop} \tag{7.1}$$

$$AND$$

$$Road\ number_{accident} = Road\ number_{road\ work} \tag{7.2}$$

Lastly, the Excel sheet containing the information on accidents was converted into a semi-colon separated text-file along with the file containing the road works and both were then imported into ArcGIS.

Modification of the road network: In order for the program to have far less information, the road network needed to be reduced. The road network consisted of about half a million of road segments and all of these needed to be joined together by using the function *'Dissolve'* in ArcGIS, which is illustrated in *Figure 7.2*. Since the name of the road was to be used in a later stage of the analysis the roads had to have their road name kept in the attribute list and not consist of one large road throughout the entire Sweden. This was managed by dissolving the roads by their names, which not only kept the name of the roads but also merged the roads that had the same name and as such decreased the amount of data allowing for an increased computer performance.

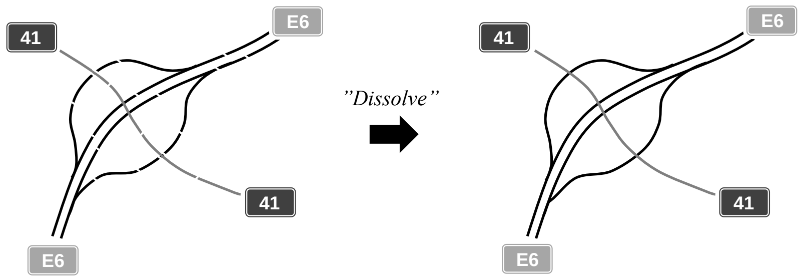


Figure 7.2 Dissolving the road network based on the attribute 'Road name'.

Removal of deficient road work and accident data: Some of the accidents and road works that have been reported into STRADA and FIFA were located outside of the road network, which made it hard to link those to a specific road that they occurred at. It was therefore decided to create a buffer of 40 metres along the entire road network. This road buffer was used in order to connect the accidents and the road works to the specific road that they were located at. The accidents and road works that were not located within this 40 metre buffer were removed completely since they could not be used in an analysis if they were not linked to a specific road. This buffer was also created in order to remove those accidents and road works that did not occur at any of the government owned roads, this operation is shown in *Figure 7.3*. The points removed from the analysis have been depicted as *"Null"* in the figure.

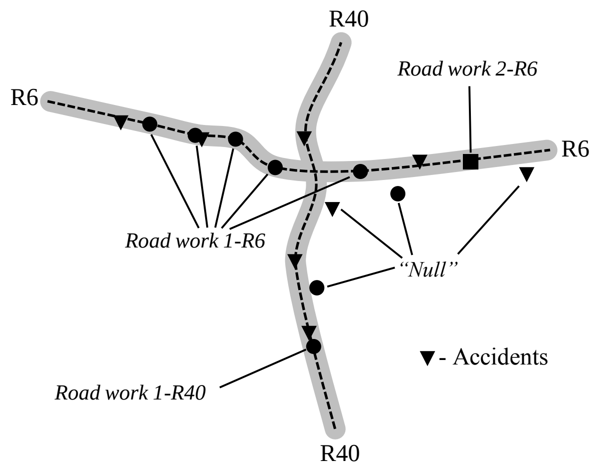


Figure 7.3 *The road works inside the grey buffer zone are linked to the conventional naming system of Swedish roads. Road work 1 is separated into two as it has been performed on two different roads. The road works and accidents that were not within the buffer zone were classified as 'Null' and consequently removed.*

Standardisation of road names: The road buffer that was created inherited the road's attributes and was now containing the same information that the road network had. This was important since both the road works and accidents needed to be given the same information on the road's name that they were located on. They were given this information by the use of the function *'Spatial Join'* in ArcGIS. This command added an extra attribute to the points that were within the buffer, for instance the name of the road. An example of this can be seen in the previous *Figure 7.3*. As can be seen in the figure, *"Road work 1"* was split up with *'Spatial Join'* because it was located at both road 6 and road 40 and as such was handled as two separate road works. Both the road works and the accidents originally contained information on which roads that they had occurred at, it was however discovered that this information was sometimes deficient since the construction firm, police or the medical personnel may have filled in the road's name wrong when they reported the road work or accident into FIFA or STRADA respectively.

The road names could also have been named in a different convention depending on which of the two sources that had reported them into STRADA. The same road could for instance be named *"E6.20"* by the police and *"Hisingsleden"* by the health care. When the function *'Spatial Join'* was performed it ensured that ArcGIS gave a consistent name to both the accidents and road works that had happened at the same road. It was imperative that the naming convention of the roads was the same since it was used in a later stage.

Creation of initial buffer zones (400 metres): In order to conduct a spatial analysis and decide whether or not an accident had occurred near a road work or not, a buffer with a radius of 400 metre was created around the road works. This buffer was created at every one of the road work nodes and was dissolved depending on their ID and at the road they were located at. The operation is illustrated in *Figure 7.4*. These buffer zones were working as an initial influence radius of 400 metres for the road works. This value was chosen due to the V3-principle described in Chapter 3, Section 3.3.1, where *V1 – Alert* was the distance that informed the road users of the upcoming road work. 400 metres was chosen as a model value because the alert distance is not a set value, but can instead vary between road works.

Furthermore, this means that if any accident had happened inside these buffer zones it could potentially mean that it was caused by the road work, provided that it occurred within the time frame that the road work was active and at the same road.

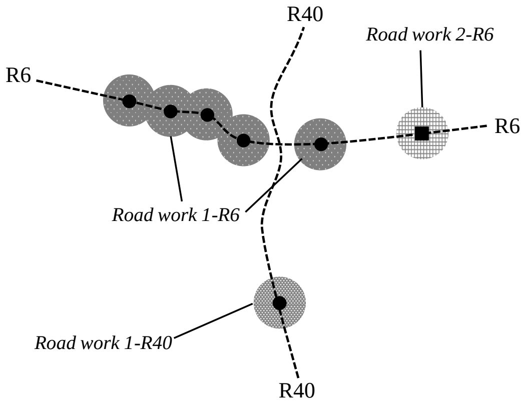


Figure 7.4 The initial buffer of 400 metres created around every single node of the road works. Note that Road work 1-R6's buffer zones are dissolved because of their ID and location.

Adding road data to road works: However, some road works that were located on roads with large traffic flows could potentially have a larger influence on the surrounding traffic. By using the function *'Spatial Join'* between the road network and the closest starting point of a road work, all the necessary data in order to calculate a new theoretical queue length build-up was located within the file of the road works. The data joined to these starting points was the annual daily traffic (ADT), width of the road, type of road, et cetera.

Only the starting point of every road work was given this data since it would have been too much information if all points would have gotten it. The spatially joined shape-file containing the road works with road data was then exported as a semi-colon separated text-file from ArcGIS into Excel where the calculations for the influence radius were to be made.

7.3.2 Creating buffers for theoretical queue length

The whole process for the ArcGIS part of this section is illustrated in *Appendix B, Section B.2*.

- Tasks
    - Calculation of variable buffer zones
    - Creation of variable buffer zones

Calculation of variable buffer zones: The first action that was performed in order to calculate this new influence radius – the theoretical queue build-up – was to calculate how many traffic lanes there were at the starting point for each road work. As stated in Chapter 2, Section 2 a road usually consists of lanes and shoulders. Therefore, the width of the road was divided by 3.5 metres (the standard width of a traffic lane) and then subtracted by 0.25 metres (the standard width of a shoulder) which gave the number of traffic lanes at each of the road works.

$$Lanes_{total} = \frac{Width_{road}}{3.5} - 0.25 \quad (7.3)$$

The capacity was then examined by using the method described in Chapter 2, Section 2.2 where the number of lanes and road type need to be known. If one or more traffic lanes were to be closed off during the road work – this was listed in the road works' descriptions – it would lead to a decrease in capacity for the road. This new capacity was calculated by multiplying the original capacity with a percentage of the available lanes after the road work was started. The formula for this is shown below.

$$C_{new} = C_{original} \times \frac{Lanes_{total} - Lanes_{closed\ off}}{Lanes_{total}} \ , \ \ C = Capacity \left[\frac{veh}{h}\right] \quad (7.4)$$

The peak hour factor (PHF) in the day depends on what road type that is being analysed. If it is a vacation road, a commuter's road or city road that is being analysed then the percentage for the maximum hour of the total ADT is different (Hasselblom, 2015). These percentages for the different road types are in the span between nine and twelve per cent. A model value for PHF of 10 % of the ADT was used at every road segment in order to see how many cars would potentially contribute to a queue build-up during the peak hour.

Lastly, the queue length was calculated by subtracting the maximum flow from the capacity, multiplying it with one hour, which is the maximum flow hour, and then, with an assumption that every car took up a space of seven metres, resulting in the final length of the build-up. These seven metres that a car occupy are just a model value, it is in reality depending on several factors such as speed, number of lanes, type of traffic, et cetera (Hasselblom, 2015).

$$Length_{queue} = \frac{|ADT \times PHF - C_{new}| \left[\frac{veh}{h}\right] \times 1\ [h] \times 7 \left[\frac{m}{veh}\right]}{Lanes_{total}} - 400\ [m] \ , \ [m] \quad (7.6)$$

As can be seen in the formula above, it subtracts the 400 metres for the initial buffer zones. This is done in order to optimise the model later on, as it is unnecessary to perform a buffer of zero metres of length for those who road works that had no queue build-up. In ArcGIS, these were separated from those that had acquired a theoretical queue build-up by the usage of *'Select by attribute'*.

Creation of variable buffer zones: The theoretical queue length was saved as a semi-colon separated text-file and once again exported to ArcGIS. Buffer zones were then created depending on the calculated queue length, how these buffer zones are illustrated in ArcGIS is shown in *Figure 7.5*, the final buffer zones are the ones with thick black borders. As can be seen in the same figure, *"Road work 1-R40"* was not assigned a variable buffer zone since it was calculated to not have any queue build-up.

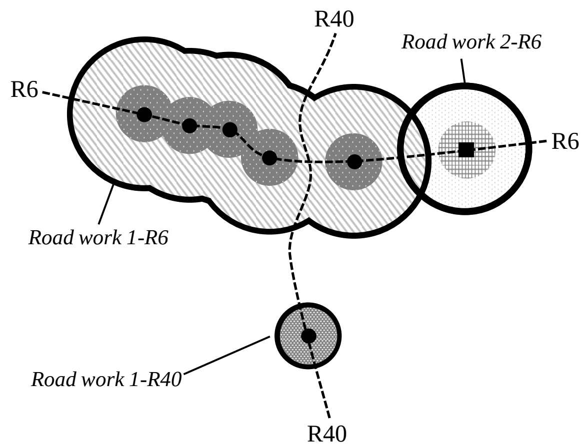


Figure 7.5 A variable buffer can be seen around road work 1-R6 and 2-R6 as they were calculated to have an extensive influence on the road network. Road work 1-R40, however, kept only its initial buffer.

The new buffer zones around some of the road works were created based on the calculation of queue lengths done in Excel. These buffer zones were only created around those road works that had a theoretical queue build-up, which means that they were in a different shape-file than the other ones with a buffer zone of 400 metres. By using the function *'Merge',* these two shape files were combined and one file was created containing the both instances.

7.3.3 Selection and extraction of relevant data

The whole process for this section is illustrated in *Appendix B, Section B.3*.

- Tasks
    - Joining the road works' data to the accidents
    - A description of time-steps
    - Creating the time-steps
    - Selection model

Joining the road works' data to the accidents: As per the selection criteria with *Equation 7.1* and *Equation 7.2*, the road work's active time period and road number needed to be included in the accidents' attribute list in order to perform a *'Select by attributes'*. This data was transferred by performing a *'Spatial join'* with the accidents and the road works' buffer zones. This *'Spatial join'* with a *'One-to-many'*-relationship create large amounts of data since every accident occurring within multiple buffer zones duplicates, just as explained in Chapter 6, Section 6.3. This was the reason why the accidents were given the road works' data and not the other way around. It was also more convenient to do it this way since it would have required more computational power to create duplicates of polygons instead of points, however, the amount of data created was still too large for the computer to handle.

In order to decrease this quantity of data, Sweden was divided into seven different segments for which a *'Select by location'* was then performed for the accidents within each of the segments. By doing this, accidents were divided into seven different shape-files which made the data more easily manageable in ArcGIS. These segments can be seen in *Figure 7.6*. Several '*Spatial join'* was then performed with the accidents and the road works that had happened within each of the seven segments.

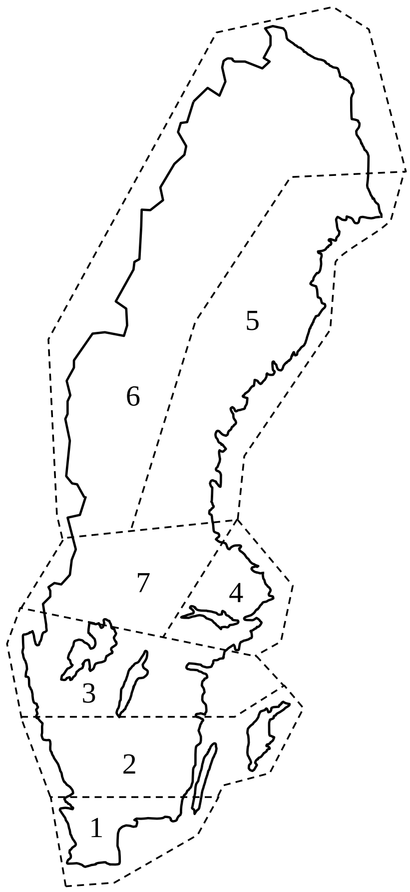


Figure 7.6 The division of Sweden into seven different segments.

A description of time-steps: In order to make an analysis of how many more accidents that happened at locations where a road work was performed – in comparison to the same period but for another year – something called time-steps had to be created. Figuratively, these symbolise the normal situation of a road segment when there is no road work performed on-site. The time-steps therefore had to be created with some calculation models in the application *'Model builder'* in ArcGIS.

Creating the time-steps: First, ten empty attributes named start1, stop1, start2, stop2, et cetera was created for the accidents in each of the seven segments – as these were to be filled in a later calculation step. For readability's sake, the following part of the text will not consider the seven segments as separate; instead they are referred to as one.

The first model had a logical expression that checked when a road work started and ended. For instance, if a given road work had an initial start date (*"start0"*) of 20090101 and a stop date (*"stop0"*) of 20090130, then the same period of year was created but for the consecutive years. For each of the accidents within the buffer zones the output from the model gave ten different values. With the previous example in mind the time-step *"start1"* would then be equal to 20100101 and *"stop1"* equal to 20100130. This continued for the four consecutive years, until *"start5"* and *"stop5"* was filled with values. How this works is illustrated in *Figure 7.7*.

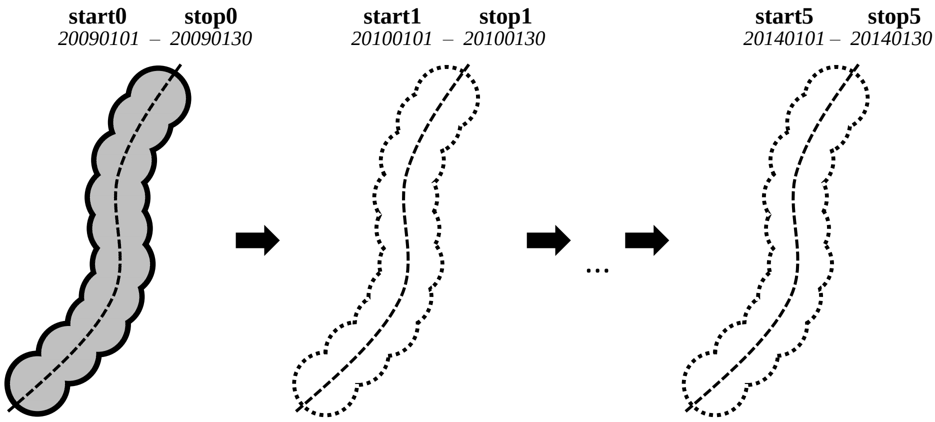


Figure 7.7 The principle behind the transformation of a road work's active period into consecutive time-steps.

For those road works that started in the years 2010-2014, a special criteria was written that prohibited the years to be larger than 20150101 and made them start over at 2009 again. This ensured that the period always was in the range between 2009 to the end of 2014.

Following, for those road works that had an active period that started in one year and ended in another, a second logical expression was written into ArcGIS's *'Model builder'*. *Figure 7.8* shows how these cases are handled along with an illustration of the model. The first oval circle to the left in the figure was the input into the model, this input contained the spatially joined accidents with the road works' data. The connected lower rectangle was the calculation of *"stop1"*, *"stop2"*, *"stop3"*, et cetera. The figure is, however, only showing one of the iterations; in reality the model continued with four identical calculations, but with changed logical expressions depending on the year. The upper rectangle connected to the input oval was the calculation of *"start1"*, *"start2"*, *"start3"*, et cetera. The logical expression that calculated the starting time ranges, is depicted above the rectangle, filling the attribute lists for the different time-steps as well as constraining the years to always be in the range of 2009-2014. The same logical expression was used for the calculation of the stopping time ranges except that it had a slightly different string with *"stop0"* as its input instead.

The output from the calculation of the different start times was then the input into the next calculation step. This calculation step had a logical expression that checked if the start time had a higher value than the stop time in each of the time-steps. This was true for those road works that started in one year and ended in another. If for instance, a road work started 20131201 and ended 20140130, then in the next time-step the start time would be 20141201 and the stop time 20090130, which is an illogical result. The expression that was used to prevent this is shown in *Figure 7.8*.

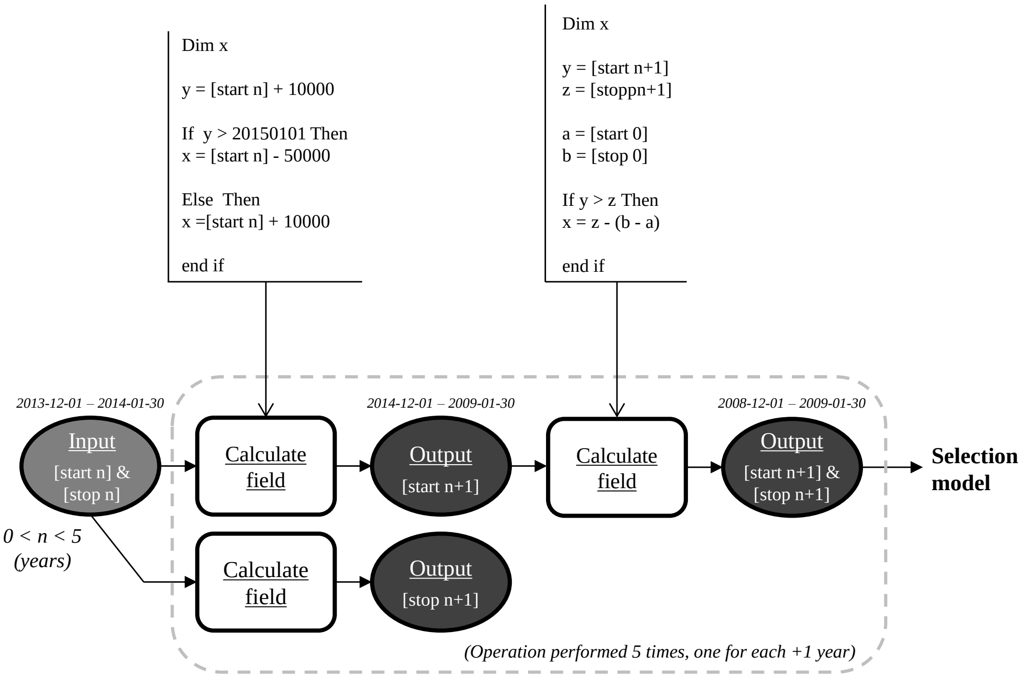


Figure 7.8 *A model in ArcGIS that calculates five time-steps to a road work occurring between 2013-12-01 to 2014-01-30.*

Selection model: The analysis then continued with a "*Select by attributes"* model where those accidents that fulfilled the selection criteria were exported to a text-file. This model is shown in *Figure 7.9* and it identified those accidents that have occurred within the road works' buffer zones – at their active period – along with those accidents that happened within the same buffer zone but within the time-steps.

The uppermost rectangle, with n = 0 years, exported those accidents that had occurred during the actual road work, which is why it was kept separate from the others. The rest of the rectangles used a similar logical expression but for their respective time-steps. The last step shown in the figure is the *'Merge'* between the selected accidents from the different time-steps.

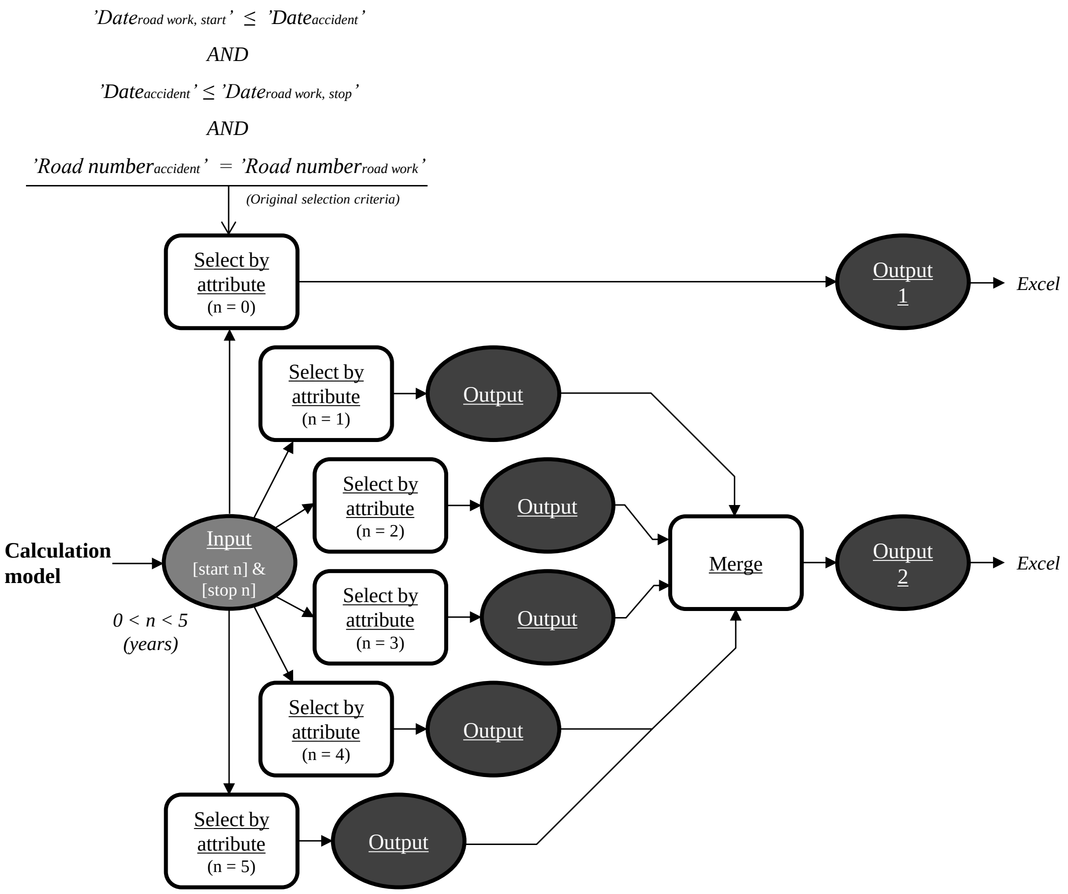


Figure 7.9 *A model in ArcGIS which identifies accidents that occurred within a road work's time-period and corresponding time-steps. This is done by the function 'Select by attributes' and is later on exported to Excel.*

How the *'Select by attributes'*-function works is illustrated theoretically in *Figure 7.10*. It can be seen that the accidents *a*, *b*, *c*, *e* and *f* all happened within any of the buffer zones, whereas accident *d* was not within any.

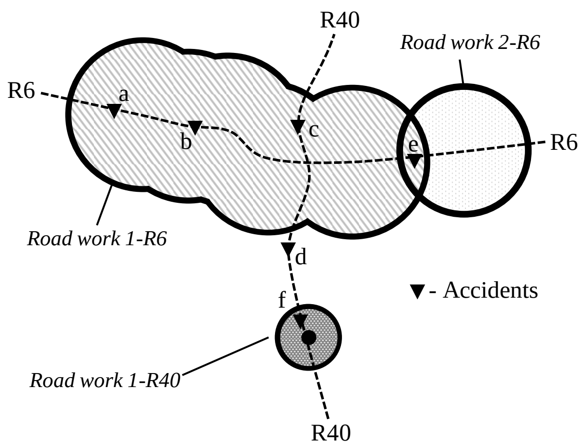


Figure 7.10 *An illustration of the theoretical principle behind the selection model. Bear in mind that this is not how it looked in ArcGIS, instead the accidents contained linked road work's spatial data in an attribute table, seen in Table 7.1.*

The selection process for the accidents in *Figure 7.10* will now be described in more detail, as well as showing the linked spatial data in *Table 7.1*. First, accident *d* was discarded since it did not happen within any of the buffer zones. Then accident *b* and the duplicated *e* were removed as it did not fulfil the first part of the selection criteria, more specifically *Equation 7.1*. Lastly, accident *c* did not fulfil the second part of the selection criteria – *Equation 7.2* – and was because of that also removed.

As noticed, accident e falls within two buffer zones, which could potentially mean that it would be counted twice in the later stage of the analysis. How this was solved is, however, addressed in the next section.

The result from the selection model can be summarised as being made up of two text-files, one containing accidents that happened at a road work and the second one containing accidents that happened at the same place as the road works, however, within the time-steps. These two text-files were then imported into Excel where the final part of the project analysis was to be completed.

Table 7.1 Accompanying attribute table for the accidents shown in Figure 7.10. The accidents that fulfil the selection criteria are highlighted in grey.

ID accident	**Date** accident	**Road no.** accident	**ID** road work	**Start** road work	**Stop** road work	**Road no.** road work
a	20120413	6	1	20120215	20120716	6
b	20110211	6	1	20120215	20120716	6
c	20120316	40	1	20120215	20120716	6
d	20100523	40	-	-	-	-
e	20120626	6	1	20120215	20120716	6
e	20120626	6	2	20090114	20090817	6
f	20110921	40	1	20110215	20111216	40

7.4 Analysis and extraction of data in Excel

- Tasks
  - Compilation of the three different quantification cases
  - Quantification of *Case (i)*
  - Quantification of *Case (ii)*
    - Elimination of potential error sources
  - Quantification of *Case (b)*

Compilation of the three different quantification cases: As stated in Section 7.3.3, the intent was to compare the accident frequency of road segments when there is a road work on-site, as opposed to the segment's normal state. Because of this, three different quantifications had to be performed. One was counting the amount of accidents that occurred within active road works (*i*), one was counting the accidents in those road works' subsequent time-steps (*ii*) and the last one was counting accidents that occurred within non-active road works' subsequent time-steps (*b*).

These three different cases can be summarised by the statements in the following bullet list. Illustrated in *Figure 7.11* is a type of Venn diagram showing the theoretical sets that relates to the statements.

1. Accidents **within** a road work's buffer zone.
    a. Accidents **within** an active road work's buffer zones or its time-step.
        i. Accidents *during* an active road work.
        ii. Accidents *during* its time-steps.
    b. Accidents **within** a non-active road work, but *during* its time-steps.
2. Accidents <u>outside</u> a road work's buffer zone.

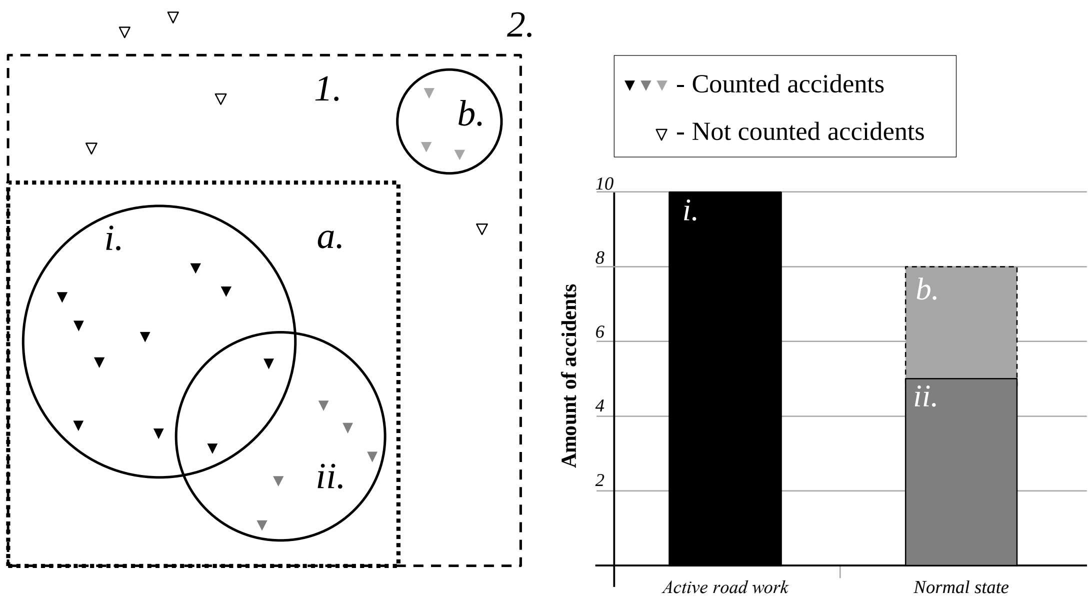


Figure 7.11 The three different quantity cases and how they relate to the compilation diagram.

As the accidents were counted, they were added to a diagram, similar to the one found to the right in the figure. Note that that the amount of accidents within the sets corresponds to the same amount for their respective bar.

Lastly, the percental difference between the two bars can then be stated to represent how many more accidents that occur on a road segment with an active road work, compared to when it has not any.

Quantification of Case (i): The first part of the analysis was to quantify the amount of accidents that had fulfilled the selection criteria for the road works' active time period, *Case (i)*. More specifically, these are the road works that was exported as '*Output 1*' from the selection model in *Figure 7.9*. These were therefore separated into the year when they had occurred and sorted accordingly. By using the built-in function '*Remove duplicates*' in Excel, all the accidents that fulfilled the criteria for multiple road work buffer zones were removed and a value was returned of how many unique accidents that remained. The value was then noted for each year and added to a diagram that will be presented in Chapter 8, Section 8.1.

Quantification of Case (ii): The next step was to quantify the amount of accidents that fulfilled the same selection criteria, but for the surrounding time-steps – i.e. '*Output 2'* in *Figure 7.9*. This was done in order to be able to make a comparison between the normal situation of a road segment and when there is a road work active on-site. For this purpose *Figure 7.12* is re-used in order to explain the background for the comparison. Following the figure, the active time period of 20090101 to 20090130 would have counted four accidents compared to the similar period a year later, which would have registered only one. The process was then done for each location in the years between 2009 to 2014, excluding the year when the road work was active.

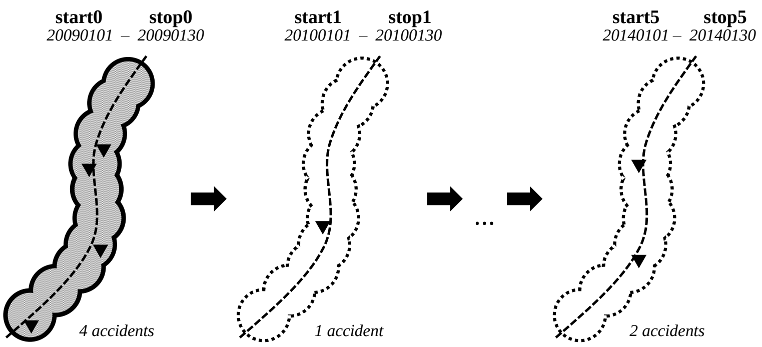


Figure 7.12 Accidents that are identified as having occurred within either an actual road work's time period or during its subsequent time-steps.

The point of this comparison is that it should in theory show a rise of the percentage of accidents at those road segments when there is an actual road work on-site, as compared to the normal situation.

In order to quantify this data, the IDs of the linked road works from the previous quantification *(i)* were used. A search was then conducted using a built-in Excel function named *'VLOOKUP'* among the time-step accidents. This function checks if there is a match between the two lists of linked road works IDs and returns either a '*True*' or '*False*' if there is a match or not. Those road works that were found to exist in both lists were then exported and quantified with the same process described in *Case (i).*

Elimination of potential error sources: When this was done, some modifications had to be done as there might be some cases where an accident is located within two road works, but is only selected for one based on the selection criteria in *Equation 7.1* and *Equation 7.2*. The problem arises if the accident later on fulfilled the criteria for the other road work's time-steps. This could result in that the same accident is counted twice; once as possibly having occurred during an active road work and once as having occurred during the normal situation. As the purpose is to examine the ratio of accidents that might have happened because of road works, this is not preferable and therefore the accidents that had these kinds of duplicates were removed from the time-step quantification.

These double-counted accidents are the ones that are within the intersection of *cases (i)* and *(ii)* in *Figure 7.11*.

Quantification of Case (b): For those road works that did not contain any accidents during their active periods, a quantification was made for their surrounding time-steps. This was done in order to count the total amount of accidents that had occurred within road segments where a road work had been performed 2009-2014.

Nonetheless, in order to be able to quantify these accidents, some operations had to be performed in Excel. Using the same '*VLOOKUP*' function as in *Case* (*ii*), those accidents with linked road works that was found to return a '*False*' statement was separated and quantified accordingly.

7.5 Comparison between free-text search and project model

- Tasks
  - Free-text search in Excel
  - Identifying accidents in ArcGIS

Free-text search in Excel: In order to find out whether or not the analysis made in ArcGIS was valid for the given problem, a free-text search based on the Swedish Transport Administration's method was performed. However, this was done for accidents that had occurred in 2009-2014, with the given accident types listed in Table 4.1 in Chapter 4, Section 4.2.3. For this analysis about 90 000 accidents were used.

Those accidents that were found by performing the free-text search can then be assumed to have been caused by road works, which means that if the ArcGIS analysis works, then those accidents should be identified in the model. Additionally, accidents were marked whether they were reported in by the police, hospitals or both.

Identifying accidents in ArcGIS: The accidents confirmed to have been caused by road works were then exported to ArcGIS and the project's model was then run with two different settings. The first setting had the same criteria as the regular model, with the initial 400 metres buffer zones as well as the variable ones. The second setting was slightly different; it added 1000 metres to all of the buffer zones in the model, which means that the smallest buffer zones now had a radius of 1400 metres. This last setting was done in order to provide material for a potential error source analysis.

When all this was complete, the number accidents that were identified in ArcGIS were noted for the results.

8 Results

This chapter describes the results from the two different analyses that have been performed both by free-text search in the accident database and in ArcGIS by the usage spatial and time data.

8.1 Accident identification in the ArcGIS-model

Those accidents identified to have occurred at road works within the ArcGIS-model were compared with those accidents happening within the time-steps, as explained in Chapter 7, Section 7.4. This quantification of accidents is illustrated in *Figure 8.1*. The black bars represent the number of accidents that have occurred within the road works' buffer zones and during their active time period per *Case (i)* in Chapter 7, Section 7.4. These are noted for the respective year that the accident occurred.

Meanwhile, the dark grey bars represent the mean amount of accidents quantified for *Case (ii)*, as described in the same chapter and section. As can be seen, they are put up against each of the years 2009 to 2014, where they can be stated to represent the mean amount of accidents that occurred in the active road work's subsequent time-steps. From *Case (b)*, described in the same chapter and section, the light grey bars were added with their mean values as well.

As can be seen in *Figure 8.1,* the different quantifications for each of the years do not have a consistent result. For three of the years (2010, 2011 and 2012), the accidents at road works outnumber the accidents on the normal state of the road (the time-steps) and in three years they do not (2009, 2013 and 2014).

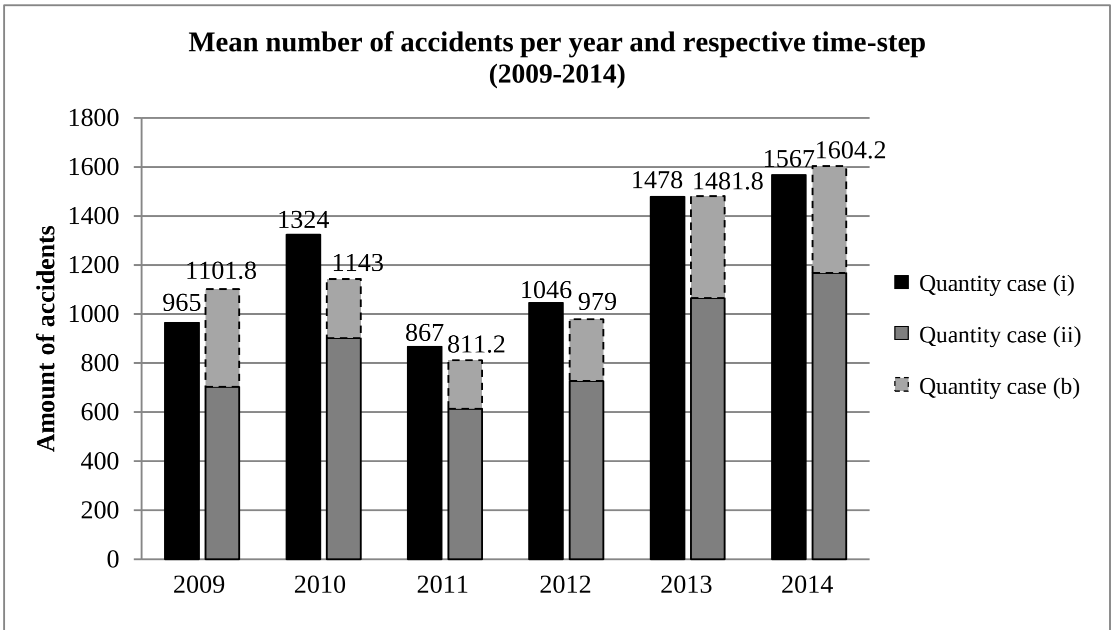


Figure 8.1 Results from the ArcGIS-model. The layout of the diagram is built upon the method described in Chapter 7, Section 7.4. The grey bars are a mean value of the accidents that happened for the each of the years' respective time-steps.

From the bars in the figure, the following table has been created in order to show amount of accidents that occurred within active road works, as well what can be expected during the road segment's normal state.

Table 8.1 *The amount of accidents that occurred on road segments during active road works and at their normal state for 2009-2014.*

Years	**Amount of accidents**	
	Active road work, *(i)*	**Normal state, *(ii)* + *(b)***
2009	965	1101.8
2010	1324	1143
2011	867	811.2
2012	1046	979
2013	1478	1481.8
2014	1567	1604.2
Total	7247	7120.2
Mean	1207.83	1186.7

By dividing the 7247 accidents occurring within active road works from the table with the initial amount of 90 054 accidents that was used as input data in ArcGIS, a percentage was calculated. This percentage shows that *potentially* 8 % of the accidents that occurred in 2009-2014 have been caused by road works per the project's limitations. All of the values can be found in *Appendix D*.

The mean value for each of the quantifications was then compiled into the diagram in *Figure 8.2*.

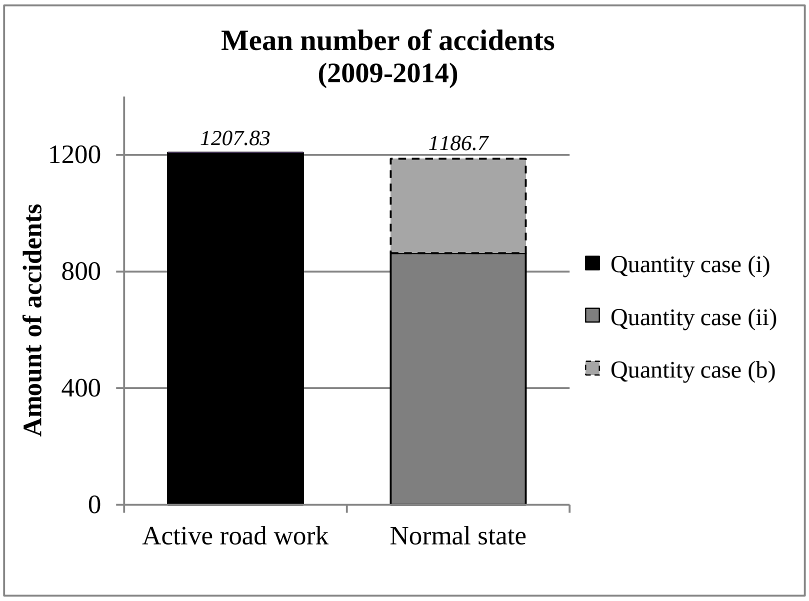


Figure 8.2 *Results from the ArcGIS-model. The two bars represent the mean number of accidents from the three different quantifications described in Chapter 7, Section 7.4.*

According to this result from the model, it was found that in the years 2009-2014, 1.74 % more accidents occurs on a road segment when a road work was active on-site as compared to its normal state.

8.1.1 Distribution of accident types

A distribution analysis – comparable to the two in Chapter 5, Section 5.1 and 5.2 respectively – showing the occurrence of different accident types was carried out as well based on the results from the model. In it can be seen that *single-car* and *rear-end accidents* together make up about 70 % of the total amount of accidents, excluding those involving vulnerable road users. In contrast, the other accident types are evenly distributed around 3–8 %.

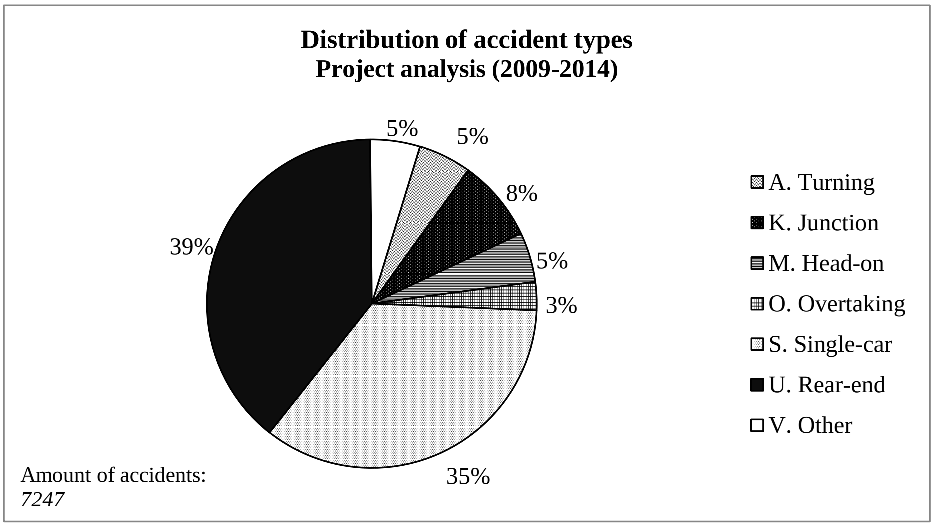


Figure 8.3 *Accident distribution on road segments where a road work took place for the years 2009–2014. Included types in the analysis are those listed from Table 4.1.*

8.1.2 Distribution of injured among age groups

Additionally, the amount of injured between the different age groups was compiled in a diagram, as shown in *Figure 8.4*. The age group which distinguished itself for all the ISS classes were 25–34 year olds. However, among those severely injured, persons in the ages 55–64 made up the largest proportion. Whereas 35–44 year olds was involved the most in accidents that had a fatal outcome.

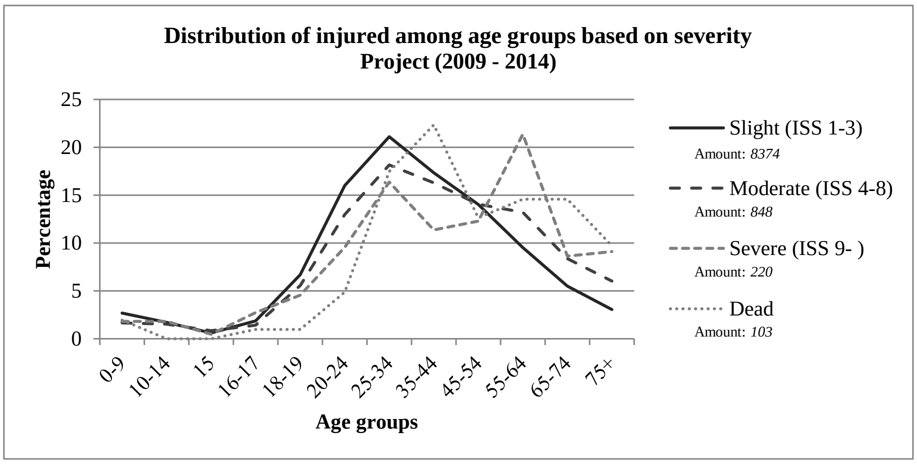


Figure 8.4 *Age group distribution of injured for the years 2009-2014 based on results from the model in ArcGIS.*

8.1 Free-text search for 2009-2014

The free-text search in STRADA was performed by the use of 26 different words that could have any correlation with road works. The results of the analysis can be seen in *Table 8.2*. The free-text search was only conducted with those indexes in STRADA given in *Table 4.1* in Chapter 4, Section 4.2.3. The total number of unique accidents that were taken out from STRADA was about 90 000 and the total number of unique accidents found by performing the free-text search was 751. This leads to the conclusion that about 0.83 % of traffic related accidents occur at road works with the given indexes listed in *Table 4.1* in Chapter 4, Section 4.2.3.

Note that the total amount of accidents in *Table 8.2* is not 751 as stated; this is because in the same accident both police and emergency hospitals could have reported two different people. This means that the accident will be counted twice, once by the police and once by the hospitals. This error applies to the *"Reported as both"*-column as well, where a third person could have been reported by both the police and the hospitals which means that the same accident will have been counted three times, which is incorrect when counting the number of unique accidents.

Table 8.2 The result of the free-text search separated by which authority that have reported the accident.

Years	Reported by the police	Reported by the hospitals	Reported by both
2009	62	40	7
2010	47	45	6
2011	58	59	5
2012	45	46	5
2013	46	47	6
2014	50	55	6
Total	308	292	35

8.1.1 Validation of the ArcGIS model

Those accidents identified to have happened at road works by performing the free-text search were, as stated earlier, imported into the ArcGIS model in order to evaluate the precision of the model. The model was run with two different settings, one with initial buffer zones of 400 metres and the other with an increase of buffer width by 1000 metres. These 751 unique accidents were run through the model and the results are illustrated in *Table 8.3*. Since the results vary depending on the settings of the model, an additional logical expression was used. This logical expression was performed in both cases and ignored the requirement that the accidents and the road works had to be located at the same road. As can be seen in *Table 8.3* all the accidents were not identified within the ArcGIS model. Only 285 accidents were identified from the initial 751 when using the full selection criteria explained in Chapter 7, Section 7.3.1. This number is reached when the settings were set to a minimum of 400 metres wide buffer zones. When ignoring that the accidents and the road works had to be located at the same road, 296 accidents were identified instead. For the second setting where minimum buffer zones were set to 1400 metres 377 accidents were identified when using the initial criteria, which is less than half of the total 751 accidents. When ignoring the road names with this setting a total of 387 accidents were identified.

Table 8.3 *The results from the comparison between the free-text search and the ArcGIS model. 751 accidents were imported into ArcGIS and depending on the settings and logical expression different results were achieved.*

Buffer zones	Input	Selection criteria	Output	Percentage of input [%]
Original model	751	Date + road number	285	38
		Date	296	39.4
+1000 metres	751	Date + road number	377	50.2
		Date	387	51.5

8.1.2 Distribution of accident types

Another distribution chart, seen in *Figure 8.5*, was created using the assumed road work accidents that was acquired from the free-text search in Excel. In the figure, it is shown that *rear-end* and *single-car accidents* again make up the majority of all the accidents that occurred. Furthermore, the five other accident types – *turning, junction, head-on, overtaking* and *other* – are distributed around 1–8 %.

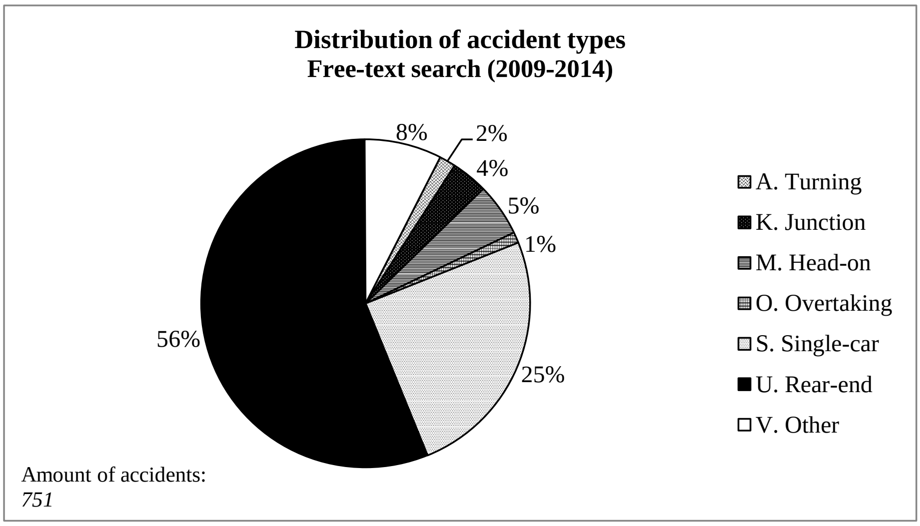


Figure 8.5 *Accident distribution on road segments where a road work took place for the years 2009–2014. Included types in the analysis are those listed from Table 4.1.*

8.1.3 Distribution of injured among age groups

Lastly, a distribution chart was created for the amount of injured among the different age groups, as shown in *Figure 8.6*. Note that the amount of injured is adding up to more than 751, this is due to the fact that more than one person can be involved in one accident alone. Noticeable this time is that among those who were severely injured, the 55–64 year olds make up nearly 40 %.

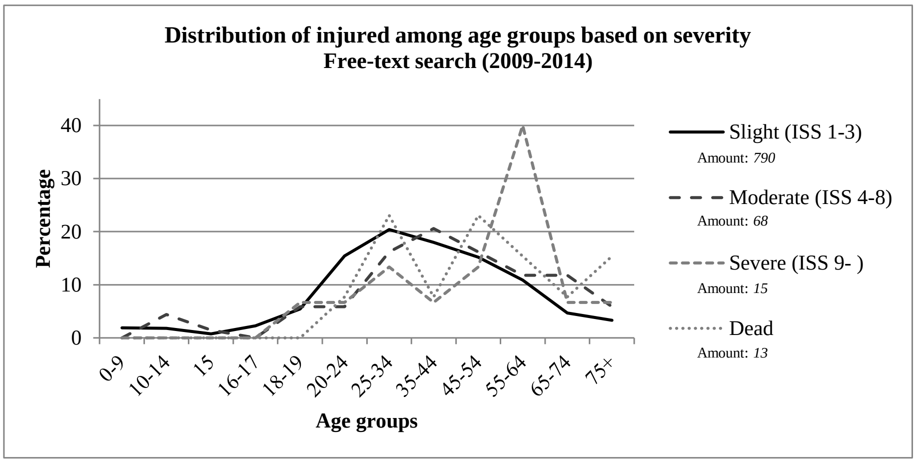


Figure 8.6 *Age group distribution of injured for the years 2009–2014 based on results from the free-text search in Excel.*

8.2 Comparison between the results and previous studies

For the final part of this chapter a comparison is presented for the results that have been acquired from the different analyses throughout the project. In this context, *Government owned roads* refer to how the distribution of accidents and injured is at Swedish roads' normal state.

8.2.1 Accident types

Throughout the thesis, the distribution of accident types has mostly looked the same, a majority of them have consisted of *rear-end collisions* or involving a *single-car*, as can be seen in *Table 8.4*. However, when examining the percental relationship between the two accident types, it can be seen that it is nearly inverted when comparing *Government owned roads* with the three *Free-text search* analyses. The distribution of the accident types from the *Project* is on the other hand somewhere in-between the other two distributions.

Table 8.4 *A compilation of the distribution of accident types among the three different analyses. For 'Government owned roads' and 'Project', the analysed years are 2009–2014. Included types in the analysis are those listed from Table 4.1.*

Accident type	**Government owned roads**	**Project**	**Free-text search**		
			2003-2007	**2003-2011**	**2009-2014**
Rear-end	28	39	50	48	56
Single-car	41	35	25	29	25
Miscellaneous	31	26	25	23	19
Data amount	90 054	7247	668	1682	751

8.2.2 Age groups

When comparing the different distributions among the age groups, it can be seen in *Figure 8.7* that there is no stable correlation between them. However, generally the number of injured increases for people of the ages 18–19 year olds, to reach a peak at around 45-54 year olds and then decrease again.

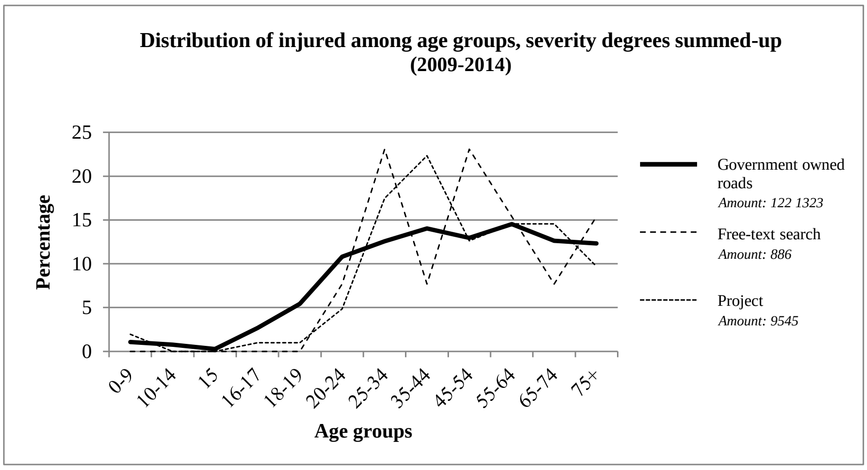


Figure 8.7 Age group distribution of injured for the years 2009–2014 based on the three different distributions that has been performed. Note that all severity classes is added together for this figure.

9 Discussion

As results have been acquired, some general discussions will now follow in order to tie the project back to the aims and research questions stated in the beginning of this thesis. Structurally, the chapter will first evaluate the model in ArcGIS and then the free-text search method. A discussion will then follow surrounding the results from the validation of the ArcGIS model, by the usage of the free-text search results.

9.1 ArcGIS model

The modelling consisted of three main parts, the first was the compilation of input data, next the general structure of the model and lastly the results that was acquired from it. This is also how the following discussion will be arranged.

9.1.1 Input data into the model

The input data was, as stated earlier, submitted into STRADA and FIFA between the years 2009-2014. This input did however contain errors from both sources. This was discovered when the accidents were imported into ArcGIS and some of them were located in Norway, the Atlantic Ocean, Finland and the Baltic Sea. These accidents show that errors do occur when they are submitted into STRADA, which could be a reason to why the results turned out the way they did. It was discovered that the road work data contained fewer spatial errors than the accident data did, however, some of the road works still had coordinates that were zero on both x and y. A reason that the road works contained fewer errors than the accidents can be due to the fact that they are submitted from two dissimilar sources and that the data are reported onto platforms with two very different focuses. Spatial data is for instance very important when constructing a new road, while accidents have other factors that are equally important.

When looking at the time-span of some of the road works that was used in the model some errors were found. One road work was for instance reported to have started in 1899, which is not likely to have happened and because of this the results could have been affected. When the submitted road works have these kinds of errors there is a possibility that other road works also are reported wrong, which could be a reason why the model is not identifying all the accidents that have been confirmed to have happened because of a road work.

Roads that have been built could have been relocated between the years 2009-2014. This means that if a road was relocated in 2014 there were no accidents that could have happened during its normal state since they are reported to have happened at the old location of the road. This does affect the results of this project since it aims to make a comparison between number of accidents that happen on a road during its normal state and during a road work. Also, new roads could have been built which means that the same situation arises.

Those road work types that were not part of either the ArcGIS-model or the free-text search were those connected to winter road maintenance such as ploughing and salt spreading. These were not included in the analysis since they did not contain any set spatial data for a specific location, which means that they are not reported into

STRADA and in extension was not included in the received data from the Swedish Transport Administration. These types of road works were however included in the reports written by the Swedish Transport Administration in 2008 and 2012, which could be a reason to why the results differ between the project and the results from those reports.

A reason to why no accidents regarding vulnerable road users were used in any of the two analyses is because these types of accidents more often occur at municipal owned roads rather than on governmental owned which this project regards. This is why they were removed from the input data to this project and the values in the reports made in 2008 and 2012 were adjusted. The reason why only government owned roads were chosen to be a part of the analysis, was because this data was easy to acquire since it was compiled by only one authority. If the same level of information would have been gathered for the municipal road network then it would have meant a more thorough analysis, but taken a longer time to collect. Information regarding the performed road works at municipal owned roads would also have needed to be collected, which would have made the project more time-consuming and not applicable because of the limited time-period of this project.

9.1.2 Design of the model

The most time-consuming step of the project was to set up the model in ArcGIS, which took time because of the many actions performed. Because of the large amount of data that were being handled, many of the actions took time and each function created additional data that made the program work even slower. In retrospect, this could have been solved by initially only analysing a smaller amount of data. This could have saved time since every action within ArcGIS would have worked faster compared to handling the entirety of data. When the needed approach was proven to work, it could then have been added into the *'Model-builder'*, along with the rest of the data. Of course, when handling these large amounts of data automatically there is a risk that some of it may be neglected or not compatible with the actions performed, which could affect the results anyway.

A source of error with the ArcGIS-model is the buffer zones. They are created as connected circles along their nodes, which implies that even the set buffer zones contain different lengths of road stretches. This can be explained with same size circles drawn around two different roads, where one of the roads is a straight motorway and the other is a small highway with many curves. Dependent of the curvature of the roads the circle will cover a larger or smaller length of the two roads, which is the case in this project. An alternative would have been to use some function that instead of calculating a linear distance from each road work, uses the available road network as a constraint. Due to the large amount of data however, this may not be realistic, as it would have involved even more factors to take into account.

In one of the variable buffer zones, the influence radius of a road work was calculated to be more than 40 kilometres long which in theory should not be likely. It has been reported in the media that queues of tens of kilometres have formed due to road works, these do however only form on roads with high traffic flows, such as Stockholm's ring road. Those road works that had large buffer zones were also found to be located at road segments with a high traffic flow. It was difficult to evaluate how large buffer zones that should have been the *correct* influence radius since the road network was very extensive. Some of the buffer zones should maybe have been larger and some of them smaller, but on the other hand to have precise calculations on every buffer zone throughout the entire road network of Sweden would have been very time-consuming.

As the calculations for the theoretical queue build-up needed the number of traffic lanes at the road work's location, it would have preferable if those were included as well in the input data received for the road network. The lanes were now calculated by using model values that may not represent the reality for each of the road segments.

The *'Select by attribute'*-function had a criterion which stated that both the road work and the accident should be located on the same road in order for the accident to be identified as to have happened at the road work. This criterion is true for minor road works that does not affect the traffic flow in an extensive manner. However, for large road works performed in urban areas – with many access roads – this criterion could be incorrect. If long queues would form, this could mean that they are building up backwards and start to form even at intersections and their connecting access roads, which could cause accidents on roads other than the ones that have work performed on them.

9.1.3 Results and performance of the model

It can be seen in the results that there is a tendency that 1.74 % more accidents occur at road works than during the normal state of the road. This is, however, not a conclusive answer since it is difficult to know for certain that the accidents were actually caused by road works. This result does also lie within a margin of error, which means that the value could be a coincidence. Road works are believed to have a more frequent occurrence of accidents than the normal state of the road, which is what this value shows. If the percentage had been reversed, meaning that fewer accidents occur on road segments with active road works, then the performance of the model would have been more questionable.

If the results from the model are believed to be correct, then a value of 8 % of the total amount of accidents are believed to be caused by road works with the given conditions limiting this project. In the study made by the Swedish Transport Administration in 2008, less than 1 % of the accidents were related to road works. That report did, however, only take the police reports into account as well as analysing the entire road network, not only the government owned roads. The difference between the two results are believed to be that the ArcGIS-model included accidents that did not occur due to the road work along with that it also identified accidents that have not been reported as a road work related accident, but still was. The difference in results could also be because the input data on both the road networks and the accident data varied in quality.

The value of 8 % should probably be lower. This is assumed since accidents could be identified in the model even though they were not caused by road works. It is also difficult to confirm distinct links between the accidents and the road works by investigating the descriptions of them both, since these are often sparse. Even if the accidents occurred within the buffer zones and during the road work's active period, it would still be difficult to find any link by reading the accident description.

9.2 Results from the validation of the ArcGIS model

As stated earlier, only 0.83 % of the accidents were found to have occurred at road works by performing a free-text search in STRADA and with the given conditions for this specific project. The reason why the percentage is that low could be because no regards were taken to either vulnerable road user or to those accidents occurring at municipal owned roads. There is no telling how the value would have changed if these were taken into account as well.

The accidents that were identified to have occurred at road works by using free-text search were imported and run through the ArcGIS-model. This was done in order to investigate if the model would identify them. It did partially, as about 38 % of the accidents were identified. This is why the selection criterion, which stated that both the accidents and the road works had to be placed at the same road, was ignored to investigate a potential error source. When running the model again, only a few more accidents were identified. A second set-up was therefore staged, which expanded the road works' buffers with 1 000 metres each. This time, it did find more accidents but not sufficiently enough, as now only about 50 % of them were identified within the model. When running this set-up without the road number selection criterion, only a few more accidents were identified.

So, why is the model unable to find the confirmed 751 accidents and why does it only identify 38 % with the original selection criteria and 50 % with the expanded buffer zones? It is believed that there might be something wrong with the input data and not the model in itself, as the simplifications that have been performed throughout has been taken into consideration with the generous buffer zones.

When investigating the number of found accidents with the expanded setting in ArcGIS, just over half of them are identified. Some of the accidents that are not identified are located in rural areas where the traffic flow often is low, which means that they should be located within 1 400 metres from the start of the road works. 1 400 metres are believed to be a relatively large theoretical queue length considering that not many road works would cause queues this long in a rural environment, due to smaller traffic flows. Even if a road work would create these long queues they are located at a segment that have high traffic flow, which means that they already have a buffer size dependent on their traffic flow and as such are larger than 1 400 metres.

The reason to why the model does not identify all of the 751 accidents could be that the road works' locations and dates were reported wrong. However, a more believed error is that the accidents were reported incorrectly, either spatially or in time. It was investigated if they had the same reported road number in STRADA as the one that it was later given when performing a '*Spatial Join*' together with the road network. It was then found that some of the accidents did not have the same road number that it was reported to have occurred at. In those cases, where the names did not match, the local name was given instead of the official name used by the Swedish Transport Administration.

In some other cases they were located in an intersection and duplicates were created where only the accidents matching the road works' criteria were identified. The only error that then remains is the time-stamp of the road-works and the accidents. These are difficult to investigate if they are correct or not.

Another error that also could be the reason why the model is unable to identify the accidents could be that the road works were never reported into FIFA, that the data obtained is incomplete. According to Jimmie Sjöberg, at the Swedish Transport Administration, it does happen that some road works are not reported into FIFA. This is not likely since it would mean that a very large part of the road works would have to be reported into FIFA incorrectly.

9.3 Comparison between the results and previous studies

When reviewing the two comparative studies that were performed regarding accident type frequency and injury distribution among the age groups, it can be seen that a correlation may appear for the former but not for the latter.

Regarding the accident types, it was noted in the results that the inverse relationship can be seen for the *single-car* and *rear-end accidents*, between the government owned roads and the three free-text search analyses. As the free-text search ones reveal, the latter accident type is the most frequent one for those studies. There is reason to believe that this is due to the queue build-up that often occurs at road works. During normal conditions on a road segment – e.g. when no road work is being performed – the traffic flow on the road is usually below the calculated capacity, meaning that no queues are created. This can also be seen in the analysis done for government owned roads, where other accident types makes up a larger proportion, especially the ones which involves a single car. Straddling in between is the distribution of accident types acquired from the project's GIS model and as previously mentioned, this model may include accidents that was not caused by road works as well as lacking those which are.

Following that, it was noted that there is not really any correlation of the injury distribution among the age groups. This is thought to be caused by the varied amounts of data that was used for each of the distributions, where the *government owned roads* is nearly ten times the amount of the *project's* data, which in turn is also ten times the data of the *free-text search*. It is therefore believed, that the first distribution is the most stable one.

Furthermore, it is shown that the amounts of injured increases greatly when looking at the adult age groups. This is also natural, since people below 18 years old are not allowed to drive in Sweden and spends thus less time in traffic with high flows than those from the other age groups.

9.4 Safety work in the road construction process

After what has now been discussed the question arises in how the safety aspect could be even further integrated within the current road construction process.

Should an accident occur during a larger construction work, then the contractor along with its appointed BAS-U covers the documentation of the accident internally within their organisation today. As was mentioned in the earlier sections of this chapter, there is a lack of preciseness in the accident reports from the police and the hospitals regarding the course of event and how the road work might have contributed to the accident. Therefore, there should be a priority to add a field on the accident reports that the police and/or health care fill in today, which states that a road work could have caused the accident.

When the Swedish Transport Agency later compiles the accident reports and finds that the accident potentially could have been caused by a road work, they contact the responsible BAS-U for the project. The BAS-U would then be required to contribute information about the road work in question, which in extension could later be used as a way to evaluate how a certain TA-plan might be prone to cause more accidents. In the end, this would create a body of knowledge that would serve as recommendations for how both the BAS-P and BAS-U could develop the safety work within the road construction process.

10 Final Remarks

The objective of this master's thesis has been to examine if traffic accidents at road works can be located and quantified by the use of a GIS-software, namely ArcGIS, between the years 2009-2014. Furthermore, a free-text search was conducted in a traffic accident database (STRADA) to be able identify those accidents that are confirmed to have happened at road works. These accidents were then used as an input in the ArcGIS model in order to investigate the validity of the model.

Additionally, this project's results have been compared with those established from previous studies in order to further investigate if the ArcGIS model is valid for the given problem.

10.1 Conclusion

It was found in this master's thesis that the use of ArcGIS could be a way to locate and quantify those accidents that have occurred at road works. Since the model gives a rough estimate of the number of accidents at road works, mainly because of the large amounts of data, the results have to be considered as estimations and tendencies of how the *true* results actually are. Even though accidents can be linked both spatially and in time to road works, it is not certain that the road work was the cause of them.

It was found that about 1.74 % more accidents occur at road segments with an active road work on-site, compared to the natural state of the road. Additionally, 8 % of the total amount of accidents is potentially identified to have occurred because of the road works within the given conditions of this project. This value is believed to be lower in real-life since the model identifies every accident within a certain distance from the road works, along with the criterion that it occurred within the time span that the road work was active. This means that accidents are identified that have not been caused by the road works as well. Even when an accident is identified in the model, it is difficult to find out what actually caused the accident by also reading the accident description.

Only 38 % of the accidents that were confirmed to have occurred at road works in the free-text search were identified in the ArcGIS model. By expanding the search radius with 1 000 metres around the road works about 50 % were identified instead. The reason why the model does not identify all of the accidents is believed to be because either that the input data is incorrect or that accidents may have occurred further from road works than predicted. When reports are submitted mistakes or errors could be done, which applies to both the entrepreneurs – as they report planned road works – and the police and health care when they are reporting traffic accidents.

A way to provide a more consistent reporting about road work related accidents could be to add a road work field in the accident reports for both the police's and health care's form. If it is filled in, then the responsible BAS-U should be required to provide more extensive information about the road work in question.

Regarding accident types, rear-end collisions are over-represented for accidents that have occurred at road works when studying the results from the free-text search. This is believed to be caused by the build-up of queues that often arises due to change of traffic capacity that the road works cause. The ArcGIS model also identifies this type of accident as the most frequent, but not as much as the free-text search method where all accidents are confirmed to have happened because of a road work. Again, this is assumed to be like this because of the model's tendency to include non-relevant accidents as well.

In conclusion, is the usage of GIS a reliable method in providing proper statistics regarding traffic accidents at road works? The answer is yes, if the analysis is performed on a more limited area, either within counties or municipalities. This, together with a more detailed description of the accidents and road works' effect on the traffic in the reports would make it possible to perform a more accurate model of how many accidents that actually occur due to road works.

By presenting to the construction business that there are tendencies where up to 8 % of the total amount of traffic related accidents are occurring due to road works, hopes are that the safety aspect should have a much larger influence during the planning and design of future road works.

10.2 Further studies

Even though the results were found to not be decisive enough to draw any final conclusions, there exists a number of ways to improve the model in order to be able to provide a more stable result. Firstly, a recommendation is to minimise the research area as this allows for a more detailed analysis to take place. A way to achieve this would be to choose a small sample which could be said to represent Sweden as a whole. This allows for less time to be spent on compiling and filtrating input data and more focus can then be given to evaluating the quality of it instead. An additional effect is that the added time can be used to acquire data from the municipalities as well, contributing to a more comprehensive result.

By also including those accidents that are occurring at municipal owned roads into the model, a wider representation of the problems at road works can be investigated. When doing this, the accidents involving vulnerable road users can be a part of the model with the intention on improving the road works' safety measures. However, this can only be performed in a more limited area since it would have been too much data to handle if the entirety of Sweden would have been included in the model.

Lastly, as the linked data of the accidents and road works from this analysis has been saved and stored, a number of further studies could be performed. An important one would be to look into the confirmed road works accidents which were not identified in the GIS model. As stated in the previous section, there is reason to believe that there might exist some sources of error regarding the way the data is reported into FIFA and STRADA respectively. It would therefore be interesting to do an in-depth analysis of that data in order to find those potential error sources.

11 References

Adriazola, C., Duduta, N. & Hidalgo, D., 2013. *Saving Lives with Sustainable Transport: Traffic safety impacts of sustainable transport policies.*, s.l.: World Recources Institute, EMBARQ.

Arbetsmiljödepartementet, n.d. *SFS 1977:1160.* Stockholm: s.n.

Arbetsmiljöverket, 2013. *Arbetsskador 2013,* Stockholm: Arbetsmiljöverket.

Arbetsmiljöverket, 2014. *Arbetsorsakade besvär 2014,* Stockholm: Arbetsmiljöverket (2014:4).

Arbetsmiljöverket, 2015. *Av.se.* [Online]
Available at: http://www.av.se/lagochratt/aml/
[Accessed 9 February 2015].

Esri, 2012a. *ArcGIS Help 10.1.* [Online]
Available at:
http://resources.arcgis.com/en/help/main/10.1/index.html#//00s50000002q000000
[Accessed 15 April 2015].

Esri, 2012b. *ArcGIS Help 10.1.* [Online]
Available at:
http://resources.arcgis.com/en/help/main/10.1/index.html#//00s500000021000000
[Accessed 15 April 2015].

Esri, 2012c. *ArcGIS Help 10.1.* [Online]
Available at:
http://resources.arcgis.com/en/help/main/10.1/index.html#//002w00000001000000
[Accessed 6 April 2015].

Esri, 2013a. *ArcGIS Help 10.1.* [Online]
Available at:
http://resources.arcgis.com/en/help/main/10.1/index.html#//00170000005n000000
[Accessed 15 April 2015].

Esri, 2013b. *ArcGIS Help 10.1.* [Online]
Available at:
http://resources.arcgis.com/en/help/main/10.1/index.html#//001700000055000000
[Accessed 15 April 2015].

Esri, 2013c. *ArcGIS Help 10.1.* [Online]
Available at:
http://resources.arcgis.com/en/help/main/10.1/index.html#//000800000019000000
[Accessed 15 April 2015].

Esri, 2014. *ArcGIS Help 10.1.* [Online]
Available at:
http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00080000000q0000
00
[Accessed 15 April 2015].

Foner, E. & Garraty, J. A., 1991. *The Reader's Companion to American History.* Boston: Houghton Mifflin Harcourt Publishing Company.

Forward, S. & Samuelsson, P., 2007. *STRADA - blev det som det var tänkt,* Linköping: VTI - Statens väg- och transportforskningsinstitut.

Fransson, M., 2008. *Byggnadsarbetaren.* [Online]
Available at: http://www.byggnadsarbetaren.se/2008/10/byggnadsarbetare-radda-att-skadas-pa-jobbet/
[Accessed 27 January 2015].

Friberg, F., 2007. *Lugnare arbete på väg - Internationell utblick för säkrare vägarbeten med inriktning mot olycksanalys och nya metoder för trafiksäkra hastigheter,* Borlänge: Vägverket.

Hasselblom, S., 2015. *How to calculate road capacities* [Interview] (6 March 2015).

Johnson, F., 2015. *Discussions during mentor meetings* [Interview] (- - 2015).

Joint Research Center, 2010. *Travel time to major cities: A global map of Accessibility.* [Online]
Available at: http://bioval.jrc.ec.europa.eu/products/gam/sources.htm
[Accessed 4 May 2015].

Liikennevirasto, 2010. *Liikennevirasto.* [Online]
Available at:
http://www2.liikennevirasto.fi/julkaisut/pdf3/esite_2010_road_planning_process.pdf
[Accessed 25 February 2015].

Liljegren, E., 2008. *Trafikolyckor vid vägarbeten 2003-2007,* Borlänge: Vägverket.

Liljegren, E., 2011. *Plötsligt var det ett vägarbete!: En studie av trafikolyckor vid vägarbeten 2003-2009 med speciellt fokus på upphinnandeolyckor,* Borlänge: Trafikverket.

Liljegren, E., 2012. *Skyddad för livet?,* Borlänge: Trafikverket.

Liljegren, E., 2014. [Online]
Available at:
http://tra2014.traconference.eu/papers/pdfs/TRA2014_Fpaper_17924.pdf
[Accessed 10 February 2015].

Mattson K, J. C., 2014. *Trafikskadade rapporterade via STRADA sjukvård,* Göteborg: SWECO.

OECD, 2013. *Long Run Trends in Car Use,* Paris: OECDpublishing.

QGIS, 2015. *A Gentle Introduction to GIS.* [Online]
Available at: http://docs.qgis.org/2.6/en/docs/gentle_gis_introduction/
[Accessed 14 April 2015].

Sandström, J., 2015. *Practicalities of a BAS-P* [Interview] (6 February 2015).

Sjöö, B. & Ungerbäck, A.-C., 2007. *Nytt nationellt informationssytem för skador och olyckor inom hela vägtransportsystemet,* Borlänge: Vägverket.

SOU 2014:24, 2014. *Olycksregister och djupstudier på transportområdet: Delbetänkande av Utredningen om översyn av bestämmelserna om undersökning av olyckor,* Stockholm: Elanders Sverige AB.

Svensk Byggtjänst, 2015. *Studenthandledning Byggprocessen,* Stockholm: Svensk Byggtjänst.

Sveriges Kommuner och Landsting, 2014. *Arbete på väg,* Stockholm: LTAB.

Söderström, L., 2004. *Väg- och trafikmiljönsbetydelse för whiplashskador vid påkörning bakifrån,* Lund: Lunds Tekniska Högskola.

The Economist, 2012. Seeing the back of the car. *The Economist*, 22 September.

Trafikverket, 2012. *Vision, verksamhetsidé och värderingar.* [Online]
Available at: http://www.trafikverket.se/Om-Trafikverket/Trafikverket/Vision--verksamhetside
[Accessed 28 January 2015].

Trafikverket, 2013. *Från planering till byggande.* [Online]
Available at: http://www.trafikverket.se/Privat/Vagar-och-jarnvagar/Fran-planering-till-byggande/
[Accessed 10 April 2015].

Trafikverket, 2014a. *För din och trafikanternas säkerhet,* Borlänge: Trafikverket.

Trafikverket, 2014a. *Strategic Choice of Measures,* Borlänge: Trafikverket.

Trafikverket, 2014b. *FIFA Handbok,* Borlänge: Trafikverket.

Trafikverket, 2014b. *Åtgärdsvalsstudie.* [Online]
Available at: http://www.trafikverket.se/atgardsval
[Accessed 10 April 2015].

Trafikverket, 2014c. *Bygg om eller bygg nytt - Kapitel 4 Tillgänglighet,* Borlänge: Trafikverket.

Trafikverket, 2014. *Trafikverket.* [Online]
Available at:
http://www.trafikverket.se/contentassets/1184eb1d113a413fbda74557c7743f7c/fifa_2014-4_handbok_sokande_handlaggare.pdf
[Accessed 5 February 2015].

Trafikverket, 2015. *Sveriges vägnät.* [Online]
Available at: http://www.trafikverket.se/Privat/Vagar-och-jarnvagar/Sveriges-vagnat
[Accessed 10 April 2015].

Transportstyrelsen, 2013. *Årsstatistik.* [Online]
Available at: https://www.transportstyrelsen.se/sv/vagtrafik/statistik-och-register/Vag/Olycksstatistik/Polisrapporterad-statistik/Nationell-statistik/Arsstatistik
[Accessed 10 February 2015].

Transportstyrelsen, 2014. *Ny STRADA uttagswebb lanseras med ökat fokus på sjukvårdens skadedata.* Tylösand, Transportstyrelsen.

Transportstyrelsen, 2014. *STRADA – informationssystem för olyckor och skador i trafiken.* [Online]
Available at: https://www.transportstyrelsen.se/sv/Vag/STRADA-informationssystem-for-olyckor-skador/
[Accessed 27 January 2015].

Transportstyrelsen, 2015a. *Olycksstatistik.* [Online]
Available at: https://www.transportstyrelsen.se/sv/vagtrafik/statistik-och-register/Vag/Olycksstatistik/
[Accessed 28 January 2015].

Transportstyrelsen, 2015b. *Sjukhusrapporterad statistik.* [Online]
Available at: https://www.transportstyrelsen.se/sv/vagtrafik/statistik-och-register/Vag/Olycksstatistik/Sjukhusrappoterad-statistik/
[Accessed 4 March 2015].

VTI, 2013. *Cyklisters singelolyckor analys av olycks- och skadedata samt djupintervjuer,* Linköping: VTI.

Vägverket, 2007. *Trafikverket.* [Online]
Available at:
http://www.trafikverket.se/contentassets/8f406c2d76db4efabccc4f8d55256703/en_vag_blir_till.pdf
[Accessed 25 February 2015].

Appendix A Extract from Work Environment Act

"*Chapter 3, section 6*

The person who orders execution of building or construction work shall

1. during each phase of the planning and projecting see to that work environment viewpoints are considered when it is applicable to the building phase as well as future usage,
2. appoint a suitable building work environment co-ordinator for the planning and projecting of work with the information that is stated in Section 7 a, and
3. appoint a suitable building work environment co-ordinator for the execution of work with the information which is stated in Section 7 b and Section 7 f.

The person who orders the execution of a building or construction work can appoint him/herself or someone else as building work environment co-ordinator. If someone else has been appointed the person who orders the execution of a building or construction is however not released from responsibility for such work information as is stated in the first paragraph of 2 or 3.

Chapter 3, section 7

During each phase of the planning and projecting of a building or construction work shall architects, design engineers and others who assist, within the framework of their assignments, see to that work environment viewpoints are heeded when it is applicable to the building phase as well as the future usage.

Section 7a

The building work environment co-ordinator who in accordance with Section 6 first paragraph 2 has been appointed for the planning and projecting of a building or construction work, shall co-ordinate the application of relevant work environment rules which during each phase of the planning and projecting shall be followed on the question of the building phase as well as the usage of the building or construction. It is especially applicable when questions on the planning of building operations which shall be performed at the same time or after each other are concluded and when the time expenditure for such a building phase is calculated. The building work environment co-ordinator shall observe such a work environment plan and other documentation as stated in Chapter 4 Section 8 second paragraph 1–3.

Section 7b

The building work environment co-ordinator who in accordance with Section 6 first paragraph 3 has been appointed for the execution of building or construction work, shall see to co-ordinating the work with preventing risks for ill-health and accidents on the worksite and implement the information which appears in Section 7 e 2–5, and

1. co-ordinate the application of relevant work environment rules when technical or organisational questions about the planning of building operations which shall be executed at the same time or after each other are concluded and when the time expenditure for such a phase is calculated,

2. co-ordinate the application of relevant work environment rules in order to ensure that he/she or they who operate activities at the worksite apply these rules in a systematic way as well as follow a work environment plan,
3. co-ordinate measures to supervise that the building or construction work with respect to the work environment is executed in a correct manner,
4. take necessary measures to ensure that only authorised persons are furnished access to the worksite, and
5. organise the information which in accordance with Section 7 g reaches those operate activities at a common worksite.

Section 7c

That which is stated in Section 6 and in regulations which have been notified in virtue of Chapter 4 Section 8 if he/she who orders the execution of building or construction work shall instead apply to one of these retained employees to the extent

1. the employee has received in the assignment to independently be responsible for planning and projecting or the work's execution, and
2. it has been agreed to in writing that the information which is stated in Section 6 and in the connection the regulations shall rest with the employee upon the execution of the commission.

If the Consumer Services Act (1985:716) shall be applied upon commission in accordance with first paragraph 1, it shall as is stated in Section 6 and in regulations which have been notified by virtue of Chapter 4 Section 8 apply to the employee upon the work's execution. It is therefore not applicable if it has been agreed to in writing that it as is stated in Section 6 and in connection with the regulations instead shall apply to the person who orders the execution of the building or construction work.

Section 7d

If a permanent site of operations is a common worksite for several activities, the person who is in charge of the worksite is responsible for the co-ordination of work environment questions. If a ship is a common worksite for several activities the ship owner is responsible for the co-ordination. If a ship has been taken in at a shipyard in Sweden it is however the person responsible for shipyard activities who is responsible for the co-ordination. The responsibility for co-ordination of protective measures which are brought about by that a ship is being loaded or unloaded in a Swedish harbour rests with the employer who has the responsibility for this work.

The responsibility for the co-ordination in accordance with first paragraph can be transferred to anyone who operates activities at the worksite or, in the question of loading or unloading of a ship in a Swedish port, at the harbour or the ship owner.

On the question of other common worksites than what is stated in the first paragraph the persons who operate activities there can agree that one of them shall be responsible for the co-ordination. This does not apply to the worksite for building or construction work.

Section 7e

The person who is responsible for the co-ordination of work environment questions in accordance with Section 7 d shall see to that

1. the work with preventing risks for ill-health and accidents is co-ordinated at the common worksite,
2. work is scheduled in a manner which is needed to prevent risks for ill-health and accidents due to that various activities are occurring at the worksite,
3. general safety measures are established and maintained and general protection rules for the worksite are issued,
4. the responsibility for the special safety measures which can be needed for a certain job to be prepared, and
5. personnel space and sanitary conditions are established at the worksite to the extent necessary.

Section 7f

If a building or construction work comprises a common worksite with other business is stated in Section 7 d, Sections 6-7 c in the question of the building or construction work as well as Sections 7 d and 7 e in the question of the other activities are applied.

The person who is responsible for the co-ordination of work environment questions in accordance with Section 7 d shall with the purpose that produce satisfactory protection conditions to see to it consultation will be implemented with the person who has been appointed to building work environment co-ordinator in accordance with Section 6.

Section 7g

Those who simultaneously, or in timely connection to each other, operate activities at a common worksite, shall consult each other and jointly act in order to produce satisfactory protective conditions.

Each and every one of them shall see to their own activities and provisions at the common worksite not lead to anyone who works there being exposed to risk for ill-health or accidents.
Those who operate activities or work at the common worksite shall follow directions from a building work environment co-ordinator in question of building or construction work and from the person who is responsible for the co-ordination of work environment questions in general cases.

Section 7h

The person who manufactures prefabricated buildings or constructions shall see to it that work environment viewpoints are considered in the event of projecting when it is applicable to the building phase as well as the future usage and see to it different parts of the projecting are co-ordinated."

Chapter 4, section 8

The Government or the authority decided by the Government is allowed to notify regulation as

1. preliminary application

   a) when it is regarding obligation for the person who orders execution of building or construction work to make sure that such a report is submitted to the supervisory authority, and

b) in general, and

2. the obligation to in general make a report or provide information to supervisory authorities or to store files which are of significance from a protection point of view.

The Government or the authority decided by the Government is further allowed to notify regulation as

1. that the person who orders execution of building or construction work shall make sure that

    a) a work environment plan is established,

    b) for the project type suitable documentation is drawn up to be considered at following jobs, and

    c) the work environment plan and documentation is appropriate in consideration to how the work proceeds and the eventual changes that have taken place,

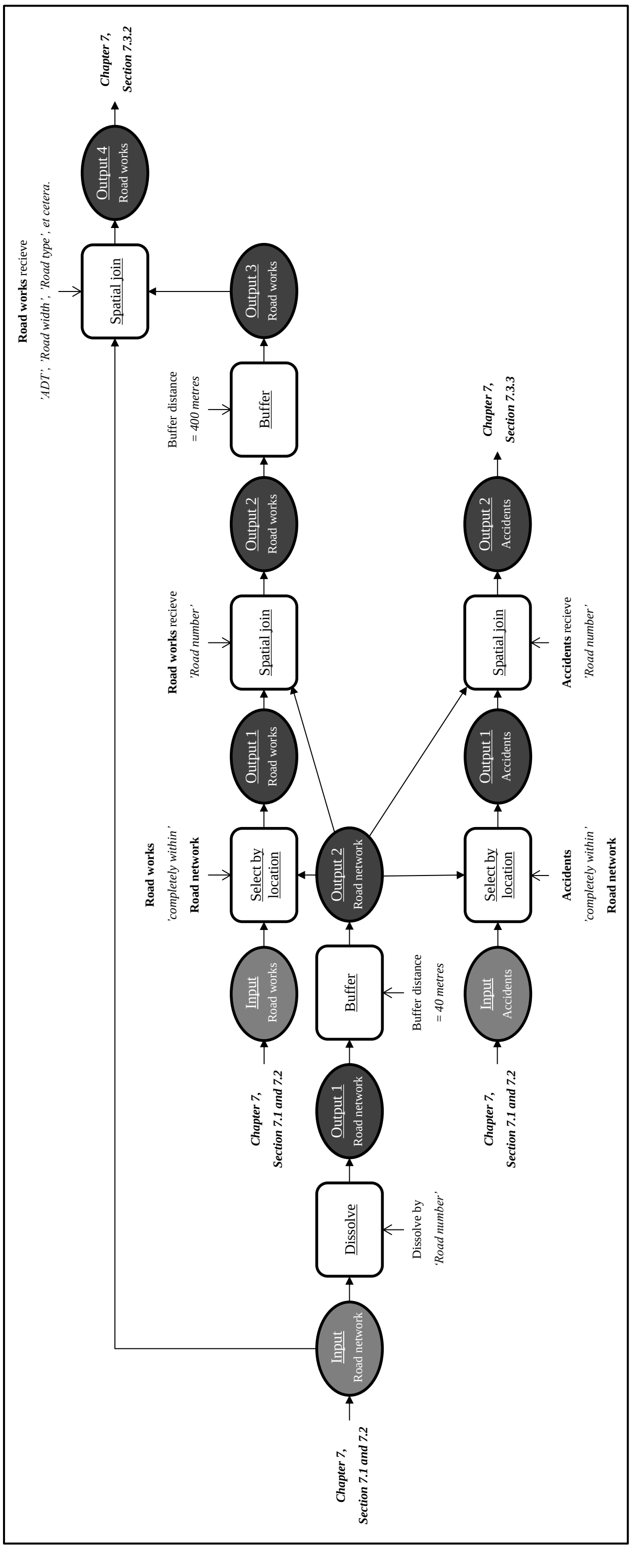
2. that a building work environment co-ordinator in accordance with Chapter 3 Section 7 a shall establish or have established a work environment plan and draw up such documentation as mentioned in 1,

3. that a building work environment co-ordinator in accordance with Chapter 3 Section 7 b shall see to that necessary adjustments of a work environment plan and such documentation as mentioned in 1 are carried out, and

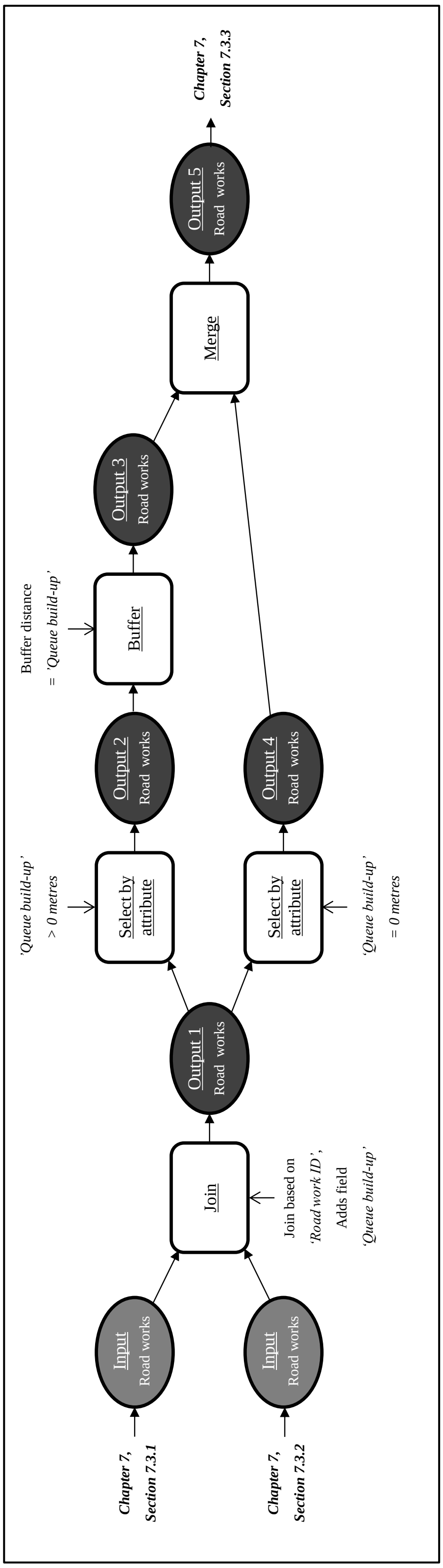
4. obligation to in general establish documents which are of significance from a protection point of view." (Arbetsmiljöverket, 2015)

Appendix B Model in ArcGIS

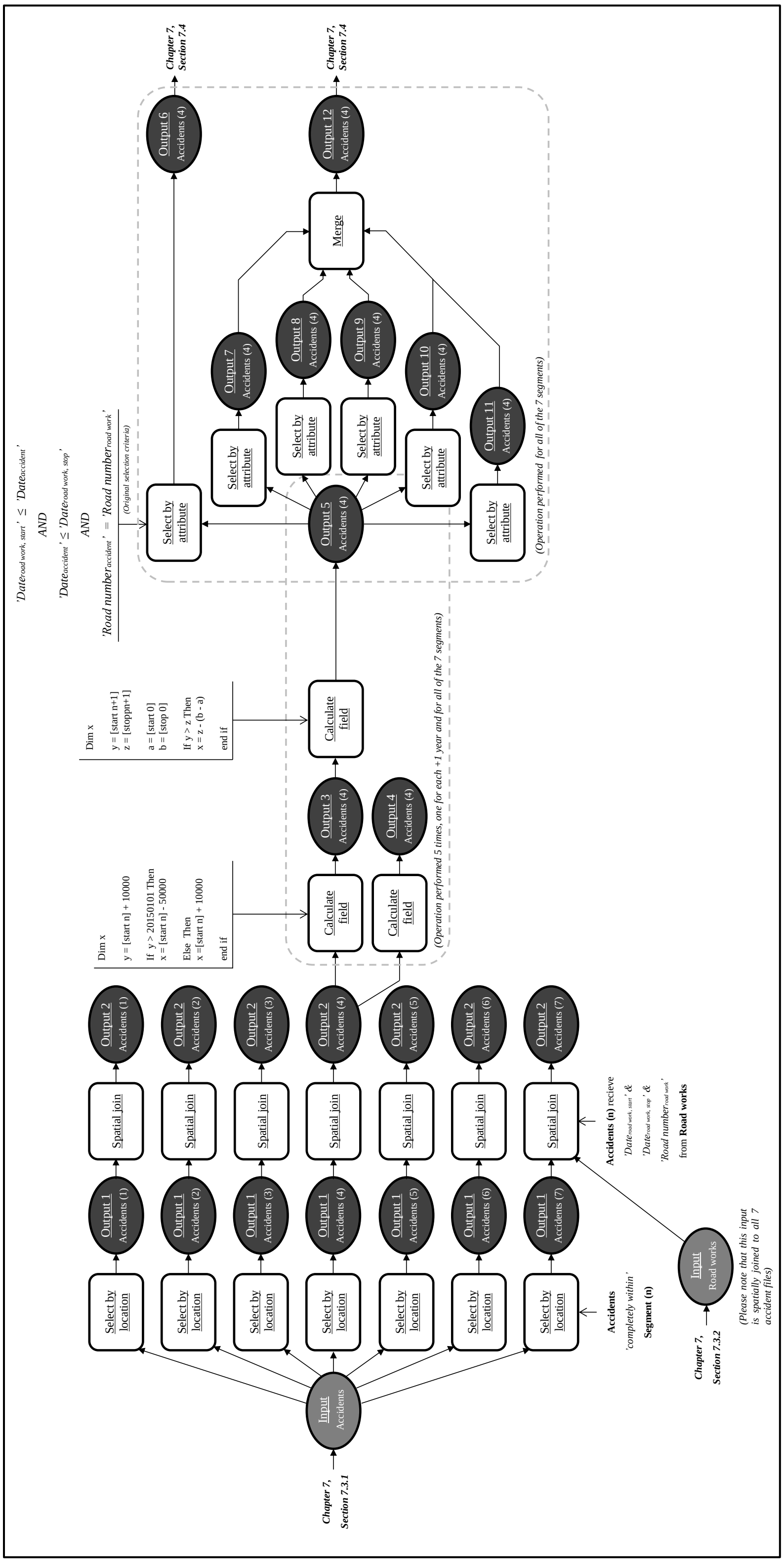
B.1 Section 7.3.1



B.2 Section 7.3.2



B.3 Section 7.3.3



Appendix C Matlab code for removing duplicates in the initial road works file

```
%% Import data from text file.
% Script for importing data from the following text file:
%
%    /Users/fredrik_svensson_/Documents/MATLAB/untitled folder/Tsy.txt
%
% To extend the code to different selected data or a different text file,
% generate a function instead of a script.

% Auto-generated by MATLAB on 2015/03/12 17:39:57

%% Initialize variables.
close all
clear all
clc

tic

filename = '/Users/fredrik_svensson_/Documents/MATLAB/untitled folder/Tsy.txt';
delimiter = ';';
startRow = 2;

%% Read columns of data as strings:
% For more information, see the TEXTSCAN documentation.
formatSpec = '%s%s%s%s%s%s%s%s%s%[^\n\r]';

%% Open the text file.
fileID = fopen(filename,'r','n','UTF-8');
% Skip the BOM (Byte Order Mark).
fseek(fileID, 3, 'bof');

%% Read columns of data according to format string.
% This call is based on the structure of the file used to generate this
% code. If an error occurs for a different file, try regenerating the code
% from the Import Tool.
textscan(fileID, '%[^\n\r]', startRow-1, 'ReturnOnError', false);
dataArray = textscan(fileID, formatSpec, 'Delimiter', delimiter, 'ReturnOnError', false);

%% Close the text file.
fclose(fileID);

%% Convert the contents of columns containing numeric strings to numbers.
% Replace non-numeric strings with NaN.
raw = repmat({''},length(dataArray{1}),length(dataArray)-1);
for col=1:length(dataArray)-1
    raw(1:length(dataArray{col}),col) = dataArray{col};
end
numericData = NaN(size(dataArray{1},1),size(dataArray,2));

for col=[1,2,3,4,6,7,8]
    % Converts strings in the input cell array to numbers. Replaced non-numeric
    % strings with NaN.
    rawData = dataArray{col};
    for row=1:size(rawData, 1);
        % Create a regular expression to detect and remove non-numeric prefixes and
```

```
        % suffixes.
        regexstr                =               '(?<prefix>.*?)(?<numbers>([-]*(\d+[\,]*)+[\.]{0,1}\d*[eEdD]{0,1}[-
+]*\d*[i]{0,1})|([-]*(\d+[\,]*)*[\.]{1,1}\d+[eEdD]{0,1}[-+]*\d*[i]{0,1}))(?<suffix>.*)';
        try
            result = regexp(rawData{row}, regexstr, 'names');
            numbers = result.numbers;

            % Detected commas in non-thousand locations.
            invalidThousandsSeparator = false;
            if any(numbers==',');
                thousandsRegExp = '^\d+?(\,\d{3})*\.{0,1}\d*$';
                if isempty(regexp(thousandsRegExp, ',', 'once'));
                    numbers = NaN;
                    invalidThousandsSeparator = true;
                end
            end
            % Convert numeric strings to numbers.
            if ~invalidThousandsSeparator;
                numbers = textscan(strrep(numbers, ',', ''), '%f');
                numericData(row, col) = numbers{1};
                raw{row, col} = numbers{1};
            end
        catch me
        end
    end
end


%% Split data into numeric and cell columns.
rawNumericColumns = raw(:, [1,2,3,4,6,7,8]);
rawCellColumns = raw(:, 5);


%% Replace non-numeric cells with NaN
R = cellfun(@(x) ~isnumeric(x) && ~islogical(x),rawNumericColumns); % Find non-numeric cells
rawNumericColumns(R) = {NaN}; % Replace non-numeric cells

%% Allocate imported array to column variable names
obhistid = cell2mat(rawNumericColumns(:, 1));
verseqno = cell2mat(rawNumericColumns(:, 2));
verhistid = cell2mat(rawNumericColumns(:, 3));
OrderNo = cell2mat(rawNumericColumns(:, 4));
ta = rawCellColumns(:, 1);
Xcoord = cell2mat(rawNumericColumns(:, 5));
Ycoord = cell2mat(rawNumericColumns(:, 6));
GEHistId = cell2mat(rawNumericColumns(:, 7));

A=obhistid;
B=verseqno;
C=verhistid;
D=OrderNo;
E=ta;
H=GEHistId;
NyaXcoord=round(Xcoord/100);
F=NyaXcoord;
NyaYcoord=round(Ycoord/100);
G=NyaYcoord;

Matrix=vertcat([A D F G]);
```

```
size(Matrix,1);

Duplicatremoval = unique(Matrix, 'rows');
size(Duplicatremoval)

%%
txtfil=fopen('fredriksfixadetxtfil.txt','wt');
aa=['obhistid' ';' 'OrderNo' ';' 'NyaXcoord' ';' 'NyaYcoord'] ;
fprintf(txtfil,[aa ' \n']) ;
for j= 1:length(Duplicatremoval)
aa=[num2str(Duplicatremoval(j,1)) ';' num2str(Duplicatremoval(j,2)) ';' num2str(Duplicatremoval(j,3))
';' num2str(Duplicatremoval(j,4))];
fprintf(txtfil,[aa ' \n']) ;
end
fclose(txtfil);

toc
%% Clear temporary variables
clearvars filename delimiter startRow formatSpec fileID dataArray ans raw col numericData rawData
row regexstr result numbers invalidThousandsSeparator thousandsRegExp me rawNumericColumns
rawCellColumns R;
```

Appendix D Results from extraction of ArcGIS model

	Amount of accidents						(ii) & (b)	
Year \ Time-step	2009	2010	2011	2012	2013	2014	Total	Mean
2009	965	1059	1105	1154	1126	1065	5509	1101.8
2010	1104	1324	1024	1257	1233	1097	5715	1143
2011	899	783	867	849	812	713	4056	811.2
2012	1097	1164	921	1046	891	822	4895	979
2013	1466	1654	1601	1575	1478	1113	7409	1481.8
2014	1572	1762	1570	1767	1350	1567	8021	1604.2

	(i)	*(ii) & (b)*
Mean	1207.83	1186.7