A data driven approach for automating vehicle activated signs
A data driven approach for automating vehicle activated signs

Diala Jomaa
Vehicle activated signs (VAS) display a warning message when drivers exceed a particular threshold. VAS are often installed on local roads to display a warning message depending on the speed of the approaching vehicles. VAS are usually powered by electricity; however, battery and solar powered VAS are also commonplace. This thesis investigated development of an automatic trigger speed of vehicle activated signs in order to influence driver behaviour, the effect of which has been measured in terms of reduced mean speed and low standard deviation. A comprehensive understanding of the effectiveness of the trigger speed of the VAS on driver behaviour was established by systematically collecting data. Specifically, data on time of day, speed, length and direction of the vehicle have been collected for the purpose, using Doppler radar installed at the road. A data driven calibration method for the radar used in the experiment has also been developed and evaluated.

Results indicate that trigger speed of the VAS had variable effect on drivers’ speed at different sites and at different times of the day. It is evident that the optimal trigger speed should be set near the 85th percentile speed, to be able to lower the standard deviation. In the case of battery and solar powered VAS, trigger speeds between the 50th and 85th percentile offered the best compromise between safety and power consumption. Results also indicate that different classes of vehicles report differences in mean speed and standard deviation; on a highway, the mean speed of cars differs slightly from the mean speed of trucks, whereas a significant difference was observed between the classes of vehicles on local roads. A differential trigger speed was therefore investigated for the sake of completion. A data driven approach using Random forest was found to be appropriate in predicting trigger speeds respective to types of vehicles and traffic conditions. The fact that the predicted trigger speed was found to be consistently around the 85th percentile speed justifies the choice of the automatic model.

**Keywords:** Optimal trigger speed, vehicle activated sign, mean speed, standard deviation, calibration, driver behaviour, data driven approach, automatic model

Diala Jomaa, School of Technology and Business Studies
Högskolan Dalarna, SE-791 88 Falun, Sweden, djo@du.se
List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANFIS</td>
<td>Adaptive neuro fuzzy inference system</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>DMS</td>
<td>Dynamic message signs</td>
</tr>
<tr>
<td>CF</td>
<td>Correction factor</td>
</tr>
<tr>
<td>CRD</td>
<td>Completely randomised design</td>
</tr>
<tr>
<td>DSLS</td>
<td>Dynamic speed limit signs</td>
</tr>
<tr>
<td>NRMSE</td>
<td>Normalised root mean square error</td>
</tr>
<tr>
<td>RF</td>
<td>Random forest</td>
</tr>
<tr>
<td>SL</td>
<td>Speed limit</td>
</tr>
<tr>
<td>SID</td>
<td>Speed indicator devices</td>
</tr>
<tr>
<td>SOM</td>
<td>Self-organising maps</td>
</tr>
<tr>
<td>TS</td>
<td>Trigger speed</td>
</tr>
<tr>
<td>VAS</td>
<td>Vehicle activated signs</td>
</tr>
<tr>
<td>VSLS</td>
<td>Variable speed limit signs</td>
</tr>
<tr>
<td>VMS</td>
<td>Variable message signs</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to thank numerous people for their contributions and support in this PhD thesis and also during the process of this thesis. First and foremost, I would like to thank my main supervisors Professor Mark Dougherty and Associate Professor Siril Yella for giving me the opportunity to pursue my PhD under their guidance, support and encouragement and for allowing me to grow as a research scientist. My deepest gratitude is to you Mark, for always being ready to listen to me and to my problems. I am really happy to have had the opportunity to know you. You have given me freedom to explore on my own and at the same time guided me to recover when I faltered. In a few words, I can say that what I am today and what progress I have made is because of you. I have learnt a lot from your ideas and thoughts during these five years.

I would also like to thank you Siril for your advice on both research as well as on my career, which have been priceless. You were always ready to help and to make time for discussion. I will never forget our discussions and your comments and the way in which you explain things. I thank you for your patience and motivation right from the beginning to the end of this research.

I am thankful to my MS thesis supervisor Associate Professor Hassan Fleyeh for encouraging me and developing my interest in research. I am happy that I worked with you in teaching and in research. I am also thankful to Dr. Karin Edvardsson for supervision and co-authoring my papers and for giving advice and motivation regarding the publication of the work.

Besides my supervisors, I would like to thank everyone in SafeX system for providing me all the necessary technical support through the work in this thesis. In particular, I thank Lars Söderström, and Markus Johnsson for their generous help in collecting the data. I would also like to thank ITS Dalarna for supporting this research. In particular, I would like to thank Johan Hallgren, Annica Roos and Tom Ramstedt for their support, without which a substantial amount of work in this thesis would not have been possible. I thank you Annica for your presence, your care and your endless support and your encouragement and motivation during these past five years.
I wish to thank all my colleagues at the faculty of data and information science at Dalarna University for providing me with a nice and friendly environment for pursuing research: Dr. Jerker Westin, Hans-Edy Mårtensson, Hans Jones, Thomas Kvist, Pär Douhan, Pär Eriksson, Anders Forsman, Dr. Anders Avdic, Hans Rosendahl, Ulrika Artursson, Anna Skogsberg, Michael Nilsson, Daniel Wikström and Rickard Stridbeck. Special thanks to Karl Hansson and Somayeh Aghanavesi for sharing the office with me and for your friendship; and to Dr. Roger Nyberg for all those chats, which have supported me a lot. Thanks to my fellow PhD students Kristin Svensson and Xiaoyun Zhao.

I wish to thank other colleagues from other departments: Professor Kenneth Carling and Associate Professor Johan Håkansson as research leaders for the research profile in Microdata analysis. Professor Lars Rönnegard for your great teaching in statistics and Dr. Moudud Alam for your guidance and help in writing this thesis. Special thanks to you Annette Lenne for your great care during all these years.

I would also like to thank all my friends: Pascal, Julia, Karl, Alexis, Helena, Catia. Thanks for all the nice times after work. Zuzana and Yasemin, my best friends, I am really glad to be so close to you. Zuzana, I cannot find a person like you who cares and supports me continuously. Yasemin, with your kind and wise heart, I really enjoy the time when we are together and I really appreciate your support to the end.

I thank my father Rizkaillah Jomaa for being a source of encouragement and motivation. I would also like to acknowledge the love and warmth of my mother, Nassima; my sisters Suzanne and Hala and brothers Naser and Mohamad and particularly my sister Lama for loving me constantly and deeply and to my lovely Dana.

Finally, I would like to express my deepest love to my children Ali and Fatima. I am grateful to you both for your patience during times when I had deadlines. I will now answer your question that you asked several years ago, ‘When you will finish’? I am finished now. I love you all!

Diala Jomaa, Borlänge, May 2016
## Contents

1  **Introduction** ........................................................................................................... 1  
   1.1  **Problem formulation** ......................................................................................... 3  
   1.2  **Objective and Aims** .......................................................................................... 4  
   1.3  **Research contributions** ...................................................................................... 5  
   1.4  **Outline of the thesis** ........................................................................................... 5  

2  **State of the art (Paper I)** ....................................................................................... 7  

3  **Data acquisition and experimental design** ............................................................ 9  

4  **Methods (Papers II-VI)** .......................................................................................... 11  
   4.1  **Paper II** .............................................................................................................. 11  
   4.1.1  Standard calibration ............................................................................................ 11  
   4.1.2  Data driven calibration ......................................................................................... 12  
   4.1.3  Experimental calibration ...................................................................................... 12  
   4.2  **Paper III** ............................................................................................................ 12  
   4.2.1  Completely randomised design ........................................................................... 13  
   4.2.2  Site selection ........................................................................................................ 13  
   4.2.3  Trigger speed selection ......................................................................................... 13  
   4.2.4  Statistical methods ............................................................................................... 14  
   4.3  **Paper IV** ............................................................................................................ 15  
   4.3.1  Design and trigger speed selection ..................................................................... 16  
   4.3.2  Statistical methods ............................................................................................... 16  
   4.4  **Paper V** ............................................................................................................. 17  
   4.4.1  Data partitioning ................................................................................................. 18  
   4.5  **Paper VI** ............................................................................................................ 19  
   4.5.1  Data Analysis ...................................................................................................... 20  
   4.5.2  Data partition and prediction ............................................................................. 20  

5  **Results** ................................................................................................................... 22  
   5.1  **Paper II** .............................................................................................................. 22  
   5.2  **Paper III** ............................................................................................................ 22  
   5.3  **Paper IV** ............................................................................................................ 23  
   5.4  **Paper V** ............................................................................................................. 24  
   5.5  **Paper VI** ............................................................................................................ 25  

6  **Conclusions and Discussion** .................................................................................. 27  
   6.1  **Discussion** ......................................................................................................... 29  
   6.2  **Limitations and future work** .............................................................................. 31  

References ...................................................................................................................... 33  

Appendix A: Appended papers ...................................................................................... 37
1 Introduction

Excessive or inappropriate speed is a major contributing factor in traffic related accidents and fatalities. Various safety signs are deployed to guide drivers and aid in adjusting their speed to current road and traffic conditions. Variable message signs (VMS) are a typical example of such safety signs. Depending on the circumstances, they are also referred to as changeable message signs (CMS), dynamic message signs (DMS), variable speed limit signs (VSLS) or dynamic speed limit signs (DSLS) (Nygårds, 2011). VMS are currently used as a key part of dynamic traffic management systems (Figure 1). Depending on the traffic situation, variable message signs inform, warn and guide drivers. This type of sign is often used on highways and expressways. According to previous studies, these signs contribute to shorter queues, travel times, uniform and lower speeds thereby resulting in fewer incidents. Providing the drivers with essential and realistic information such as expected speed, expected travel times or queues increases the driver’s comfort level, while simultaneously reducing the driver’s stress and irritation level (Davidsson et al., 2007). However, these signs are demanding, equipment dependent and expensive to operate.

![Figure 1 Example of Variable message sign (source: Kronborg 2001)](image)

The term VMS includes Vehicle activated signs (VAS) and Speed indicator devices (SID). VAS and SID are activated by the speed of approaching vehicles (referred to as trigger speed from this point forward). The signs are usually powered by electricity; however, battery and solar powered signs are also commonplace. The signs display a warning message when drivers exceed a particular threshold i.e. the trigger speed. Note that the warning messages displayed by VAS and SID signs are slightly different. VAS display a warning message, typically ‘slow down’ in combination with the current speed limit, whereas SID display drivers’ speed in green or red colours in order to alert
drivers of their current speed (Figure 2). Usage of either a smiley face or a sad face on SID is also not uncommon (Walter, & Knowles, 2010). It should be noted that some literature consider SID as a type of vehicle activated sign, and both belong to the category of VMS, as a higher class of sign. Hence, in the current thesis, VAS are considered as a separate type, where the display and operation of VAS is somehow different to both VMS and SID.

VAS and SID are relatively cheaper than VMS and require little maintenance (Winnett, & Wheeler, 2002). Therefore, VMS are usually installed on highways where the risk of accidents is potentially high. Both types of signs can be installed quickly and at relatively low cost. They are most effective at locations that have been the scene of collisions caused by drivers who travel at inappropriate speeds. Apart from this similarity, VAS and SID are deployed and used in different ways. VAS are usually permanent installations, whereas SID are usually placed onsite for not more than 2 to 3 weeks, i.e. SID are generally temporary and relocated from site to site. Previous research studies indicated the effect of both signs with regard to speed reduction, as being between 2 and 7 mph (Winnett, & Wheeler, 2003; Walter, & Knowles, 2008; Pesti, & McCoy, 2001; Sandberg et al., 2006), but it is not clear which type of sign should be considered when discussing usage and effectiveness on drivers speeds.

Studies concerning the trigger speed of VAS have not been carefully examined. Nygård and Helemers (2007) and Nygård (2011) reviewed effects of VAS and VMS, in articles published between 2000 and 2009. The articles that were published during that period of time concentrate mainly on the influence of the sign upon human behaviour. The main conclusion drawn from their reviews is that relevant information that is displayed by the sign plays an essential role in influencing driver behaviour. The trigger speeds of these signs have not been emphasised in their reviews. In addition, large number of studies has been conducted in order to test the effectiveness of VAS in reducing vehicle speed and improving safety. A frequent measure on
effectiveness of VAS on driver behaviour is the reduction in the mean speed of vehicles as they pass the sign. However, some recent studies (Quddus, 2013; McMurtry et al., 2009; Garber & Earhart, 2000) have indicated a positive relationship between the standard deviation of vehicle speeds and number of traffic accidents. To the best of my knowledge, it is not clear whether the optimal trigger speed of VAS has an effect on mean and standard deviation of vehicle speeds or not and is worthwhile investigating.

1.1 Problem formulation

In practice, the trigger speed of the VAS is set to a constant value that corresponds to the traffic agencies’ recommendation for a particular road segment. Such practices, however, do not consider the fact that an optimal trigger speed might exist, i.e. a trigger speed that has the most beneficial impact on driver behaviour. It is my belief that each sign may need a different trigger speed for different road segments. It can be argued that VAS have a limited short-term effect on driver behaviour and that excessive use of these signs might, in the long-term, lead to a situation in which drivers tend to disregard the warnings. This might be the case when an inappropriate trigger speed is employed in the VAS. If drivers encounter a substantial number of incorrectly configured signs, there is a likelihood that drivers will begin to disregard all such warning signs. Therefore, it is essential that drivers perceive the message displayed by VAS as relevant and credible with regard to the traffic and road conditions. This is why setting of trigger speed is required in VAS.

Another important problem that is worth mentioning here concerns the power consumption of battery driven VAS. As mentioned earlier, battery driven VAS are not uncommon (see section 1). Operation of battery driven VAS is deemed to be poor without an appropriate trigger speed. This is because lack of an appropriate trigger speed threshold coupled with high traffic flow implies more frequent activation, thereby leading to quicker depletion of power. This is true also when using solar powered VAS, which are often preferred due to their low installation and operation costs. Whilst solar operated VAS are more desirable than battery driven ones, setting an appropriate trigger speed is still an issue. In order to keep the sign operational, the sign should be triggered only when necessary.

Setting the trigger speed according to traffic dynamics by taking into account vehicle type, road and traffic conditions is yet another challenge. As an example, behaviour of heavy vehicles is often different to that of cars. Furthermore, in terms of traffic volume, cars are the most common vehicle type.
Thus, ignoring the above-mentioned issues might cause VAS to fail to properly warn different vehicle classes under different weather conditions.

Apart from the problem of setting the trigger speed of the VAS, another important problem is the sensitivity in acquiring data. This is because acquiring speed data is a challenge in itself. A commonly used method for acquiring speed and traffic data is the Doppler radar. However, data collected using the Doppler radar is sensitive to the calibration of the radar. The radar should be mounted in an overhead position, pointing down towards the roadway at approximately 30 degree angle from the horizontal direction. Often in reality, radar calibration is carried out subjectively by installation personnel, thus greatly increasing possibility of introducing error in the collected speed data, which in turn results in incorrect calculation of the trigger speed.

1.2 Objective and Aims

The main objective of this thesis is to investigate an appropriate trigger speed for VAS by taking into account vehicle type, road and traffic conditions. In particular, this thesis first evaluated the effectiveness of the trigger speed of the VAS on driver behaviour in order to find its optimum value. An automatic model designed to predict an appropriate trigger speed according to the conditions was also developed. The aim of this thesis is to fulfil the above-mentioned objective by investigating the following research questions.

(R1) What is the trigger speed of the VAS that was used in previous studies? How effective were the VAS with regard to assessing driver behaviour?

(R2) How should the data be collected?

(R3) What is the effect of trigger speed on driver behaviour, and what is the optimal trigger speed?

(R4) Is the effect of VAS comparable to that of SID on driver behaviour according to the trigger speed, time of day and type of site or not?

(R5) How can an optimal trigger speed of the solar VAS be found in order to reduce drivers’ speed, while simultaneously optimising power consumption?

(R6) How can an automatic trigger speed for VAS, which responds to traffic conditions and to different type of vehicles be assessed? Are different VAS trigger speeds required for different vehicle types?
1.3 Research contributions

In order to answer the above questions, the research contributions in this thesis are presented in the six appended papers as follows:

**R1** A review study on the effectiveness of VAS was conducted. The review focused on parameters used for the assessment of the effectiveness of the signs. Furthermore, parameters used for the configuration of the trigger speed of the VAS were also studied (paper I).

**R2** A data driven calibration system was applied to correct vehicle velocity data collected by Doppler radar. At this point, it is worth mentioning that velocity of the vehicle was used instead of speed to be able to account for vehicle direction. Data were initially pre-processed to be able to match individual vehicles, and the correction factor (CF) for vehicle velocities detected by radar was determined. Different calibration measures were used to validate accuracy of the obtained CF (paper II).

**R3** A comprehensive understanding of the effectiveness of the trigger speed of the VAS on driver behaviour was gained by systematically collecting speed data whilst varying the value of the trigger threshold of the VAS. This study investigated the criteria for determining the optimal trigger speed that gave the best compromise between reducing the mean speed and the standard deviation of speed (paper III).

**R4** A comparative study between VAS and SID was done in order to determine the effectiveness of these signs with regard to driver behaviour in terms of the nature of the site (local roads as opposed to a highway), time of day (day as opposed to night) and trigger speed chosen for the sign (paper IV).

**R5** A data driven approach was developed to optimise the trigger point that achieves the best compromise between speed reduction and power consumption of solar VAS. The algorithm was deployed in two stages: first stage involved exploration of the proprieties of the data and then, in the second stage, trigger speed of the VAS was determined using a clustering technique (paper V).

**R6** A data driven model was investigated to predict the appropriate trigger speed of the VAS depending on type of vehicle and time of day (paper VI).

1.4 Outline of the thesis

The work in this thesis has been segmented into five sections respectively. Section 2 provides a review of previous studies conducted regarding the effectiveness of VAS. The review focuses on the parameters used for assessment of the effectiveness of the signs. Section 3 presents data collection for
each experimental design in the different papers included in this thesis. The methods used in the different papers included in this thesis are listed in Section 4. In Section 5, a summary of the five papers (paper II–paper VI) presents the aim of the thesis along with the corresponding results. Finally, Section 6 presents discussion and concluding remarks.
2 State of the art (Paper I)

Previous research on VMS and VAS was reviewed to be able to answer research question (R1); what is the trigger speed of the VAS signs that was used in previous studies? How effective were the VAS signs with regard to assessing driver behaviour?

Several methods for measuring the effectiveness of the VMS and VAS are reported in the literature. Most frequently, the effectiveness of the signs is measured with the reduction of the vehicle mean speeds, using vehicle speed data collected by loop detectors, radars or by using micro-simulation techniques. In this context, the mean speed of the vehicle is usually examined. However, the vehicle mean speed might not always be a reliable measure of a sign’s effectiveness, as shown by Kathmann (2001) in his study. The author argues that the effect of the signs might not affect the average speed of the vehicles and instead, the whole vehicle speed distribution should be examined.

Another measure of the effectiveness that has been proposed in a handful of studies is the relationship between the presence of VMS and VAS and the number of traffic accidents. Taylor et al. (2002) presented a robust predictive model based on extensive accidents and speed databases. The authors deduced that the reduction in the number of traffic accidents was 4.5% to 7.5% for each 1 mph reduction in the mean speed of the vehicle. While Taylor et al. (2002) did not consider the severity of the traffic accidents, Nilson (2004) reported that the reduction of 1 mph in the mean speed of the vehicle led to a 6.6% reduction in all injury-causing accidents, a 0.7% reduction in accidents that cause serious injuries and a 12.7% reduction in fatal accidents. The main advantage of using the relationship between presence of the signs and the traffic accidents is that the direct impact of the signs on traffic safety is measured. However, the drawback of this method is that the data collection usually needs to be done over a longer period of time, since traffic accidents are rather rare, for example, Taylor et al. (2002) used a database covering a 5-year period in their study. The third way of measuring the effectiveness of the VMS and VAS, proposed by some researchers, is by conducting interviews with drivers who have passed the signs. Luoma et al. (2000) argued that the effectiveness of the signs need not be assessed with only statistical measures, e.g. mean vehicle speed. Instead, other behavioural changes, such as the focus of a driver’s attention or cautious overtaking, might give good estimate of the effectiveness of a sign. Another measure of the effectiveness, suggested by Charlton and Baas (2006), is the sign visibility and comprehensibility.

A majority of the studies reported on the effectiveness of VMS and VAS; however, only a few studies attempted to deal with the technical configurations of these signs, in particular, the trigger speed. There is no clear consensus in the literature regarding the level at which the trigger speed should be set. Thus, Winnett et al. (1999) reported a trigger speed being set to 4mph.
under the posted speed limit, Winnett and Wheeler (2002) 5mph above the posted speed limit and Mattox et al. (2007) 3mph above the speed limit.

Lastly, a large number of studies have evaluated the effectiveness of the VMS and VAS on driver behaviour, but only a few studies have considered the effects of these signs under different road and weather conditions (Rämä, 1999; Rämä & Kulmala, 2000; Luoma et al., 2000; Jiang et al., 2011). These studies reported a small beneficial reduction in mean speed but increases in the homogeneity of driver behaviour. In Rämä (1999), there was more of an effect on mean speed and standard deviation in summer than in winter.

Studies concerning the trigger speed values of the signs have not been carefully examined. Past studies have reviewed the effects of VAS and VMS (Nygård & Helemers, 2007; Nygårds, 2011). Such studies have mainly concentrated on the influence of VAS and VMS upon driver behaviour, where the mean speed of drivers passing the sign was generally reduced. The main conclusion drawn from their reviews is that relevant information that is displayed on the sign plays essential role in influencing driver behaviour. However, the effect of the trigger speed of the VAS and its influence on driver behaviour has not been carefully studied. A large number of studies have been conducted to test the effectiveness of VAS in reducing vehicle speed and in improving safety. A frequent measure of the VAS effectiveness on driver behaviour is the reduction in the mean speed of vehicles as they pass the sign. More recently, studies have indicated a positive relationship between the standard deviation of vehicle speeds and number of traffic accidents (Quddus, 2013; McMurtry et al., 2009; Garber & Earhart, 2000). At this stage, it is worth pointing out that the optimal trigger speed of VAS should have a combined effect on both the mean and standard deviation of vehicle speeds.
3 Data acquisition and experimental design (Papers II-VI)

Relevant literature within the field indicated that when selecting a site the notoriety of a given road segment for speeding or a high accident rate (where inappropriate speeds were the initial problem) ought to be considered (Wiinnett, & Wheeler, 2002). Test sites were chosen after duly considering factors such as speeding areas, change from high to low speed limit, presence of school, bus stop and so on, all of which might affect data collection.

Traffic speed data were collected by Doppler radars installed onsite at four different sites. First\(^1\) and second sites\(^2\) (site-1 and site-2) were located on a local road in Borlänge and were both restricted to 40 km/hr with speed limit change from 50 km/hr to 40 km/hr. The third site\(^3\) (site-3) was also located on a local road in Borlänge but was restricted to 60 km/hr but a location where the speed limit changes from 60 km/hr to 50 km/hr. Fourth site\(^4\) (site-4) was located on highway E16 between Borlänge and Djurås in central Sweden and was restricted to 60 km/hr where there was a speed limit change from 90 km/hr to 60 km/hr.

In this context, the vehicle activated sign was installed at the beginning of the road segment with a lower speed limit to remind drivers. According to the aim for each experiment conducted in this thesis, data were collected and prepared in different ways. The first experiment in paper II was aimed at dealing with potential problems due to faulty calibration of the radar. For this purpose, a data driven calibration method was developed and evaluated to provide accurate data for this research study. For this experiment, individual speed data (referred to as data 1) were collected 24 hours a day from the first site in Borlänge using three Doppler radars: radar installed 100m before the sign, at the sign and 60m after the sign.

The experiment done in paper III studied the effect of the trigger speed of the VAS on driver behaviour. Data used in this experiment (referred to as data 2) were collected from the first two sites using two radars. The vehicle activated sign was also installed and equipped with a data logger and a modem in order to be able to remotely alter the trigger speed settings of the sign. Such a setup facilitates alteration of trigger speeds thereby permitting the study to compare the effect of different trigger speeds on driver behaviour, particularly on the mean and standard deviation of vehicle speeds. The data collection was done over a period of 16 weeks. The data were collected 24 hours a day at two points: 100m before the sign, at the sign and 60m after the sign.

The experiment conducted in paper IV studied the effect of the trigger speed

---
\(^1\) Site-1 (Latitude: 60.476904, Longitude: 15.464145)  
\(^2\) Site-2 (Latitude: 60.462058, Longitude: 15.467076)  
\(^3\) Site-3 (Latitude: 60.497165, Longitude: 15.452249)  
\(^4\) Site-4 (Latitude: 60.558988, Longitude: 15.137701)
of the VAS compared to SID i.e. another type of VMS. Data (referred to as data 3) used in this experiment were collected from the first and the fourth site where the road characteristics were completely different. At each site, data were collected continuously before the SID and the VAS were installed and while the signs were in operation. At each site, the data collection was also done at two points: 100m before the sign and at the sign. The data collection period was one week.

The experiment described in paper VI aimed at building an adaptive sign that detects and records vehicle speed and predicts a trigger speed respective to previous traffic conditions. Relevant data (referred to as data 4) were collected using single radar installed 100 m before the sign. Data were collected 24 hours a day at the four sites (although not simultaneously) for the sake of comparison.
4 Methods (Papers II-VI)

This chapter presents the methods used in each of the papers found in appendix A.

4.1 Paper II

This paper has investigated the research question (R2): How to collect data using VAS? See section 1.2. Vehicle speed data (Data 1) were collected, and an appropriate correction factor was determined. Data were collected using a continuous wave Doppler radar (named Siersega) to measure the speed of vehicles passing the radar. A successful measurement requires that the radar has a direct view of the vehicles and requires careful setup. Therefore, the radar was installed in a side fire overhead position and set on a road post-lamp at a fixed height between 2.25 to 3.25 metres above the ground. Additionally, the radar was facing the oncoming traffic inclined at 30 degrees parallel to the roadway and tilted downward 20 degrees. The cosine of the angle between the radar unit and its target determines the magnitude of error. During the data collection stage, vehicle speeds were gathered using the two radars. While the first radar was positioned at the same place as the vehicle activated sign, the second radar was located either 100m ahead of the sign or 60m after the sign. The rationale behind moving the second radar back and forth relative to the sign was to be able to test the accuracy of the CF.

Data were pre-processed using two approaches, resulting in two different data sets. As per the first approach, missing data and outliers were identified and removed which resulted in a ‘clean’ data set. In the second approach, missing data and outliers were constructed by filling in missing values whilst keeping all the outliers resulting in a ‘complete’ data set. Results of calibration are summarised in chapter 5 (see section 5.1). What follows next is a description of the calibration methods employed in the current work.

4.1.1 Standard calibration

The standard calibration of the radar involves subjectively installing the device relative to the manual specifications. The collected velocities were then multiplied by a certain CF based on the inverse of the cosine error of the two angles between the radar and the travelling direction of the vehicle. In this study, the correction factor (CF) proposed by radar is expressed in equation 1:

\[
CF = \frac{1}{\cos 30 \cos 20}
\]
4.1.2 Data driven calibration

The data driven calibration method was based on the inverse of the ratio between the actual distance and the travelled distance between two Doppler radars installed on the road. The travelled distance was extracted from the matched velocities recorded by both radars. The main idea for matching velocities was done by finding the time difference between the two radars, which are supposed to be initially synchronised. The time difference included a certain time delay \( \partial \) due to disturbances such as overtaking, double counting or the radar may be occupied with another vehicle coming from the other direction. The detection of a time delay was extracted by using a numerical algorithm, which starts at any point and recursively approaches to an approximate solution. The CF proposed in this thesis was calculated by the following equations where \( n \) is the total number of matched vehicles, \( v \) is the mean velocity and \( \Delta t \) is the time difference for the same individual vehicle recorded by the two radars (see equations 2 and 3):

\[
CF = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\text{actual distance}}{\text{travelled distance}} \right)_{i} 
\]

\[
\text{travelled distance}_{i} = v_{i}(|\Delta t| - \partial) 
\]  

4.1.3 Experimental calibration

The experimental correction factor was based on several runs carried out by driving a car in cruise control mode on the road segment where the two radars were installed. The velocities shown in the speedometer of the car and time were recorded. The experimental CF was the baseline ratio between the velocities reported by the radars and the velocities shown in the speedometer of the car. It is worth mentioning that to determine the vehicle used in the experiment, several runs were done for a particular time period where the traffic flow was very low.

4.2 Paper III

This paper has investigated the research question \((R3)\): What is the effect of trigger speed on driver behaviour, and what is the optimal trigger speed? (See section 1.2) The effect of the trigger speed of the VAS on driver behaviour was mainly based on the experimental design chosen for the data collection as well as the value of the trigger speed that activates the sign. Speed data (data 2) were collected, and a completely randomised design (CRD) has been employed. This design is mainly based on the comparison of the values of the vehicle speed based on the different levels of the trigger speed to assess how the VAS trigger speed affected driver behaviour.

Data were analysed by both descriptive and inferential statistics. Mean speed, standard deviation and the coefficient of variation were used in the descriptive statistics. The Bartlett test, the One-way test, the pairwise t and
pairwise F-tests were conducted in the inferential part of the data analysis. While the Bartlett test was performed in order to check the homogeneity of the variances between trigger speeds, the One-way test was used to check the mean speed's homogeneity, when the assumptions of equal variances were relaxed. Furthermore, the pairwise tests (t-test and F-test) were also used to perform a pairwise comparison for mean speeds and speed variances, respectively.

Results of this investigation are summarised in chapter 5 (see section 5.2). What follows next is a description of the completely randomised design, site and trigger speed selection and the statistical methods employed in paper III.

4.2.1 Completely randomised design

A completely randomised design (CRD) was employed in this paper to study the effect of the trigger speed of the VAS on driver behaviour. CRD is simple designs based on randomisation and replicated with test subjects assigned at random treatment levels of the primary factor and eliminate the effect of other secondary factors to ensure the accuracy of the data (Yau, 2013). In this paper, the primary factor was considered as the trigger speed for the VAS, and all other factors such as traffic intensity and traffic conditions were assumed to be secondary factors. The main idea is to define each hour as one experimental unit, and the trigger speed should be randomly assigned at each hour. Due to the difficulty in changing the trigger speed every hour, the trigger speed was randomly assigned to each week and then from each week; one working day was randomly chosen for further analysis.

4.2.2 Site selection

The test site was mainly selected based on the following criteria that may influence speed choice:

- Site located in an area known for speeding or collisions
- Site located where there is a change in the speed limit from high to low
- Site located where there is no sharp bends, roundabouts or pedestrian crossings
- Site located in close proximity to an external power source such as a power lamp
- Site should have sufficient space with at least 150 m to monitor drivers’ speed before and after activating the sign

4.2.3 Trigger speed selection

The trigger speed values chosen for the experiment were based on the following:

- Activate the sign for all vehicles (0 km/hr)
- Do not activate the sign (150 km/hr)
- Activate the sign for all vehicles that exceed the speed limit posted on the road (40 km/hr)
- Activate the sign for all vehicles that exceed the 15th percentile of the speed of vehicles travelling on the road (42 km/hr for site 1 and 46 km/hr for site 2)
- Activate the sign for all vehicles that exceed the speed limit posted on the road plus 10% of the speed limit (44 km/hr)
- Activate the sign for all vehicles that exceed the 50th percentile of the speed of vehicles travelling on the road (47 km/hr for site 1 and 49 km/hr for site 2)
- Activate the sign for all vehicles that exceed the 85th percentile of the speed of vehicles travelling on the road (50 km/hr for site 2 and 52 km/hr for site 2)

4.2.4 Statistical methods
The statistical methods used in thesis consisted of descriptive statistics and inferential statistics. Time mean speed, standard deviation and the coefficient of variation were used in the descriptive statistics. Time mean speed $\mu$ was used to describe the central tendency, where $n$ is the total number of vehicles and $S_i$ is the spot speed for vehicle $i$ (equation 4); standard deviation $SD$ was used to find the speed variation for the spread of the data (equation 5); and the coefficient of variation $CV$ was used to describe both central tendency and spread of the data, as provided in equation 6 (Mousa, 2005; Muchuruza et al., 2005).

$$\mu = \frac{1}{n} \sum_{i=1}^{n} S_i$$ \hspace{1cm} (4)

$$SD = \sqrt{\frac{n \sum_{i=1}^{n} (S_i - \mu)^2}{n-1}}$$ \hspace{1cm} (5)

$$CV = \frac{SD}{\mu}$$ \hspace{1cm} (6)

In the inferential part of the data analysis, the Bartlett test, the One-way test, the pairwise t and pairwise F-tests were applied to study if there were changes in mean speed and standard deviation between different trigger speeds. In general, the procedure of the hypothesis testing is based on setting up two hypotheses: a null hypothesis and an alternative hypothesis. Under the null hypothesis, it is assumed that there is no statistically significant relationship among the variables, whereas the alternative hypothesis challenges this assumption. The goal of hypothesis testing is usually to reject the null hypothesis in favour of the alternative hypothesis. The hypothesis testing procedure involves two additional parameters: a probability value $p$ and a confidence level $\alpha$. The $p$-value is the probability of the observed outcomes occurring
under the assumption that the null hypothesis is true. The confidence level $\alpha$ is the threshold value that the researcher selects to indicate at what point the null hypothesis will be rejected. If the p-value is less than or equal to $\alpha$, then the null hypothesis will be rejected (Wackerly et al., 2008; Dalgaard, 2008). The null hypothesis for Bartlett’s test indicated that the speed variances $\sigma^2$ between trigger speed values were equal (equation 7).

$$H_0: \sigma_0^2 = \sigma_{40}^2 = \sigma_{44}^2 = \sigma_{46}^2 = \sigma_{47}^2 = \sigma_{49}^2 = \sigma_{50}^2 = \sigma_{52}^2 = \sigma_{150}^2$$

(7)

The null hypothesis for one-way test indicated that the time mean speed was equal between different trigger speeds, while relaxing the assumptions of equal variances (equation 8).

$$H_0: \mu_0 = \mu_{40} = \mu_{44} = \mu_{46} = \mu_{47} = \mu_{49} = \mu_{50} = \mu_{52} = \mu_{150}$$

(8)

As with Bartlett’s test and one-way test, the pairwise t-test and the pairwise F-test were used to check variances and mean equality between different trigger speed levels, but only as a pairwise comparison. In the latter test, the comparison establishes which of the trigger speed levels differs from the others.

4.3 Paper IV

This paper has investigated the research question (R4): Is the effect of VAS comparable to the effect of SID on driver behaviour in terms of trigger speed, time of day and type of site (see section 1.2)? A comparison between VAS and SID in terms of the effectiveness on driver behaviour was the aim of paper IV. The comparison was done by determining the effects of VAS and SID in terms of site characteristics, time of day and trigger speed. More specifically, applying another design to the experiment and a positioning of the signs as well as road geometry were taken into consideration. The research conducted over the course of writing this paper was somehow related to the topic in paper III, but the experiment was designed in a shorter time period when the effect of these signs were found to be greater for the first week compared to the second week; this is most likely due to sign novelty.

Data 3 were collected and further divided into day/night periods, where the day period was defined as 07:00 to 19:59 and the night period was defined as 20:00 to 06:59. Due to variation in the number of hours of daylight in Sweden between different seasons, it was hard to divide the time according to the number of daylight hours. Therefore, in this paper, the respective division of the time was based on the difference in traffic flow volumes at different times of the day. Results of effectiveness on behaviour in terms of trigger speed, time of day and type of site are summarised in chapter 5 (see section 5.4). What follows next is a description of the design of the experiment, the trigger speed used and the statistical methods employed in the following work.
4.3.1 Design and trigger speed selection

The design of the experiment was controlled in respect to the same assumption used in the previous study (paper III), but site selection and time of the data collection were different. The sites were selected from two different road geometries: local roads and highway roads with different speed limits. Compared to the experiment in paper III, the data collection was performed in one week. One day period of installation of each sign with a particular trigger speed was randomly assigned to weekdays in order to reduce the possible effect of the external factors i.e. traffic and weather conditions. The trigger speed values chosen for the experiment were only based on the speed limit posted on the road segment and to the 85th percentile of the speed of passing vehicles. The trigger speed values for the two signs were determined as follows:

- Activate SID for all vehicles that exceed the speed limit 40 km/hr at site 1
- Activate SID for all vehicles that exceed the 85th percentile of the speed of vehicles travelling at site 1
- Activate SID for all vehicles that exceed the speed limit 60 km/hr at site 4
- Activate SID for all vehicles that exceed the 85th percentile of the speed of vehicles travelling at site 4
- Activate VAS for all vehicles that exceed the speed limit 40 km/hr at site 1
- Activate VAS for all vehicles that exceed the 85th percentile of the speed of vehicles travelling at site 1
- Activate VAS for all vehicles that exceed the speed limit 60 km/hr at site 4
- Activate VAS for all vehicles that exceed the 85th percentile of the speed of vehicles travelling at site 4

4.3.2 Statistical methods

Two statistical measures were used: effect on mean speed and effect on standard deviation of vehicle speed. To exclude the chance of the external random effect for the variability between sites and differences between speeds, the first measure was based on the difference between the mean speed of vehicles before and during the operation of the sign (Walter and Broughton 2011). The first measure used in this study is presented in equation 9; where \( \delta_2(t) \) is the effect for time period t at the sign, \( \mu_2 \) is the mean speeds at the sign and \( \mu_1 \) is the mean speed 100m before the sign.

\[
\delta_2(t) = (\mu_2(t) - \mu_2(before)) - (\mu_1(t) - \mu_1(before)) \quad (9)
\]

The second measure was based on the coefficient of variation, which is an-
other appropriate way to find the extent of variability in relation to the mean speed. This measure is presented in the following equation (10), where $C_v(t)$ is the coefficient of variation for time period $t$ at the sign, $\sigma_2(t)$ is the standard deviation at the sign for time period $t$, $\mu_2(t)$ is the mean speeds at the sign (see equation 10).

$$C_v_2(t) = \left( \frac{\sigma_2(t)}{\mu_2(t)} \right)$$  (10)

ANOVA analysis and pairwise t-test were also used here to compare means for multiple independent groups of data. One-way ANOVA is an extension of the t-test to 3 or more samples focus analysis on group differences. Two-way ANOVA (and higher) focuses on the interaction of factors.

Table 1. Effect of mean speeds respective to the null hypotheses based on the combination of different factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Null hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS</td>
<td>$H_0: \mu_2(\text{TS}=\text{SL}) = \mu_2(\text{TS}=85^{\text{th}} \text{percentile})$</td>
</tr>
<tr>
<td>Time</td>
<td>$H_0: \mu_2(\text{time}=\text{day}) = \mu_2(\text{time}=\text{night})$</td>
</tr>
<tr>
<td>Site</td>
<td>$H_0: \mu_2(\text{site}=\text{local road}) = \mu_2(\text{site}=\text{highway})$</td>
</tr>
<tr>
<td>TS: Site</td>
<td>$H_0: \text{no interaction between trigger speed and site}$</td>
</tr>
<tr>
<td>TS: time of day</td>
<td>$H_0: \text{no interaction between trigger speed and time of day}$</td>
</tr>
<tr>
<td>Site: time of day</td>
<td>$H_0: \text{no interaction between site and time of day}$</td>
</tr>
</tbody>
</table>

In fact, two-way ANOVA is used to determine whether the main effects and interaction effect are statistically significant. The null hypothesis for a main effect is that the response mean for all factor levels are equal. The null hypothesis for an interaction effect is that the response mean for the level of one factor does not depend on the value of the other factor. To assess each null hypothesis, p-values for each term were compared to a significance level 0.01. The table below presents the effect of mean speeds respective to the null hypotheses based on the combination of three factors: trigger speed (speed limit or 85th percentile), site (local road or highway) and time of day (night or day).

4.4 Paper V

This paper has investigated the research question (R5): How can the optimal trigger speed of the solar VAS be found in order to reduce drivers’ speed while simultaneously reducing power consumption? See section 1.2. The paper investigated a solar powered sign to be able to report its usage in the absence of access to direct electricity to power the sign. In other words, effi-
ciently configuring the trigger speed might help the battery of the sign from running out of power. In this work, the trigger speed of the VAS is based on two objectives: safety and energy. For the safety objective function, the effect of the trigger speed of the VAS was figured by calculating the variation of the mean speed of a vehicle travelling before triggering the sign (100m before the location of the vehicle activated sign) as well as after triggering the sign (at the location of the sign). For the energy objective function, the energy consumption was calculated respective to the sum of all flashes for the speed of all vehicles that were detected by the radar. The length of the flash for each vehicle could vary from full brightness 1.8A down to ambient brightness to under 0.5A.

Traffic data were clustered into traffic patterns using appropriate clustering techniques to be able to extract the trigger speed for each traffic pattern. The data (data 3) were initially pre-processed by excluding motorcycles and long trucks on the assumption that their speed might distort the results of clustering. A two stage clustering algorithm was utilised to extract the trigger speed of the VAS. The algorithm first employed a SOM to visualise and explore the proprieties of the speed data and second, clustered the same data into two different clusters that had the same speed characteristics, as determined through the utilisation of K-means. Results of clustering are summarised in chapter 5 (see section 5.4). What follows next is a description of the data partitioning methods used in this work.

4.4.1 Data partitioning

Self-organising maps (SOM) and k-mean clustering were used to explore and partition the data into different clusters that have similar speed characteristics. K-means clustering is an unsupervised clustering algorithm, which classifies a given data set through a certain fixed number of clusters k a priori. The main idea is to define the k centroid for k clusters. The centre of the cluster k is the mean of the data items within the cluster. K-means algorithm proceeds by first randomly selecting k of the items where each selection is done by partitioning data items into k initial clusters. Each item is assigned to the cluster to which it is the most similar, based on minimising the distance between the item and the cluster mean. It then computes the new mean for each cluster and assigns the new mean as the new cluster centre. This process iterates until a stopping condition is reached. In fact, this algorithm aims to minimise a simple objective function \(d(x_i, x_j)\) known as the dissimilarity measure, where \(||x_i - x_j||^2\) is the squared absolute Euclidian distance between each data point \(x_i\) and each data centre \(x_j\). The criterion of the within point cluster \(W(C)\) is minimised by assigning N observations to the K clusters in such a way within each cluster that the average dissimilarity of the observations from the cluster mean is minimised (see equation 11). Note that \(x' = (x'_{1k}, ..., x'_{pk})\) is the mean vector of the \(k^{th}\) cluster and \(N_k\) :

\[
W(C) = \sum_{k=1}^{K} N_k \sum_{c(i)=k} ||x_i - x'_k||^2
\]  

(11)
Self-organising map (SOM) is an artificial neural network that is trained using an unsupervised learning algorithm (Chen et al., 2013) to construct a map that discretized the input space of the training samples into low dimensions. The network often consists of two layers of artificial neurons: an input layer and an output layer. Every input neuron is connected to every output neuron by a weighting value. The Euclidian distance is calculated between the input vector and the incoming weighted vector for each output. The output neuron with the smallest distance is declared as the winner and its weights modified to be closer to the input vector. In fact, SOM is different from other neural networks as they apply competitive learning as opposite to the error correction based on the gradient’s descent. The competitive learning is basically an iterative process where the connections’ weights are modified according to the following equations (12) and (13) (Watts, & Worner, 2009; Shukla et al., 2012), where $w(t)$ is the connection weight at time $t$, $x(t)$ is the input vector, $h(t)$ is the neighbourhood function, $\alpha$ is the learning rate, $d$ is the Euclidian distance between the winning unit and the current unit and $\sigma$ is the neighbourhood width parameter.

\[ w(t + 1) = w(t) + h(t)(x(t) - w(t)) \]  \hspace{1cm} (12)

\[ h(t) = \alpha \cdot e^{-\frac{d^2}{2\sigma^2(t)}} \] \hspace{1cm} (13)

4.5 Paper VI

This paper investigates the research question (R6): How can an automatic trigger speed for VAS, which responds to traffic conditions and to different types of vehicles be assessed? Are different VAS trigger speeds required for different vehicle types?

Given the size, weight and behavioural differences between cars and trucks, a uniform VAS trigger speed might be inappropriate. This paper studied the speed characteristics between all types of vehicles within different periods of the day, particularly at night and during the day. Further, an automatic trigger speed system was presented. The automatic system consisted of two steps: in the first step, SOM partitioned the input data into separate clusters that have similar characteristics and then in the next step, RF and ANFIS predicted the 85th percentile speed within each cluster with the assumption that the 85th percentile was, in general, the appropriate trigger speed.

Results of clustering are summarised in chapter 5 (see section 5.5). A description of the analysis and data partitioning and prediction methods used in this work follows next.
4.5.1 Data Analysis

The work done in this study was analysed using field data collected at four sites (data 4). The data were grouped into two ways:

- Type of vehicles
- Time of day

To identify the type of the vehicle that passed the VAS, a simple classification was done by comparing the length of vehicles to a threshold, recommended by traffic engineering. For instance, the length of cars was considered between 21dm to 60dm and the length of vans and trucks was between 61dm and 94dm. In this context, the analysis was done respective to four classes: motorcycle, cars, vans/trucks and long trucks/busses. For further analysis, data on each day was split into day/night periods of time; data acquired between 06:00 and 19:59 was considered as the day period and data acquired 20:00 and 05:59 was considered as night period.

Hypothesis testing was used to prove whether there are differences in mean speed and standard deviation within the vehicle classes and time of day. Pairwise t-test was selected to analyse the mean samples between the pairwise group and F-test used to compare the two variances.

4.5.2 Data partition and prediction

Based on the SOM algorithm described in the previous section (4.4.1), the SOM network was trained with 3 dimensional inputs: speed, time of day and type of vehicle. In this study, the type of vehicles is mainly based on the length of vehicle detected by the radar. A brief description of the classifiers is provided for the sake of completion.

4.5.2.1 Random Forest

A random forest (RF) is an example of ensemble machine learning, proposed by Leo Breiman, for building tree predictors and letting them vote for the most popular class. The algorithm for inducing a random forest is based on bootstrap aggregation or so called bagging. For bagging, given a training set $X = x_1, \ldots, x_n$ with response variables $Y = y_1, \ldots, y_n$, B times selects ntree a random sample with replacement (bootstrap sample) from the original training set. For each bootstrap sample, an unpruned regression tree grows where at each node a random sample mtry of the predictors is also selected to choose the best split from among those variables. After training, predictions can be made by averaging the predictions from all ntree regression trees. However, the main advantage of this algorithm is the ability to improve stability and accuracy by reducing variance and helping to avoid over-fitting, which is common within other machine learning algorithms (James et al., 2013). RF differs in only one way from the general bagging process. The algorithm uses a modified tree learning algorithm that selects, at each candidate split in the
learning process, a random subset of the features. This process is also called feature bagging. The reason for this is to overcome the correlation of the trees from an ordinary bootstrap sample. A detailed description can be found in (Breiman, 2001; Breiman et al., 1984). For the implementation of the RF, the number of trees and the number of selected features were experimentally tuned. As the number of trees increases, the error converges to a limit where there is no presence of over fitting as in the case of other learning algorithms. The most important parameter is to decide upon the number of features to test at each split. A common practice is to start with a large number and then either increase or decrease the number of features until the minimum error for the prediction is obtained.

4.5.2.2 Adaptive Neuro Fuzzy Inference System

An adaptive neuro fuzzy system is a powerful system that combines the concepts of two approaches into one integrated system, where ANN learning algorithms are used to determine the parameters of the fuzzy inference system to share data structures and knowledge representations. A typical ANFIS structure, proposed first by Jang (1993), consists of 5 layers of nodes. In the first layer, the input membership functions are mapped to a set of rules in the second layer, which, in turn, are mapped to a set of output characteristics in the third layer. The output characteristics are mapped to output membership functions in the fourth layer and finally, the output memberships functions are mapped to a single output (Khoshnevisan, 2014; Chang, & Chang, 2013; Jang, & Sun, Ullah, & Choudhury, 2013). The fuzzy inference system is based on Takagi Sugeno’s methodology where the output membership functions have a constant value. The network is trained using a hybrid learning algorithm based on two steps. In the first step (forward pass), the premise parameters i.e. network parameters are kept fixed and the information is propagated forward in the network using the least square method to identify the consequent parameters for the current cycle through training. In the second step (backward pass), the error is propagated backward while the premise parameters are modified using the gradient descent method by keeping the consequent parameters fixed. The rule extraction method first uses the Fuzzy c-means (FCM) clustering function known as ‘genfis3’ to determine the number of rules and membership functions for the antecedents and consequents. Fuzzy c-means (FCM) clustering technique (genfis3) was also used to optimise the results by extracting a set of rules that models the data and generates an initial FIS for ANFIS training.
5 Results

This chapter summarises the results of paper II through to paper VI.

5.1 Paper II

Standard calibration, data driven calibration and experimental calibration methods were compared in this study. Two data sets were used to be able to identify the CF (see section 4.1). The results obtained in paper II are summarised in table 2. The data driven CF obtained using the ‘clean’ data was superior to the CF retrieved using the ‘complete’ data. The CF for the shorter distance (60 metres) was closer to the experimental CF than the CF for the longer distance (100 metres). At this point, it is worth mentioning that the experimental CF in the current case is equal to 1.18. The obtained results were validated by calculating normalised root mean square error (NRMSE), which is basically the square root of the mean square error. It can be concluded from the investigation in paper II that CF obtained using the ‘clean’ data set collected at a shorter distance between the radars could be reported as an appropriate factor while correcting speed data.

Table 2. Comparison between different correction factors obtained

<table>
<thead>
<tr>
<th>Data type</th>
<th>Distance = 100 metres</th>
<th>Distance = 60 metres</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean</td>
<td>1.26 (0.26)</td>
<td>1.20 (0.19)</td>
</tr>
<tr>
<td>Complete</td>
<td>1.78 (0.48)</td>
<td>3.42 (0.70)</td>
</tr>
</tbody>
</table>

Note: Values in the parentheses are RMSE’s

5.2 Paper III

The results from the hypothesis testing showed that there was strong evidence of a difference in mean speeds between a trigger speed that is set to the 85th percentile and other trigger speeds (see table 3). The effect of trigger speeds on mean speed was found to be significant at both sites (Site-1 and Site-2). There was a greater reduction in the number of speeders when the vehicle activate sign was triggered at the 85th percentile; 4% of the speeders reduced their speed when the TS was set to the 85th percentile as opposed to the speed limit. Additionally, the standard deviation and the coefficient of variation were shown to decrease when the trigger speed was increased. Meanwhile, the lowest standard deviation obtained was also when the trigger speed was set to the 85th percentile. According to these results, an interesting finding was that the optimal trigger speed was approximately near the 85th percentile, which had the desired effect of lowering the standard deviation. Another interesting finding was that standard deviation was high when the trigger speed was set near the speed limit.
Table 3. Tukey multiple comparisons of difference between means and p-value with 95% confidence level at site-1 (85th percentile speed is 50 km/hr)

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>40</th>
<th>44</th>
<th>46</th>
<th>47</th>
<th>49</th>
<th>50</th>
<th>52</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>-0.63</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p=0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>0.98</td>
<td>1.60</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p=0.19</td>
<td>p&lt;0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>0.78</td>
<td>1.41</td>
<td>-0.20</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.42</td>
<td>p&lt;0.01</td>
<td>p=1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>1.72</td>
<td>2.34</td>
<td>0.74</td>
<td>0.94</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p=0.62</td>
<td>p=0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>1.60</td>
<td>2.26</td>
<td>0.62</td>
<td>0.82</td>
<td>-0.12</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p=0.82</td>
<td>p=0.42</td>
<td>p=1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>2.74</td>
<td>3.36</td>
<td>1.76</td>
<td>1.96</td>
<td>1.02</td>
<td>1.14</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p=0.01</td>
<td>p=0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>1.53</td>
<td>2.16</td>
<td>0.55</td>
<td>0.75</td>
<td>-0.19</td>
<td>-0.07</td>
<td>-1.21</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p=1.00</td>
<td>p=1.00</td>
<td>p=0.02</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>4.12</td>
<td>4.75</td>
<td>3.15</td>
<td>3.34</td>
<td>2.41</td>
<td>2.52</td>
<td>1.38</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
<td>p&lt;0.01</td>
</tr>
</tbody>
</table>

### 5.3 Paper IV

An overall reduction in mean speed and standard deviation was observed at both sites and with both signs. The results in table 5 showed that the trigger speed of the SID had no detectable effect on driver mean speed. In contrast, the trigger speed of the sign had an effect on drivers’ speed, particularly on the local road. Further, a lower coefficient of variation was observed when the trigger speed was set to the 85th percentile speed. On the local road (Site-1), the speed indicator device had a greater effect than the vehicle activated sign; however, in contrast, their effectiveness was comparable when tested on highways (Site-2). The results obtained from ANOVA tests proved the hypotheses that when studying the effect of VAS and SID on drivers’ speed, time of day and type of site are significant. Another interesting finding was that the main effects of vehicle speeds cannot be interpreted without considering the effect of interaction between time of day and type of site.
Table 4. Coefficient of variation and effect of VAS and SID at Site-1

<table>
<thead>
<tr>
<th>Time</th>
<th>Sign</th>
<th>CV at sign</th>
<th>δ at sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>SID-SL</td>
<td>0.14</td>
<td>-5.4</td>
</tr>
<tr>
<td></td>
<td>SID-85</td>
<td>0.15</td>
<td>-5.6</td>
</tr>
<tr>
<td></td>
<td>VAS-SL</td>
<td>0.17</td>
<td>-3.3</td>
</tr>
<tr>
<td></td>
<td>VAS-85</td>
<td>0.14</td>
<td>-2.4</td>
</tr>
<tr>
<td>Night</td>
<td>SID-SL</td>
<td>0.14</td>
<td>-5.3</td>
</tr>
<tr>
<td></td>
<td>SID-85</td>
<td>0.16</td>
<td>-5.8</td>
</tr>
<tr>
<td></td>
<td>VAS-SL</td>
<td>0.20</td>
<td>-2.6</td>
</tr>
<tr>
<td></td>
<td>VAS-85</td>
<td>0.16</td>
<td>-1.9</td>
</tr>
</tbody>
</table>

Table 5. Coefficient of variation and effect of VAS and SID at Site-2

<table>
<thead>
<tr>
<th>Time</th>
<th>Sign</th>
<th>CV at sign</th>
<th>δ at sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day</td>
<td>SID-SL</td>
<td>0.14</td>
<td>-8.2</td>
</tr>
<tr>
<td></td>
<td>SID-85</td>
<td>0.15</td>
<td>-8.2</td>
</tr>
<tr>
<td></td>
<td>VAS-SL</td>
<td>0.14</td>
<td>-8.0</td>
</tr>
<tr>
<td></td>
<td>VAS-85</td>
<td>0.13</td>
<td>-8.6</td>
</tr>
<tr>
<td>Night</td>
<td>SID-SL</td>
<td>0.15</td>
<td>-8.4</td>
</tr>
<tr>
<td></td>
<td>SID-85</td>
<td>0.18</td>
<td>-8.0</td>
</tr>
<tr>
<td></td>
<td>VAS-SL</td>
<td>0.14</td>
<td>-8.1</td>
</tr>
<tr>
<td></td>
<td>VAS-85</td>
<td>0.16</td>
<td>-9.0</td>
</tr>
</tbody>
</table>

Note:
SID-SL: Speed Indicator Device with a message “Your Speed” and with trigger speed set to speed limit
SID-85: Speed Indicator Device with a message “Your Speed” and with trigger speed set to 85th percentile speeds
VAS-SL: Vehicle Activated Sign with a message “Slow Down” and with trigger speed set to speed limit
VAS-85: Vehicle Activated Sign with a message “Slow Down” and with trigger speed set to 85th percentile speeds

5.4 Paper V

The results showed that the best clustering of the data was with the choice of \( k = 3 \), based on partitioning the time of day into three clusters that correspond to the number of vehicles passing the sign. The partition of night time was identified at both sites within the cluster between 22:00 and 06:59. The day time was partitioned into two clusters that were not similar at the two
sites, but the centroid mean speed and the centroid standard deviation of the two clusters were slightly different. For each cluster, the centroid mean speed was the expected trigger speed of the vehicle activated sign. In this context, the day clusters could be combined when triggering the sign with the centroid mean speed.

Table 6. Speed reduction and power consumption achieved by setting the trigger speed to 15th, 50th and 85th percentiles at Site-1 and Site-2 compared to the trigger speed determined by clustering

<table>
<thead>
<tr>
<th>Site</th>
<th>Trigger speed set at</th>
<th>Speed reduction at day time (km/hr)</th>
<th>Speed reduction at night time (km/hr)</th>
<th>Power consumption (Ah)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15th percentile speeds</td>
<td>4</td>
<td>7</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>50th percentile speeds</td>
<td>6</td>
<td>6</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>85th percentile speeds</td>
<td>3</td>
<td>8</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>determined by clustering</td>
<td>6</td>
<td>8</td>
<td>1.75</td>
</tr>
<tr>
<td>2</td>
<td>15th percentile speeds</td>
<td>7</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>50th percentile speeds</td>
<td>8</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>85th percentile speeds</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>determined by clustering</td>
<td>2</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

The effect of the trigger speed of the clustering models was validated by comparing the speed to other static trigger speeds based on 15th, 50th and 85th percentile speed of the vehicle. The results showed that at night time the clustering model had the greatest speed reduction, which was similar to the 85th percentile speeds at one site and to 15th percentile at another (see table 6). The energy consumed was based on the total number of activation of the sign. In this study, the energy consumed by the clustering model was actually near the 50th percentile trigger speed and larger than 85th percentile speed. Triggering the sign with 85th percentile at night could be the best compromise to reduce the power consumption of the VAS while increasing a reduction of vehicles’ speed.

5.5 Paper VI

Hypotheses testing confirmed that there were significant differences between mean speeds and speed variances for different type of vehicles and at different time of the day i.e. the trigger speed of a VAS cannot be static and must be altered depending on the type of vehicle, time of the day and its location. At each site (data 4), four different traffic patterns were clustered using SOM. A SOM network was trained with 3 dimensional inputs based on speed, time of the day and length of vehicle. However, within the four clus-
ters obtained by SOM, and at each of the sites, the results of the predictive models showed that RF performed better than ANFIS (see table 7).

Table 7. Performance respective to the RMSE and time for the two models RF and adaptive neuro fuzzy systems (ANFIS) within the four cluster clusters at the four sites

<table>
<thead>
<tr>
<th>Site</th>
<th>Clusters</th>
<th>RF</th>
<th>ANFIS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>Computational time (seconds)</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.07</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.09</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.14</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.11</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.22</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.13</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.06</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.11</td>
<td>0.57</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.29</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.10</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.12</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.11</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.11</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.07</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.09</td>
<td>0.34</td>
</tr>
</tbody>
</table>

As it can be seen the prediction accuracy of RF and ANFIS tracks the actual speed profile smoothly, i.e. excluding the inappropriate speeds (too fast or too slow) from the dataset (see figure 3). Given these results, it is worth pointing out that RF was an adequate model to predict the trigger speed for a VAS in terms of computational performance (shorter calculation time) and in terms of efficiency (lower root mean square error).

![Figure 3 (a) Speed prediction for next hour using ANFIS; (b) RF at Site-4](image-url)
6 Conclusions and Discussion

This thesis has investigated the effectiveness of the trigger speed of the VAS on driver behaviour. Studies aimed at automatically triggering the VAS according to traffic and road conditions were also investigated. In order to accomplish the objectives, the investigation in this thesis pursued six research questions R1 – R6 (see section 1.2). Conclusions from the individual investigations are as follows.

(R1) What is the trigger speed of the VAS signs that were used in previous studies and furthermore, how effective were the VAS signs with regard to assessing driver behaviour?

The effects of the trigger speed of the VAS were not considered in previous studies. There is no consensus on the setting of the trigger speed of these signs. Additionally, the effect of various road and weather conditions was also not clearly studied. The effectiveness of VAS on driver behaviour was extensively established in two frequent measures: reducing mean speed and reducing standard deviation.

(R2) How should the data be collected from the VAS?

Doppler radar was found to be suitable (subject to appropriate correction) method to collect speed data. CF calibrated in a data driven manner based on the distance between two radars was more accurate than the standard method. This is due to the distance between radars, which could be easily measured and adjusted. It is challenging to validate the proposed correction when there is no absolute measurement for accuracy. Findings from paper II indicate that the experimental CF was an appropriate reference factor for data collection. Another finding is that data cleaning could help in finding a CF that is much closer to the reference CF. Removing outliers and excluding all missing values without any data reconstruction was found to provide more accurate results. An additional finding is that the CF for the radar obtained from the shorter distance was much closer to the reference CF than the factor obtained through the longer distance.

(R3) What is the effect of trigger speed on driver behaviour, and what is the optimal trigger speed?

The trigger speed of VAS has an effect on driver behaviour, as it can be observed by a reduction in mean speed as well as standard deviation. Findings from paper III indicate that setting the trigger speed relative to the speed limit is not a good practice. Standard deviation of vehicle speed is high when the trigger speed is set near the speed limit. The optimal trigger speed near the 85th percentile speed had the desired effect of lowering the standard deviation. It is recommended that the trigger speed be aimed at lowering the standard deviation as opposed to lowering the mean speed of vehicles.
(R4) Is the effect of VAS comparable to the effect of SID on driver behaviour in terms of trigger speed, time of day and type of site?

Findings from the comparative study between the two signs in paper IV showed that the trigger speed of SID did not have any effect on driver behaviour. This finding is related to the type of warning messages displayed on this sign. SID showed the speed of each driver passing the sign. The colour of the message (in red) was the only difference in the information presented to the drivers who exceeded the trigger speed. Another finding showed that the type of road is an essential factor to be considered when studying the effect of these signs on drivers’ speeds. The signs had a variable effect at different sites; SID had a greater effect than the VAS on local roadways but their effectiveness was comparable when tested on highways. At this point it is worth mentioning that the operational and capital costs involved in installing and maintaining either type of signs are substantial.

Time of day was found to be another important factor when assessing the effect of VAS on driver behaviour. Lowest coefficient of variation was observed during the day when the VAS were activated with trigger speed equal to the 85th percentile. In contrast, a low coefficient of variation was observed during the night when the trigger speed was set to the posted speed limit on the road segment. A general finding from this study is that we cannot interpret the main effects of vehicle speeds without considering the interaction effect between time of day and type of site.

(R5) How can the optimal trigger speed of the solar VAS be found in order to reduce drivers’ speed while simultaneously reducing power consumption?

Findings from paper V indicate that the night cluster had an increase in mean speed and standard deviation when compared to the day cluster, indicating the need for a different trigger speed depending on the time of day. This study further reinforces the findings in paper III that the VAS should be individually configured and adapted to the location and traffic conditions. The optimal trigger speed of the VAS was set to the 85th percentile speed, which had the desired effect of lowering the standard deviation. For power consumption, the energy consumed by the clustering model was actually near the 50th percentile trigger speed and larger than 85th percentile speed. To summarise, the trigger speed should be set to the 50th percentile to be able to reduce power consumption and at the 85th percentile to be able to lower the standard deviation.

(R6) How can an automatic trigger speed for VAS, which responds to traffic conditions and to different type of vehicles be assessed? Are different VAS trigger speeds required for different vehicle types?

Findings from paper VI indicate that different classes of vehicles behave differently given different road and traffic conditions. On a highway, the mean
speed of cars differs slightly when compared to the mean speed of light trucks (or vans) and heavy trucks, whereas on local roads there are differences between the classes in both mean speed and standard deviation. The optimal trigger speed will need to be pre-determined according to the nature of the site and to the type of vehicle. A differential trigger speed of the VAS should be considered particularly when there is a difference in mean speed or standard deviation between the different classes. RF was found to appropriately predict the trigger speed of the VAS. The fact that the predicted trigger speed was found to be consistently around the 85th percentile speed reinforces previous findings.

6.1 Discussion

The optimal trigger speed was approximately near the 85th percentile speeds, which had the desired effect of lowering the standard deviation. In other words, the 85th percentile had the desired effect of lowering the crash rate. According to previous work, when driving speeds and standard deviation are inappropriate to traffic conditions, the risk of traffic crash becomes high (Wegman, & Aarts, 2006). Following high speed variation leads to a higher crash rate, also to an increase in number of severe injuries and traffic fatalities. The relationship between vehicle speeds and risk of crashing is often referred to as being ‘U-shaped’. This was first described by Solomon (1964). Referring to Solomon study, crashes happened when driver travelled with speed either above or below the average speed under normal conditions. Garber and Gadiraju (1989), Quddus 2013, McMurtry et al. 2009, Garber and Earhart 2000 also showed that a higher speed variation results in higher crash rates.

Setting the trigger speed of VAS could be related to the major problem in setting the appropriate speed limit for a specific road segment. There are numerous guidelines and methodology used throughout the world for setting the appropriate speed limit, but there is no consensus concerning the methods and approaches that should be used to select the most appropriate speed limit. Initially, in Europe, speed limits were set in accordance to the 85th percentile to reflect a typical driver’s behaviour. It was argued that setting the speed limits lower than the 85th percentile speed does not encourage driver compliance with the posted speed limit (Forbey et al., 2012). In Sweden, the setting of the speed limit is however based on a safety and injury system approach that aims to reduce the number of road fatalities to zero (Box, & Bayliss, 2012). In this approach, avoiding death and injury is the absolute priority considered when setting the speed limit. Following the speed limit is identified according to the crash types and the tolerance of the human body during a crash. Compared to other countries in Europe, this approach results in a lower speed limit.

Setting the trigger speed of VAS to type of vehicle could also be related to the general problem of the strategy of a different speed limit for trucks. A number of authorities have encountered this general problem and used the DSL
strategies by either setting the maximum speed for trucks near a certain threshold value, such as 10–15 km/h lower than cars, or by requiring all trucks to be equipped with maximum speed limiters (Harwood, 2003; Garber et al., 2005). The safety effects of such settings have been inconclusive in previous studies. Some studies found no difference between DSL and uniform speed limit for cars and trucks (Harkey, & Mera, 1994; Hall, & Dickinson, 1974; Idaho transportation, 2002; Freedman, & Williams, 1992). Other studies found that DSL can be a better policy choice (Duncan et al., 1998; Ghods et al., 2012). In this thesis, it was proved that the speed of different types of vehicles might be highly correlated with the type of road. On highways, cars and long trucks were slightly different in speed but on local roads trucks travelled at lower speeds than cars. Apart from that, in this thesis we found the effect of the trigger speed on highways to be lower than on local roads although the traffic flow between the two roads was not comparable. In this context, a differential VAS trigger speed should be considered, particularly when trucks and cars have different speed characteristics. A uniform optimal trigger speed could be more appropriate on highways, and a differential trigger speed could be a good recommendation on local roads particularly where the trigger speed for trucks is at least 10 km/hr lower than for cars.

Regarding traffic conditions, time of day is proven to be an essential factor when studying the effectiveness of VAS on driver behaviour. The time of day could be considered in greater period round the year, but in this context the time of the day will incorporate the weather conditions. In a previous Swedish study, it was reported that the variation of number of deaths in traffic, respective to time of day, particularly in different months and on different days are large. It was also shown that the highest number of people killed was on Saturdays in July and August, but the lowest number of deaths was in March, followed by January and February. In this thesis, time of day was considered as day and night periods were somewhat challenging when the number of hours of daylight varied to a great extent. Therefore, based on the traffic flow, a clustering technique was an appropriate technique to divide the time of day into two main periods: low flow for the night time and high flow for day time.

The type of road is often the main factor considered in Sweden for setting speed limit for a specific road segment. Thus, finding the optimal trigger speed of VAS could be an important tool for setting appropriate speed limit, particularly for road segments known for speeding. One recommendation from this study is to use this type of sign before lowering or raising the speed limit posted for a specific road segment. Bearing in mind that changing the trigger speed of this sign might confuse drivers when the message of this sign consisted of “Reduce Speed” with the posted speed limit; therefore, a good recommendation is to eliminate the speed limit from the display of these signs.
Similar to the VMS approach, dynamic VAS could be an appropriate approach for pre-determining according to road and traffic conditions. At the same time, developing VAS to be dynamic signs requires fast processing, analysing and storing of a large amount of traffic data, making the system expensive and challenging to be implemented. To cope with real time traffic management and time lags, an accurate predictive model based on historical traffic speed data was proposed in paper VI. Using RF with a simplistic input with one hour back into the past could smoothly predict the speed in the present. Additionally, such models excluded the inappropriate speeds of occasional vehicles travelling very fast or slowly because such occasional vehicles can distort speed predictions. It is worth mentioning that this thesis did not take advantage of the power of RF when the input number was not high. If research is expanded to consider such factors, including the geometry of the road and weather conditions, RF may produce similar promising results.

6.2 Limitations and future work

To generalise our findings, data collection at different sites particularly on highways needs to be explored. Data collection in this thesis was the most challenging task, which took much effort from the research team because of a number of problems:

- The installation of the VAS and the SID at highway sites was challenging due to the need for agreement between several traffic agencies. The effect of these signs at highway locations was unknown. Such signs are often installed on local roads.

- The safety impacts of a differential trigger speed for trucks were not validated due to the problem faced in the data collection. The selection of a site with high flow of trucks is required to provide the 85th percentile speed at different time periods. It was experienced that the flow at the two periods: night and day was not balanced. However, a highway is supposed to have a higher rate of traffic flow than local roads and could offer better potential for assessing the effectiveness of trigger speed respective to vehicle type.

- The VAS used in this study were equipped with two radars and a data logger to detect and record vehicle speed and a modem to facilitate communication to the radar. Such a setup facilitates alteration of trigger speed, thereby permitting the study to investigate the effect of different trigger speeds on various driving speeds. This equipment embedded in the sign provides the limitation of the sign to be powered by solar panels. The solar panels were not in use in this study due to their limited power capacity. Therefore, the restriction of the position and direction of the sign was limited to the presence of a power supply.

- The experimental CF i.e. driving and recording speed was not an easy task for the calibration of the radar. More work is needed in the valida-
tion for the data driven calibration of the radar used in the experiment.

- Long-term effects of these signs, particularly at highway locations, were not observed in this thesis. The sign should be installed for a longer time at the road, and the data collection should be done at different periods of time.

- The effect of a differential trigger speed was not validated in this study due to time. Driver behaviour was measured by the mean and standard deviation of speed. Other measures should be explored with regard to this issue.

- Other predictive models such a Gaussian process based on kernel based learning could be useful in speed prediction, which is based on non-linear regression problems.
References


Appendix A: Appended papers