The application of vehicle classification, vehicle-to-infrastructure communication and a car-following model to single intersection traffic signal control

Joel Andrew Dodsworth

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Institute for Transport Studies

August 2018

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is their own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The first algorithm described in Chapter 4 is based on work published in:

Dodsworth, J., Shepherd, S. and Liu, R. 2014. Real-time single detector vehicle classification. *Transportation Research Procedia*. **3**, pp.942-951.

Chapter 5 is based on the following work currently being considered for publication in *Transportmetrica B*:

Dodsworth, J., Shepherd, S. and Liu, R. 2018. Optimizing Isolated Signal Control using V2I and Vehicle Classification Information – A Hybrid Model.

The original work contained within these papers is all the candidate's own work with guidance provided by Shepherd and Liu.

Acknowledgements

All work has been undertaken by the candidate under the specific guidance of Professor Simon Shepherd and Professor Ronghui Liu. The contrasting contributions of Simon and Ronghui have been interesting, challenging and encouraging and have helped to shape this thesis.

All simulation work has been undertaken using the PTV-Vissim software package. Special thanks to PTV for supplying the software at no cost for such a long period of time, it is hugely appreciated.

Many of the ideas in my work were formed during the first years of working as a practitioner in the design, commissioning and validation of traffic signal controlled junctions. The mentoring provided by Andy Poole, particularly in providing the opportunities to work with on-line responsive control strategies, was invaluable to developing my understanding. I also thank WSP for sponsoring the initial years of my study and allowing me the flexibility of working part-time.

Finally, I thank my family and friends for giving me the time and space to complete my work. This study would not have been possible without your enormous patience and support.

Abstract

On-line responsive traffic signal optimization strategies most commonly use data received from loop detectors to feed information into an underlying traffic model. The limited data available from conventional detection systems has dictated the way that current 'state-of-the-art' traffic signal control systems have been developed. Such systems tend to consider traffic as having homogenous properties to avoid the requirement for more detailed knowledge of individual vehicle properties. However, a consequence of this simplification is to limit an optimizer in achieving its objectives.

The first element of this study investigates whether additional data regarding vehicle type can be reliably extracted from conventional detection to improve optimizer performance using existing infrastructure. A single detector classification algorithm is developed and it is shown that, using a modification of an existing state-of-the-art optimization method, a modest improvement in performance can be achieved.

The emergence of connected vehicle technology and, in particular, Vehicle-to-Infrastructure (V2I) communications promises more comprehensive data. V2I-based optimization methods proposed in literature require a minimum penetration rate of V2I equipped vehicles before performance matches existing systems. To address this problem, the second part of the study focuses on the development of a hybrid detection model that is capable of simultaneously using information from conventional and V2I detection. It is demonstrated that the hybrid detection model can begin to realise benefits as soon as V2I data becomes available. V2I-based vehicle classification is then applied to the developed hybrid model and significant benefits are demonstrated for HGVs.

The final section of the thesis introduces the use of a more sophisticated internal traffic model and a new optimization method is developed to implement it. The car-following model based optimization method addresses the lack of modelled interaction between vehicles and is shown to be capable of reducing vehicle stops over and above the developed (vertical queue based) hybrid model.

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

Demand has outstripped supply for decades when considering the capacity of road networks in many major cities in the UK and across the world. The fact that commercial traffic data providers such as TomTom (2018) and INRIX (2018) make headlines by producing annual congestion rankings for major cities is symptomatic of the problem.

Transport policy in the UK, and elsewhere to varying degrees, has historically focused on providing ever increasing capacity to meet demand, often creating barriers between communities to accommodate the increasingly imposing infrastructure. However, it is now starting to be recognised, at least in some cities, that this policy is not sustainable and that economic growth should be pursued in a more inclusive manner (Leeds City Council, 2018). Furthermore, providing additional capacity tends simply to induce further demand and, as air quality becomes an increasingly prominent topic, this is not a desirable outcome. Papageorgiou et al. (2007) provides a well presented reasoning for why the continuous expansion of transportation infrastructure cannot be the only solution to increasing demand. The discussion, as part of a review of traffic control strategies, focuses on the impact of congestion on economic competitiveness.

The proportion of road traffic miles driven in the UK by each vehicle class has changed considerably in recent years with a 5% increase in the share of 'light commercial vehicle' miles from 1995 to 2015 after a roughly constant 10% share in the preceding 30 years (Department for Transport, 2018). This change largely reflects changing consumer habits and is an example of how cities must adapt

management and control of traffic to cater for changes in travel behaviour, whether policy-led or driven by external factors.

This research considers the often studied topic of traffic signal control in the context of the shifting attitudes of policy makers. Initially this introductory chapter provides a background to the use of traffic signal control and discusses issues that can limit the performance of optimizers in achieving chosen objectives. The emergence of connected vehicle technology is then discussed before the scope and objectives of the research are set out. Finally, an outline of the thesis is presented and summarised.

1.1 Motivation

Traffic signals have become an increasingly important tool for managing traffic flow as demand has increased. The use of traffic signal control offers a means of improving safety at junctions with poor visibility between conflicting vehicle movements and at locations where pedestrians come into conflict with other road users. However, aside from safety implications, it can also be used to manage traffic flow more strategically to achieve desired policy objectives.

Since the first traffic signal control installations, methods of control have evolved to accommodate increasing demand and to take advantage of advances in technology. The concept of Urban Traffic Control (UTC), i.e. managing traffic from a centralized system, has become widespread in the past half century to the point that most cities and major towns across the world manage traffic signals using a UTC system. Alongside the technological developments in communications and computing, there has also been a variety of control strategies developed in order to optimize the operation of the signals. Some strategies have been developed specifically for isolated junctions whilst others can be applied to much larger networks.

A comprehensive review of the history of traffic signal control is presented by Hamilton et al. (2013) whilst other reviews by Heydecker (2004) and Papageorgiou et al. (2007) also provide useful insights. Papageorgiou et al. (2007), for example, summarise the difficulties involved in optimizing traffic signals and conclude that the solution to the optimization problem is infeasible for more than one intersection, thus requiring simplifications to be introduced. Unfortunately, the simplifications often reduce the effectiveness of strategies in saturated conditions.

Often control strategies consider network (or area) wide optimization, and work by synchronizing signal cycle times across either a group of junctions or the whole network. The green splits and offsets between adjacent junctions are then adjusted, either dynamically or via offline optimization, usually with the aim of providing coordination and minimizing delay to traffic on strategically important routes. At isolated junctions a more dynamic approach to optimization can often be taken without fear of decisions in the optimization process causing adverse effects at upstream or downstream junctions.

Minimizing delay is a common aim of traffic signal control but is just one of the many possible objective functions that can be considered. Shepherd (1994) lists minimization of delays to public transport, minimization of delays to pedestrians, maximizing reliability and minimizing environmental impacts as examples of objective functions. The effectiveness of an optimization method in achieving its objective is determined by the quality of available data, the representativeness of the traffic model it is using and the efficiency of the optimization algorithm itself.

'Intelligent' methods of traffic signal optimization came to prominence with the development of the 'offline' tool TRANSYT (Robertson, 1969). TRANSYT was developed as a tool to optimize fixed-time plans for networks of junctions and,

although very successful and still widely used today, it was recognized that there are various limitations to fixed-time control. The performance of plans naturally deteriorates over time due to changes in traffic patterns and demand (Bell and Bretherton, 1986). Updating fixed-time plans to address performance degradation requires large amounts of traffic flow information to be gathered at regular intervals for generating new plan timings. Regular revalidation exercises on that scale are an unrealistic prospect for many Local Authorities. Fixed-time plans are also limited in that they cannot take into account day to day fluctuations in demand or the random nature of arrivals meaning that even optimal fixed-time plans can take a long time to recover from perturbations.

Some Local Authorities automate changes to sets of fixed-time plans by monitoring flow and occupancy on the network. This enables timings to be adapted in the case of events and in response to changes in traffic conditions (Reid, 2007). Introducing different fixed-time plans in response to network conditions can provide some benefit compared to fixed-time alone but is still limited by the need for regular revalidation of the plan timings and meticulous configuration of the flow and occupancy thresholds.

Many of the ideas of TRANSYT were subsequently incorporated into the 'on-line' optimization tool SCOOT (Hunt et al., 1982). SCOOT (Split Cycle Offset Optimization Technique) makes use of vehicle detectors to provide real-time flow information to the optimization process, addressing the previously described problem of plan degradation. SCOOT was primarily developed as a network-wide optimization tool (although optimization is actually split into smaller 'regions'). In SCOOT the green splits, cycle time and offsets for each region are updated at regular intervals to adapt the signal timings to the current traffic conditions. SCOOT is used widely in the UK and in many cities

internationally but there are various other network focused control strategies that have been developed. These will be reviewed in more detail in Chapter 2.

Network-wide control strategies are popular in dense road networks as they enable coordination between closely spaced junctions. However, the benefit of coordination reduces as the distance between junctions increases, dependent on factors that lead to platoon dispersion such as the 'friction' along the connecting links as well as, for example, the homogeneity of vehicle mix/driver behaviour. Where platoon dispersion is significant, junctions may be considered in isolation. The single isolated junction is a simpler problem to solve than for a network but is not entirely straightforward.

In the UK it is still common to operate the standard System-D vehicle actuated (VA) strategy (Highways Agency, 2005) at isolated junctions. System-D VA is a demand responsive system that enables a green signal to be terminated early if all traffic has been cleared. The system is not capable of systematic optimization but remains popular at junctions where, for example, signal control has been introduced to improve safety but demand does not necessarily exceed capacity.

Subsequent developments such as the popular Microprocessor Optimised Vehicle Actuation (MOVA) control strategy (Vincent and Peirce, 1988) introduced systematic optimization. In MOVA, the optimizer considers the intensity of approaching traffic rather than simply whether there are vehicles present or not. The MOVA strategy is very popular in the UK with over 1000 junctions operating the system in 2005 (Crabtree, M.R.K., J.V., 2005).

1.2 Summary of limitations to optimization

One of the historical limitations to traffic signal optimization has been the information available to the internal traffic model to build an accurate

representation of traffic conditions. Ordinarily, for adaptive control strategies, traffic detectors provide a snapshot of vehicles passing a fixed point from which an on-line model is constructed. It is expensive to install and maintain the inductive loop detectors and, consequently, the number of detectors required for most strategies is kept to a minimum; often just a single detector per lane of an approach.

It has been shown that humans (and human learning algorithms) can outperform 'conventional' algorithms at traffic signal control (Quinn et al., 1988; Box and Waterson, 2012). One of the reasons for this is that humans are able to view the wider traffic situation and make a more informed decision on when to change the signals. The emergence of more sophisticated and reliable video detection and the on-going development of cooperative systems, including vehicle to infrastructure (V2I) technology, offers the possibility of more accurate representation of traffic conditions.

A consequence of the limitations of conventional vehicle detection is that various aspects of traffic behaviour must be approximated or homogenised. For example, the vehicle mix at traffic signals can significantly affect the discharge rate but is usually only represented in control strategies as a fixed, user-defined value because information on vehicle type is not available. In reality the vehicle mix will vary from one cycle to the next and, consequently, so will the rate of discharge. Other parameters such as travel-time (i.e. from an upstream detector to the traffic signal stop line) are fixed values based on observations by traffic engineers over a limited period of time. Ultimately, the variation of actual values compared to the configured fixed values limits the ability of an optimizer to reach an optimal solution.

As already discussed, the ability (or lack thereof) to discern vehicle type on an approach to a junction can affect the performance of an optimizer. A lack of knowledge of vehicle type from conventional detection also results in all vehicles being treated equally by optimization strategies when considering vehicle stops. Unfortunately, this methodology can have inequitable results as vehicles with poorer braking/acceleration performance (and generally larger headways to preceding vehicles) are more likely to be stopped by optimization strategies that consider the efficiency of flow. It is worth noting that this is not a desirable outcome given the disproportionate volume of emissions from larger vehicles compared to the number of miles travelled (Figure 1.1)

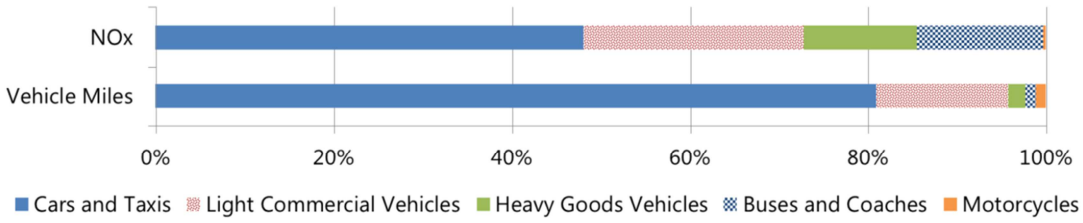

Figure 1.1: Comparison of urban vehicle miles driven by each vehicle class compared to urban NOx emissions in the UK for 2015 (Department for Transport (2018) and NAEI (2018)).

1.3 Emergence of connected vehicle technology

The discussion so far has identified that there are limitations with conventional detection that necessitate simplifications in the internal traffic models of traffic signal optimizers. As a consequence of those simplifications, the effectiveness of an optimizer in achieving objectives related to minimizing emissions is limited.

As a simple example, if vehicles with high emissions can be stopped less frequently then it is likely to be beneficial in terms of air quality. However, in order to stop specific vehicle classes less frequently requires more information than is currently received by traffic signal control strategies from conventional detection. Additionally, it is not desirable for vehicles with low acceleration to be

positioned at the front of a queue as it is likely to constrain the performance of the vehicles behind it at the start of the following green but, as described previously, this impact cannot be evaluated with simplified traffic models.

Connected vehicle technology is an emerging technology that is often referred to as connected and autonomous vehicle (CAV) technology. However, to combine the two elements is to disregard the benefit that can be provided by connected vehicles with little or no autonomy. Despite much attention in the media, it is likely to be at least 2025 before fully autonomous vehicles will begin to operate with full integration on public highways (Adams, 2017) and possibly another decade before autonomous vehicles are a significant proportion of the vehicle fleet. In the meantime, connected vehicle technology offers the opportunity to realise significant benefits in the shorter term.

Vehicle-to-infrastructure (V2I) technology, an aspect of connected vehicle technology, is the wireless transfer of information from a vehicle to roadside infrastructure such as traffic signals. V2I introduces the potential to obtain information on vehicle position, speed and type (and many more properties if the information is made available) to utilise in an optimization strategy. Consequently, V2I provides an opportunity to address the shortcomings of conventional detection that limit the effectiveness of traffic signal optimizers. It has been predicted that almost all new cars will be connected by 2020 (Gissler, 2015) although the term ‘connected’ is often used to refer to any connection rather than V2I specifically. Currently there are very few V2I enabled vehicles and it is therefore important to ensure that traffic signal control strategies are capable of maintaining existing performance levels whilst taking advantage of V2I data where available.

1.4 Scope and objectives of research

This thesis tackles issues related to the limitations of traffic signal optimization caused by the quality (and nature) of available data and the simplified representation of traffic conditions. In particular, in the context of the discussed air quality issues surrounding road transport, the thesis seeks to quantify the benefit in terms of stops and delay that can be obtained by providing a more detailed representation of traffic to an optimizer. The overall **objectives of the study** can be summarised as follows:

1) To investigate the feasibility of supplying vehicle type information to an optimizer, using only existing infrastructure, in order to test whether knowledge of individual vehicle type can be used to improve performance of an existing signal optimizer in terms of minimizing vehicle stops and delay;
2) To understand whether any performance benefit can be achieved by applying V2I detection to an existing optimization strategy, taking into account the need to accommodate initially low V2I penetration rates;
3) If the use of V2I detection provides a performance benefit, then a subsequent objective is to assess whether any further benefit can be provided by including vehicle type information, derived through V2I, in a modified optimizer; and
4) To develop an optimization method capable of using a more sophisticated representation of traffic to assess whether it can provide a benefit in terms of overall optimizer performance compared to a conventional method.

The study has been conducted using a Simulated Environment, described in Chapter 3, which enables comparison of optimizer performance in repeatable conditions.

The initial investigation of this thesis (Chapter 4) is focused on whether the appropriate information (such as vehicle speed and type), required to enable an optimizer to accurately estimate delay and the impact of stops, can be reliably extracted from conventional detection using existing infrastructure. This aspect of the thesis is undertaken to understand whether, within the constraints of conventional detection, any significant benefit can be achieved without the introduction of additional hardware.

Additional hardware in this context refers to hardware that would be required to estimate speed and vehicle type from conventional detectors. The reason that additional hardware has not been considered (although relevant literature has been reviewed), is that it has been assumed such investments would not be realistic given the emergence of connected vehicle technology and more sophisticated video detection techniques.

In the second element of the thesis (Chapter 5), the benefit to the optimizer of introducing connected vehicle technology by simulating the use of V2I communication is demonstrated. First, V2I based detection is applied to an existing optimizer by developing a traffic model capable of using conventional detection and V2I data simultaneously. The existing optimization method, a representation of the popular MOVA strategy, is introduced and described in detail in Chapter 3. Subsequently, vehicle type information (assumed to be supplied through V2I) is applied by modifying the MOVA representation.

In the third part of the thesis (Chapter 6), a new optimization method is developed to enable a more sophisticated traffic model to be accommodated. The construction of the optimizer is described and the optimization method is compared to the MOVA representation with and without V2I based detection.

This research focusses on on-line signal optimization at isolated junctions. That focus certainly does not preclude the principles of this study from being applied to network control strategies. However, by focusing on isolated junction control, it has allowed a representation of the existing, well-established, MOVA control strategy to be developed within the timescales of this study. The MOVA representation provides a robust benchmark for all subsequent testing within the thesis.

1.5 Outline of thesis

Figure 1.2 shows the outline of this thesis. A more detailed structure of how the various aspects of the study interact is provided in Chapter 3.

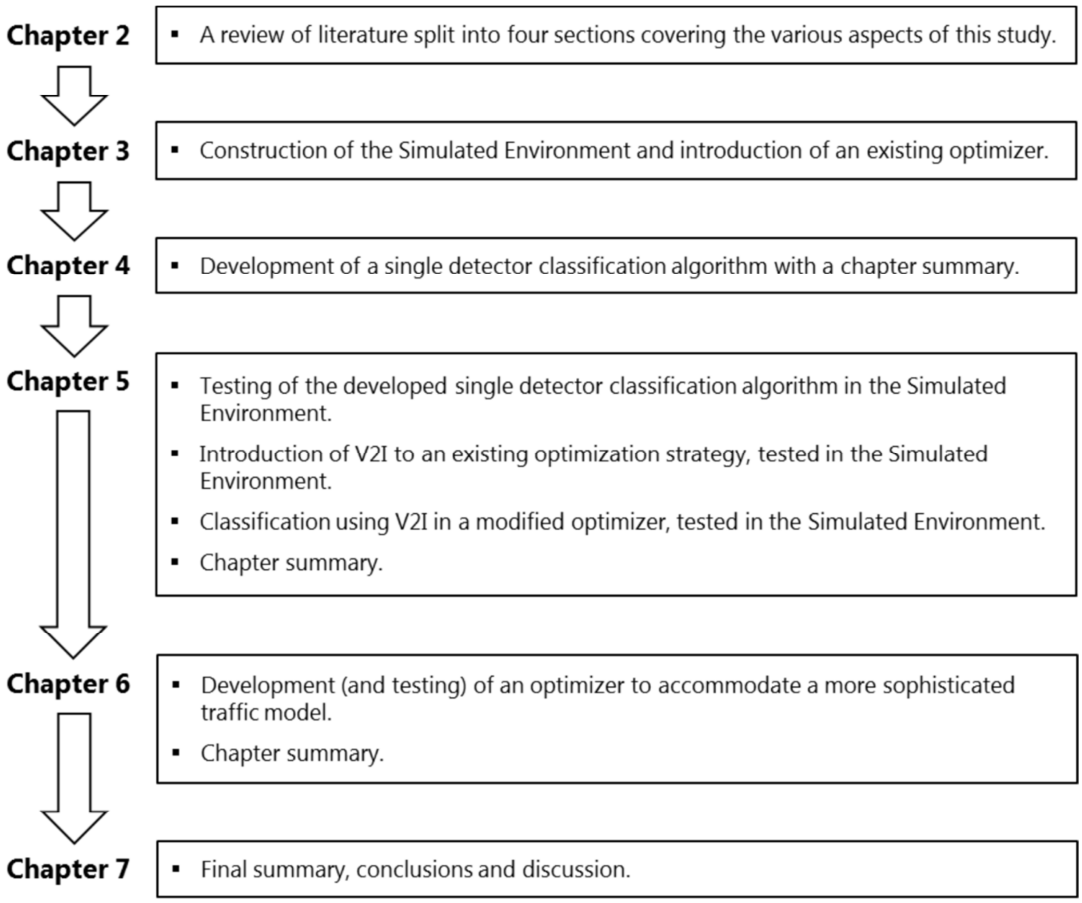

Figure 1.2: Outline of thesis.

At the end of each of the main chapters (i.e. Chapter 4, Chapter 5 and Chapter 6) there is a summary of the outcomes. Chapter 7 draws the conclusions from each chapter together and discusses the findings and potential opportunities for further work.

Chapter 2
Literature review

2.1 Introduction

This chapter is divided into four sections that cover:

- Single detector vehicle classification;
- Vehicle classification in a traffic signal control context;
- Connected vehicle technology in traffic signal control systems; and
- On-line traffic models.

Each of the four sections is a distinct discipline but together they form the foundation of this research.

At the start of the second section there is a review of traffic signal control strategies that provides a background for the subsequent literature review. At the end of each section is a discussion and identification of a research gap from which objectives are set out. Finally, there is a summary of the objectives set out in each of the sections.

2.2 Single detector vehicle classification

Speed estimation and vehicle classification (i.e. the identification of vehicle type based on characteristics such as length) from a single detector has been the subject of many research papers (Oh et al. (2002), Sun and Ritchie (1999), Ki and Baik (2006), Meta and Cinsdikici (2010), Ye et al. (2006), Hellinga (2002), Wang and Nihan (2003), Coifman (2001), Zhanfeng et al. (2001)) with complex filtering methods often being proposed in an attempt to improve accuracy. Accurate

classification from a single detector is an enticing prospect for practitioners as it can provide redundancy at traditional monitoring sites, such as MIDAS (Tucker S., 2006) in the UK, where two closely spaced detectors are employed at a known distance to calculate speed (thus allowing length to be derived) and occupancy. Indeed, the vast majority of the literature regarding single detector classification focuses on freeway applications where such monitoring stations are predominantly located.

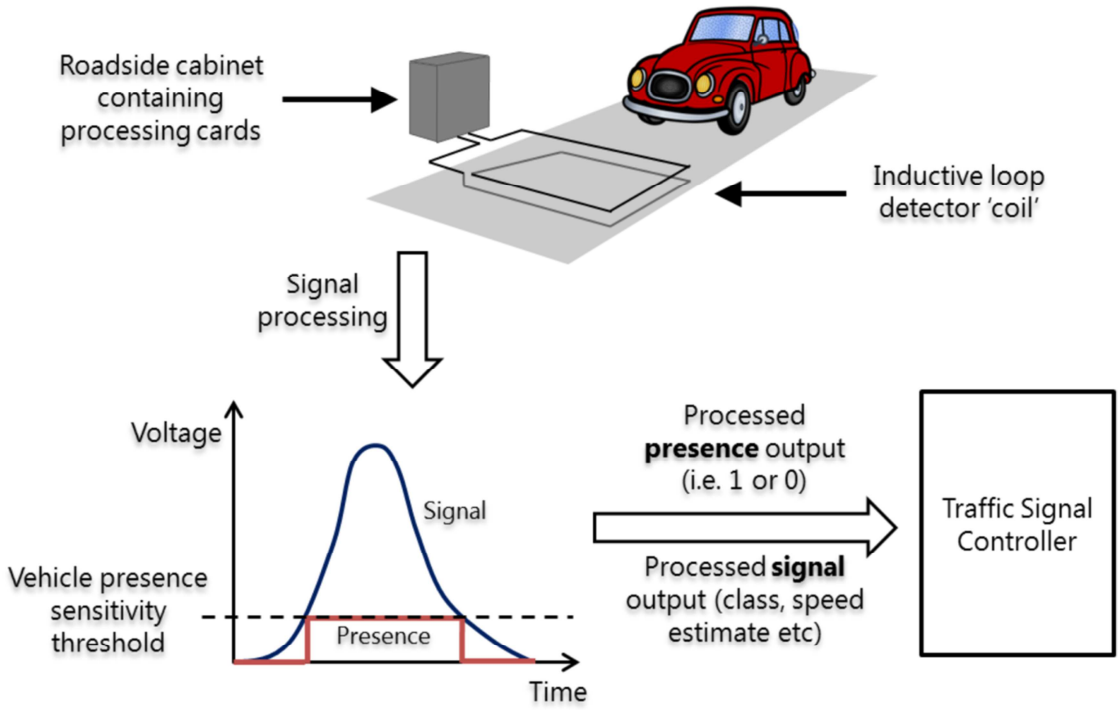

Figure 2.1: Example of an inductive loop detector installation providing the raw signal data or processed presence output to a control/monitoring application (own work).

There are two principal techniques of estimating speed (and classifying vehicles) from a single detector. The first technique makes use of the signature profile created by a vehicle chassis passing through the electromagnetic field of an inductive loop (or a magnetic sensor) to classify vehicles. Examples of this technique are Oh et al. (2002), Sun and Ritchie (1999), Ki and Baik (2006), Meta and Cinsdikici (2010). The second draws on the relationship between flow and occupancy to calculate a space mean speed over a specified time interval using

processed vehicle presence data. Examples of this technique are Wang and Nihan (2003), Coifman (2001), Zhanfeng et al. (2001).

In this chapter the first technique will be referred to as the 'profile method' and the second as the 'flow-occupancy method'. Figure 2.1 is an example of an inductive loop installation providing an output in different formats to a control and/or monitoring application.

2.2.1 The profile method

The profile method of estimating vehicle speed and/or classifying vehicles exploits the waveform profile of the electrical pulse created by a vehicle chassis when passing over an inductive loop detector (ILD). The electrical pulse is created as the vehicle passes through the electromagnetic field formed by the inductive loop detector. The widely used detector packs in the UK such as the Siemens SLD4 (Siemens Plc, 2016b) are primarily used to output a bivariate signal of either 0 or 1 depending on whether a vehicle presence sensitivity threshold has been exceeded. For some applications, such as vehicle actuated traffic signal control methods, the simpler vehicle presence output is sufficient. However, there are detector packs (RTEM, 2013) that explicitly manipulate the waveform profile, processed at a high scan rate, to derive vehicle speed and class.

Oh et al. (2002) exploit the profile method, setting the detector scan rate at 7 milliseconds as a trade-off between sampling rate and the ability to manipulate the waveform profile. Oh et al. (2002) observed that the profiles for each vehicle exhibit different features. The maximum magnitude, magnetic length and area of a vehicle profile are all examples of vehicle specific features that can be used

to determine vehicle type. However, the profiles are also functions of vehicle speed.

Previous research by Sun and Ritchie (1999), using a simple linear regression equation between vehicle speed and vehicle profile slew rate (the slope of the waveform profile at half the maximum magnitude) showed promising results. However, it was found to result in a significant increase in speed estimation error for longer vehicles when applied to a new dataset containing a more diverse vehicle mix.

To address the increase in error, Oh et al. (2002) presented three neural network architectures for grouping vehicles. The reasoning for grouping the vehicles was so that different speed estimation models could be applied to each group. Oh et al. (2002) chose to use four vehicle groups incorporating a wide range of vehicle lengths and the results showed a correct classification rate of 90% for the group with the largest training data set. For the second vehicle group (i.e. small buses, vans and pickup trucks), the results showed a 78% correct classification rate. The results of the research suggest that the presented method suffers from one of the same issues associated with the flow-occupancy method in that, where vehicle lengths are close to the threshold of neighbouring vehicle groups, the error in grouping the vehicles increases.

Very similar research, using a back-propagation neural network (BPNN), was presented by Ki and Baik (2006). In this case, vehicles were grouped into five classes to include motorcycles. The vehicle specific characteristics used were the variation rate of the frequency, waveform of the magnetic profile and occupancy. The results were very similar to that presented by Oh et al. (2002) in that the method performed well for all vehicle classes (particularly for trucks

where the recognition rate was 100%) but the results for the second vehicle group, incorporating vans, were not so accurate (approximately 79%).

Meta and Cinsdikici (2010) identified that the Ki and Baik (2006) method did not solve the problem of noise and it was also suggested that the relatively few vehicle specific features extracted from the waveform profile (fewer than those used by Oh et al. (2002)) limited the ability to distinguish between vehicle type. To address these shortcomings, Meta and Cinsdikici (2010) applied a discrete Fourier transform to clear the noise before applying a feature extraction technique. A statistical feature, the local maximum parameter, was extracted from the waveform in addition to other features. The local maximum parameter exploits the fact that the undercarriage of different vehicles can vary significantly in height. For example, a waveform associated with a heavy goods vehicle (HGV) is likely to incorporate more local maxima than a car. Finally, a BPNN was employed to classify the vehicles. The results showed an improvement on previously presented methods with a 98% recognition rate for cars, 90% for vans and 93% for HGVs. Buses and motorcycles were classified with 100% accuracy, albeit with smaller vehicle samples (approximately 30 each).

2.2.2 Flow-occupancy method

The flow-occupancy method of estimating vehicle speed draws on the fundamental relationship between traffic flow, speed and occupancy (as a proxy for density). Using this method, a space mean speed is calculated over a specified time interval using processed vehicle presence data from a vehicle detector. An advantage of this method is that any type of detection technology capable of providing a presence output could be used.

The space mean speed of a sample of vehicles can be derived from the relationship between speed, flow and occupancy. The derivation is described by Kwon et al. (2003) and is summarised in Appendix A using similar terminology. This relationship enables the speed of a sample of vehicles to be estimated on the assumption that vehicle speeds within the measured sample interval are constant. A summary of the derivations is described below.

The calculated flow in any sample time period is given by the number of recorded vehicles divided by the length of the sample time period (Equation A.1 in Appendix A). The occupancy is given by the sum of the so-called 'on-time' of each recorded vehicle divided by the length of the sample time period (Equation A.2 in Appendix A). The on-time of a vehicle is defined as the time spent travelling over a conventional detector.

The space mean speed for a particular sample of vehicles is defined by the sum of each vehicle speed divided by the total number of vehicles (Equation A.3 in Appendix A). The speed of each vehicle is, of course, unknown in this case but can be estimated using the relationship between speed and length of a vehicle (Equation A.4 in Appendix A). Through substitution and re-arrangement, the space mean speed of a particular sample can be redefined as being approximately equal to the flow multiplied by the space mean length divided by the occupancy. This enables the speed to be approximated based on the known parameters of on-time and flow along with an estimate of vehicle length (Equation A.8 in Appendix A).

Equation 2.1 reiterates the relationship between parameters for reference given the various terminologies in the related literature reviewed here.

$$\overline{L_i} = MEVL = \frac{1}{n_i} \cdot \sum_{k \in K} L_{tot} = \frac{1}{g} \tag{2.1}$$

where:

$\overline{L}_i$ = Space-mean length (metres).

n = Number of vehicles in the specified vehicle sample.

$MEVL$ = Mean effective vehicle length (metres).

g = Reciprocal of MEVL.

For a given sample time period, Equation A.8 in Appendix A can also be expressed as

$$\overline{v} \approx \frac{n}{T \times O \times g} \tag{2.2}$$

as is the case in Wang and Nihan (2003). However, Athol (1965) referred to the *g*-factor as the *K*-factor.

Both components of the MEVL, vehicle length and effective detector length, can be estimated but, whilst the effective detector length can be considered constant, the vehicle length is not.

In the case of inductive loops there are numerous factors that can contribute to L_{det} varying from one detector to another even if the design length of each loop is identical. These include variations in the buried depth of the cable, the length of the cable run from the detector to the roadside cabinet and the sensitivity of the monitoring equipment. However, at a specific site this value can be assumed constant as the value does not change significantly over time. Wang and Nihan (2003) recognised that the effective detector length can change from one detector site to the next and proposed a correction coefficient to the *g*-factor to account for changes in loop sensitivity. This could also be extended to other types of detection (such as radar or video detection) where the size of the detection zone may vary slightly from one site to another.

The length of each vehicle is more difficult to calibrate as it cannot be measured directly from the detector and is therefore usually given as an average from observed data for the associated link. The value of the MEVL or its inverse the, so called, *g*-factor has been subject to various research including Wang and Nihan (2003), Coifman (2001) and Zhanfeng et al. (2001) with its value subject to extensive debate. Further methods have been presented in literature by Mikhalkin et al. (1972), Pushkar et al. (1994) and Dailey (1999) but this review focuses on more recent literature that encompasses and builds on some of the methods presented previously.

It is clear that the *g*-factor is not in fact constant as it depends on the vehicle mix in any measured sample time period which can vary significantly from a mean length measured over a longer time interval. As the length of a sample time period increases it becomes more likely a representative vehicle mix will be captured but it also increases the likelihood that the actual vehicle speed during the interval is not constant, particularly in saturated conditions, thus severely reducing the effectiveness of the speed estimate.

Zhanfeng et al. (2001) demonstrated empirically that the *g*-factor can vary by up to 50% between detector sites and can also vary significantly at the same site due to changes in the vehicle mix over time. In recognition of the fact that a constant *g*-factor is therefore invalid, Zhanfeng et al. (2001) employ a strategy of specifying a free-flow speed and identifying free-flow conditions by selecting an occupancy threshold value. This enables the *g*-factor to be calculated for each sample time period (in this case 5 minutes). In congested periods this method does not hold and, instead, the *g*-factor is estimated using historical data.

Wang and Nihan (2003) take an alternative approach and separate sample time periods with long vehicles (i.e. HGVs) from those without. Time intervals of 5 minutes are used and split into 20 seconds sample periods. The 20 second

sample time periods are sorted by average occupancy per vehicle and a ratio threshold between occupancy and effective vehicle length used to select sample periods with no long vehicles. The ratio threshold used is based on trial and error but enables the average speed for the 5 minute time interval to be calculated. The remaining sample time periods are used to estimate the volume of long vehicles.

This method relies on the presence of sample time periods with no long vehicles in order to estimate speed. It is also susceptible in periods of high volume traffic (where vehicle speeds can drop significantly) to discarding legitimate sample time periods of slow moving short vehicles that it identifies as long vehicles. The method can feasibly be used to estimate vehicle speed in 'real-time', that is, in 5 minutes time intervals (often considered real-time in traffic applications) but, as in the case of Zhanfeng et al. (2001), it is not feasible to use on a vehicle-by-vehicle basis.

Coifman (2001) recognises that the purpose of developing a flow-occupancy method is to produce a solution that could be deployed on a simple processor. Indeed, this reasoning can be extended to the legacy communications architecture employed by most UTC systems as it enables the vehicle presence data to be transmitted by existing methods with minimal software changes. Transmitting the data to a central system provides the advantage of enabling algorithm improvements to be quickly applied to every detector site.

Coifman (2001) extends on the work by Zhanfeng et al. (2001), selecting a free-flow occupancy threshold, below which a pre-specified free-flow speed is used. Coifman (2001) proposes the use of an exponential filter to dynamically update the MEVL value between sample time periods. An important update to Coifman (2001) introduces the concept of using the median (rather than mean) vehicle 'on-time' value to estimate a median velocity for each sample time period

(Coifman et al., 2003). The idea of using a fixed number of vehicles rather than a fixed sample time period is also explored but it is noted that the fixed sample time period is more practical as it is easier to observe break down in traffic flow.

Coifman and Kim (2009) discuss two techniques, including the previously introduced median method, the sequence method of Neelisetty and Coifman (2004) and introduces a third that is referred to as the distribution method. In the updated study, the median method is implemented as a rolling sample of 33 vehicles, of which the speed of the centre vehicle is estimated. Interestingly this contradicts Coifman et al. (2003) in that, presumably, the benefit of using a sample with a fixed number of vehicles outweighs the issues associated with observing flow break down.

The sequence method incorporates the median method but also assumes that the speed of successive vehicles rarely differs significantly, even under congested conditions, and that it follows that the ratio of vehicle on-times can be used to deduce vehicle lengths where the successive vehicle lengths differ significantly. However, it was quickly realised that, in fact, the assumption of successive vehicle speeds rarely differing significantly does not hold true in congested conditions, particularly when vehicle speeds drop below approximately 10mph (i.e. stop/start conditions).

The distribution method also incorporates the median method but attempts to address the shortcomings of the sequence method by introducing a bi-modal filter. Coifman and Kim (2009) identify that the distribution of observed vehicle lengths (and on-times) exhibits a bi-modal property. That observation is incorporated into the distribution method by analysing the on-time distribution of each sample of vehicles. In many cases, due to the relatively small sample size, the distribution will exhibit a single peak from which the median method is

used to estimate speed. If a bi-modal distribution of on-times is identified then the median on-time of the dominant peak is used to estimate speed.

This technique suffers from the same issues associated with Wang and Nihan (2003) in that a sample could be considered uni-modal but may be comprised solely of HGVs. In that case, deducing whether it is indeed a sample of HGVs or a sample of slow moving short vehicles is more difficult. In this case, three additional tests are applied including the already described low occupancy threshold, the on-time variance and the estimated speed from the previous sample.

Ultimately, Coifman and Kim (2009) conclude that the distribution method offers the best performance. Length based classification into three categories, approximately translating to <8.5 metres (Class 1), 8.5-14 metres (Class 2) and >14 metres (Class 3), produced results in excess of 99% correct classification rate for Class 1, 73-76% for Class 2 and 94-97% for Class 3. This demonstrates that, as with the profile method, vehicles with lengths between 8.5-14 metres are difficult to classify accurately.

The reviewed literature would suggest that there is a limit to vehicle classification performance from a single conventional detector, whether using the profile or flow-occupancy method. However, that limitation may matter less if classifying more broadly into, say, two classes (i.e. long or short). However, it does suggest that, if more accurate classification of multiple vehicle classes is required then an alternative approach is necessary. Installing a second conventional detector is one possible option but is costly and there are potential alternatives that could provide more comprehensive data. Those alternatives, such as V2I data and video detection are discussed in more detail later in this chapter and in Chapter 7 respectively.

2.2.3 Other methods

A more recently introduced vehicle detection technology uses magnetic sensors. The wireless detector sensors (Clearview Intelligence, n.d.), approximately the size of a coffee mug, are core drilled into the road surface and monitor changes in the Earth's magnetic field. This type of sensor offers advantages to practitioners in that they can be installed in locations where ILDs are not feasible. This may include locations where the road surface is in a poor condition or where the ducting infrastructure, required to connect the ILDs to a traffic control system, would be difficult to install. However, the currently available systems are more complex to configure than conventional ILDs and the requirement to monitor battery life of wireless equipment has perhaps limited its popularity in the UK.

Magnetic sensors could be used to estimate speed and classify vehicles with the flow-occupancy method in the same way as ILDs. However, magnetic sensors are also capable of measuring changes to the Earth's magnetic field along multiple axes (i.e. the direction of the lane and vertically). This offers the possibility of applying the profile methods for ILDs, such as that proposed by Meta and Cinsdikici (2010), to the vehicle signature produced by the magnetic sensor. Cheung et al. (2005) presented results from limited experiments that provided promising results but the overall classification accuracy was low (approximately 60%).

He, Y. et al. (2012) applied a methodology more akin to that proposed for ILDs by Meta and Cinsdikici (2010) with much improved classification accuracy, although the dataset was small and included an unusually high proportion of buses, for which the profile method has been shown to classify accurately. Likewise, Tafish et al. (2016) applied a discrete Fourier transform and a feature

extraction technique. Tafish et al. (2016) used a much larger sample size with a HGV proportion of approximately 17% (for reference to the work by Oh et al. (2002) with ILDs) and classified into three vehicle groups. The results showed very similar results to those of Oh et al. (2002) with approximately 88% classification accuracy for group 1 (cars and vans), 74% accuracy for group 2 (rigid HGVs) and 91% accuracy for group 3 (articulated HGVs). The results demonstrate that using magnetic sensors is clearly a valid approach to vehicle classification when compared to ILDs.

2.2.4 Discussion and research gap

The recurring feature of the flow-occupancy methods, with the exception of Coifman et al. (2003) and Coifman and Kim (2009), is that they all estimate speed over a specific sample time period ranging between 20-30 seconds to, in some cases, minutes in length. In the case of Coifman and Kim (2009), a minimum sample size of vehicles (33) is required. All the algorithms have been produced to estimate freeway speed where it is likely that, for the majority of the time, conditions will be free-flow and, consequently, the assumption that speeds are constant during each sample interval holds.

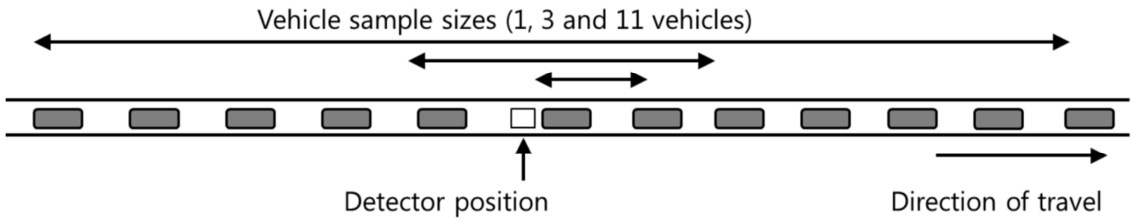

Figure 2.2: An example of vehicle sample sizes over a detector.

The methods discussed are useful for their applications and pose no problem if the data is being collected offline for later analysis or even if the data is to be used to detect congestion or incidents at a reasonable reporting frequency. However, for the purpose of this thesis it is necessary that each vehicle is

classified instantly upon leaving the detector so that the information can be fed into an on-line traffic model.

Figure 2.2 demonstrates that the vehicle crossing the detector is in the centre of the vehicle sample. The estimated speed (and subsequently classification) data cannot be utilised until the last vehicle in the sample has crossed the detector. For the 11 vehicle sample, assuming headway between vehicles of approximately 2 seconds in the busiest conditions, it would take 10 further seconds (after the centre vehicle has crossed the detector) to 'complete' the sample. Most conventional detection at traffic signals is located no more than 10 seconds travel time from the junction. Consequently, even an 11 vehicle sample would be ineffective for a traffic signal application as the centre vehicle would, in most cases, have crossed the stop line by the time information became available. The larger vehicle samples recommended by Coifman and Kim (2009), would compound this problem. A delay of, say, 2-3 seconds may be less problematic but there is no guarantee, even for a 3 vehicle sample, that the last vehicle will cross the detector within 2-3 seconds of the centre vehicle.

Furthermore, the proposed traffic signal application for the speed estimation algorithm developed in this work is for traffic signal controlled junctions where speeds are likely to fluctuate significantly over short periods of time, particularly in oversaturated conditions. This fluctuation in speed significantly reduces the accuracy of methods that use longer sampling periods as they assume speed is constant for the duration of each sampling interval. In the case where a vehicle has been stationary – or is moving very slowly – over the detector, it will heavily influence all the 'rolling' samples either side of the stationary vehicle, causing consistent underestimation of speed.

The alternative method for estimating speed from a sample of vehicles, proposed by Coifman et al. (2003), of using the median detector on-time in a

sample produces the opposite problem. Using the median on-time effectively excludes the long on-time caused by the stopped, or almost stopped, vehicle and therefore over-estimates vehicle speeds over the detector when queues form from the traffic signals. In the freeway application, for which the Coifman et al. (2003) method was developed, this is unlikely to be an issue but it requires an alternative solution for this application.

The use of a sample size larger than one vehicle necessarily introduces latency into the classification process where the vehicle being classified is in the centre of the sample. In practice, data from detectors on freeways is often recorded to report to a central database at, say, 5 minute intervals for use in strategy selection by traffic managers or to inform road users of traffic conditions. For such applications, the requirement for the 'second half' of the vehicle sample to be collected (16 further vehicles in the case of Coifman and Kim (2009)) before the 'centre' vehicle can be classified is not significant. However, for this application the usefulness of the classification data deteriorates rapidly once a vehicle has left the detector and it is therefore not viable to implement such methods.

The profile method, whether applied to ILDs or magnetic sensors, provides more promising results in the sense that it could provide 'instantaneous' speed estimation and vehicle classification as a vehicle leaves the detector. Instantaneous classification is essential for the traffic signal control optimizer application in this thesis. However, current commercial products (i.e. RTEM RTEM (2013)) require a bespoke detector card to perform the profile method classification. For new or refurbished traffic signal controllers the bespoke cards can be specified but for existing sites the use of the profile method would require additional expenditure to replace the incumbent cards. The practitioner

is also dependent on specialists to configure the detector card to extract optimal performance.

There are many other methods of estimating vehicle speed and classifying vehicles such as using video detection, radar and even infrared/ultrasonic sensors that have not been included in this literature review for the sole reason that they all require additional infrastructure. However, some of these solutions will be discussed later in Chapter 7. The methodology used in this thesis is to first explore the maximum that can be extracted from existing infrastructure before investigating how V2I technology could be used to improve the performance of traffic signal control. The review of literature for single detector speed estimation and vehicle classification has therefore focused on methods that could be delivered without additional infrastructure requirements as it has been assumed that V2I data will provide accurate vehicle speed and type information.

In summary, the review of literature for single detector vehicle classification has demonstrated two primary methods for classifying vehicles from existing infrastructure (namely ILDs). The profile method showed more useful results for the application to traffic signal optimization due to the ability to provide instantaneous classification as a vehicle leaves the detector. However, the requirement for replacement detector cards does not meet the initial aim of this research to extract the maximum performance from existing infrastructure.

The flow-occupancy method offers much reduced data requirements, making the use of multiple types of detector (as long as it is capable of producing a vehicle presence output) feasible. This in itself makes the flow-occupancy method attractive in that it could potentially be easily applied using a range of existing detector infrastructure.

Reviewed methods use vehicle presence data collected at 60Hz (Coifman et al., 2003), a relatively high frequency. However, controller scan rates are usually no more than 5-10Hz (Siemens Plc, 2015), outstation transmission units (OTUs) 20Hz (Steel, 2012). Most existing UTC communications architecture (i.e. communicating from the OTU back to a centralised system) limits detector data transfer to 4Hz as a legacy from the early development of the SCOOT traffic control system with the communications infrastructure available at that time. Any developed method would therefore need to prove effective in using detector data at a lower frequency than the reviewed methods.

The main drawback of the flow-occupancy method is that all the reviewed methods are for freeway applications and do not provide the instantaneous classification required in this thesis. There is therefore an identified research gap to investigate whether the flow-occupancy method of classifying vehicles could be adapted to the stop/start conditions of an urban setting and whether it is feasible to provide an instantaneous vehicle classification output. The objective of the research is to investigate whether a method can be developed that is capable of providing an adequate performance for the application of traffic signal optimization using existing infrastructure. The use of existing infrastructure will necessitate using data supplied at a lower frequency than the reviewed methods.

2.3 Vehicle classification in a traffic signal control context

2.3.1 Signal optimization strategies

There are many 'on-line' responsive traffic signal control strategies in use across the world today. These control strategies make use of real-time detector data to adapt signal timings in response to traffic conditions. SCOOT is a particularly successful strategy (in terms of popularity at least) but there are a number of

other strategies for network-wide control that have experienced various degrees of commercial success including SCATS (Sims and Dobinson, 1980) and UTOPIA/SPOT (Mauro and Di Taranto, 1989). Other examples are PRODYN (Henry et al., 1983), OPAC (Gartner, 1983), RHODES (Mirchandani and Head, 2001) and TUC (Diakaki et al., 2002). ALLONS-D (Porche, Isaac et al., 1996) is a highly decentralized method and to some extent sits between network-wide and isolated control.

Network-wide control strategies can be split, generally speaking, into centralised (SCOOT, TUC) or decentralised/hierarchical (SCATS, UTOPIA/SPOT, PRODYN, OPAC) where more optimization decisions are taken at a local level but informed by a strategic central function. This categorization is not entirely clear-cut as some systems control the network from a central location but operate in a decentralized manner. For example, although SCOOT operates from a central system, it does so with independent 'regions' of adjacent, similar, junctions that are defined during configuration of the system by a UTC operator. More recent versions of SCOOT (Bodger, 2011) include the ability to operate individual junctions independently when the regional optimizer judges that there is more benefit than providing coordination, thus introducing a more decentralized approach. However, although decentralized to some degree, the SCOOT system still considers at least region-wide optimization with a methodology based on the off-line TRANSYT tool.

SCOOT operates by calculating green durations (i.e. stage lengths) 4 seconds before each stage change is scheduled to occur. At this point the optimizer decides to either end the stage early, retain the original decision or change 4 seconds later (Department for Transport, 1995). If changing earlier or later a smaller change (1 second) is made to the permanent stage length for the following cycle. This method allows the timings to follow longer term trends

whilst adapting to shorter term ‘noise’. Offsets are calculated every cycle and cycle times calculated at intervals of no less than 2.5 minutes. The SCOOT philosophy of small changes to timings is designed to limit perturbations in the network but in doing so it does limit the general applicability of the system. For example, in the UK it is now commonplace to operate SCOOT during the peak periods with MOVA (or System-D VA) run outside the peaks when coordination is less critical to provide more dynamic control (Witts D., 2013). The switch between SCOOT and MOVA is based on heuristic rules, usually applied to flow and occupancy data collected from detectors.

More decentralized systems such as SPOT/UTOPIA provide a more defined hierarchical control method. In this case, the SPOT control system provides the individual junction level control. Optimization is performed over a time horizon of 120 seconds, repeated every 3 seconds. Communication on a local level between SPOT controllers enables traffic forecasts and counts to be used from neighbouring junctions. UTOPIA provides area level forecasts of public transport arrivals that are then included in the local optimization objective function. The SPOT objective function includes delay, stops, queue lengths and also considers the deviation from the reference plan (provided at area level). The latter element enables the degree of freedom within which the individual level control can operate (Fox et al., 1998).

For isolated junctions (with no interaction with adjacent junctions) the optimization problem is simpler, although certainly not trivial. MOVA and LHOVRA (Peterson et al., 1986) are examples of strategies developed specifically for isolated junctions. LHOVRA, popular in Sweden, is phase (signal group) based and uses multiple detectors on an approach to drive a series of functions that can provide priority to heavy vehicles and vehicles on the main road, and aim to reduce accidents. The timing logic for LHOVRA is relatively simplistic and

does not attempt real-time optimization. An alternative method, referred to as approximate dynamic programming has been developed by Cai et al. (2009) and has shown promising results when compared to fixed-time plans but it is yet to be evaluated against the currently considered 'state-of-the-art' MOVA system in either simulated or real-world conditions.

MOVA has become a very popular method of control in the UK (Highways Agency, 2005). The MOVA control strategy is based on the Miller (1963) approach to optimization that employs a win/loss technique to calculate whether or not it is beneficial to extend the current green signal for a specified time-step. The Miller technique is, in effect, a rolling horizon strategy with a horizon of one cycle (not fixed) that considers in detail the effect of extending a green signal for vehicles within 8-10 seconds of the stop line (depending on the upstream detector location) on delay. Miller proposed a time-step of 2 seconds but MOVA uses a time-step of 0.5 seconds (Vincent and Peirce, 1988). In over-saturated conditions MOVA switches to a different optimization method in order to maximize capacity. In that mode of operation, MOVA redistributes available green time according to efficiency of use.

An alternative approach to traffic signal control is presented by Box and Waterson (2013) in the form of a neural network based junction controller. The junction controller can learn strategies through supervised learning (by a human expert) or reinforcement learning by temporal difference. The temporal difference trained neural network matched the performance of the human trained controller at an isolated junction. The human trained neural network (Box and Waterson, 2012) had previously been shown to outperform MOVA by up to 25% when probe data (i.e. from GPS/LIDAR etc) is accurate.

2.3.2 Applying vehicle classification to traffic signal optimization

Vehicle classification data is usually collected for the purposes of off-line analysis. Common methods of data collection include manual traffic surveys, video detection and automatic number plate recognition (ANPR) cameras. The classification data, amongst other applications, is often collected by survey companies and subsequently analysed for transport planning purposes, often to provide data for transport models. Such applications do not require the vehicle classification to be performed in real-time and so video detection is often employed simply to record images for later manual analysis by human operators. In the UK, ANPR data can be submitted to the Driver and Vehicle Licensing Agency (DVLA) to determine vehicle type, providing accurate data for modelling purposes (Driver & Vehicle Licensing Agency, 2018).

Applying vehicle classification to a traffic signal optimizer introduces more onerous requirements as it requires that classification data is available for use immediately. In fact, the closer a vehicle travels to the traffic signal stop line before the optimizer is provided with the classification data, the less useful it becomes as there is less time available for the optimizer to react.

Whilst it is possible that ANPR data could be used to provide instantaneous vehicle type information, to do so would likely require that a database is used to store raw number plate data along with the vehicle details in order to rapidly recall the vehicle type information. If that were the case then it may introduce data governance issues, although this might be possible to overcome. An ANPR solution would be costly but is also likely to be accurate and thus useful if the vehicle type data could be acquired at a reasonable distance from the traffic signal stop line. It would, however, require additional infrastructure to be installed on street. To date, the author has found no literature on the use of this technique for the proposed application.

As discussed in the previous section, there are various techniques for classifying vehicles from ILDs but there are also different types of detection that are capable of providing similar information. Historically, traffic signal control has most commonly employed single detectors rather than pairs of detectors on the approach to a junction. The use of single detectors limits the cost of installing the ILDs and, for many signal control strategies, the primary purpose of a detector is to count or provide occupancy data which negates the requirement for detector pairs. An exception to this is Speed Discrimination (SD), used in the UK in conjunction with the System-D VA control strategy. SD uses a pair of loops to determine the speed of approaching vehicles. If the speed is above a pre-defined threshold then the currently green signal and following safety period is extended to minimise the likelihood of vehicle conflicts (Highways Agency, 2002).

2.3.3 Transit signal priority (TSP)

A basic example of vehicle classification in a traffic signal control context is the use of selective vehicle detection (SVD) to reduce delay to buses (and/or trams). Traditionally, this was implemented using ILDs, either by placing a pair of ILDs in a dedicated bus lane approximately 10-12 metres apart to ensure only longer vehicles are detected, or by using the profile method described in the previous section. A profile method commonly employed in the UK for the bus priority application was the now defunct PRISM system (Hertfordshire County Council, 2009). An alternative method for detecting specific vehicles is the Sietag system that uses a radio signal antenna located in the road surface (Whittingham, 2004). The relevant vehicles are fitted with Radio Frequency Identification (RFID) transponders that, when matched with a database of tag type and serial numbers, enables a trigger to be generated in the traffic signal controller to provide priority.

Automatic Vehicle Location (AVL) is, again, primarily deployed on buses and is a popular method of providing bus priority centrally without the need for physical infrastructure at the roadside. One method of AVL is to use radio beacons to transmit vehicle location to a central system. This method requires equipment to be installed on buses to transmit the GPS based location when travelling through one of a series of stored 'virtual' detection zones or trigger points (Hounsell and Shrestha, 2005). An alternative method, becoming increasingly common, is for bus operators to use the ticket machines already installed on buses to perform a similar operation but using mobile data communication (i.e. 3G/4G). This method negates the need for additional equipment on the buses and takes advantage of the fact that bus operators in many cases already transmit Real-Time Passenger Information (RTPI) data from their ticket machines at regular intervals for the purpose of tracking buses and predicting arrival times at bus stops (FirstGroup Plc, 2018).

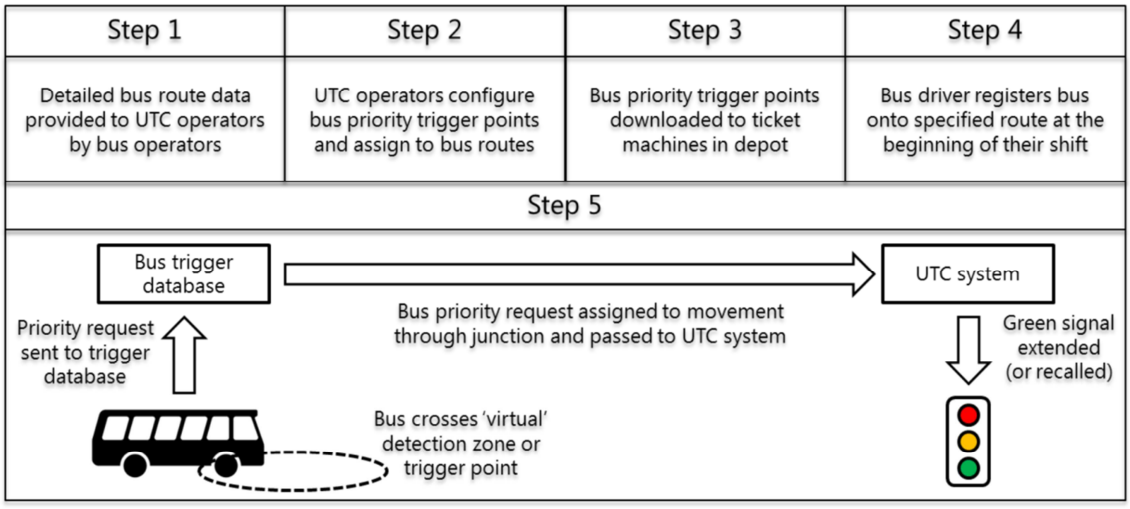

Figure 2.3: An example implementation of AVL for a bus priority application (own work).

Both AVL methods require UTC operators to configure bus trigger points that are downloaded to buses on a regular basis (as new junctions are added or layouts modified etc). However, the accuracy of GPS location in many devices is not high enough to be lane specific (Ordnance Survey, 2018). Instead, an

example solution is for UTC operators to configure route specific bus priority trigger points. The movement of each bus route through a junction is known and so each trigger point will be configured only to apply to a specific movement through a junction. Buses do not necessarily operate the same route from one day to the next and bus drivers register their bus service on the ticket machine at the start of their route. The ticket machine, now aware of the route it is operating, can transmit a priority request (with a route reference) to a central trigger database system. Finally, the trigger database can assign a movement through the junction from the trigger request and pass it to the signal control system (see Figure 2.3).

It is clear that implementation of AVL is not entirely straightforward and much of the process described in Figure 2.3 is undertaken using bespoke systems or even spread-sheet tools. The process varies between Local Authorities with multiple AVL methods in use across Europe (Hounsell and Shrestha, 2005). In addition, bus routes are changed periodically and so the assignment of bus triggers to routes must be reviewed on a regular basis. However, despite that, an AVL system does allow the progress of vehicles to be compared with that predicted by timetable and hence differing levels of priority can be provided according to bus lateness. This is a useful tool for UTC operators as it prevents priority being given to a bus that is early (at the expense of other traffic), only for it stop at a downstream bus stop for a lengthy period.

2.3.4 Other applications for vehicle classification

Bus priority continues to be a significant focus for Local Authorities as there is political pressure to reduce bus journey times and increase bus patronage. As an example, the Transport Strategy for Leeds City Council (LCC) aims to double bus patronage within ten years and LCC is investing £173.5 million in bus travel

with the Leeds Public Transport Programme over three years in an attempt to improve performance (Leeds City Council, 2016).

However, air quality is becoming an increasingly prominent issue, particularly since legally binding targets were introduced by the European Union (EU) with the 2008 Ambient Air Quality Directive (European Union, 2008). As part of a wider range of measures to address air quality issues, some cities have chosen to ban diesel vehicles. In Hamburg a ban has already been introduced that excludes all diesels older than a Euro 6 (or Euro VI for HGVs) standard specification (Chazan, 2018). However, the ban is very limited in its geographical extents and residents are exempted. In the UK, London plans to introduce an Ultra-Low Emission Zone in a central area of the city from April 2019 that will ban all pre-Euro 6 standard diesel vehicles (Transport for London, 2018). In five other UK cities, a Clean Air Zone (CAZ) is to be implemented by 2020 (DEFRA, 2015) in line with the Department for Environment, Food and Rural Affairs Clean Air Zone Framework (DEFRA, 2017). In Leeds, the proposed CAZ will cover a large proportion of the city and will introduce a charge for vehicles with pre-Euro 6 diesel engines. However, private cars are planned to be exempt from the charge.

It is worth noting that there is continuing legal action against various governments within the EU in an attempt to force further action which suggests that the air quality issue will remain prominent for some time (Cuff, 2018). As a consequence, it is appropriate that the objectives of traffic signal optimization are reviewed and that there is an increased focus on investigating ways to minimize emissions as part of the overall objective function.

2.3.5 Methods of utilising vehicle classification within a control system

So far, the application of vehicle classification to provide bus priority through traffic signal control has been reviewed and the possibility of alternative applications for vehicle classification, such as for reducing vehicle emissions. However, the review has mainly focused on the method of identifying selected vehicles (such as buses) and generating a priority request that is subsequently passed to a traffic signal control system to make a decision. This section focuses on how various traffic signal control systems actually utilise vehicle classification data.

When considering public transport (or emergency vehicle) priority, there are various techniques that can be used to implement it depending on the control strategy that is being operated. In the UK, many cities still operate fixed-time plans and, in London, the Selective Priority Network Technique (SPRINT) strategy was developed to incorporate bus priority into a fixed-time network (Hounsell et al., 1996). The SPRINT strategy enables green signal extensions and recalls subject to a set of constraints that a UTC operator can decide on. The constraints include the maximum time difference from the base plan and the maximum allowable levels of saturation to permit green extensions and recalls (Fox et al., 1998). The SPRINT strategy also includes a recovery period following a priority event to compensate general traffic.

Selected Vehicle Priority in a UTMC Environment (SPRUCE) is an alternative system for fixed-time networks. SPRUCE is currently operated in Leeds, Sheffield, Bradford and Edinburgh and is a tool for UTC operators to write bespoke strategies rather than a strategy itself (Collier, 2017). Bus priority strategies can be developed to include all the features of SPRINT but can be modified and made as simple or complex as necessary by the UTC operator.

A common theme of SPRINT and SPRUCE is that the method of handling priority requests, particularly in the recovery phase, is based on a set of heuristic rules, taking into consideration average traffic flow and saturation flow parameters. Optimization techniques that incorporate an on-line traffic model can provide more optimal recovery strategies. In theory, a strategy in SPRUCE could be developed to incorporate an on-line model but it would potentially be difficult to implement and replicate without an object oriented programming environment.

In road networks where adaptive traffic signal control is in operation there are various techniques for providing bus priority. Again, the method of implementation depends on the control strategy. In the case of the SCOOT system, green extension and recall facilities are provided through the Bus SCOOT module (Bretherton et al., 1996). The extent to which priority is granted to buses depends on the degree of saturation (DOS) targets for non-priority links set by the UTC operator. The higher the DOS target, the greater the extent of priority afforded to buses but with greater delay to general traffic. Reductions in delay vary from as high as 50% in low flow conditions to 5-10% when the degree of saturation is high (Department for Transport, 2000).

The SPOT/UTOPIA system implements transit priority by providing the local junction controller (SPOT) with forecasts of the bus or tram arrivals. Each transit vehicle is effectively treated within the optimizer as multiple private vehicles to provide a greater weighting (i.e. making a green extension more likely). A very high weighting can be applied to provide buses (or more commonly trams) with absolute priority (Fox et al., 1998). Providing the local controller with a forecast of the transit arrival time means that the junction offset can be slowly adjusted in advance of the vehicle arriving so that it coincides with a green signal. This is, in effect, an automated version of a strategy developed by Sheffield City Council

in SPRUCE for fixed-time control to ensure that trams arrive at a green signal (Collier, 2017).

In MOVA, bus priority has traditionally been applied in a similar manner to SPOT but with buses always being provided absolute priority (if arriving during a green signal) by effectively suspending the optimizer until the bus has crossed the stop line. Providing priority to buses outside the 'visibility' of the optimizer introduces the possibility of causing additional delay to vehicles on opposing approaches (including any buses on those approaches) that may ultimately outweigh the benefit to the first bus. However, the latest version of MOVA (MOVA 8) introduces a bus priority feature that enables buses to be included in the optimization process with a higher weighting than other traffic (TRL Limited, 2018). This provides the opportunity to apply differential priority to transit vehicles depending on whether or not they are late running.

The ALLONS-D optimization method (Porche, Isaac et al., 1996) incorporates the capability to use minimization of person delay as an objective function, primarily as a tool to provide 'passive' bus priority. As with SPOT and MOVA, passive priority is used here to refer to providing vehicles with a higher 'importance' than other vehicles in the optimization process making it less likely they will be stopped.

In practice, the objective function of minimizing person delay was not used as tests were performed by weighting by vehicle type rather than occupancy. The tests concluded that a large reduction in delay for buses (~30%) could be expected when the level of background demand – in this context the 'non bus' demand – was low with negligible increase in overall delay (Porche, I. and Lafortune, 1997). At a higher level of background demand there was a larger (~5%) increase in overall delay on the network to achieve a similar benefit for

buses. However, the comparative effects on person delay were not reported and the level of bus demand is also unclear.

Outside of transit priority applications, SCOOT attempts to incorporate some degree of general vehicle classification by using profiling to convert each consecutive quarter second of presence data into a Link Profile Unit. However, this type of method suffers from reduced accuracy with the onset of congestion and does not determine between individual vehicles or vehicle type so cannot be applied to all types of optimization traffic model.

2.3.6 Discussion and research gap

Of the on-line control strategies reviewed there are some that provide the facility to implement differential vehicle priority. That is, some vehicles can be given a higher weighting within the optimization process to make it more likely that they will receive a green signal upon arriving at the junction. Each reviewed control strategy that incorporates the ability to provide differential priority (i.e. SCOOT, SPOT, MOVA and ALLONS-D) does so in slightly different ways but the common theme between them is that the differential priority has been applied to transit vehicles with a policy designed to reduce journey times and increase journey time reliability of public transport.

The application of vehicle classification to bus or tram priority is a special case in that the objectives for prioritising public transport vehicles often conflict rather than complement those for general traffic. For example, in many cases, extending a green signal for a bus will incur additional delay for general traffic on opposing approaches that worsens the overall performance index (as in the case of Porche, I. and Lafortune (1997)), particularly if considering vehicle delay rather than person delay. This is not necessarily a bad policy if it supports an increase in bus patronage. However, there can be significant performance differences between other vehicles such as private cars and HGVs (i.e. rate of

acceleration and braking) that, if considered explicitly, could have an even greater influence on overall optimizer performance.

No reviewed strategy explicitly considers vehicle class of general traffic. SCOOT implicitly considers vehicle length to improve queue length estimation accuracy but, as already discussed, the method used suffers in congested conditions. It is of particular relevance, given the previously discussed topic of air quality becoming an increasingly prominent issue, that methods of reducing stops for the most polluting vehicles are investigated.

The identified research gap in this section is the consideration of differential priority for general traffic. The objective of the research will be to explicitly classify vehicles with the aim of applying higher weightings in the optimization process for larger vehicles (that roughly correlate to those with the highest emissions). The effect on the objective functions of individual vehicle classes as well as the overall performance index will be evaluated.

2.4 Connected vehicle technology in traffic signal control systems

The term ‘connected vehicle technology’ is often applied to various aspects of information exchange from vehicle-to-vehicle (V2V) and between vehicles and infrastructure (V2I). The terms car-to-x (C2x) and vehicle-to-x (V2x) are alternative terms used that incorporate V2V and V2I. The type of information exchange referred to by the term V2x can range from data as simple as receiving traffic alerts or transmitting an emergency signal in the case of an accident to receiving advance warning of queues from downstream vehicles and communicating with traffic signal controllers.

Communication with traffic signal controllers can be with the objective of smoothing the trajectory of approaching vehicles by enabling drivers to adapt their speed to arrive at a green signal. It can also be to provide a traffic signal optimization strategy with more accurate information on vehicle position and speed (for example) to improve efficiency. In fact, research into the potential benefits of connected vehicle technology for improving the performance of traffic signals is loosely split into three categories:

- Optimization of vehicle trajectory to minimize acceleration/braking;
- Providing more accurate spatial data to improve the efficiency of traffic signal optimization; and
- Removing the requirement for physical traffic signals.

A preferred communications technology for connecting vehicles has not yet emerged but a communications protocol, sometimes referred to as ITS-G5 (ETSI, 2010) or Dedicated Short Range Communications (DSRC), has been developed based on the IEEE 802.11p wireless standard (Jiang and Delgrossi, 2008). The range of DSRC is stated to be approximately 1000 metres (DEVPOST, 2018), although that distance is likely to vary depending on influencing factors in the environment such as buildings and trees. Stahlmann et al. (2016) evaluated ITS-G5 based GLOSA systems and reported 'real-world' results. The findings suggested that, as with other wireless technologies, the predicted range was optimistic when considering environmental factors such as foliage and buildings. Even so, the relatively long range compared to standard WiFi introduces the possibility of utilising a single beacon to serve multiple junctions in dense road networks thus reducing the cost of deployment.

An alternative to ITS-G5 is the use of cellular networks such as 3G/4G and, in the near future, 5G. Cellular technology has the advantage of being deployed in a very large number of mobile devices and so implementing V2x solutions is

relatively straightforward. The downside is that the latency of current cellular networks is not low enough for safety critical applications, although that may change in the future.

In the short term then at least, there is likely to be a mix of dedicated and cellular communications technology depending on the application. For the purposes of this research it will be assumed that data can be supplied to a traffic signal control optimizer at a minimum frequency of 1Hz. In reality, if using DSRC, it is possible that the data could be supplied considerably more frequently (i.e. up to 20Hz) but, if collating data from various sources, it is likely that the data will require some degree of processing to smooth the data before passing it the traffic signal optimizer at a lower frequency (Alessio, 2017).

Communications infrastructure aside, the structure and format of the data to be broadcast from traffic signals to vehicles is also an important consideration. The International Organization for Standardization developed Signal Phase and Timing (SPaT) protocol provides signal timing data to vehicles and Map Data (MAP) describes the physical geometry of a junction (Amsterdam Group, 2015). The combination of SPaT/MAP data enables vehicles to map themselves onto a detailed layout of the junction and thus interpret the associated phase and timing data correctly.

2.4.1 Optimization of vehicle trajectory

In the first use case, the focus is on optimizing vehicle trajectories to reduce delay and emissions (by minimizing stops), usually with fixed signal timings that are known to the vehicles. There are two principal methods proposed for optimizing trajectories:

- A **passive** system – the driver of a vehicle is presented with information on the time until the next signal change. The response to the information is within the control of the driver; and
- An **active** system – the, presumably autonomous, vehicle automatically adjusts its trajectory in response to the data received from the traffic signals.

An example of a passive system is Green Light Optimal Speed Advice (GLOSA) where signal timings are broadcast to approaching vehicles and a suggested speed is presented to the driver in order to arrive at a green signal (Bodenheimer et al., 2014). The difficulty associated with this method is that the current traffic conditions and behaviour of preceding vehicles can affect the speed at which a vehicle can approach the junction.

Eckhoff et al. (2013) investigated the potentials and limitations of GLOSA and found that it provided significant benefit in under-saturated conditions (up to 11.5% lower CO_2 emissions) but that it could result in longer waiting times and higher CO_2 emissions for non-equipped vehicles in congested conditions. Katsaros et al. (2011) reported a reduction in average fuel consumption of up to 7% from a GLOSA simulation with a much larger reduction in 'stopped time' (up to 89%) although this does not necessarily suggest a similar reduction in delay. The optimal distance for activating GLOSA was found to be 300 metres.

Kamalanathsharma and Rakha (2016) is an example of 'active' vehicle trajectory optimization and uses SPaT data provided from the traffic signals through Infrastructure to Vehicle (I2V) communication. It is unclear what operational strategy the traffic signals employ but it is assumed to be fixed-time. Fuel savings from 5% (at approach speeds of 30km/h) to 23% (at approach speeds of 90km/h) are reported.

Similarly, He, X. et al. (2015) focus on optimization of vehicle trajectories through traffic signals but include queue consideration. Traffic signal timings are predicted based on assumed readily available historical data. Fuel savings of up to 29% are reported.

For both the passive and active cases reviewed here, it is assumed that the traffic signal control systems operate fixed time plans as there is no mention of dynamic control. It is plausible that some degree of adaptive signal control could be developed whilst still retaining the ability to optimize vehicle trajectory. However, at 500 metres (Eckhoff et al.), or even 300 metres (Katsaros et al.), optimal distance for activation of GLOSA messages, the travel time to the junction is likely to be in the region of at least 30 seconds in many cases. The scope for adapting signal timings would therefore be significantly reduced as the next change of the signals would have to be fixed well in advance of the arrival of the vehicle at the stop line.

It is unclear whether the benefit of providing optimized trajectories outweighs the diminished flexibility of signal control (and associated loss of performance) required to do so.

2.4.2 Improving the efficiency of traffic signal optimization

The second category, in which the research undertaken in this thesis can be placed, uses the additional information (i.e. vehicle position and speed) available to the traffic signal optimizer to develop new, or improve existing, optimization strategies.

Feng et al. (2015) developed a state estimation algorithm to address low penetration rates of connected vehicle technology. For the state estimation algorithm, an approach link is modelled as three regions, a queuing region, a slow-down region and a free-flow region (Figure 2.4). The regions are defined

based on information from V2I equipped vehicles. In the queue region, the queue length is based on the queue propagation speed. However, in the slow-down region, the algorithm uses the Wiedemann (1974) car-following model to predict where an unequipped vehicle is likely to be. The algorithm compares the actual acceleration of the known vehicle with the expected acceleration from the Wiedemann (1974) model. If, for example, the known vehicle is decelerating more harshly than would be expected given the modelled positions, it can be assumed that the vehicle may be reacting to an undetected unequipped vehicle. In that case, a new vehicle is inserted into the model. In the free-flow region the number of equipped vehicles is divided by the V2I penetration rate to calculate the total number of vehicles.

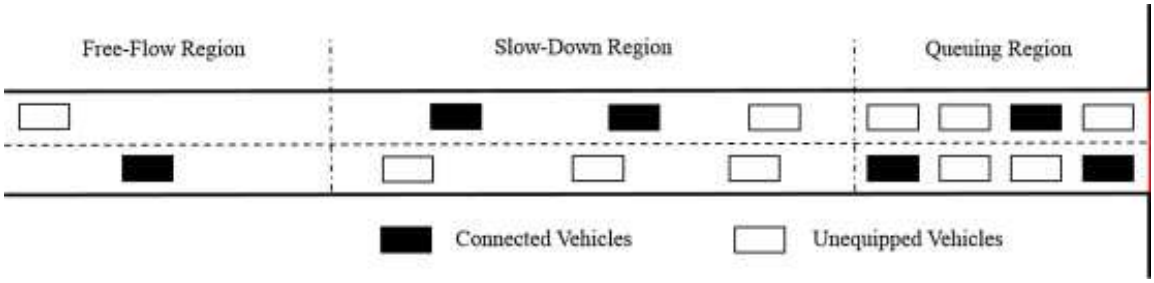

Figure 2.4: Diagram showing how an approach link is split into three regions (Feng et al. (2015)).

An adaptive control algorithm is then developed, based on the Controlled Optimization of Phases (COP) algorithm (Sen and Head, 1997) used in RHODES, that utilizes the connected vehicle data. Minimization of delay and queues are used as objective functions and reductions in delay of approximately 16% at a higher demand scenario and 10% at a lower demand, compared to a 'well-tuned' fully actuated control strategy, are reported at a 100% connected technology penetration rate. Reduced benefits are still reported at a 50% penetration rate. At a penetration rate of 25% the performance is worse than the conventional COP strategy for the lower demand scenario suggesting that

V2I begins to provide a benefit between 25-50%. At a higher demand level, V2I provides a benefit even at 25% penetration rate.

Lee et al. (2013) develop a cumulative travel-time responsive algorithm utilizing Kalman filtering. At 100% penetration rate a reduction in delay of 34% is reported although this is compared to a generic VA strategy based on optimized timings from TRANSYT rather than a more sophisticated on-line optimized strategy. The paper reports that at least 30% of vehicles are required to be equipped with V2I to achieve benefits.

Yang et al. (2016) develop a control method for connected and automated vehicles that integrates the three stages of technology development. The control method switches between methods using heuristically defined thresholds of technology penetration rate. The control method is compared to a basic VA strategy with conventional detection and it is found that the developed algorithm requires approximately 50% penetration rate of V2I vehicles before it outperforms the conventional method.

2.4.3 Removing physical traffic signals

The third category assumes that all vehicles must be equipped with V2I and is an extension of optimization of vehicle trajectories to the point that physical traffic signals can be dispensed with altogether. The premise of this research theme is that, if all vehicles are equipped with V2I (and a high level of autonomy) then it will be possible to manipulate the trajectories on the approach to a junction to the point that the vehicle paths will interleave and there will be no, or very little, requirement for vehicles to stop. In this case, physical traffic signals would no longer be required.

Ahmane et al. (2013) assume a 100% penetration rate and use Timed Petri Nets with Multipliers to enable vehicles to negotiate a right of way through a

junction. The reported reduction in average stopped times compared to traffic signals is impressive (between 40% and 60%) but the stopped time is not the full delay that the vehicle experiences and the traffic signal control strategy used for comparison is not reported. Lee et al. (2013) also propose a new 'Connected Vehicle Intersection Control' system that manipulates the speed of approaching vehicles to negotiate a safe path through the junction. The reduction in average stopped time (again, not the delay) is reported as 99% compared to generic AC with a 44% reduction in fuel consumption expected. Travel time, perhaps a truer indication of delay, was reduced by 33%.

The promise of significant reductions in delay and stops is an appealing prospect but there are some practical issues with this method. Firstly, the requirement for 100% V2I coverage (and a high level of autonomy to accurately manipulate trajectories) is possibly unrealistic, at least in the short-medium term, given that there is likely to be resistance to a fully autonomous vehicle fleet. Secondly, even if it were possible to fully automate the vehicle fleet, cyclists and pedestrians will still need to be catered for in most cases.

2.4.4 Discussion and research gap

In each of the three research areas there have been interesting studies and promising algorithms developed. For trajectory optimization applications, the research has generally been undertaken with fixed, or near-fixed, traffic signal control. The use of fixed-time control raises the question of whether the benefits provided by GLOSA applications outweigh the performance difference between fixed-time and dynamic signal control. This is probably more the case at isolated junctions than in networks where signal timings are generally more constrained. The research also suggested that GLOSA applications are most effective in under-saturated traffic conditions and perform less well during congestion.

The traffic signal optimization application of V2I is effectively the alternative approach to finding performance gains at traffic signals. This approach allows traffic signals to operate dynamically with the performance gains being realised through improved prediction of queue lengths and vehicle arrival times.

The study of Feng et al. (2015) is perhaps the most representative study as it builds on an existing vehicle actuation control strategy (COP) with a state estimation model for connected vehicles. Yang et al. (2016) also use a car-following model to predict where conventional vehicles should be but compares to a less sophisticated vehicle actuation control strategy. However, although Yang et al. (2016) develop a method for switching between control methods, depending on V2I/autonomous vehicle penetration rate, there is still an acceptance that the strategies do not provide benefits until the V2I penetration rate is at least approximately 30%. That leaves an, as yet unknown, period of time that such strategies cannot realistically be implemented.

There are two identified research gaps in this section. The first is that the reviewed literature focuses on developing control methods that solely use connected vehicle data. More than one method (Feng et al. (2015) and Yang et al. (2016)) use the connected vehicle data to infer the presence of unequipped vehicles but none attempt to use existing conventional detection to provide data to assist the algorithms. The second gap is that there is very little mention of V2I range and how the varying range of the wireless communication technology can affect the performance of the proposed control methods.

The first objective for this research will be to develop a control method capable of incorporating a hybrid detection traffic model that can make use of conventional and connected vehicle data. The second objective will be to assess the impact of V2I range on the performance of a developed control method.

2.5 On-line traffic models

On-line traffic models are used by traffic signal optimization strategies to construct a representation of traffic conditions on the approaches to a junction. The traffic model enables the optimizer to predict the arrival times of vehicles at the junction and to make decisions on when junction approaches should be served.

There are various types of traffic model that have been developed for off-line and on-line traffic signal optimization applications. Macroscopic models tend to describe traffic behaviour at a high level of aggregation using properties such as traffic density, mean speed and flow. Models that are based on hydrodynamic theory (Lighthill and Whitham, 1955) treat traffic flow in a similar way to a fluid and tend to be used to represent traffic behaviour in large networks. Microscopic models describe the movement of individual vehicles through the network using various techniques but generally based on the interaction between a following and preceding vehicle to maintain a safe following distance.

2.5.1 Vertical queue model

The vertical queue model, a form of macroscopic model, is sometimes referred to as a point-queue model and provides a representation of vehicle progress along a link that does not consider interaction between vehicles. It is widely used in on-line traffic models and, in its simplest form, vehicles are modelled as travelling at a constant speed until they reach the stop line. At the stop line they join a 'vertical' queue if the signal state is red or the queue has not yet fully discharged. The MOVA control strategy uses a simple vertical queue model in the form of a shift register (Vincent and Peirce, 1988). In the shift register, the

link is split into multiple segments with the length of each segment based on the average free-flow speed of vehicles along the link and the model time-step.

In the case of SCOOT, the detector state (i.e. whether or not a vehicle is present) is evaluated at quarter second intervals. The presence data is converted to a Link Profile Unit which is moved a segment closer to the stop line every modelled time-step. However, SCOOT also incorporates the Platoon Dispersion Model (PDM) developed in the off-line optimization tool TRANSYT. The PDM is a more sophisticated vertical queue model that takes into account the tendency for platoons of vehicles to become less defined as they travel along a link (Figure 2.5). For network optimization this is an important feature as assuming no platoon dispersion could provide overly optimistic results.

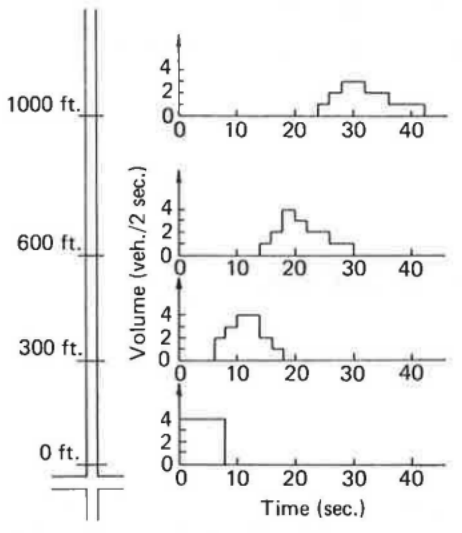

Figure 2.5: Representation of platoon flow pattern (McCoy et al., 1983).

An advantage of the vertical queue model for on-line applications is that it requires very little computational power compared to more complex models. However, the modelled vertical queue has no physical dimensions and is therefore limited in its ability to fully capture the impact of queue spill-back at upstream junctions. As a consequence, this type of model performs reasonably

well in low to moderate congestion but is less successful in congested or over-saturated conditions.

2.5.2 Cell transmission model

The cell transmission model (CTM), first presented by Daganzo (1994), is a macroscopic model that provides a convergent approximation to the Lighthill-Whitham-Richards model. CTM is a spatial model that splits links into a number of cells based on the distance a typical vehicle will travel during one model time-step. Queuing is incorporated into CTM by applying the constraint of a maximum number of vehicles being present in a cell at a particular time (i.e. jam density) as well as the maximum number of vehicles that can flow into the cell in the next time-step (i.e. capacity).

The effect of bottlenecks and queue spill-back at upstream junctions can be more accurately modelled by CTM than with a vertical queue model and it is therefore more useful for congested conditions. Attempts have been made to incorporate CTM into an on-line control strategy (Lo, 1999) but it is worth noting that Van den Berg et al. (2003) and Aboudolas et al. (2007) have questioned the applicability of CTM to real-time control applications given the computationally intensive nature of the model.

For this research, however, the applicability of CTM on computational grounds is not the main consideration as it does depend to some extent on the optimization method being employed. For example, car-following models that are discussed in later in this section can be equally, if not more, computationally intensive. Instead, the general applicability of macroscopic models to the problem of minimizing vehicle stops or emissions is the focus. Neither vertical queue models nor the cell transmission model consider the interaction between individual vehicles. Consequently, the impact at a microscopic level of the interaction between vehicles with different acceleration characteristics cannot be

accurately evaluated. For this reason, it is not considered the optimal choice for this application where minimizing stops by vehicle type will be part of the investigations of the study.

2.5.3 Store-and-forward model

The store-and-forward model is another macroscopic model that simplifies the representation of the network compared to a cell transmission model on the basis that accurate representation of link-internal traffic flow has limited significance in an interrupted traffic flow network (i.e. networks with signal control). In the store-and-forward method, the model describes a continuous average outflow from each network link, as long as there is sufficient demand upstream and sufficient space downstream (Figure 2.6, where *S* is the saturation flow, *G* is the green time, *C* is the cycle-time and *u* is the modelled outflow).

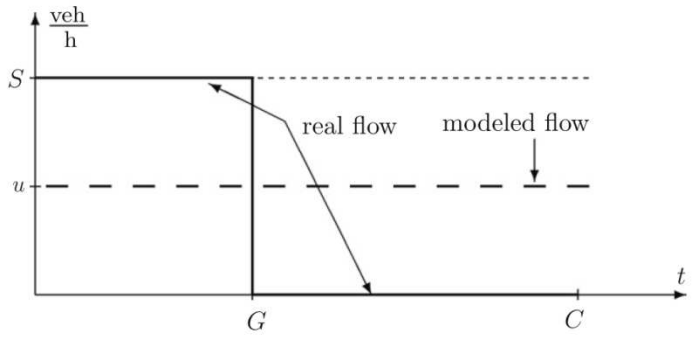

Figure 2.6: Simplified store-and-forward outflow model (Aboudolas et al., 2009).

Store-and-forward methods, most notably the TUC strategy, have shown promise when applied to congested networks. As Aboudolas et al. (2009) note, the most suited control objective under congested traffic conditions is to minimize the risk of oversaturation and spill-back of link queues. However, even if that is achieved at an area level by applying a store-and-forward method, it could still be beneficial at a local level, if also attempting minimization of stops and/or emissions, to consider the microscopic interaction between vehicles.

2.5.4 Car-following model

Car-following models address the primary issue raised with the previously discussed types of model by considering the interaction between vehicles at a microscopic level. There have been many car-following models developed but two of the most well-known models are implemented in the Aimsun (Gipps, 1981) and PTV-Vissim (Wiedemann, 1974) micro-simulation software packages.

The Gipps (1981) model is one of a class of collision avoidance car-following models. In this type of model, a following vehicle attempts to maintain a safe distance so as to avoid a collision. The Gipps (1981) formula uses two terms to describe constraints associated with the acceleration, maximum speed and deceleration of a following vehicle. These terms are used together to determine the speed of a following vehicle.

The first term describes the following constraints:

- A vehicle should not exceed the desired speed of its driver; and
- The acceleration of a vehicle should first increase with speed as engine torque increases and then decrease to zero as the vehicle approaches the desired speed (Aghabayk et al., 2015).

The second term ensures that a following vehicle maintains a safe distance to the preceding vehicle. That is, the following driver must be sure that their vehicle will stop safely if the leading vehicle brakes suddenly.

The (Wiedemann, 1974) model is a psycho-physical car-following model based on drivers' perception of the difference in speed between their vehicle and the preceding vehicle. The model assumes that a driver is in one of four states, either free driving, approaching, following or braking. The basic premise is that the driver of a faster vehicle begins to decelerate as they reach the point which they perceive the speed difference. However, because drivers cannot accurately

determine the exact speed of a preceding vehicle, the following vehicle will fall further behind the preceding vehicle until the point that the driver perceives the difference and begins to accelerate again. This results in a sub-conscious iterative process of acceleration and deceleration which results in oscillation of the gap between the vehicles.

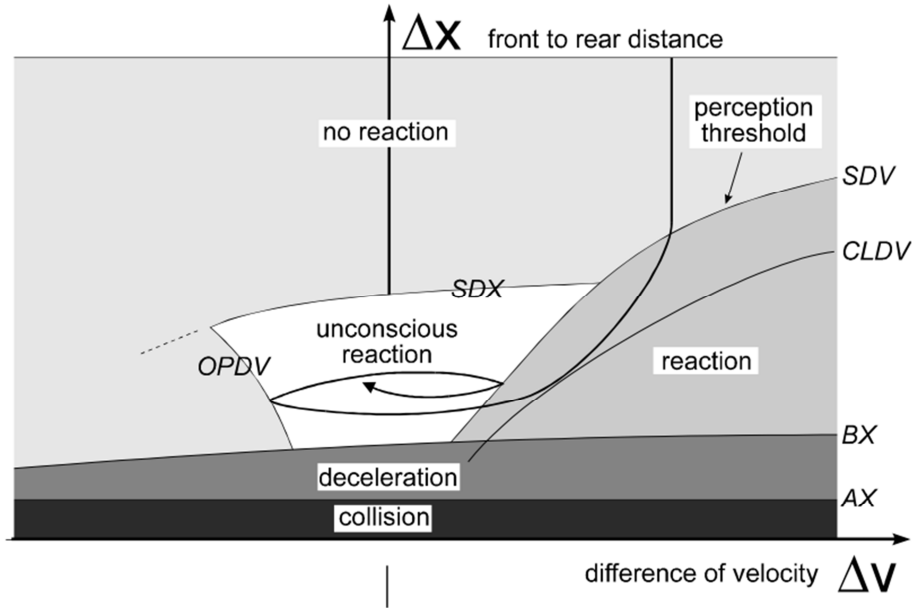

Figure 2.7: Typical car-following behaviour of a vehicle (PTV-Vision, 2011).

Figure 2.7 shows the typical behaviour of a vehicle as it approaches a preceding vehicle, first crossing the SDV threshold (the point at long distances where drivers perceive a speed difference when approaching a slower vehicle). At this point the driver decelerates but then decelerates more severely when reaching the CLDV threshold (the point at short distances where drivers perceive their speed is higher than the preceding vehicle). Finally, the vehicle enters the state of unconscious reaction where the gap between the vehicles oscillates between the SDV and OPDV (the point at which a driver perceives they are travelling at a lower speed than the preceding vehicle) thresholds (Aghabayk et al., 2013).

2.5.5 Discussion and research gap

The traffic models reviewed in this section are a selection of the most widely used on-line traffic models. As discussed, the macroscopic models, although useful for larger network simulation, do not consider the interaction between individual vehicles. However, computationally, they are generally much more efficient than microscopic models and the vertical queue model has been used extensively in on-line control strategies.

To model the more complex interaction of vehicles with different acceleration characteristics, the traffic model is required to model the behaviour of individual vehicles and that can be achieved by using a car-following model. It is acknowledged that this significantly increases the computation required to evaluate the model but when applied to an individual junction, and depending on the method used, it is possibly less significant.

The identified research gap is the application of a traffic model, capable of considering the interaction between individual vehicles, to an on-line traffic signal optimizer. The objective for this research is to identify whether the use of a microscopic model can provide a benefit to the performance of an optimizer in terms of its ability to minimize stops.

2.6 Summary

The review of literature in each of the four areas has identified various research gaps, from which some clear objectives have been identified. The objectives of this research, based on those set out in Chapter 1, are summarized below:

1) To investigate whether a single detector vehicle classification method can be developed, using only existing infrastructure, which is capable of explicitly classifying HGVs. Then, to assess whether the accuracy of such a

method is adequate to provide a performance benefit (in terms of minimizing delay and stops) by modifying an existing traffic signal optimization technique;

2) To develop a control method capable of incorporating a hybrid detection traffic model that can simultaneously make use of conventional and connected vehicle (i.e. V2I) data in order to test whether benefits of V2I can be realised sooner than the, approximately, 30-50% penetration rate described in current literature;
3) Extend the hybrid model to utilise V2I-based vehicle classification data and investigate the effect on optimizer performance;
4) Apply a microscopic traffic model to an on-line responsive traffic signal optimizer to identify whether the more detailed representation of vehicle interaction can provide a benefit to the performance of an optimizer in terms of its ability to minimize stops; and finally
5) Assess the impact of V2I range on the performance of the developed control method.

Chapter 3
Methodology

3.1 Introduction

This chapter describes the construction of a Simulated Environment that incorporates a state-of-the-art control algorithm for isolated junctions. The Simulated Environment is employed throughout subsequent chapters to test various modifications to, and ultimately a replacement for, the state-of-the-art control algorithm introduced in the following sections.

The state-of-the-art traffic signal control algorithm for isolated junctions, described in this chapter, will be used as the benchmark in later comparisons. The Miller method has been selected for this purpose. The Miller method is in common use in the UK as a component of the popular MOVA strategy and incorporates a reasonably sophisticated optimizer, providing a more useful comparison than a generic VA strategy.

Modifications are made to the Miller algorithm in this chapter to provide a representation of MOVA. The development of a MOVA representation was required, as opposed to simply using the MOVA software package, because the core components are modified in chapters 4, 5 and 6. MOVA is proprietary software and, consequently, those modifications would not have been possible.

A 20 site MOVA trial demonstrated an average reduction in delay of ~13% compared to up-to-date System-D VA timings (Peirce and Webb, 1990). Any performance gains provided by enhancements or alternatives to the representation of MOVA presented in this chapter (proposed in chapters 5 and 6) are further to this reduction.

Figure 3.1 provides a flow chart demonstrating the interaction between the various components of the MOVA representation and developments made in chapters 4, 5 and 6. The next section describes the first component of the MOVA representation, the Miller algorithm.

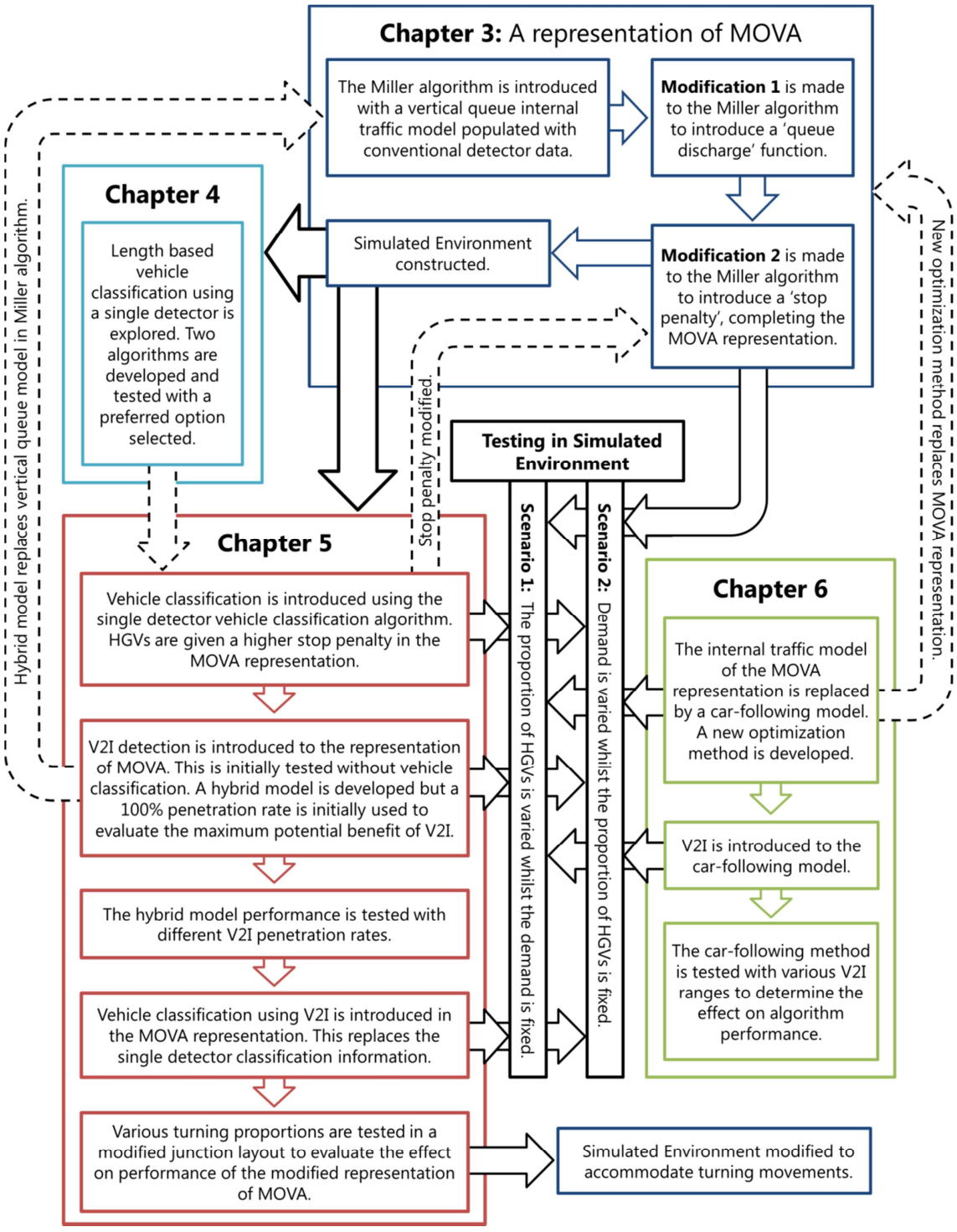

Figure 3.1: A flow chart demonstrating the interaction between various elements of the thesis.

3.2 A state-of-the-art control algorithm

Generic VA strategies are capable of providing a benefit over fixed-time plans at isolated intersections, particularly during under-saturated conditions. The efficiency of operation is improved by the ability to re-allocate available green time to alternative approaches by monitoring gaps in the traffic. System-D VA is an example of a VA strategy and is widely used in the UK. A typical detector layout is shown in Figure 3.2 and the basic logic is described in Figure 3.3.

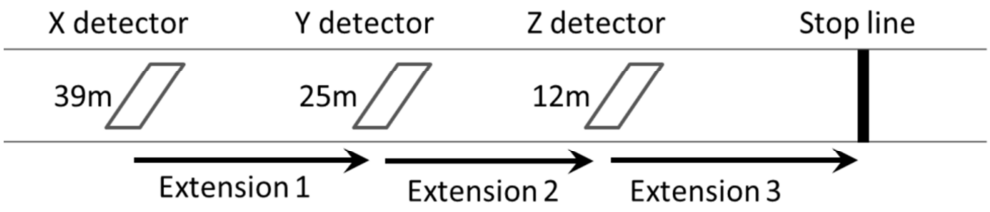

Figure 3.2: VA detection layout.

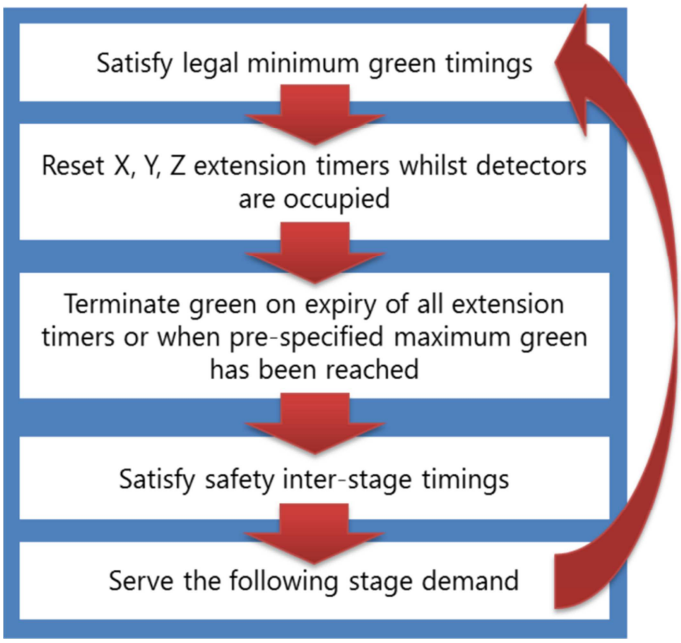

Figure 3.3: VA logic.

Each detector has an associated extension timer that is set to enable vehicles travelling at a reasonable speed to reach the downstream detector (or, in the case of the Z detector, the stop line). The green signal is terminated at the point all timers have expired on all approaches and there is demand for a conflicting stage.

The performance of the System-D VA strategy relies on appropriate maximum green times being set by practitioners, ideally based on timings derived from a well validated network model. Fine adjustments can then be made on street, based on observations. However, in much the same way as for fixed-time plans, even a well implemented system requires regular reviews of the timings to minimize degradation of performance.

Operationally, the System-D method is prone to extending the green whilst traffic is at a low intensity, even when there is a significant queue on an opposing approach. At junctions where multiple approach lanes receive green in the same stage (the vast majority of sites), the system requires simultaneous gaps in traffic across all the lanes for the green signal to be terminated.

The System-D method does not incorporate a traffic model and has no concept of the intensity of traffic on the approaches to the junction. As a result, the system does not attempt to perform optimization in any way beyond the basic logic. Therefore, this thesis considers the more complex case with an internal traffic model and an optimization approach.

3.2.1 The original Miller method

For the purposes of this research, the Miller optimization approach (Miller, 1963) that underpins the MOVA control strategy is used as the basis for testing the effect of V2I and vehicle classification. This is due to the ease of which the control method can be replicated - ensuring that comparisons are not affected

by the optimization process itself – and that it is capable of systematic optimization, rather than simply a series of heuristically derived logical steps.

The Miller method estimates the difference in delay to traffic at a junction over one cycle caused by changing the signals in 1 time-step (*h*), 2 time-steps (2*h*) etc compared to changing immediately with a suggested time-step of two seconds. The difference in delay is estimated for each time-step up to a horizon of approximately 8-10 seconds, depending on the upstream detector location. If changing the signals immediately causes the least delay then the signals will change, otherwise the current signal is extended and the process repeated at the next time-step until a pre-specified maximum green limit is reached (usually set for policy reasons).

As described in the introduction section, the following modifications are made later in this chapter, in section 3.2.2 and section 3.2.3, to mimic the 'conventional' approach as represented by MOVA:

- **Modification 1** – Identifies when the queue discharge rate drops significantly below the saturation flow rate (i.e. vehicles are travelling at an approximately constant speed) in order to begin the Miller optimization process; and
- **Modification 2** – Adds minimization of stops to the objective function.

First, the original Miller method is introduced. The method works with a variable cycle; it 'scans' the junction in a pre-determined time interval of *h* seconds and determines whether to change to the next stage or to extend the green time by *h* seconds to the stage that is currently on green. The decision is made based on minimizing the total delay of traffic through the junction. The workings of the Miller method are illustrated in a simple two stage junction with a north-south,

east-west layout shown in Figure 3.4, although as Miller noted, the same argument can be applied to any type of junction.

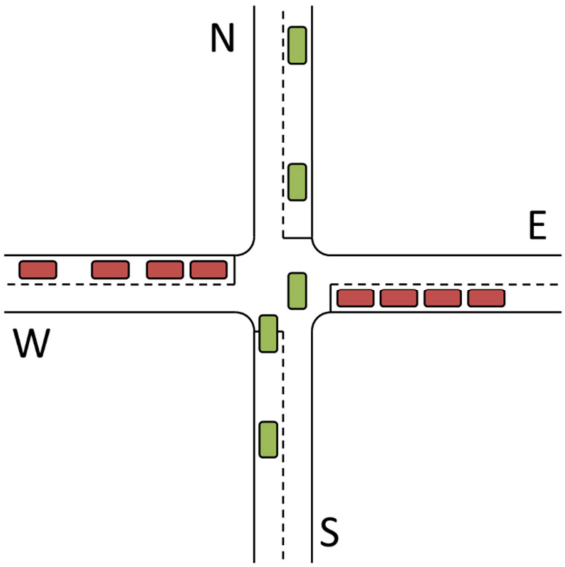

Figure 3.4: A two stage junction.

The calculations for delay are split into two parts. The first part calculates the *saving* in delay that would benefit traffic that is currently receiving a green signal if that green were to be extended *h* seconds (Miller suggested *h* = 2 seconds). For example, if the N-S stage is receiving a green signal, the effect of the green signal continuing for a further *h* seconds is considered. Miller showed that the additional number of N-S vehicles which can cross the stop-line (or at least travel to a 'point of no return') during the *h* seconds green extension would be:

$$N_{NS} = \delta_N + \delta_S - q_N \left(\frac{1 - \delta_N / s_N}{1 - q_N / s_N} \right) - q_S \left(\frac{1 - \delta_S / s_S}{1 - q_S / s_S} \right) \qquad (3.1)$$

where:

δ_N = Number of vehicles crossing the stop-line during the *h* second green extension from approach *N*.

q_N = Arrival rate of vehicles on approach *N* per *h* second green extension.

s_N = Saturation flow rate of vehicles on approach *N* per *h* second green extension.

These vehicles will no longer have to wait for the next green, and save a total wait time which includes the amber period, red period, and the lost time due to acceleration at the start of the next green. Thus the total saving of delay to the N-S traffic, if the green is extended by *h* seconds, would be:

$$D_{NS} = N_{NS}(a + r_{NS} + l_{NS}) \tag{3.2}$$

where:

a = Amber period (seconds).

r_{NS} = Length of next red phase for the *N* and *S* approaches (seconds).

l_{NS} = Start-up lost time at the beginning of the next green signal for the *N* and *S* approaches due to acceleration (seconds). This includes the red/amber period.

For each vehicle that crosses the stop-line during the green extension *h,* the following vehicle moves up by one in the queue and is therefore closer to leaving during the next green signal than if the green was not extended. However, by extending the current green signal, more vehicles will arrive before the next green signal and be subject to delay in addition to what would otherwise have been the case. The second term(s) of Equation 3.1 represent the difference between the rate at which vehicles are leaving and arriving on the link during the green extension period.

Consider the values of *N* from Equation 3.1 set out in the following table for a green extension on a single approach lane of *h* =5 seconds and saturation flow *s* = 2000 veh/h:

Table 3.1: Values of *N* from Equation 3.1 for a single approach lane.

	δ=1	δ=2	δ=3	δ=4
***q* = 360 veh/h**	1	2	3	4
***q* = 720 veh/h**	0	2	3	5
***q* = 1440 veh/h**	-4	0	4	7

The values in the table are somewhat contrived and are rounded to integer values for the purposes of clarity. However, it can be seen that where the predicted demand flow in the considered extension period is less than the number of vehicles predicted to cross the stop line, Equation 3.1 will result in a positive value (and thus a delay saving). Conversely, where the predicted demand flow is greater than the number of vehicles predicted to cross the stop line, Equation 3.1 will result in a negative value and thus cause additional delay. Where the values are equal, there is no delay saving by extending the green signal for that value of *h*.

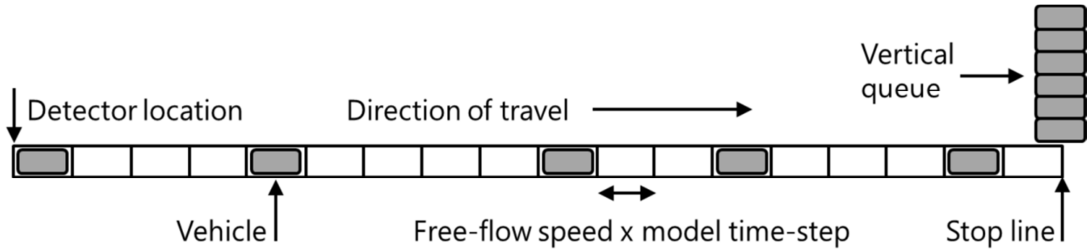

Figure 3.5: Schematic of a vertical queue model.

The Miller algorithm utilises a simple vertical queue model, although it is described by Vincent and Peirce (1988) as a shift register. Vehicles are 'shifted' along the link in cells, the number of which is defined by the calibrated free-flow speed and the model time-step (Figure 3.5). If the signal state is red then, once

vehicles have travelled the full length of the link, vehicles are added to a vertical queue (n_w in Equation 3.3). The length of the next red phase, *r,* for this thesis has been estimated by considering the discharge time required for any currently queued vehicles and any additional vehicles expected to arrive before the initial queue has finished discharging. The length of the red phase could alternatively be estimated using an estimator such as a moving average or exponential smoothing based on the length of the red phase in previous cycles.

The second step of the Miller calculation is to determine the delay *caused* to the vehicles waiting at the junction on the east and west approaches by extending the green for north-south traffic. Miller's second equation estimates the time required to discharge the initial queue on approach *W* as well as the additional queue formed during that process. The same process is also applied to approach *E.* Equation 3.3 calculates the estimated time to discharge traffic on the opposing approach *W* if the decision to change the signals is taken in the next time-step (of 2 seconds in this case). The value of *k* is the smallest integer such that:

$$n_W + \sum_{i=1}^{k_W} q_W - \sum_{i=2+l_W/h}^{k_W} s_W \leq 0 \qquad (3.3)$$

where:

n_W = Number of vehicles already waiting on approach *W*.

k_W = Number of units of *h* seconds of time until the queue on approach *W* is discharged.

l_W = Lost time on approach *W* before saturation flow is achieved (seconds).

Then the total delay to the E-W traffic caused by extending N-S green time by *h* seconds is:

$$D_{EW} = h\left[n_W + n_E + \sum_{i=1}^{k_W} q_W + \sum_{i=1}^{k_E} q_E \right] \quad (3.4)$$

The value of *k* will almost certainly differ between approaches. However, vehicles arriving after a queue has discharged will not experience delay and so only the vehicles in the initial queue and those added to the queue during the period of queue discharge are considered in Equation 3.4.

Equation 3.4 is then subtracted from Equation 3.2 to provide a quantity defined by Miller as the 'test quantity T'; the difference between delay *saved* and delay *caused* by extending the green. A negative value of T implies that extending the green is not beneficial and that the signals should change immediately. However, values of T for green extensions of *2h*, *3h* etc must also be calculated as it is possible that an extension of the current stage beyond the next time-step may cause less delay. The value of T is calculated for each time-step up to the upstream detector location (usually 8-10 seconds from the stop-line).

A junction with multiple stages would operate in the same manner except that for each additional stage, a further instance of Equation 3.4 would be calculated for the relevant approach lanes, with the lost time l_w, in Equation 3.3 modified for each stage to account for the length of time required to discharge the traffic in the stages that precede it.

3.2.1.1 Constant speed assumption

The Miller method, using conventional detection, relies on vehicles travelling at an approximately constant speed to be valid due to the absence of live, vehicle specific, speed data.

Consider Equation 3.1, used to calculate the number of vehicles that will experience a saving in delay if the green is extended by a time-step *h*. At the beginning of the green, in most cycles, vehicles on the approach will be stationary and so it could be concluded that no vehicles will cross the stop-line

during any extension of the green signal (i.e. no benefit in extending the green signal). If that were the case then, with no intervention, the Miller algorithm would simply switch the signals at the end of the legal minimum green time to clear any queue on opposing approaches.

Additionally, if Equation 3.1 was employed at the beginning of the green signal then, even if the situation described above was avoided, some vehicles will be accelerating and, potentially, others decelerating until any queue has fully discharged. In this case, the predicted trajectory of the vehicles using the vertical queue model described in Figure 3.5 is likely to be inaccurate, leading to sub-optimal decisions on green extensions. As such it is necessary to first ensure vehicles are travelling at an approximately constant speed by clearing any queue on the link before entering the delay minimising process. Section 3.2.2 describes the modification made to the original Miller algorithm to address this problem.

3.2.1.2 Impact of vehicle stops

The delay saving calculated in Equation 3.2 considers the delay that would have been incurred by a vehicle if it had been stopped. The delay includes the amber period, red time and the lost time at the start of the following green due to acceleration. However, it is unclear in Miller (1963) whether the lost time value in Equation 3.2 incorporates the full delay incurred by a vehicle stop or whether it only includes the time required for the vehicle to cross the stop-line. If it is the latter then there is an additional delay saving for the vehicle if it does not stop that is not accounted for by Miller (Figure 3.6).

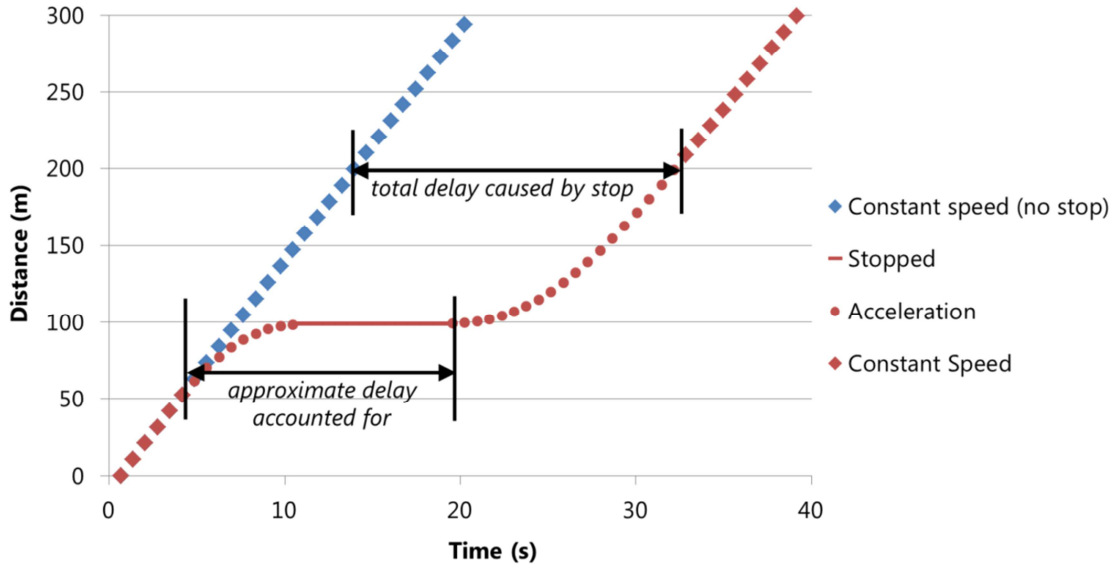

Figure 3.6: A space-time diagram showing delay caused by a vehicle stop.

It is also the case that the Miller approach uses the single objective function of minimizing delay. The single objective function is a limitation if, for example, a practitioner wishes to place a greater emphasis on, say, air quality. In that case it may be beneficial to reduce the number of stops at the expense of a small increase in delay. Section 3.2.3 explains how vehicle stops have been incorporated into the Miller method to provide a combined objective function of vehicle delay and stops.

3.2.2 Modification 1: Queue discharge

Identifying the moment at which a queue has cleared with conventional detection is not completely straightforward. For example, a gap between vehicles longer than the average headway for discharging traffic could indicate that the queue has finished discharging but it could also be caused by a vehicle with slow acceleration or a lack of driver attention. The MOVA software incorporates a 'critical gap' function (Vincent and Peirce (1988) suggest a gap of approximately 3.5 seconds) using a supplementary 'X' detector sited approximately 3.5 seconds from the stop-line (Crabtree, M.R., 2011).

The critical gap is measured from the moment a vehicle leaves the X detector and is reset each time a new vehicle is detected until the pre-specified gap value is reached. Consider the example of a queue that extends to the X detector but with vehicles not positioned on the X detector itself (Figure 3.7). If the critical gap is timed from the start of the green signal then vehicles in the queue at the point of the X detector will not have started moving by the time the critical gap value has been reached. In that situation, the end of the queue discharging will have been falsely identified.

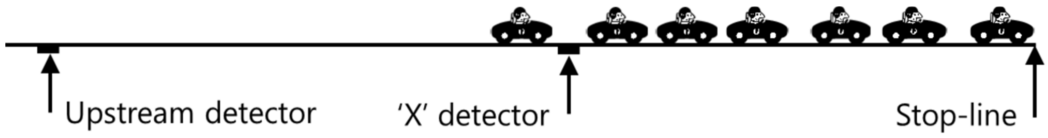

Figure 3.7: Example of queued vehicle behaviour.

To overcome this issue, a variable minimum green is used in MOVA to reduce instances of spurious gaps in traffic, early in the green, being falsely identified as the end of a queue (Crabtree, M.R. et al., 2012). A variable minimum green value is preferred to a fixed minimum green as the number of vehicles between the stop-line and X detector can vary from one cycle to the next. The number of vehicles depends on the length of the vehicles and the headway between vehicles at standstill.

MOVA is proprietary software but for this work a variable minimum green time has been calculated based on the average headway of the vehicles that have crossed the X detector during the red signal as shown in Equation 3.5. The search for the critical gap does not begin until the variable minimum green time has elapsed.

$$g_{\min} = \frac{\delta_{red}}{s_{\theta}} + l_{\theta} \tag{3.5}$$

where:

$g_{\min}$ = Minimum green (seconds).

δ_{red} = Number of vehicles to cross 'X' detector during red.

s_{θ} = Saturation flow rate on any approach θ.

l_{θ} = Lost time due to acceleration on any approach θ.

Miller (1963) does not consider the case of a queue extending beyond the upstream detector. However, it is clear that the estimation of queue discharge time could be significantly underestimated if vehicles beyond the upstream detector are ignored. A simple queue estimation algorithm has therefore been incorporated for this thesis as it is assumed that MOVA must also incorporate one. Once a queue has reached the detector furthest upstream of the stop line there is no longer any information provided to the model regarding vehicle arrivals. The queue length is approximated, for this thesis, by increasing the length of the queue each model time step by the average arrival rate (veh/s) multiplied by the model time step as shown in Equation 3.6.

$$Q_t = Q_{t-1} + (i \cdot q) \tag{3.6}$$

where:

Q_t = Length of the queue at time t (veh).

i = Model time step (seconds).

q = Average arrival rate (veh/s).

3.2.3 Modification 2: Stop penalty

In addition to ensuring the full delay caused by a vehicle stop is accounted for, the introduction of a value of delay per vehicle stop enables minimization of stops to be added to the objective function.

The MOVA control strategy makes use of fixed value 'stop penalties', and, although it is a proprietary software it is known that the weighted stops are summed with the net delay (Vincent and Peirce, 1988). In order to provide a more accurate representation of MOVA, it is proposed that the stop penalty is incorporated into Miller's model by modifying Equation 3.2 as shown below:

$$T_{NS}^{sp} = D_{NS} + SP(\delta_N + \delta_S) \tag{3.7}$$

where:

SP = Stop penalty value (seconds).

T_{NS}^{sp} = Total delay and weighted stops saving for vehicles on approaches *N* and *S* (seconds).

The stop penalty concept will be varied in later chapters as it is useful as a means to provide higher penalties (and therefore increased weighting) for HGVs.

3.3 Construction of the Simulated Environment

The following sections describe the construction of the Simulated Environment used throughout this thesis, including the software architecture of a signal controller (see Appendix D) developed for this research to enable testing of the various developed algorithms. The chapter also details the scenarios tested and the evaluation criteria used to determine the performance of the algorithms in Chapter 5 and Chapter 6.

A micro-simulation modelling environment (PTV-Vissim) was used to mimic real world conventional detection and V2I technology. PTV-Vissim is an example of a micro-simulation modelling tool that employs a car-following and lane changing model to simulate traffic behaviour at an individual vehicle level. The car-following model simulates the interaction between vehicles travelling on the road network by providing a representation of the behaviour of a following vehicle in response to the trajectory of a leading vehicle.

The PTV-Vissim software package was chosen for this thesis as it enables external signal controller modules to be developed that can integrate with the model environment to receive data from simulated conventional vehicle detectors and to control the state of the traffic signals. The software also provides the option to write live vehicle record data (i.e. position, speed, acceleration, unique identifier) to a database at a chosen frequency. This can be read by external software and used to represent the data that could be received through V2I communication.

As vehicles enter the network they are randomly assigned parameter values for performance and behaviour (i.e. desired speed, acceleration, deceleration, desired headway) from a pre-defined user selected distribution. This methodology introduces stochasticity to the car-following model, enabling it to provide a representation of real world conditions. As a consequence of the stochastic nature of micro-simulation, it is necessary to run each model multiple times using various random seeds that produce a different set of parameter values for each vehicle entering the network. The results of the multiple runs are then averaged to provide a performance indicator.

A schematic of the micro-simulation model configured is shown in Figure 3.8 for the single intersection, the layout of which is based on the original paper by Miller (1963). In the conventional model, detectors were placed at 50 metres

and 150 metres from the signal stop-lines on each approach. In the hybrid V2I model it was assumed that vehicles could be detected from approximately the same point as the conventional upstream detector. In reality, V2I information may be available further upstream (up to 1000 metres) by using Dedicated Short-Range Communications technology (DSRC) or further using 3G/4G/5G (albeit at a reduced frequency) but that was not initially modelled so as to provide a fair comparison between approaches.

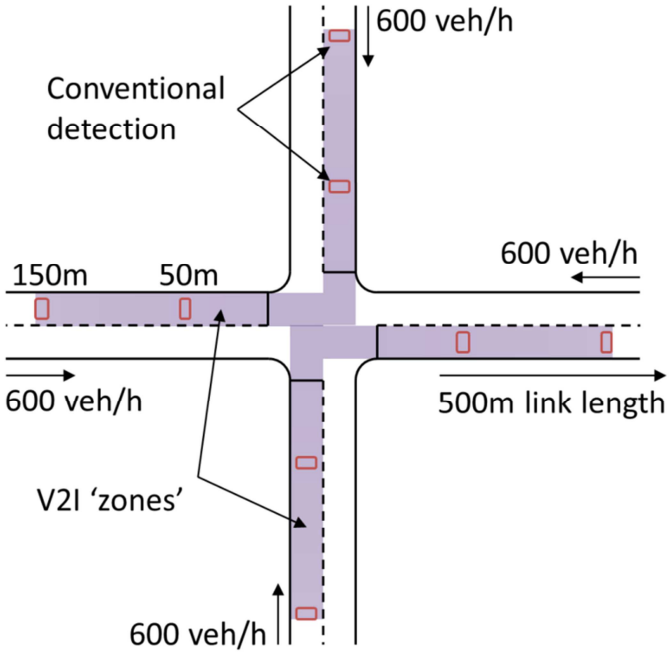

Figure 3.8: The simulated layout for the Miller method - conventional and V2I detection overlaid.

Vehicle mix and desired speed distributions for each vehicle type were input into Vissim using three vehicle 'classes' – cars/vans, rigid HGVs and articulated HGVs – based on Automatic Traffic Count (ATC) data provided by Rotherham Metropolitan Borough Council. The observed vehicle mix is replicated across all approaches of each model and the total demand on each approach is the same.

The traffic signals operate a simple two stage, north-south followed by east-west, strategy. The interstage value (5 seconds for both stages) incorporates the

leaving amber of the terminating phase and starting red/amber of the next phase.

Table 3.2: Distribution of vehicle characteristics (min and max values)

	Car/Van		HGV	
	Min	Max	Min	Max
Length (m)	4.1	8.8	10.2	18.5
Acceleration (m/s^2)		3.3		1.8
Desired Deceleration (m/s^2)	-3		-1.3	

In all the various scenarios tested for this research the results have been averaged across 10 different random seeds (consistent between scenarios). The number of random seeds was chosen as a value that provided statistical significance to the results whilst allowing the required simulations to be completed within the timescales of the study. The results include confidence intervals to demonstrate the statistical significance. The simulation period for each scenario is one hour.

The proportion of the sum of HGVs in the base scenario is 5% (4% rigid HGVs and 1% articulated HGVs), as determined from the ATC data. From herein the term HGV is used to define a combination of rigid and articulated HGVs with a 4:1 ratio. Table 3.2 shows the vehicle length distribution and values for acceleration for each vehicle class used in the simulations. Desired vehicle speed distributions, shown in Figure 3.9, were also derived from ATC data using speed data from time periods with free flowing traffic.

It should be noted that, although the data collected from the ATC site provided a continuous distribution of vehicle lengths, this was approximated in Vissim using a distribution of the closest available vehicle model types. This resulted in a small gap between car/van and HGV vehicle lengths which may affect the results described in Chapter 4 in a real-world application as, in reality, there may

be a less defined grouping of vehicle lengths. However, the percentage of vehicle lengths in the crossover region, as observed from ATC data, is small.

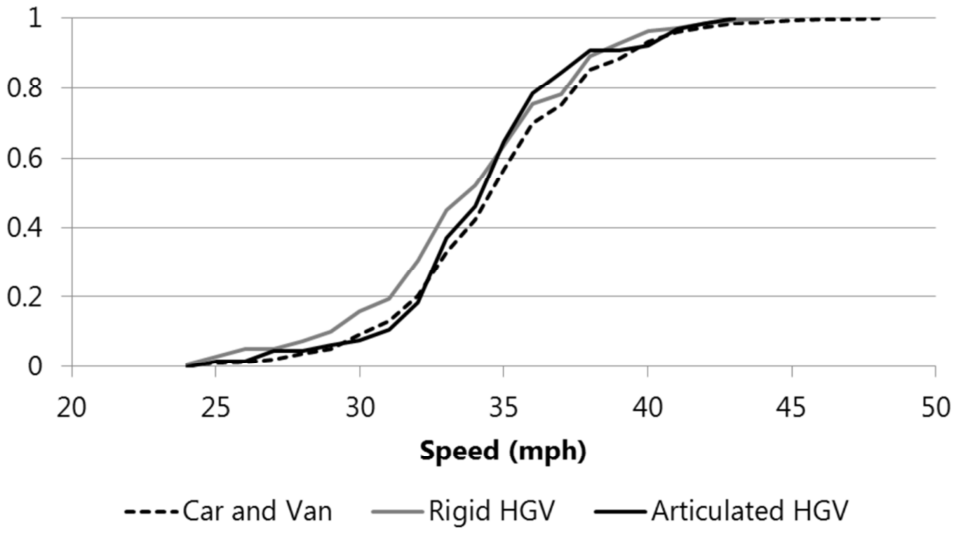

Figure 3.9: Cumulative speed distributions derived from ATC data.

In the base scenario the demand on each approach to the junction is 600 veh/h with vehicles entering the network according to a Poisson distribution. The exact nature of the implementation is unknown as it is not discussed in the Vissim documentation. Two further scenarios have been assessed:

- Scenario 1: The proportion of HGVs is varied whilst the demand is fixed; and
- Scenario 2: Demand is varied whilst the proportion of HGVs is fixed.

The degree of saturation of the junction for a given demand varies according to the vehicle mix. It has been assumed that the average number of stops provides a good indication of saturation and that a value below 0.9 stops per vehicle indicates that the junction is under-saturated. Above that value the junction is increasingly likely to be over-saturated during at least some of the simulation period and the Miller algorithm no longer valid. All reported results are for under-saturated conditions.

The delay evaluation tool available in Vissim was used to output the average total delay per vehicle. This is computed by subtracting the 'theoretical' travel time for each vehicle (i.e. travelling at its desired speed with no signals) from the actual travel time. Travel time markers were placed on the entry to the network and downstream of the signals. The delay evaluation tool also records the number of stops for each vehicle.

3.3.1 Integrating the Signal Controller with micro-simulation

To control the traffic signals in the micro-simulation model it was necessary to develop an external traffic signal controller module (using Visual Basic .NET) that receives detector data from the model every time-step (10Hz) and responds with the relevant signal state information.

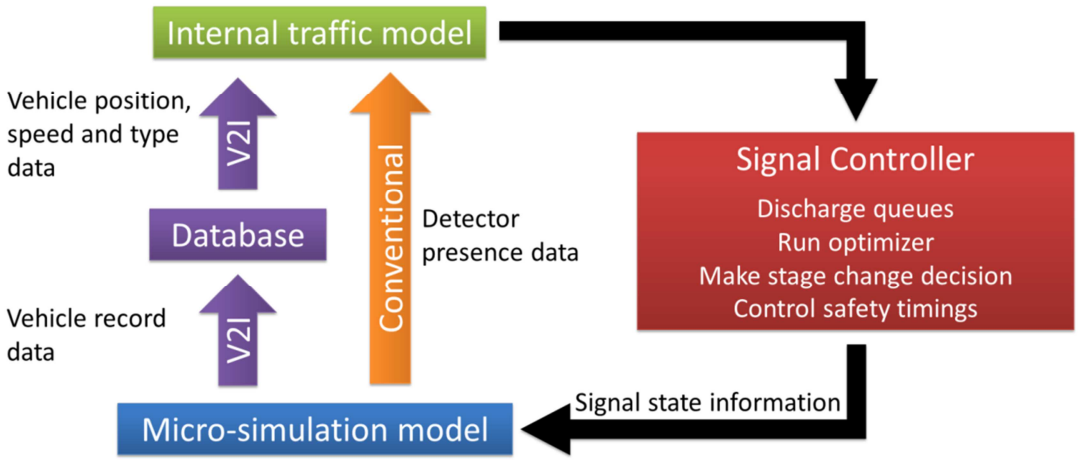

Figure 3.10: Traffic signal controller schematic – V2I and conventional.

Figure 3.10 shows how the micro-simulation model, representing the 'real-world' situation, outputs detector data, either in the form of conventional presence data or as V2I based data (or both) to an internal traffic model. The internal traffic model is then used by the optimizer, within the Signal Controller.

The Signal Controller manages several tasks. First, at the start of each green, the queue discharge process – for which the method varies depending on the

internal traffic model used (e.g. the vertical queue model in the MOVA representation) – is undertaken. When queues are deemed to have discharged on all links currently at green, the chosen optimization method is employed to make decisions on when to move to the next stage. At this point the Signal Controller assumes responsibility for managing the inter-stage process and ensuring all safety timings are adhered to. When V2I data is introduced into the MOVA representation, the vehicle record data output from Vissim into a 'live' database replaces the conventional detection method. The addition of accurate vehicle position data from the database enables the internal traffic model to become spatial thus allowing alternative methods of identifying queue discharge to be employed. Figure 3.10 provides a schematic to demonstrate the difference between how conventional and V2I data is used.

3.3.1.1 Implementation of V2I

Implementation of a V2I proxy required modification to some aspects of the Signal Controller model. Vehicle record data was written to a database from Vissim at a rate of 1Hz. In reality, V2I information may be available at a higher frequency but, given that information will be received from vehicles at different times and from different sources, it would make sense to collate the data into a database, filter it, and populate the optimizer traffic model. Using that logic it is likely that the processed information would be available at a similar frequency to that used in this research.

The external Signal Controller module was modified to enable the vehicle record data to be filtered every second from the intermediary database and the internal model updated accordingly. It has been assumed for this work that V2I is able to provide speed, position and vehicle type information. The information in the internal traffic model is updated every second from the database but the Signal Controller operates at 10Hz. For intermediary time-steps (i.e. time-steps

between database updates) the position of each vehicle is predicted based on the speed/position from previous updates. The exact nature of the prediction depends on the internal traffic model employed (e.g. vertical queue model, car-following model).

The proportion of V2I equipped vehicles within a simulation period is determined by the Signal Controller model by applying a random number generator when newly detected vehicles are entered into the internal traffic model. For example, given a random number range between 1 and 100 and a V2I penetration rate target of 20%, if the generated random number falls below 20 the vehicle will be V2I equipped. The vehicle remains V2I equipped (or otherwise) for all future time-steps.

Vehicles are equipped with V2I on a random basis, described above, that is independent of vehicle class. In reality, it is possible (although not certain) that some vehicle classes may become equipped more quickly than others. For example, if haulage companies found value in equipping vehicles with V2I, it may be that HGVs are equipped more quickly than other vehicles. It could potentially be the case that potential benefits to HGVs of V2I technology are realised more quickly than for the rest of the vehicle fleet. This could be tested in future research.

Applying V2I to some but not all of the vehicles entering the internal traffic model provided some challenges. The intermediary database is populated by Vissim with the position, speed, acceleration, time and Vissim vehicle number information for every vehicle travelling through the junction, regardless of whether it is equipped with V2I or not as this decision is made within the Signal Controller. In reality, a database collecting V2I data would only present V2I

equipped vehicles to the Signal Controller, thus removing the need for some of the logic used for this thesis (shown in Figure 3.11 as a dashed outline).

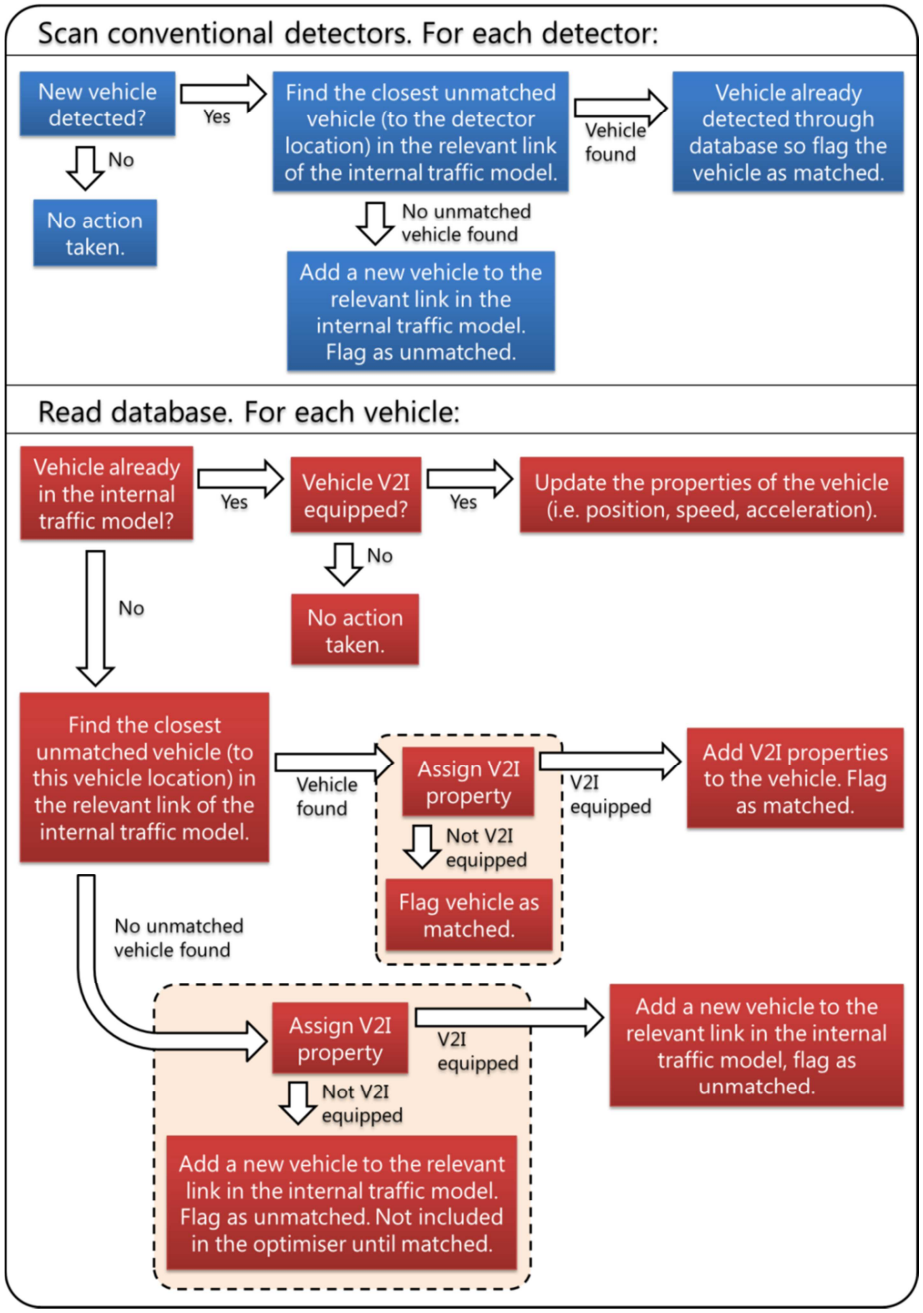

Figure 3.11: Adding new vehicles to the internal traffic model.

When the modelled V2I penetration rate is less than 100%, some vehicles will not be V2I equipped and they are instead detected through conventional detection only. However, even those vehicles that are detected conventionally must be matched with the relevant vehicle in the Vissim 'V2I' database to

prevent duplicate vehicles being entered into the model. The logic for adding new vehicles to the internal traffic model from either the conventional detectors or from the V2I database is shown in Figure 3.11.

The process of 'matching' vehicles detected conventionally and using V2I assumes that the range of V2I is at least that of the conventional detection. In doing so, vehicles can be matched at the point a conventional vehicle leaves the detector thus reducing the likelihood of internal model error.

3.3.2 Turning movements

The simulation environment described above enables the performance of each algorithm to be compared in typical conditions. However, it does not include any turning movements, in particular opposed turns. It has been assumed that unopposed left turning vehicles would have only a minor impact on performance for the algorithms described in Chapter 5 and Chapter 6 compared to opposed right turning vehicles. An additional simulation environment (Figure 3.12) was therefore developed to enable a small storage area in the centre of the junction for opposed right turning vehicles. Results from testing in this environment are shown separately in Chapter 5.

In general, the geometry of traffic signal controlled junctions is designed to accommodate waiting vehicles in the safest possible manner. In most cases, vehicles making turning movements across opposing traffic are provided with dedicated road space (i.e. a flared approach lane). TD 50/04 (Highways Agency, 2004) states that "*the storage length of the left and right turn entry lanes should be designed to meet the capacity requirements of the junction*" and that "*to avoid turning traffic blocking the adjacent lane it should be of sufficient length to accommodate the longest queue of stopped traffic*". However, in some locations there may be insufficient space to provide adequate storage capacity, or the volume of turning traffic may not be deemed sufficient to require a dedicated

lane. At junctions where such conditions exist, there is a risk that vehicles giving way to oncoming traffic may impede other vehicles.

In the modified environment the same conditions have been created for each approach in terms of turning proportions and the storage capacity within the junction. Testing in this environment has been undertaken using 600 veh/h demand on all approaches with an increasing proportion of right turners up to a proportion of 15%. The evaluation criteria used for the modified layout in Chapter 5 is derived in an identical manner to the initial testing.

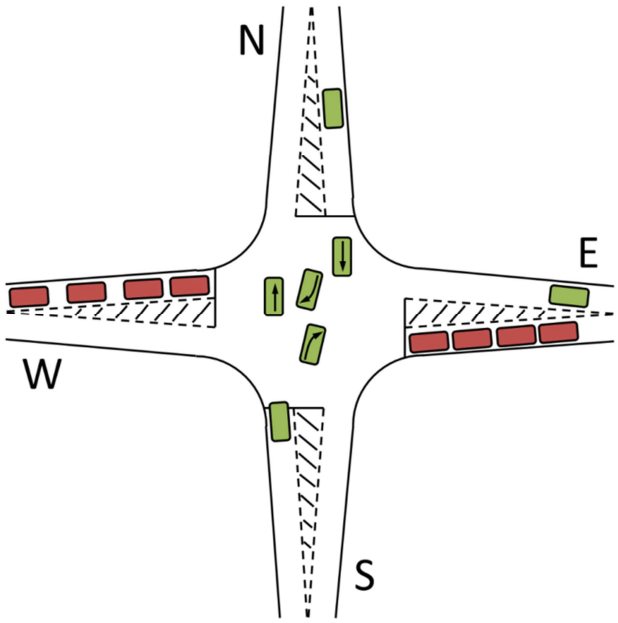

Figure 3.12: Modified layout for turning movements.

3.3.3 Evaluation criteria

A performance indicator (PI) has been calculated for each scenario to take into account delay and stops. The PI is derived by summing the Vissim model outputs of total delay with a factored number of total stops for each vehicle type (also output from Vissim). The factor for stops, or 'stops factor' for the purposes of this thesis, is an approximate value in delay of one vehicle stop.

The stops factor is used to determine the influence of stops in the PI calculation for evaluation purposes and remains constant to provide consistency between results. Robertson et al. (1980) discusses the merits of including vehicle stops in the TRANSYT PI and proposes a value of 20 seconds, the value used here, to provide a suitable balance between stops and delay that minimises fuel consumption. The result for each vehicle type is then weighted by its associated PCU value before being summed with the other vehicle types into a final PI value.

$$PI = \sum_{\alpha \in V} PCU^{\alpha} \left(D^{\alpha} + \left(20 S^{\alpha} \right) \right) / N_v \tag{3.8}$$

where:

D^{α} = Total delay (seconds) for vehicle type α

N_v = Total number of vehicles

PCU^{α} = PCU value for vehicle type α

S^{α} = Total stops for vehicle type α

V = Set of vehicle types

The stops factor of 20 seconds, used within the PI calculation, assigns a value of delay to each vehicle stop recorded by Vissim to enable stops to be considered in the PI. The stops factor remains constant whilst the PCU value provides the weighting per vehicle type. For evaluation purposes the stops factor provides a means for assessing the impact of a particular optimization method on a combination of delay and stops. The value must remain constant for all scenarios and methods to provide a fair comparison.

The value used has been chosen at 20 seconds based on the figure proposed by Robertson et al. (1980) to provide a balance between stops and delay that minimizes fuel consumption. Various alternative values have been assessed during the course of this study and have been shown not to significantly affect the outcomes for the scenarios tested.

The PI is also weighted by PCU value of each vehicle type to ensure that the size of a vehicle is factored into the overall performance (i.e. a stop is less desirable for a vehicle with inferior acceleration performance which is likely to correlate with vehicle size to some extent). In reality, the impact on emissions/fuel consumption etc is unlikely to be exactly proportional to the PCU factor. However, since the exact figures are unknown, using the PCU value would seem to provide a reasonable representation of the difference.

To investigate the sensitivity of this parameter, some alternative methods of evaluating across different vehicle types have been tested on a sample of results during the course of this study, particularly in the process of deriving optimal stop penalty values for the modified Miller method. The ratio of weighting for the evaluation criteria using the various methods is shown below:

Table 3.3: Comparative weighting per vehicle type in evaluation criteria using alternative methods.

	Car/Van	Rigid HGV	Articulated HGV
Equal weighting	1.0	1.0	1.0
PCU value (used)	1.0	1.8	3.3
Optimal stop penalty value	1.0	2.5	4.0

In Table 3.3, the alternative methods of evaluation used are:

- Applying an equal weighting to all vehicle types; and
- Applying the optimal stop penalty values for each vehicle type as calculated in Chapter 5.

The optimal stop penalty values for the sample of results tested were shown not to deviate from the values calculated using the PCU value method, even using a constant weighting, although the difference between the performance indicators decreased.

The stops factor differs from the stop penalty value, discussed in the introduction of Modification 2 earlier in this chapter. Part of the role of the stop penalty is to enable the delay that is expected to be experienced by a particular vehicle as a result of a stop to be fully accounted for during the optimization process. As a result, the stop penalty, in effect, partly contributes to the delay minimizing objective function. The optimal stop penalty value changes depending on average speed and proportion of HGV (and later individual vehicle type).

3.4 Conclusions

At the beginning of this chapter it was established that a benchmark was required to which later work could be compared. The Miller optimization method was chosen as a suitable strategy because it is relatively straightforward to implement and it is capable of systematic optimization unlike many generic VA strategies.

The methodology of the Miller method, that underpins the popular MOVA control strategy, has been explained in detail. Some modifications have subsequently been made to ensure that the benchmark used in comparisons

adequately reflects the performance that can be achieved by a 'state-of-the-art' control system, in this case MOVA.

A 20 site MOVA trial demonstrated an average reduction in delay of ~13% compared to up-to-date System-D VA timings, a generic VA strategy used in the UK. Any performance gains provided by enhancements or alternatives to the representation of MOVA presented in this chapter (proposed in chapters 5 and 6) are further to this reduction.

The construction of a Simulated Environment has then been described including the development of an external Signal Controller model that enables different traffic models and optimization methods to be tested. Finally, the evaluation criteria used in chapters 4, 5 and 6 has been explained.

Chapter 4
Single detector vehicle classification

4.1 Introduction

This chapter describes the development of a speed estimation algorithm for the purposes of instantaneous length based classification of vehicles leaving a single detector. The proposed application of the algorithm is to utilise existing detection at traffic signal controlled junctions, commonly a single upstream detector on each approach lane, to provide reliable vehicle classification data for optimization strategies. The chapter investigates whether, through the use of various filtering processes, a technique primarily employed for freeway applications can provide speed estimates that are accurate enough to classify vehicles for use in a traffic signal control strategy.

There are two main approaches to classifying vehicles from a single detector. The first approach is to use the vehicle profile produced by a vehicle chassis travelling through the electro-magnetic field created by an inductive loop detector (Figure 2.1). This shall be referred to in this chapter as the 'profile' method. The distinctive profile created by different types of chassis can be matched to a database of known vehicle chassis, thus enabling vehicles to be classified. The profile method is capable of providing a good performance if calibrated effectively but it requires specialist calibration and, in the context of a traffic signal application, additional roadside equipment compared to a standard traffic signal installation.

The second approach is to make use of the processed presence output from the detector to estimate vehicle speed, utilising the relationship between speed, flow and occupancy (described in Chapter 2). This shall be referred to as the

'flow-occupancy' method. The flow-occupancy method requires no additional roadside equipment compared to a standard traffic signal installation as the processed presence output is already passed to the Signal Controller for use in either generic VA strategies or adaptive control algorithms such as MOVA and SCOOT.

The flow-occupancy method to vehicle classification from a single detector is the focus of this chapter as it is important to first understand the extent of what could be achieved from the use of existing infrastructure. This will provide a useful benchmark from which to compare the potential benefit of V2I technology. Accordingly, this chapter investigates the accuracy that can be achieved using two alternative methods of flow-occupancy type classification.

- **Method 1** – Estimate vehicle speed over the detector independently of traffic signal control; and
- **Method 2** – Use the internal queue model of a traffic signal control optimizer to provide additional information to the algorithm from downstream of the detector.

As discussed in Chapter 2, the flow-occupancy approach to single detector classification has predominantly been used in freeway applications where the variation in speed over a detector is usually relatively small. In that context, the classification of vehicles can be performed using large sample sizes and/or over large time intervals as assumptions regarding (approximately) constant vehicle speed over a detector hold true for a large proportion of the time. However, in an urban context, particularly in close proximity to traffic signal controlled junctions, the speed of vehicles over a detector can vary dramatically from on vehicle to the next as a queue builds at a red signal and then discharges during the following green. The algorithms developed in this chapter look to address

the issues associated with estimating speed (and subsequently classifying vehicles) from a single detector in an urban setting.

First, the methodology used to develop, test and evaluate both algorithms is described in section 4.2. Next, some initial investigation into vehicle sample sizes is undertaken and some potential sources of error in vehicle presence data explained. Finally, the development of each algorithm is detailed, the results compared and conclusions drawn.

4.2 Methodology for testing

In order to develop a speed estimation algorithm, the Simulated Environment described in Chapter 3 has been utilised to provide the relevant data from conventional detectors. The Signal Controller software module is modified to enable it to record the detector presence data from the detectors furthest from the stop line on each approach (Figure 4.1).

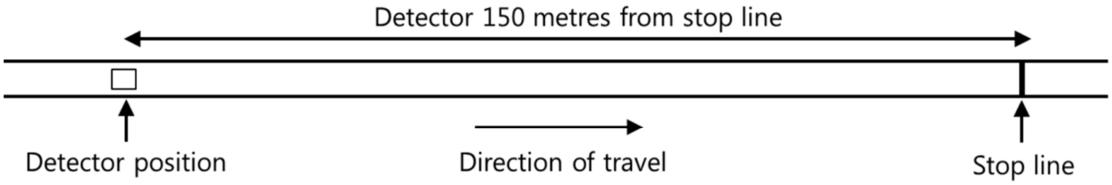

Figure 4.1: Schematic of initial model setup.

To evaluate the accuracy of the developed algorithms, the vehicle record feature in Vissim is used to write the actual vehicle speed and type information (as configured in the Simulated Environment) to a database at a 1Hz frequency. The database is subsequently filtered off-line to provide speed and type information for each vehicle at the location of the conventional detection. Figure 4.2 shows the methodology applied to provide a platform for developing and evaluating the vehicle classification algorithms. The algorithm development in this chapter

has been undertaken off-line using a spread-sheet tool but the preferred algorithm is subsequently incorporated into the Signal Controller in Chapter 5.

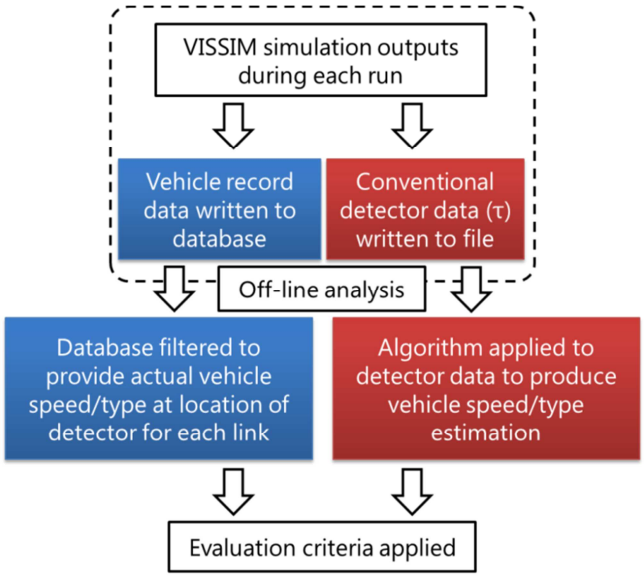

Figure 4.2: Schematic showing the testing and evaluation process.

The simulations have been run with a resolution of 10 time steps per simulation second. The vehicle presence data from each detector has been recorded at 4Hz (to mimic the data available to SCOOT) and 10Hz by the Signal Controller. The 4Hz communication rate, used to transmit processed vehicle presence data back from the roadside to a central SCOOT system, has been included in the testing as it is an existing feature of many UTC systems. In fact, the majority of UTC systems now operate with IP based communications but the software still uses legacy protocols limiting data transfer to second-by-second communication of bits (with the exception of 4Hz SCOOT data). Thus, assessing the performance of each algorithm when using data at 4Hz will determine whether the performance significantly degrades compared to 10Hz. If not, the developed algorithms could be deployed on centralised traffic signal control systems using existing infrastructure with very minor software modifications.

As shown in Figure 4.2, the Signal Controller software module is used to process the conventional detector presence data. The presence data is output in the form of vehicle 'on' times (herein referred to as τ) and is written to a file. The file is subsequently imported into a spread-sheet to enable development of the algorithms.

Early testing for this work showed that using fixed sample intervals produced a higher error rate than using a fixed sample size with a varying time interval. However, sample size and the latency of availability of the classification data are more significant factors in the development of algorithms for this application and are discussed in section 4.3.

4.2.1 Classification by length

Classification of the vehicles is performed simply by subtracting the effective length of the detector from the product of the final speed estimate and vehicle τ. As described in section 2.2.2, the effective length of the detector is not known precisely and will vary from site to site depending on various factors such as how a detector is installed. The 'average' effective length could be calibrated by site. However, the effective length will also change slightly depending on the shape and size of the vehicle chassis.

The vehicle is then classified by defining length thresholds between the desired vehicle classes (Table 4.1). As described previously, the Simulation Environment consists of three vehicle classes split into car/van, rigid HGV and articulated HGV. The algorithm has been tested with a three class bin for consistency with the vehicle classes configured in the Simulated Environment, as described in the previous chapter. However, a two class bin has also been tested to determine whether a reduced number of vehicle class bins can improve the performance of the algorithms.

Table 4.1: Length thresholds for two and three vehicle class bins.

	2 Class (m)	3 Class (m)
Car/Van	$l<=8.8$	$l<=8.8$
Rigid HGV	$l>8.8$	$8.8<l<=12$
Articulated HGV	*as above*	$l>12$

4.2.2 Evaluation criteria

The performance of each algorithm has been evaluated using three separate indicators. The first two indicators evaluate the performance of the algorithm in classifying the vehicles into either two or three class bins. The third indicator evaluates the performance when taking into account the application of the algorithm to the traffic signal optimization process described in Chapter 5.

When classifying into three bins, referred to as '3 Class' in the results section, the evaluation criterion is straightforward. The accuracy of the algorithm is simply the sum of correctly classified vehicles, when compared to the Vissim vehicle record data, divided by the total number of vehicles in the vehicle record sample:

$$\frac{\sum SV_c + \sum rHGV_c + \sum aHGV_c}{\sum veh_{vr}} \tag{4.1}$$

where:

SV_c = Correctly classified as a short vehicle (i.e. car/van).

$rHGV_c$ = Correctly classified as rigid HGV.

$aHGV_c$ = Correctly classified as an articulated HGV.

veh_{vr} = Any vehicle present in the Vissim vehicle record data.

Classification into two bins, referred to as '2 Class' in the results section, requires that a decision is made on how to define the two classes. Intuitively it would

make sense to group rigid and articulated HGVs into a single HGV class, hence the length thresholds set in Table 4.1. In this case, the requirement for the algorithm to correctly determine the type of HGV is removed. The evaluation criterion is then:

$$\frac{\sum SV_c + \sum HGV_c}{\sum veh_{vr}} \tag{4.2}$$

where:

HGV_c = Correctly classified as an HGV, either rigid or articulated.

4.3 Initial Consideration

As discussed previously, the predominant use of the flow-occupancy method for classifying vehicles from a single detector is in freeway applications. In a freeway setting, the vehicle sample sizes used to estimate vehicle speed can be relatively large (33 in the case of Coifman et al. (2003)) as, for the most part, the speed is not expected to vary rapidly over the detector from one vehicle to the next. Consequently, the speed estimation tends to follow longer term trends and is less sensitive to stop-start conditions.

In Chapter 2, the principle behind the flow-occupancy method of speed estimation was discussed. The use of a mean τ value for a vehicle sample to estimate speed was introduced using the following equation:

$$\overline{v} \approx \frac{n}{T \times O \times g} \tag{4.3}$$

Into which the mean value of τ and the MEVL can be substituted (see Appendix A):

$$\overline{v} \approx \frac{MEVL}{\overline{\tau}} \tag{4.4}$$

where:

$\overline{\tau}$ = Mean detector 'on-time' value (seconds).

$\overline{v}$ = Space-mean speed (ms^{-1}).

n = Number of vehicles in the specified vehicle sample.

T = Sample interval length (seconds).

O = Occupancy in the specified sample interval.

g = Reciprocal of MEVL.

$MEVL$ = Mean effective vehicle length (metres).

The issue of underestimating speed when the estimate is heavily influenced by the longest vehicle (i.e. in stop-start conditions) was explained in Chapter 2, section 2.2.4, along with the opposite problem of overestimating speed in the same circumstances when using the median τ value method of Coifman et al. (2003).

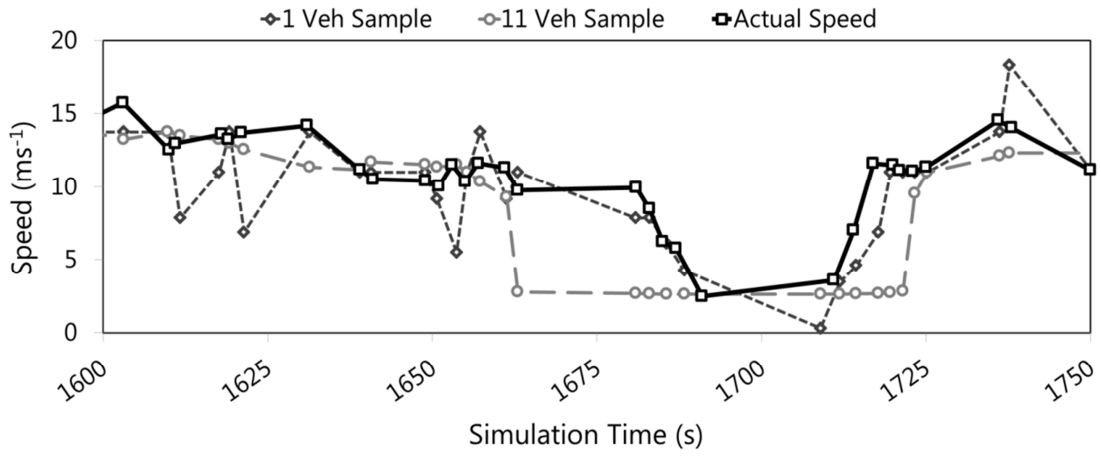

Figure 4.3: Comparison of unfiltered speed estimation with different sample sizes.

Figure 4.3 shows the results of unfiltered speed estimation from a single detector, applying Equation 4.3 with a one-vehicle sample and an eleven-vehicle sample. The detector data used in the example shown was collected at 10Hz

and is compared to the actual vehicle speeds recorded from the Vissim vehicle record data.

The plotted speeds demonstrate that the speed estimation using a larger, eleven-vehicle, sample is less sensitive to changes in vehicle length. This is expected as the sample occupancy is averaged across multiple vehicles. Consequently, the larger sample size provides consistent speed estimation that is able to follow longer term trends. However, the performance significantly deteriorates when stop-start conditions occur, such as a queue from the downstream traffic signals reaching the upstream detector (approximately 1665 simulation seconds in Figure 4.3).

In addition to the speed estimation issues, the use of any sample size larger than one vehicle necessarily introduces latency into the classification process where the vehicle being classified is not the last vehicle in the sample. This is because, as shown in Figure 4.3, every vehicle in the sample must necessarily have crossed the detector before a speed estimate can be calculated.

In practise, data from detectors on freeways is often recorded to a centralised database, or 'in-station', for use in strategy selection by traffic managers or to inform road users of traffic conditions. In that case the data is reported at, say, 1 minute intervals (NIS Ltd, n.d.). For such applications, the requirement for the 'second half' of the vehicle sample to be collected (16 further vehicles in the case of Coifman et al. (2003)) before the 'centre' vehicle can be classified may not be significant. However, for this application the usefulness of the classification data deteriorates rapidly once a vehicle has left the detector and it is therefore not viable to employ large sample sizes.

4.3.1 One-vehicle sample

The use of a one-vehicle sample eliminates the latency issues but results in the speed estimation being heavily influenced by vehicle length. In fact, for a one-vehicle sample size, n = 1 in Equation 4.3 and consequently the MEVL becomes the Effective Vehicle Length (EVL). It is therefore possible to reduce the equation further:

$$v = \frac{EVL}{\tau} \tag{4.5}$$

The speed estimation from a one-vehicle sample is less consistent at higher speeds, as shown in Figure 4.3. This is expected given the direct relationship to vehicle length. However, despite the erratic nature of the speed estimation, Figure 4.3 demonstrates that (as would be expected at low HGV proportions) the estimated speed follows the actual speed much more closely than a larger sample during periods where queuing affects vehicle speed over the detector.

This chapter investigates whether, through the use of various filtering processes, the flow-occupancy method of estimating speed (primarily used in freeway applications) can be adapted to an urban situation. If such a method can produce reasonable results then it would provide a means of classifying vehicles using existing infrastructure.

4.3.2 Potential sources of error

The detector information has been collected at 4Hz and 10Hz in order to assess whether each algorithm can still provide reasonable performance at a lower data resolution. There is an inherent measurement error resulting from the rounding of the analogue vehicle detection profile when it is processed into a digital signal that increases as the data resolution decreases. For example, an average car of 4.5 metres in length could be expected to spend approximately

0.325 seconds on a detector (with 2 metres effective length) when travelling at $20ms^{-1}$, 0.43 seconds at $15ms^{-1}$ and 0.26 seconds at $25ms^{-1}$. At 4Hz, all those timings will potentially be rounded to either 0.5 seconds or 0.25 seconds, depending on the sampling method used, potentially resulting in significant speed estimation error.

Increasing the resolution of the vehicle presence sampling reduces the influence of this error, as can be observed in Table 4.2, in which the values of τ for an identical vehicle sample, using 4Hz and 10Hz resolution processing, are compared to the actual value of τ calculated using the vehicle speed data from the Vissim vehicle record output and the known length of the detector in the Simulated Environment.

As briefly discussed above, the method of sampling (i.e. scanning for and packaging the data) for SCOOT can differ between outstation units, even from the same manufacturer. Modern outstation units are capable of scanning for data at a much higher resolution (up to 20Hz) and can therefore store data at resolutions that can be easily downscaled to 4Hz. However, the way that the data is converted to 4Hz can vary (Steel, 2012). For example, users can choose whether *all* 50ms samples within a 250ms period are required to be active (i.e. a vehicle is present) or whether *any* of the 50ms samples are required to be active to set the ¼ second reply bit to the SCOOT in-station active. The different options equate to rounding up or rounding down to the nearest 250ms and it is clear from this that there can be a significant degree of uncertainty in the information that a centralised system receives from detectors on street. In this thesis the data is assumed to be rounded up.

Table 4.2: Average error for a 3200 vehicle sample.

	4Hz	10Hz
Average τ error (s)	0.09	0.04
RMS τ error (s)	0.31	0.21

MOVA is a decentralised control strategy that operates locally at a junction. There is therefore no need to transmit data to a central in-station and, consequently, the same historical limitations to communication rate do not apply. In the case of MOVA, the detector data frequency limitations will be determined more by the roadside equipment and will depend on the manufacturer.

4.4 Algorithm development: Method 1

Coifman (2001) proposed that, for detector occupancy values below a specific threshold, an average free-flow speed should be used to improve estimation accuracy. This chapter further explores the idea that assumptions can be made based on observed data that improve the estimation of speed and, subsequently, vehicle length.

The initial, unfiltered, speed estimates are calculated using Equation 4.5 at the beginning of each pass through the algorithm based on nominal short and long vehicle lengths, in this case a vehicle length (L_{veh}) of 4.5m and 16m respectively. The value of *EVL* short and long vehicles differed in testing between 4Hz and 10Hz resolution data. For example, in Method 1 the short vehicle *EVL* was calibrated at 6.9m and 6.1m for 4Hz and 10Hz respectively. The difference in the two values can be explained by the rounding errors associated with processing the detector data into 4Hz, detailed in section 4.3.2. The rounding up of the τ value results in a consistent overestimation of τ in the 4Hz scenario that

necessitates an increase in *EVL* to compensate. This is discussed in more detail in section 4.6.2.

In both proposed methods a speed estimate is calculated using the typical short vehicle EVL and is then subjected to a series of filters. The initial short vehicle speed estimate (SVE) is adjusted based on various conditions and, depending on which criteria are met, either SVE is taken forward, or it is discarded in favour of the long vehicle speed estimate (LVE). The process is shown in more detail in Figure 4.4.

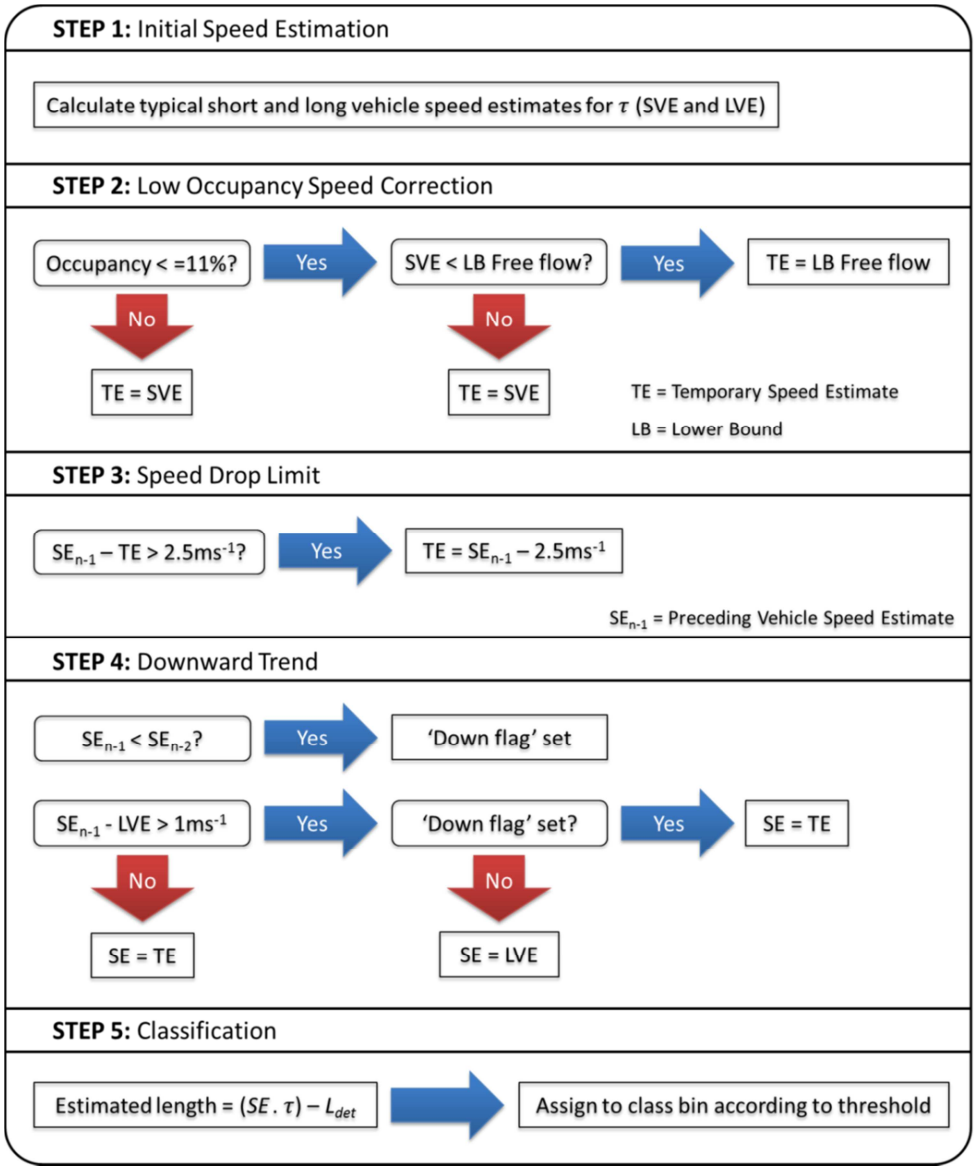

Figure 4.4: Schematic of the Method 1 Speed Estimation and Vehicle Classification Algorithm.

4.4.1 Low occupancy speed correction

The first filter presented in the algorithm is introduced to categorize the vehicle speed estimate with more certainty, based on the occupancy of a vehicle sample, into either a 'free-flow' or 'constrained' state. If the occupancy falls outside the free-flow threshold then no further deduction is made about the vehicle state at this point.

The initial speed estimate uses a one-vehicle sample. However, for the low occupancy filtering, a larger sample size of 20 vehicles is used as it provides an indication of the longer term trend of vehicle speed over the detector. It should be noted that, in this case, the sample covers the current and previous 19 vehicles rather than placing the current vehicle in the centre of the sample, as described in the Coifman et al. (2003) method. Consequently, the data can be used instantaneously. As described, the occupancy sample is simply used here to categorize the speed estimate and is not used to estimate the speed itself, as in the equations described in section 4.3. The previously described problems with larger sample size in this application are less influential as a result.

Analysis of the data output from simulation runs enabled the selection of a 'low occupancy' threshold value, in this case a 20-vehicle sample occupancy of 11% (see Figure 4.6). However, rather than assuming that *all* vehicles below this occupancy are travelling at an average free-flow speed, the filter differs from Coifman (2001) in that it only adjusts SVE_n if it is less than a free-flow lower bound speed.

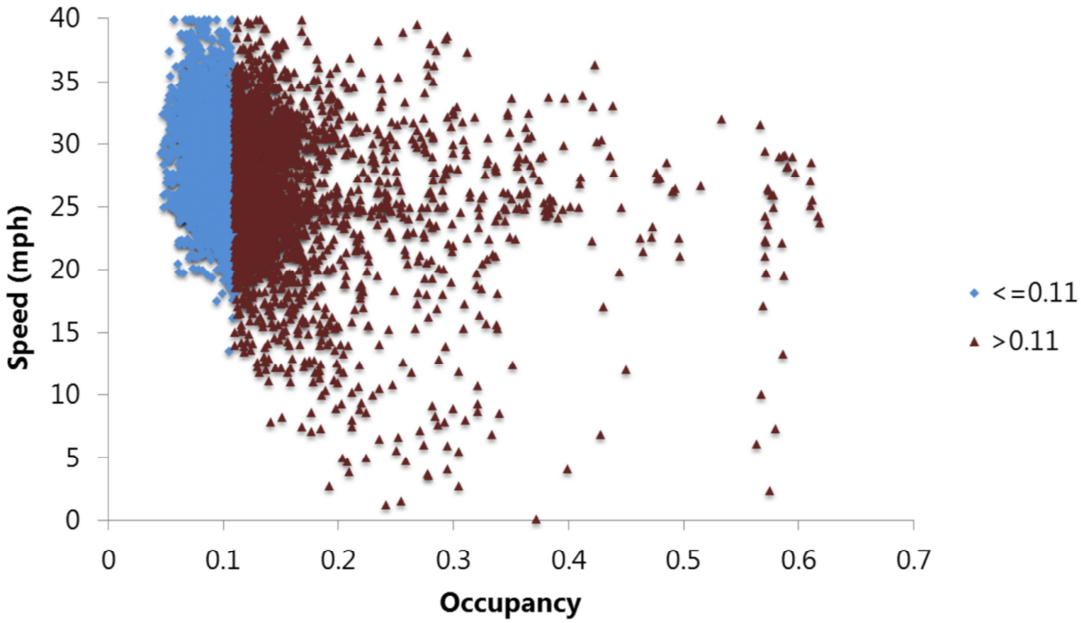

Figure 4.6: A sample of speed/occupancy relationship with free-flow threshold shown.

In this case, the free-flow lower bound speed is selected as approximately 70% of the average free-flow speed observed from the collected data. If the occupancy is less than (or equal to) 11% but the estimated speed is less than the lower bound free-flow speed then the speed is adjusted to the lower bound free-flow speed (see Figure 4.5). The reason for not adjusting all vehicles to the average free-flow speed is to preserve variation in the data.

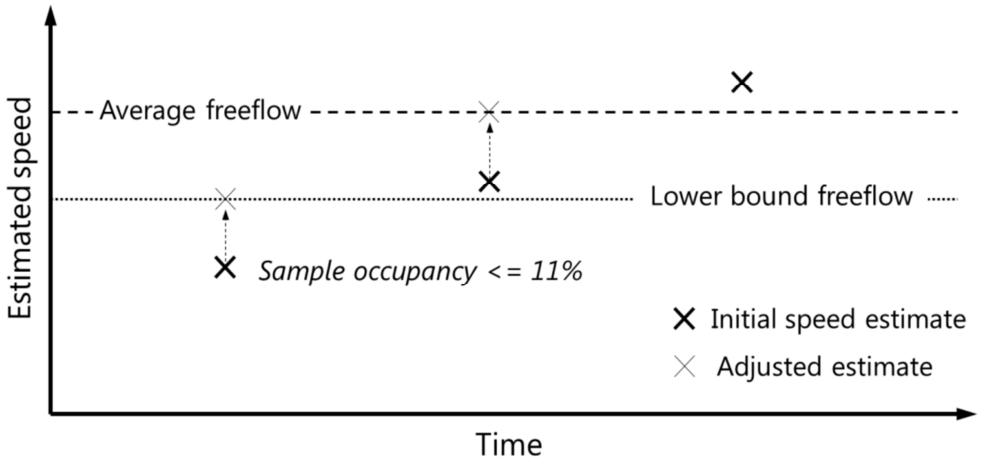

Figure 4.5: Example of adjustments made to initial speed estimate.

The optimal low occupancy threshold value for the vehicle sample used in this study was found to be 11%. This is marginally higher than the 10% used by Coifman (2001). Figure 4.7 shows the classification accuracy of the algorithm in a selection of cases with different occupancy thresholds. It can be seen that optimal threshold value for the 20% HGV proportion case is higher than the 11% value used in this study. However, it was considered that the 20% HGV case is less likely to occur than the 800 veh/h 5% HGV case and so, considering the minimal change in the base case, the value was chosen at the optimal point for the 800 veh/h case.

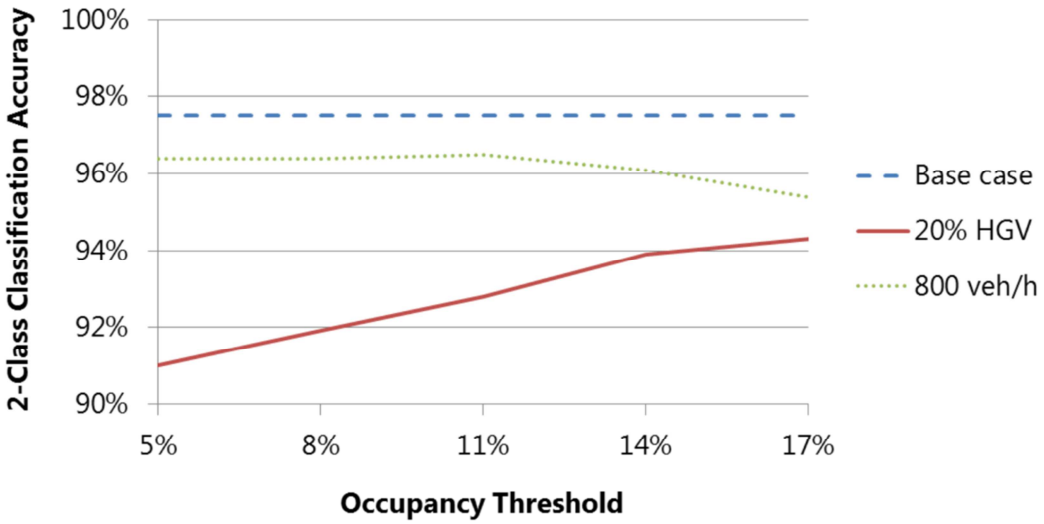

Figure 4.7: Occupancy threshold value selection.

4.4.2 Speed drop limit

It was found during the development of the algorithm that introducing a limit to how much vehicle speed could drop over the detector from one vehicle to the next provided a modest improvement to the performance of the algorithm if calibrated carefully. In this case the limit imposed is a 2.5ms^{-1} drop from the final speed estimate for the preceding vehicle. It was found that the speed rarely legitimately falls by more than this value and imposing a limit allows more variability in the estimated vehicle lengths.

4.4.3 Downward trend

The site of the detector on each approach, from which the algorithm is receiving presence data, experiences queuing from the traffic signals, even in periods of otherwise under-saturated traffic conditions due to the random nature of arrivals. It is therefore difficult to quickly and reliably determine when vehicle speeds have dropped over the detector. A longer duration of τ could represent a drop in speed but it could also be caused by a longer vehicle crossing the detector.

For this work, a 'down flag' is introduced to identify where a possible downward trend may begin. A down flag is triggered where the preceding vehicle speed estimate (SE_{n-1}) is lower than its preceding vehicle (SE_{n-2}).

If the speed drop limit is triggered and LVE, calculated at the beginning of the algorithm step, is lower than SE_{n-1} then, if the down flag is *not* present, the speed estimate can be substituted for LVE. Here, the algorithm recognises that vehicle speeds over the detector have not dropped prior to the current estimate, and that a longer vehicle crossing the detector is the more likely cause of an increase in τ than a drop in vehicle speed.

4.5 Algorithm development: Method 2

The algorithm development described in the previous section relies solely on presence data received from a single detector. The algorithm works in isolation of the traffic signal control strategy, despite the proximity of the detector to the junction. The algorithm developed in this section makes use of the traffic signal optimizer modelled queue lengths to provide additional information for speed estimation. As in the previous section, speed estimations are derived for a typical short and long vehicle and are then subject to a series of filters.

The immediate difference is that the low occupancy speed correction filter is not used in Method 2 as the general traffic conditions can instead be deduced from the traffic signal optimizer model. The process is described by the schematic shown in Figure 4.8.

4.5.1 Downward trend flag

As for Method 1, the common occurrence of queuing or decelerating/accelerating traffic over the detector requires that such conditions can be identified quickly. The introduction of the queuing model from the traffic signal optimizer enables possible downward trends to be identified more easily. Once the available storage between the last vehicle in a queue and the detector reaches less than 100 metres, the downward trend flag is set.

4.5.2 Speed drop limit

Another concept introduced in Method 1 has been modified slightly in Method 2. If the initial speed estimate falls more than $2.5ms^{-1}$ below the preceding vehicle speed estimate, the initial speed estimate can, in some cases, be substituted for the long vehicle estimate. If the speed estimate for a long vehicle falls below the preceding vehicle final speed estimate, it can be assumed that an increase in τ has been caused by a longer vehicle crossing the detector. If the long vehicle speed estimate is greater than the preceding vehicle final speed estimate then the initial speed estimate is kept but the drop in speed limited to $2.5ms^{-1}$.

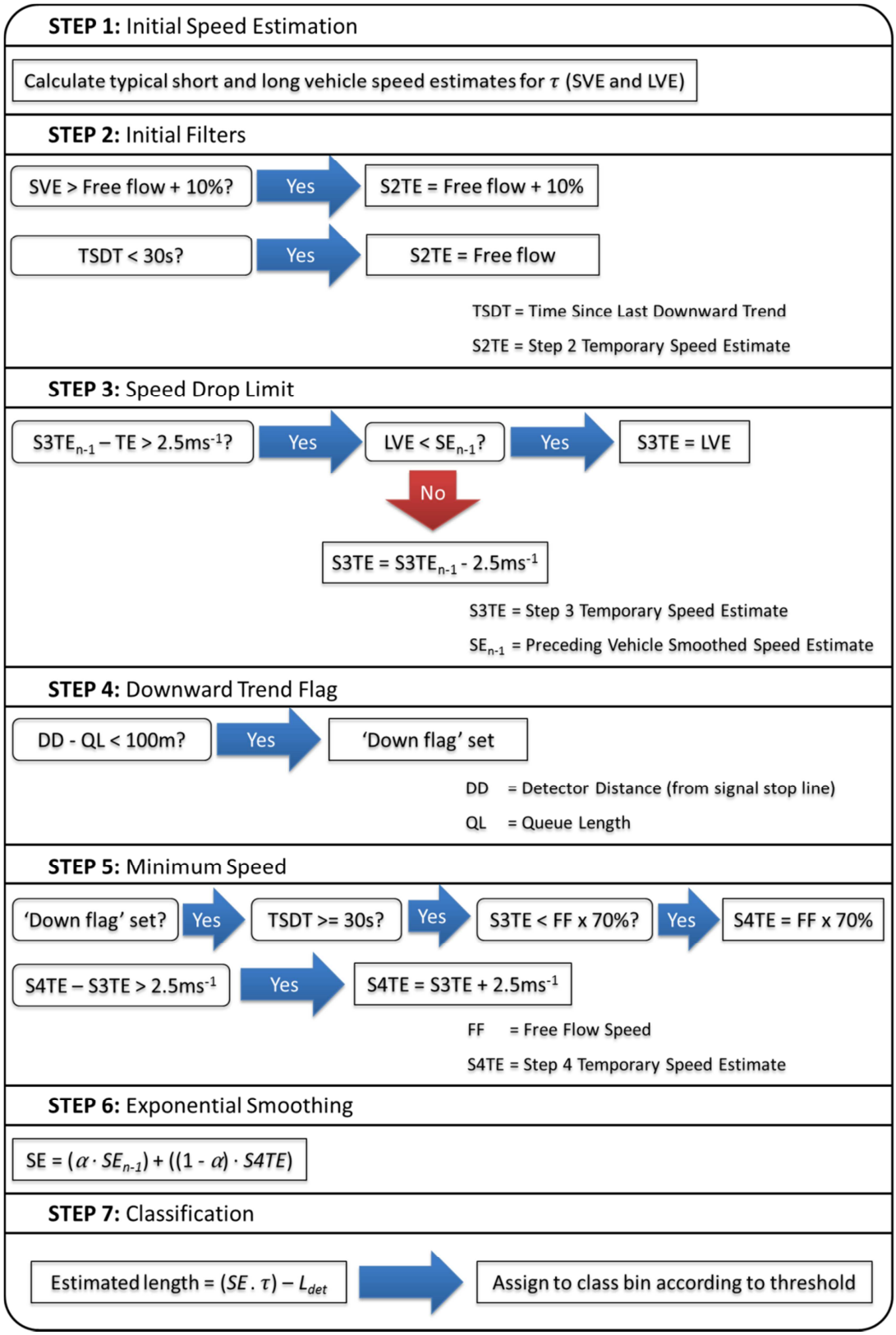

Figure 4.8: Schematic of the Method 2 Speed Estimation and Vehicle Classification Algorithm.

4.5.3 Minimum speed

A new concept, introduced for Method 2. This acts as a replacement for the low occupancy speed correction described in Method 1. If periods of free-flow over the detector can be reliably identified, then a lower bound minimum speed limit can be applied to reduce instances of longer vehicles being incorrectly identified as a slow moving short vehicle. The downward trend flag, previously described, is used to identify when vehicle speeds over the detector become constrained. Consequently, if the flag is not set, the conditions can be assumed to be free-flow.

To ensure that vehicle speeds are unconstrained, an additional condition is included that takes into account the time since the downward trend flag was last cleared. In this case, the flag must have been cleared for more than 30 seconds before this filter can be implemented.

If the conditions for free-flow are met, the minimum speed is implemented as 70% of the free-flow speed. This value is a compromise since it has been found that, with higher demand, a smaller value of 20% provides better results due to the frequent fluctuation of speed over the detector.

4.5.4 Exponential smoothing

Smoothing of the filtered speed estimation is introduced in Method 2. Exponential smoothing is utilised, as shown in Equation 4.6. However, in addition to the basic smoothing, the value of α is modified to provide a variable weighting that is influenced by the time that has elapsed since the last vehicle was detected.

$$SE_n = \alpha \cdot FSE + (1-\alpha) \cdot SE_{n-1} \tag{4.6}$$

where:

FSE = Filtered speed estimate for current vehicle.

SE_n = Smoothed speed estimate for current vehicle.

SE_{n-1} = Smoothed estimate for preceding vehicle.

When headway between vehicles is small, it is likely that changes in speed will also be relatively small. Figure 4.9 demonstrates that observation and also shows that the constraint on speed quickly reduces as the headway between vehicles increases. The basic exponential smoothing function does not take into account the rapid decay in usefulness of the preceding vehicle estimate in estimating the current vehicle speed.

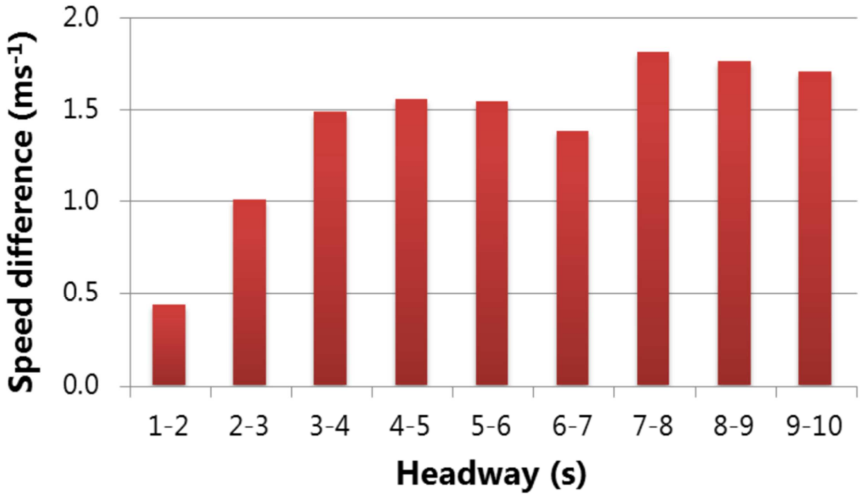

Figure 4.9: Average difference in speed between consecutive vehicles over a detector at different headways for a 3200 vehicle sample.

To address the shortcoming described above, a logarithmic function has been incorporated, shown in Equation 4.7, to change the value of α according to the headway between vehicles. The logarithmic function allows the rapid decay in usefulness of the previous vehicle speed estimate to be reflected in the smoothing.

$$\alpha = (0.075 \cdot \ln(H - 1)) + 0.45 \tag{4.7}$$

where:

H = Headway from preceding vehicle (s).

A value of 1 second was chosen as the minimum headway as it was deemed unlikely that headways less than that would occur. However, this figure could be adjusted if necessary.

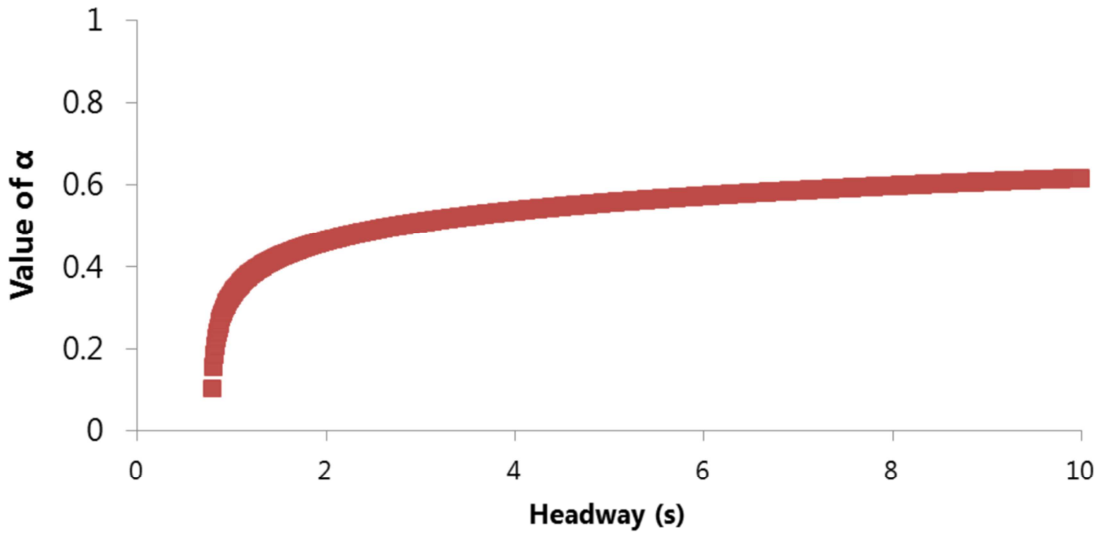

Figure 4.10: A plot of the α value for increasing headway.

Figure 4.10 plots the value of α for increasing headway using Equation 4.7. It was expected that the value of α would tend to 1, even for relatively low headway values, given that Figure 4.9 demonstrates that speed difference is only significantly constrained by headways below 3 seconds. However, it was found that retaining some of the previous speed estimate proved advantageous, particularly at higher proportions of HGV due to the increased variability in vehicle length. This is possibly a reflection of the fact that the speed difference between consecutive vehicles rarely exceeds 5mph. The final logarithmic function used is a compromise between the optimum parameters for high HGV volumes (low value of α) and high demand.

Further testing has shown that the smoothing process could be simplified somewhat by applying a constant smoothing factor either side of a headway threshold value of approximately 2 seconds. The smoothing factor for headways below 2 seconds would largely be weighted towards the previous speed estimate as the preceding vehicle would constrain the speed. Headways above 2 seconds would be given an approximately 50/50 weighting. Initial testing has shown that this method could improve classification accuracy at higher HGV proportions (due to the increased stability of the speed estimation) but at the expense of some speed and length estimation accuracy. Consequently, the development of alternative smoothing methods may depend on the application of the algorithm.

4.6 Results

4.6.1 Introduction

This section details the algorithm performance according to the evaluation criteria previously described in section 4.2.2. The algorithm results are shown for each method using the following scenarios:

- Different proportions of HGV with fixed demand of 600 veh/h (Scenario 1); and
- Different demands with a fixed proportion of 5% HGV (Scenario 2).

The results have been presented in various different formats to provide a comparison with similarly presented results in other literature. Figure 4.11, Figure 4.12 and Figure 4.13 show a sample trace of the speed of successive vehicles over a detector to illustrate the effectiveness of the algorithm (in this case Method 2) in following the variations in speed. They also provide

an indication of the volume of wrongly classified vehicles in different scenarios.

It can be seen even from a relatively small sample that errors occur during periods of rapid change in speed between successive vehicles. This observation is not unexpected since they are the periods during which it is most difficult to estimate speed accurately.

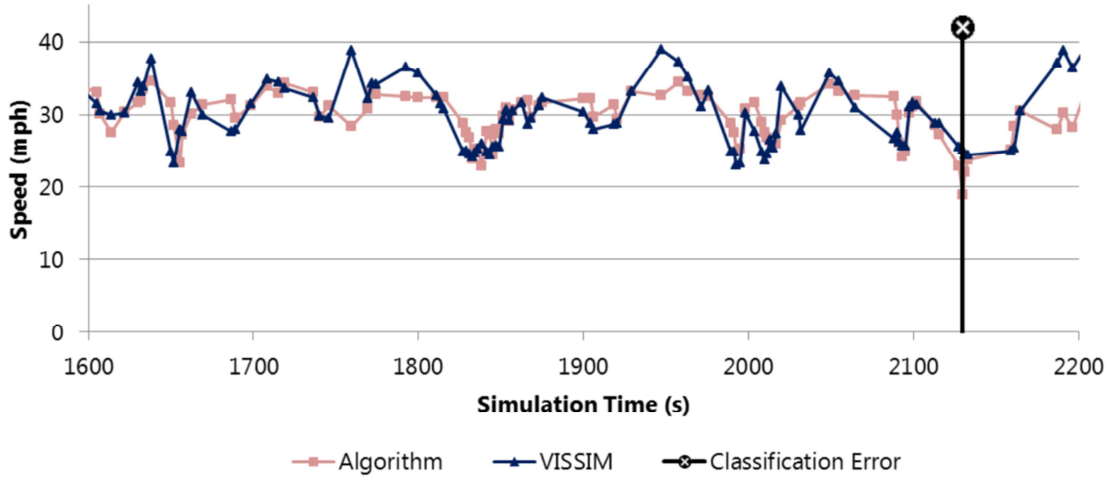

Figure 4.11: Sample of vehicle speeds over a detector (base case).

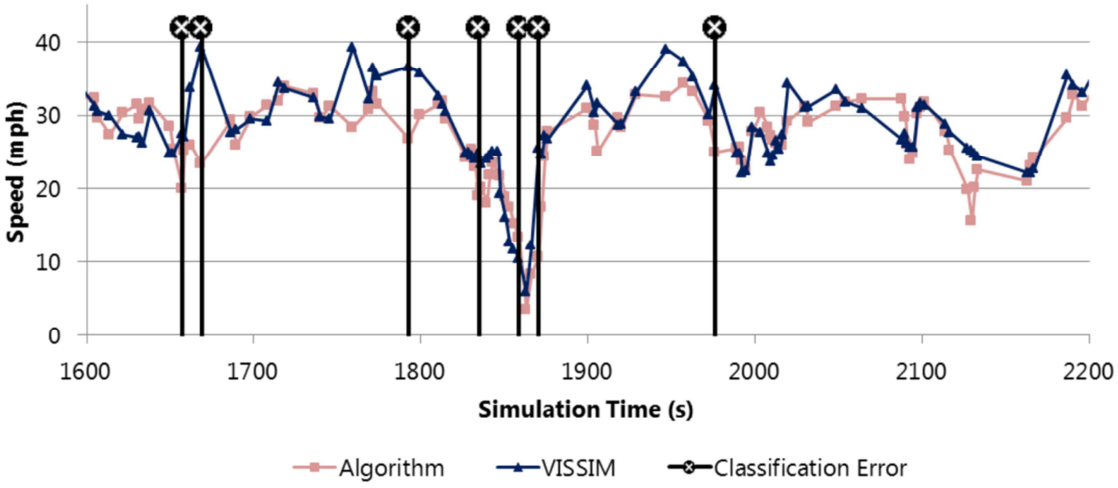

Figure 4.12: Sample of vehicle speeds over a detector (20% HGV).

At very low speeds the algorithm speed is shown to be close to zero. This is a result of a vehicle moving very slowly over the detector, or even being stationary for a period of time before the vehicle speed is estimated. As a consequence, the estimated speed is lower than the speed the vehicle is travelling by the time it leaves the detector (the point at which the Vissim speed is measured).

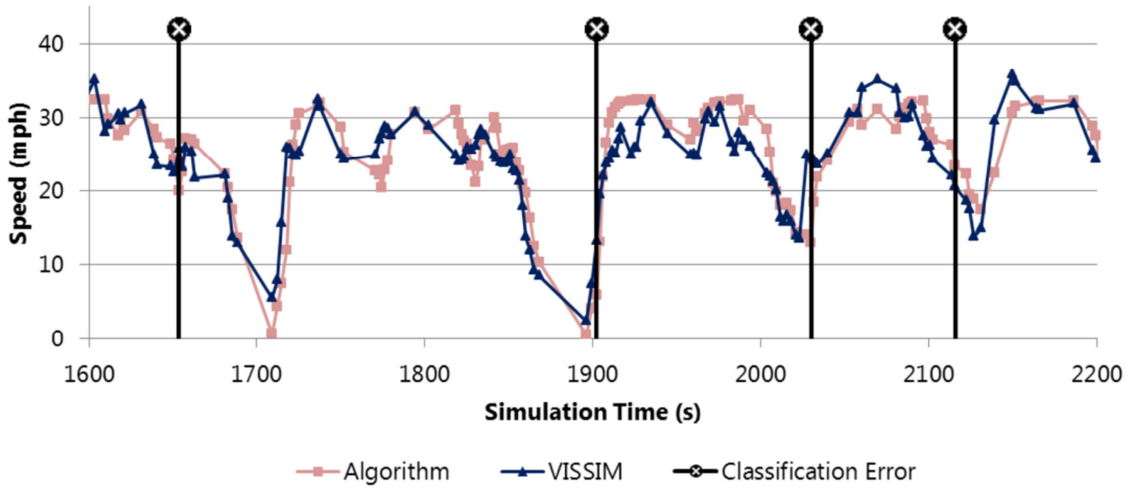

Figure 4.13: Sample of vehicle speeds over a detector (800 veh/h).

4.6.2 Method 1 results

Table 4.3 and Table 4.4 demonstrate the accuracy of the Method 1 algorithm in estimating vehicle speed and length for Scenario 1 and Scenario 2. The performance of the algorithm in estimating speed and length deteriorates as the percentage of HGVs increases, as expected. There is also a worsening in performance as demand is increased up to 800 veh/h. This is due to the increased variability of vehicle speed over the detector caused by the queue from the downstream signals reaching the detector more often.

Table 4.3: Method 1 speed and length estimation RMSE results for Scenario 1.

	5% HGV (Base case)		**10% HGV**		**15% HGV**		**20% HGV**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
Speed (mph)	4.0	3.5	4.3	3.7	4.5	4.0	4.8	4.2
Length (m)	0.9	0.9	1.1	1.0	1.4	1.2	1.6	1.4

Table 4.4: Method 1 speed and length estimation RMSE results for Scenario 2.

	600 veh/h (Base case)		**700 veh/h**		**800 veh/h**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
Speed (mph)	4.0	3.5	4.1	3.5	4.4	3.6
Length (m)	0.9	0.9	1.0	0.9	1.4	1.2

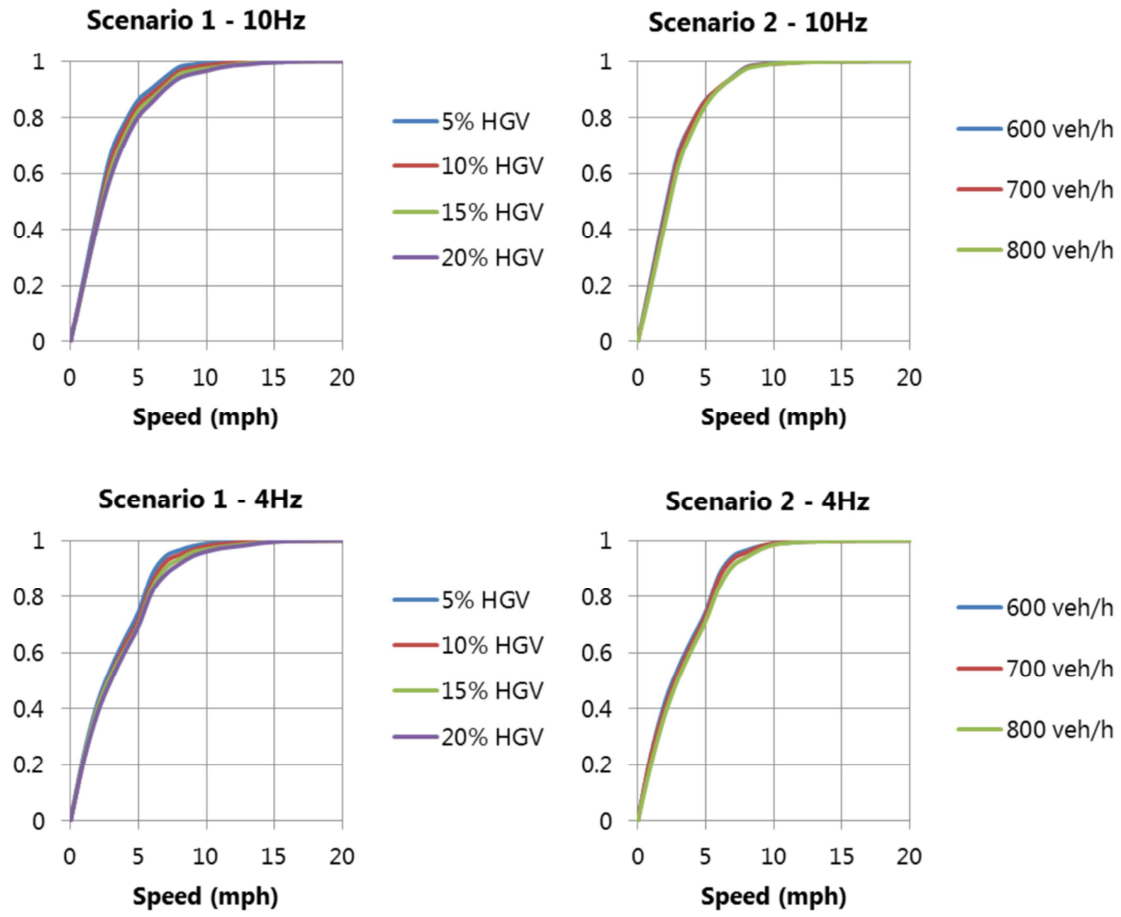

Figure 4.14: Cumulative Distribution Functions for speed estimation AAE - Method 1.

Figure 4.14 shows Cumulative Distribution Functions (CDFs) for the Average Absolute Error (AAE) in the speed estimation of Method 1 at 10Hz and 4Hz. It can be seen that the majority of estimation error is less than 5mph with a greater number of larger errors as the proportion of HGVs increases. At 4Hz there is a deviation in the CDF at approximately 5mph that indicates a cluster of errors distributed around that value.

As discussed previously, the τ value is rounded up when the detector presence data is converted to 4Hz and this produces lower speed estimation (on average) at free-flow speeds than the equivalent 10Hz data. For example, τ values of 0.3 and 0.4 seconds are most common when vehicle speeds are approximately 30-40mph. However, at 4Hz these τ values will both be rounded up to 0.5 seconds when the presence data is processed resulting in consistently lower speed estimation for τ values that are not equal to those at 10Hz.

Table 4.5: Demonstration of the effect of EVL and data frequency on error.

τ values (s)		**Speed estimate (mph)** *(5.8 metres EVL)*			**Speed estimate (mph)** *(6.9 metres EVL)*	
10Hz	4Hz	10Hz	4Hz	Difference	4Hz	Difference
0.20	0.25	65	52	13	62	3
0.30	0.50	43	26	17	31	12
0.40	0.50	32	26	6	31	2
0.50	0.50	26	26	0	31	-5
0.60	0.75	22	17	4	21	1
0.70	0.75	18	17	1	21	-2
0.80	1.00	16	13	3	15	1

To compensate, the EVL used in the algorithm is increased to provide the overall least error but, as a consequence, there is an increase in estimation error wherever the τ values are similar. Table 4.5 shows that increasing the EVL for 4Hz data reduces the overall speed estimation error but introduces an error of

approximately 5mph when the τ values are both 0.5 seconds (shown in the highlighted row). This is consistent with the Figure 4.14.

The capability of the algorithm to classify vehicles within various criteria is shown for Scenario 1 and Scenario 2 in Table 4.6 and

Table 4.7 respectively.

Table 4.6: Method 1 classification results for Scenario 1 (%).

	5% HGV (Base case)		**10% HGV**		**15% HGV**		**20% HGV**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
2 Class	97.3	97.5	96.1	96.4	94.2	94.5	92.3	92.8
3 Class	96.9	97.3	95.3	95.7	93.0	93.3	90.8	91.0

Table 4.7: Method 1 classification results for Scenario 2 (%).

	600 veh/h (Base case)		**700 veh/h**		**800 veh/h**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
2 Class	97.3	97.5	97.2	97.3	95.8	96.5
3 Class	96.9	97.3	96.8	97.1	95.4	96.2

The results show that, in the case of the 2 class and 3 class bins, the performance of the algorithm deteriorates quite significantly as the proportion of HGVs increases. Interestingly, the algorithm results slightly improve when using 4Hz data compared to 10Hz as the HGV proportion increases. This is not the case for Scenario 2.

4.6.3 Method 2 results

The results in Table 4.8 and Table 4.9 show the speed and length RMSE for Scenario 1 and Scenario 2. Only a small deterioration in speed estimation is experienced as the proportion of HGVs increases. The Method 2 algorithm is shown to outperform the Method 1 algorithm in both scenarios for speed

estimation and classification. As for Method 1, the 10Hz resolution data provides greater accuracy than 4Hz, as expected.

Table 4.8: Method 2 speed and length estimation RMSE results for Scenario 1.

	5% HGV (Base case)		**10% HGV**		**15% HGV**		**20% HGV**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
Speed (mph)	3.7	3.1	3.9	3.3	4.0	3.6	4.2	3.8
Length (m)	0.8	0.8	1.0	0.9	1.2	1.1	1.3	1.2

Table 4.9: Method 2 speed and length estimation RMSE results for Scenario 2.

	600 veh/h (Base case)		**700 veh/h**		**800 veh/h**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
Speed (mph)	3.7	3.1	3.8	3.1	4.1	3.4
Length (m)	0.8	0.8	0.9	0.9	1.2	1.0

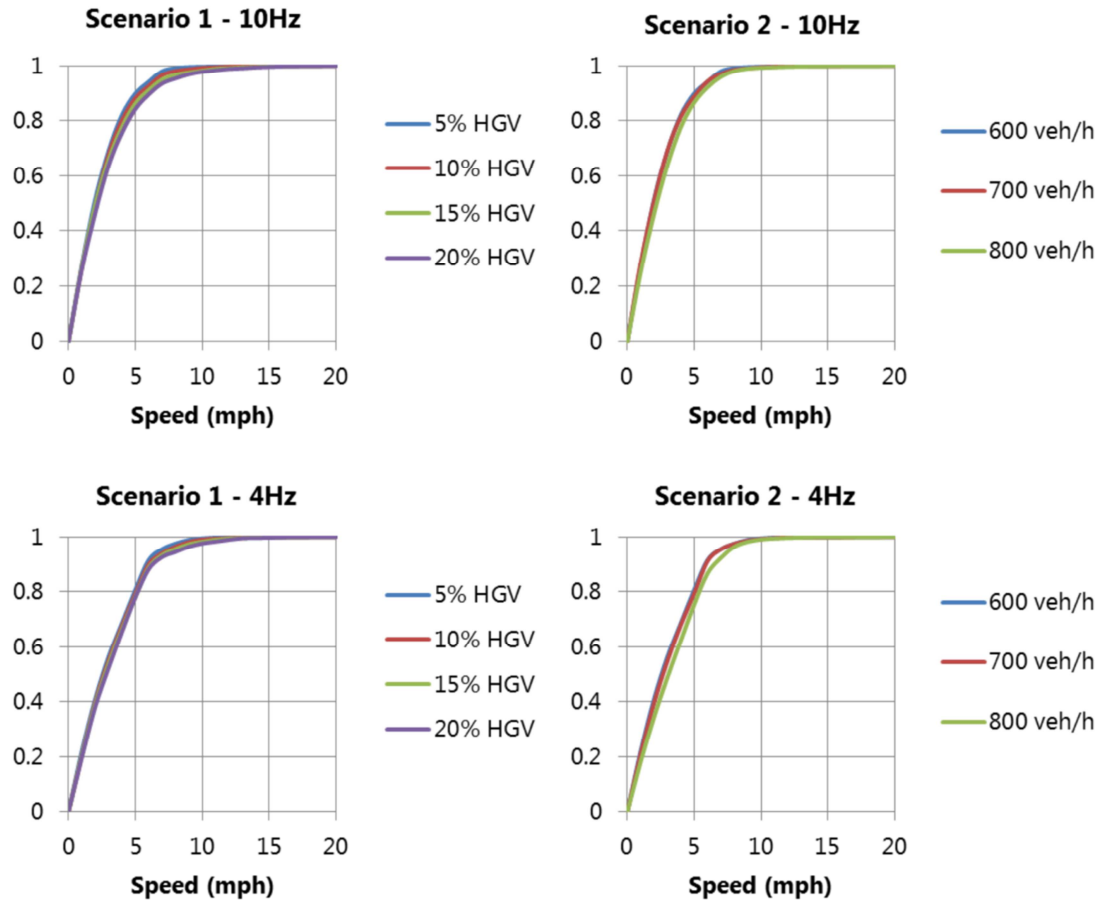

Figure 4.15: Cumulative Distribution Functions for speed estimation AAE - Method 2.

Figure 4.15 shows CDFs for the AAE in the speed estimation of Method 2 at 10Hz and 4Hz. The CDFs at 10Hz for both scenarios follow a similar trend to those shown in Figure 4.14 for Method 1. However, it can be seen that there are fewer estimation errors greater than 5mph. For the 4Hz data the CDF is comparatively smoother than that produced by Method 1. The exponential smoothing incorporated into Method 2 distributes the speed estimation error more widely and can be observed in the 'straightening' of the CDF from approximately 3-7mph. The result is approximately 10% fewer speed estimation errors greater than 5mph in both scenarios.

Wang and Nihan (2003) presented a standard error of 3.47mph for a 24 hour sample of vehicles that contained between 5-10% long vehicles during what would assumed to be the morning and evening peak periods, rising to approximately 15% during the inter-peak period. The RMSE, of 4.0mph and 3.6mph respectively, calculated for Method 1 and Method 2 at 15% HGVs (10Hz) compare favourably with that presented by Wang and Nihan (2003). This is particularly useful given that the freeway application would be expected to have less significant speed differences than an urban application immediately upstream of a traffic signal controlled junction. Comparing between standard error and RMSE values is considered reasonable in this case given the large sample size.

The CDF plots have been included to provide a comparison with those presented by Coifman and Kim (2009). The CDFs demonstrate that the speed estimation of both methods described in this chapter compares favourably. However, it is more difficult to compare the distribution of larger errors as the various methods proposed (Coifman (2001), Coifman et al. (2003), Coifman and Kim (2009)) focus on a freeway application where variation in speed at free-flow is likely to be greater than in constrained urban conditions.

The performance of the Method 2 algorithm in classification is shown in Table 4.10 and

Table 4.11. The Method 2 algorithm consistently outperforms Method 1, particularly at 4Hz. The 4Hz data resolution provides comparable results in both scenarios although the performance degrades more significantly at 800veh/h. This finding is significant as it enables the algorithm to be applied to legacy systems that use the 4Hz communication protocol with no hardware changes and very little software development.

Table 4.10: Method 2 classification results for Scenario 1 (%).

	5% HGV (Base case)		**10% HGV**		**15% HGV**		**20% HGV**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
2 Class	97.5	97.8	96.5	96.7	94.9	94.8	93.1	93.3
3 Class	97.1	97.5	95.5	96.1	93.0	93.6	90.5	91.5

Table 4.11: Method 2 classification results for Scenario 2 (%).

	600 veh/h (Base case)		**700 veh/h**		**800 veh/h**	
	4Hz	10Hz	4Hz	10Hz	4Hz	10Hz
2 Class	97.5	97.8	97.3	97.3	96.3	97.0
3 Class	97.1	97.5	96.9	97.1	95.8	96.6

For both algorithm methods the classification into three length bins is less accurate than into two bins. This is perhaps to be expected given that vehicle length distributions, as in this case, are often bi-modal (Coifman and Kim, 2009). A bi-modal distribution lends itself to classifying into two length bins as it provides a clear distinction between long and short vehicles. However, when a third length bin is introduced that spans between long and short vehicles there is a less clear distinction and a greater probability of incorrectly classifying into the neighbouring length bin.

The results demonstrate that the performance of both algorithms when classifying into three length bins is comparable with two length bins at low HGV proportions but deteriorates significantly as the proportion of HGVs increases.

4.6.4 Cumulative performance improvement

Each of the algorithms is described in this chapter as a series of steps for the purposes of clarity. Each of the steps provides an enhancement in the accuracy of speed and length estimation as well as classification that contributes to the overall performance of the algorithm. Figure 4.16 demonstrates the improvement in classification performance provided by each step of Method 2 compared to a baseline of simply classifying every vehicle as a short vehicle. The cumulative performance is shown as three distinct phases for a selection of cases where:

- Phase 1 is the initial filtering of speed estimates (step 1);
- Phase 2 is the application of intermediate filters (steps 3, 4 and 5); and
- Phase 3 is the final smoothing process (step 6).

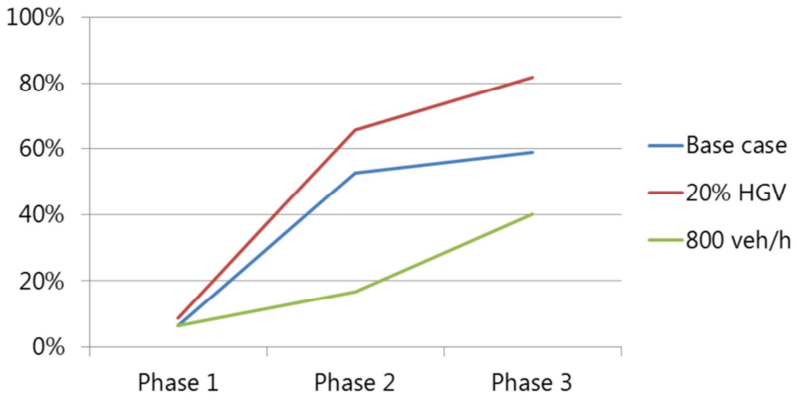

Figure 4.16: Cumulative classification performance improvement - Method 2.

Figure 4.16 shows that the initial phase is a small improvement but that the second phase provides a significant benefit for the cases with lower demand. The third phase provides a less substantial improvement for the lower demand cases but a greater improvement for the 800 veh/h case. This highlights the importance of incorporating all of the phases to provide resilience of performance in different scenarios.

4.7 Conclusions

In this chapter, two algorithms have been developed to estimate the speed of vehicles crossing a detector and to classify them by length. The performance of both algorithms is similar in both scenarios but the second algorithm, incorporating exponential smoothing and knowledge of the queue length from the downstream traffic signals, consistently outperforms the first. Consequently, Method 2 has been incorporated into the previously described Signal Controller for testing in the following chapter.

The first algorithm, Method 1, is based on previous work by Dodsworth et al. (2014) with some modifications made to reduce the number of parameters that would require calibration. The second algorithm is, to some extent, based on the first but makes use of the modelled queue from the downstream traffic signals and exponential smoothing. The algorithms have been tested with speed profiles derived from ATC data retrieved from a site in a 40mph speed limit section of road, as described in the previous chapter. The algorithms have not been tested with different speed profiles but in the previous work by Dodsworth et al. (2014) it was shown that the algorithm performed marginally better with a 30mph speed profile.

It is recognised that it is unlikely either algorithm could be used as a replacement for a dual loop detector site. However, at traffic signal junctions

where single detectors are employed, the preferred algorithm could prove useful to traffic authorities for basic monitoring of HGV proportions on the urban road network. The 'instant' nature of the speed estimation, as opposed to the five minute aggregation employed by standalone ATC sites, would also provide a quick response to drops in vehicle speed that could be used for incident detection purposes.

Neither algorithm experiences a significant deterioration in performance when receiving detector presence data at 4Hz until the demand is increased to 800veh/h. However, despite the deterioration in performance at higher demand, the preferred algorithm could still prove useful to deploy on existing systems that use a 4Hz data protocol, a common feature of many Urban Traffic Control systems in the UK.

The primary proposed application of the developed algorithm, however, is to provide classification information to a traffic signal optimizer. The modified Miller method of signal optimization, described in Chapter 3, requires that any queue that exists at the start of a green signal is discharged before optimization begins. That condition dictates that, for the majority of the time, vehicles will be moving freely over the detector during the period that classification is actually utilised in the Miller process. This provides a benefit in that, as Figure 4.17 demonstrates, the speed and length estimation error above 20mph (assumed to be relatively free-flowing) is consistently smaller than for below 20mph. The exception, speed estimation at 600veh/h, is likely to be due to the very small number of vehicles crossing the detector at less than 20mph in the base case.

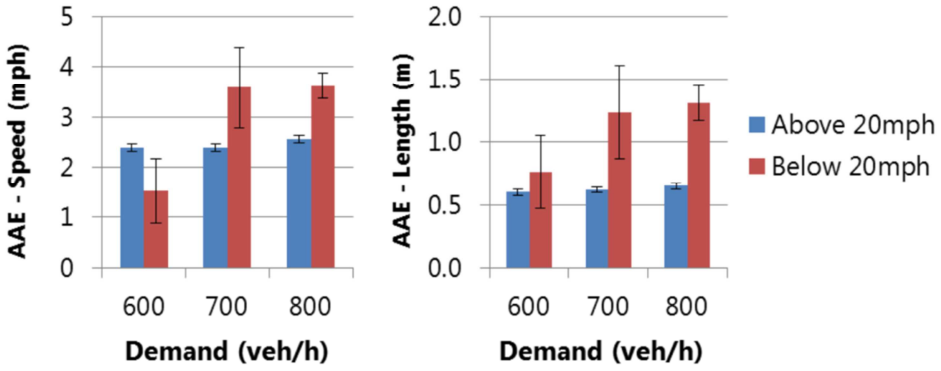

Figure 4.17: AAE for speed and length estimation - Method 2.

It is not straightforward to quantify the level of classification accuracy that is required for the proposed application of the algorithm as there are multiple variables that will dictate how much of an effect on optimizer performance the classification accuracy will have. For example, even if a classification algorithm could classify vehicles with a 100% success rate, the trajectory of a vehicle beyond the upstream detector is estimated and so there is potential error in the prediction of when a vehicle crosses the stop line.

The traffic model used in the Miller method is a simple vertical queue model but more sophisticated models such as a car-following model could introduce further error as the interaction between vehicles is explicitly modelled and so the classification accuracy is important at all speeds.

The following chapters will investigate the application of the preferred algorithm to two optimization techniques and report and provide conclusions on the results.

More investigation would be needed to prove algorithm performance on links with different characteristics and at different detector distances from a traffic signal stop line.

Chapter 5
Vehicle classification, V2I and a hybrid model

5.1 Introduction

The Miller method was first introduced in Chapter 3 and two subsequent modifications were made in order to provide a representation of the MOVA control algorithm that could be used as a benchmark for comparisons in this chapter and Chapter 6.

In this chapter, the potential benefit of differentiating vehicle classes within the traffic signal optimizer is investigated. The stop penalty modification to the Miller method, described in Chapter 3, enables the stops to be incorporated into the objective function. In the representation of MOVA the same stop penalty value is applied to each vehicle regardless of type. In reality, however, the value of delay caused by a stop can vary significantly between vehicles, particularly by vehicle type, and so a single stop penalty value is likely to result in sub-optimal decisions by the optimizer.

Consequently, this chapter investigates the application of vehicle classification to enable different stop penalty values to be applied according to vehicle type with the aim of providing a more accurate representation of the impact of stops in the optimizer traffic model. Figure 5.1 provides a summary of each element in the chapter.

Initially, vehicle classification is applied to the MOVA representation using the preferred single detector classification algorithm, described in the previous chapter. Next, V2I detection is introduced and is compared, without classification, to the MOVA representation. This enables the benefit of V2I detection alone to be quantified. A hybrid model is developed to enable

evaluation of the effect of increasing V2I penetration rates on the realisation of benefit.

Ultimately, the single detector classification algorithm is replaced with V2I classification data V2I detection and the results compared to the benefit of providing V2I alone. Finally, the Simulated Environment is modified to include opposed right turning movements on each approach. The effect on the performance of the developed algorithm at various turning proportions is evaluated.

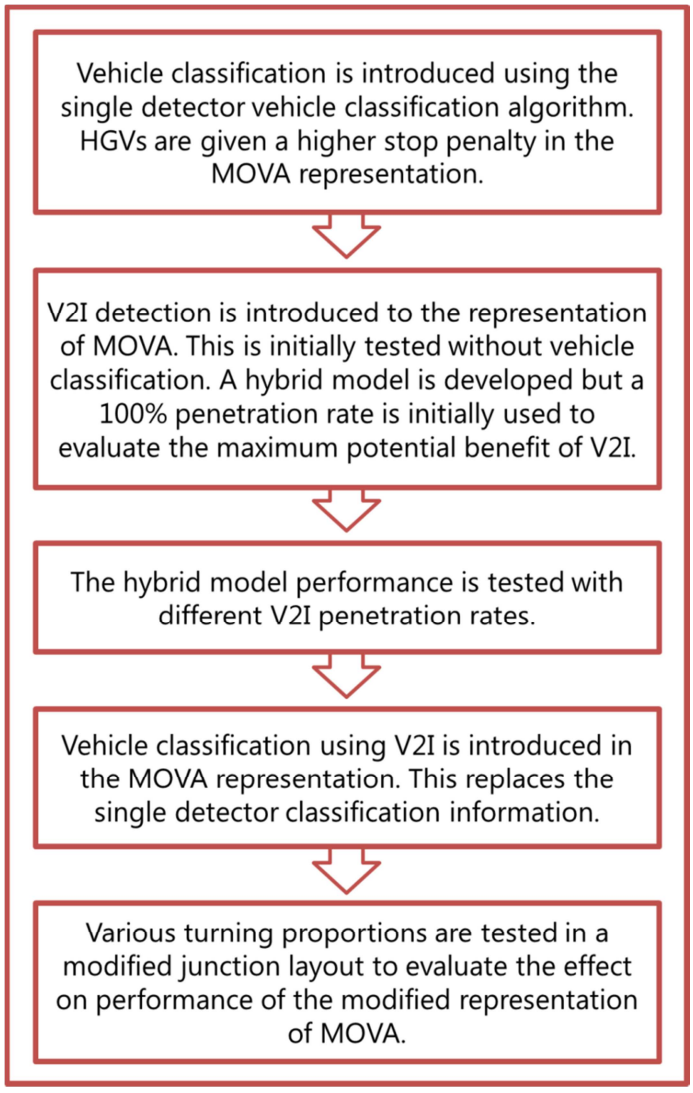

Figure 5.1: A summary of chapter 5.

5.2 Methodology

5.2.1 Incorporating vehicle classification

Vehicle classification is first introduced into the MOVA representation using data provided by the preferred single detector classification algorithm described in Chapter 4. The value of delay associated with a vehicle stop will vary significantly depending on the performance of the vehicle that is being considered. In general, longer vehicles will have an inferior acceleration performance to shorter vehicles. Clearly there are exceptions to this, not least HGVs that are 'empty runners' but the single detector length based classification algorithm offers a proxy for vehicle performance. Vehicles with slower acceleration will experience more delay through the stop/start process and, if the vehicle is at the front of a queue, the additional delay will also affect the following vehicles.

Vehicle classification enables the stop penalty value in the MOVA representation to be varied, per vehicle, according to vehicle type (based on length). Equation 3.7 (introduced in Chapter 3) has been modified as shown in Equation 5.1 to include a variable stop penalty value based on vehicle type.

$$T_{NS}^{spv} = D_{NS} + \sum_{\alpha \in V} (SP^{\alpha} (\delta_N^{\alpha} + \delta_S^{\alpha})) \qquad (5.1)$$

where:

SP^{α} = Stop penalty value (seconds) for vehicle type α.

$\delta_N^{\alpha}, \delta_S^{\alpha}$ = Number of vehicles of type α crossing the *N, S* approach stop-lines respectively during green extension.

V = Set of vehicle types.

The vehicle classification takes place instantaneously upon the vehicle leaving the detector, at which point a representation of the vehicle is inserted into the internal traffic model. The effect of applying vehicle classification is tested and compared to the MOVA representation with no classification.

5.2.2 V2I and a hybrid model

The more precise vehicle position and speed data provided by V2I to the internal controller model improves the accuracy of the decisions made by the MOVA representation. The algorithm can more accurately determine how many vehicles will cross the stop line for a given green extension as well as the length of queues on approaches currently at red. However, a secondary benefit is that the 3 second amber period (in the UK) that follows the green can be more efficiently used. That is, the amber signal in the conventional scenario is displayed later than is necessary in most cases in order to cater for slower vehicles. With V2I the amber signal can be displayed to drivers at the relevant time according to the individual vehicle speed thus allowing the signal to be changed earlier. It was assumed for this paper that, in the algorithm, 2 seconds of the amber could be used as effective green, leaving a 1 second 'buffer' for minimizing any driver uncertainty.

Figure 5.2 shows the advantage of V2I technology in determining the point at which the Miller algorithm begins. The vehicle at the back of the queue can be clearly identified and subsequently used to determine the point at which the queue has finished discharging. Various criteria were tested for determining the point at which the Miller algorithm should start, the most successful of which was simply that the last queuing vehicle must exceed a pre-specified speed and be within a pre-specified acceleration range.

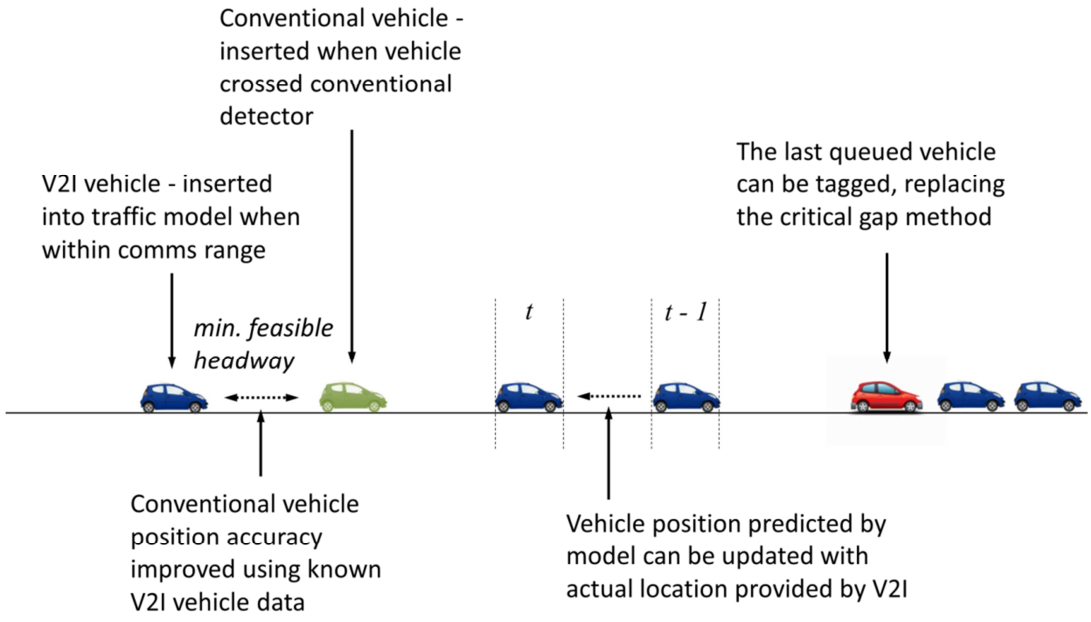

Figure 5.2: Optimizer traffic model modifications for V2I.

The internal controller model has been developed as a hybrid model capable of handling conventional detector and V2I data. Conventional detector data is used to insert a conventional vehicle into the model at a constant speed. If the position of a V2I vehicle matches the conventional vehicle closely (i.e. within 20 metres) it is redefined in the model as a V2I vehicle. The position of conventional vehicles in the model can be manipulated if required based on the known position, speed and acceleration of neighbouring, V2I enabled, vehicles whilst ensuring a safe distance is maintained.

In the MOVA representation the internal traffic model employs a shift-register method of modelling vehicle trajectory along each approach link to a junction. The shift-register is essentially a form of vertical queue model where vehicles travel along the link at a constant speed to the traffic signal stop line, at which point they are added to a 'vertical' queue. This is described in more detail in Chapter 3 section 3.2.1.

When V2I is incorporated into the model it becomes feasible to implement a spatial element to the queue model. The V2I data provides the position of an enabled vehicle along the link and the speed that the vehicle is travelling (either by directly providing the speed, or by allowing the speed to be derived from successive positional records). The accurate positional data allows the length of a queue, and subsequently the interaction of approaching vehicles with the back of the queue, to be modelled.

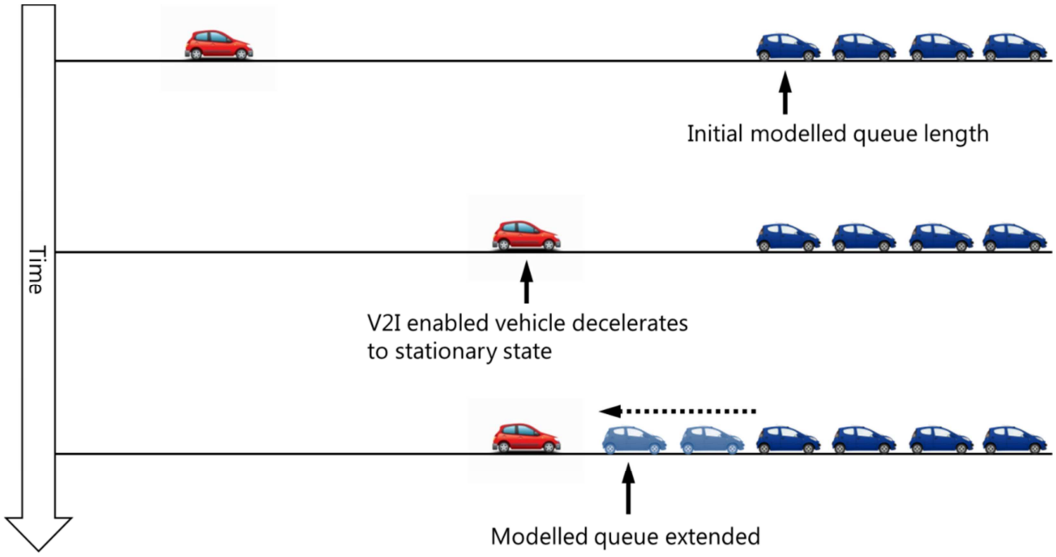

Figure 5.3: Hybrid model adjustment of queue length example.

The model has been constructed so as to enable hybrid operation depending on the number of V2I enabled vehicles on the link and where those vehicles are relative to the end of the modelled queue. The hybrid model incorporates a number of rules that determine how conventional vehicles traverse the link. The rule-based approach enables V2I data to be used to increase the accuracy of the model around those vehicles that are enabled and is, in some ways, a form of car-following model. However, whilst a following vehicle must maintain a minimum safe headway to the leading vehicle, as is the case in more sophisticated car-following models, the hybrid model does not attempt to represent the full complexity of interaction between vehicles as it is designed to

be able to switch between a vertical and spatial queue model depending on the number of V2I enabled vehicles present on the link. The basic premise is that, where a vehicle is V2I enabled and the V2I data is up-to-date, the vehicle acts as a 'ground truth' point that can be used to increase the accuracy of the estimated positions of conventionally detected vehicles around it. Figure 5.3, Figure 5.4 and Figure 5.5 provide examples of how the model is manipulated based on V2I data.

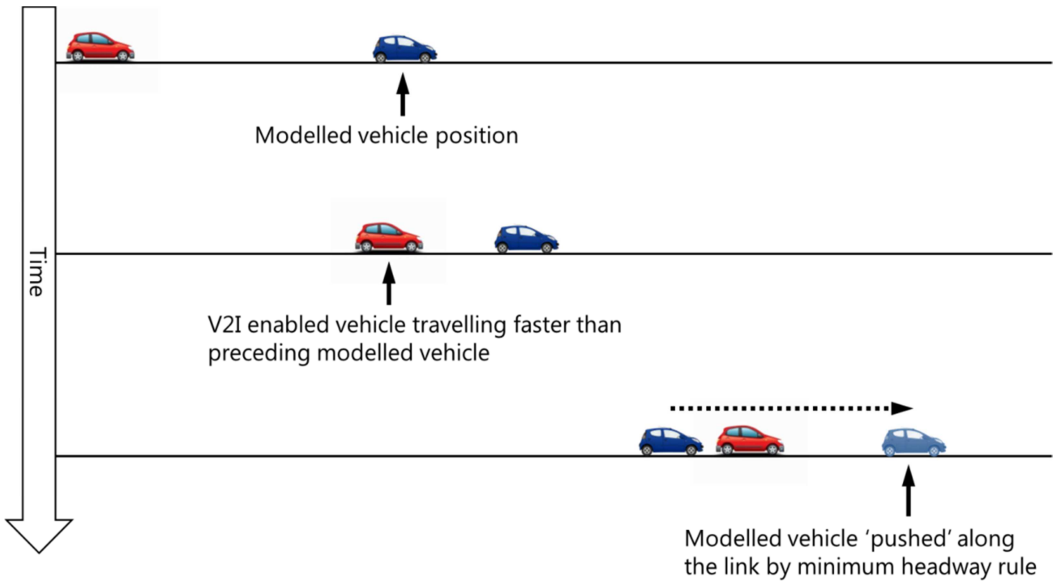

Figure 5.4: Hybrid model adjustment of trajectory example 'push'.

If a queue of conventional vehicles has been modelled but a V2I vehicle stops before (Figure 5.3) or beyond the end of the queue, the length of the modelled queue can be adjusted accordingly. When in free-flow conditions, if the distance between a following, V2I enabled, vehicle and a preceding conventionally detected vehicle falls below a value dictated by a 'minimum headway' rule, the V2I enabled vehicle effectively 'pushes' the preceding vehicle along the link (Figure 5.4).

For a multi-lane situation, if the distance between vehicles in the internal model falls below the minimum headway then it could be because the preceding,

unequipped, vehicle has changed lanes. In that case, the preceding vehicle may be 'pushed' unnecessarily. However, unless lane changing is frequent at a particular site, it is unlikely to have a significant impact on performance. If lane changing of an unequipped vehicle occurs into lanes controlled by different signal phases (i.e. in the case of a flared approach) then it would be expected that the vehicles would be detected by additional conventional vehicle detection and updated appropriately in the internal model.

If a conventionally detected vehicle is added to the link at an average free-flow speed and is following a V2I enabled vehicle travelling more slowly then, depending on the position of both vehicles along the link, the trajectory of the following vehicle may be constrained by the V2I vehicle (Figure 5.5).

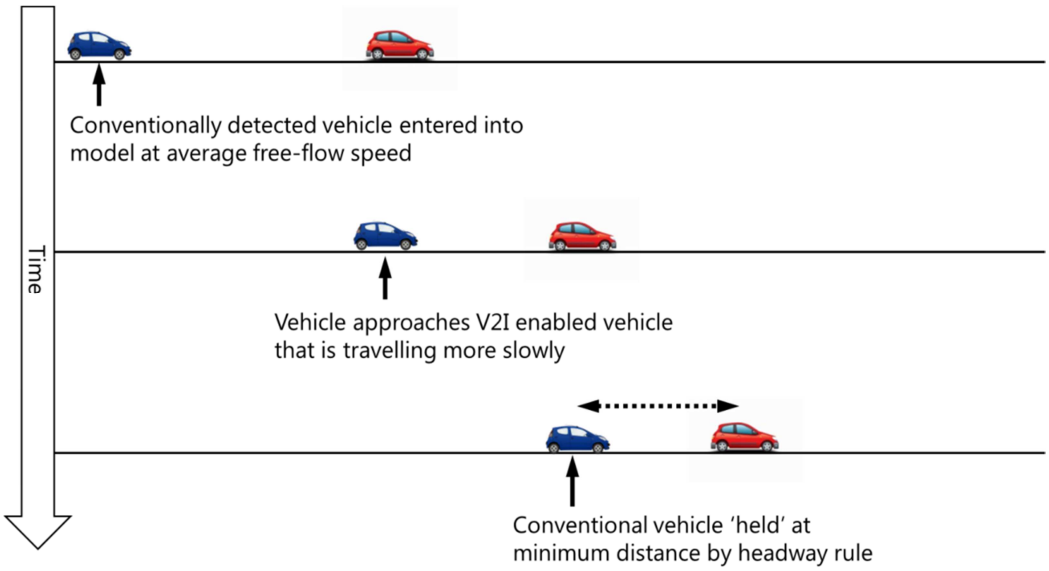

Figure 5.5: Hybrid model adjustment of trajectory example 'hold'.

As discussed, the extent to which the position of conventional vehicles can be modified based on the trajectory of a V2I enabled vehicle is governed by a parameter derived from the minimum safe distance between vehicles at standstill and the minimum headway. In free-flow conditions that parameter will be governed by the minimum headway that is based on the average saturation flow value. However, at very low speeds the minimum safe distance between

vehicles will be the dominant factor. Equation 5.2 describes how the minimum distance between vehicles is calculated.

$$x_{hw} = MIN\left[x_{safe}, \frac{v_n}{s} \right] \tag{5.2}$$

where:

x_{hw} = Minimum distance to preceding vehicle (m).

x_{safe} = Safe distance between vehicles at standstill (m).

v_n = Speed of the following vehicle (ms^{-1}).

s = Saturation flow on the approach link (veh/s).

The model will operate as a vertical queue model until the point that a V2I vehicle enters the link. This means that, until a V2I vehicle is present in the queue, the detection of free-flow conditions (the point that the optimization process begins) relies on identifying a 'critical gap' over the conventional detection. If a V2I vehicle is present in the queue then, depending on the position of the V2I enabled vehicle in the queue, the method of identifying free-flow conditions can be switched (see Figure 5.6). Figure 5.6 demonstrates that the position of the V2I enabled vehicle in the queue determines which method of identifying free-flow conditions is used.

Testing undertaken during the development of the hybrid model showed that the V2I method of identifying free-flow conditions (triggered by the last queuing vehicle reaching a minimum speed threshold) provides the optimum performance when the position of the V2I enabled vehicle is no more than three vehicles from the end of the queue. Beyond that, the uncertainty of the behaviour of the following vehicles in discharging (and consequently, the estimation of the speed of the last queuing vehicle) outweighs the advantages

of using the V2I data. The ability of the hybrid model to switch continuously between methods as new vehicles join the queue maximises the opportunity of using V2I data but ensures that, when V2I penetration rate is low, the performance does not degrade below that provided by conventional detection.

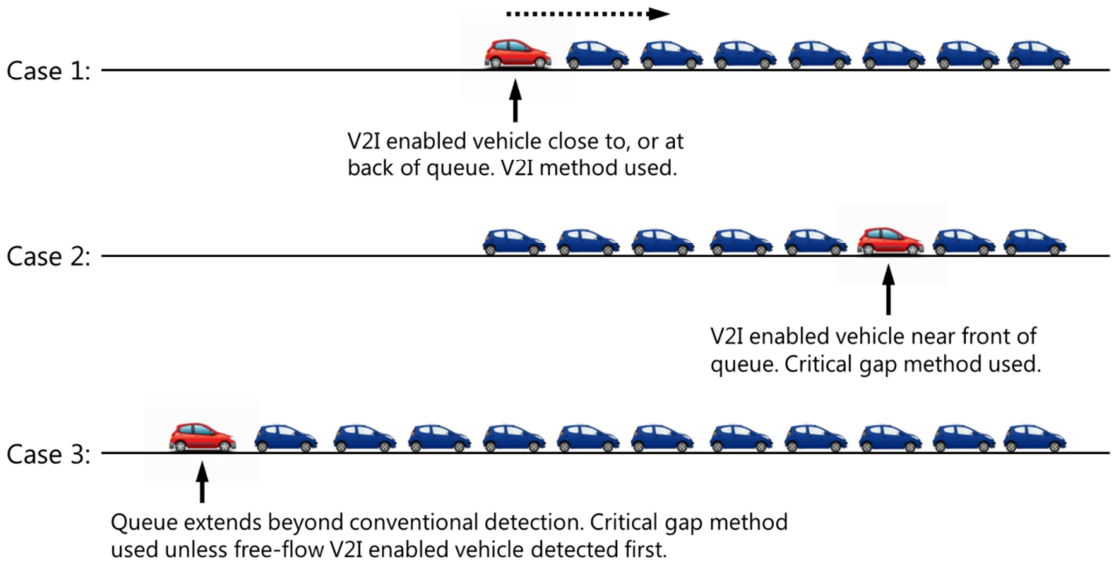

Figure 5.6: Example cases of identifying the point at which the optimization begins.

The presence of a V2I enabled vehicle introduces a spatial element to the queue model. Once a V2I vehicle is detected, and whilst it remains in the link, the interaction of following vehicles that are conventionally detected is modelled. The modelling of interaction is not as sophisticated as a true car-following model but instead is based on a set of rules such as the minimum distance between vehicles already described. The interaction of conventionally detected vehicles that are following a V2I enabled vehicle, in addition to the minimum distance rule, is governed by the driver reaction time and constant acceleration.

When stationary in a queue, the reaction time of a conventionally detected vehicle to a preceding vehicle accelerating away has been modelled as 0.7 seconds. Gipps (1981) chose a reaction time value of 0.6̇ seconds in the development of a car-following model. The Signal Controller developed for this

work operates at 10Hz and consequently the reaction time here is rounded to 0.7 seconds. As the preceding vehicle beings to move away, a reaction timer is started. Once the reaction timer expires, the following vehicle accelerates until the average free-flow speed is reached. However, the speed and position of the following vehicle is then subject to the minimum distance rule already described. The process is shown in Figure 5.7.

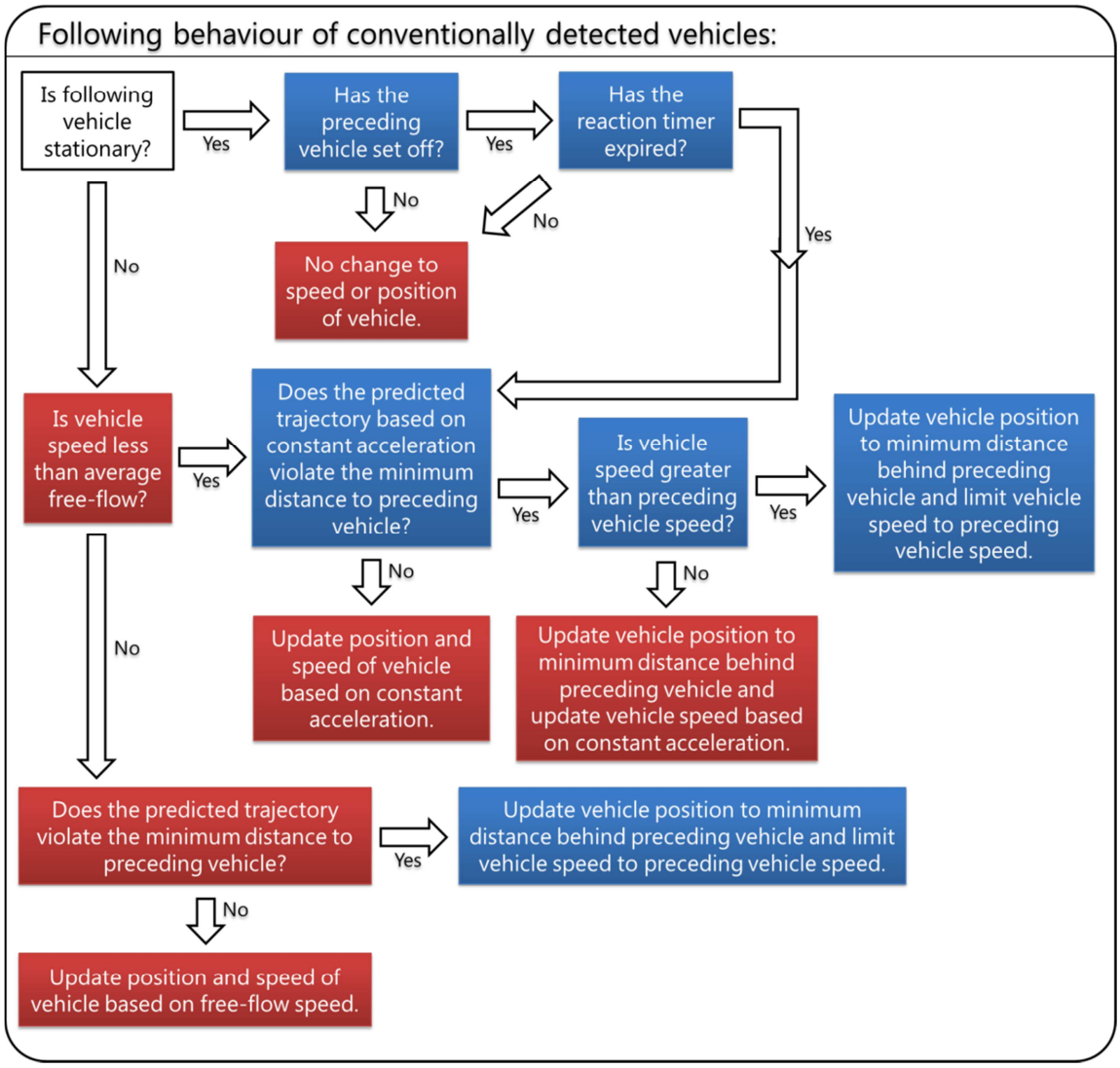

Figure 5.7: A schematic showing the methodology, repeated each controller time-step, for updating conventionally detected vehicle trajectories when a V2I enabled vehicle is present.

Testing during the development of the hybrid model, using a single acceleration value for all vehicles, suggested that a value of 2.5ms^{-1} is most appropriate in

this case. However, although a single value has been used here, this parameter is likely to change from site to site depending on junction geometry, gradient and vehicle characteristics. If the hybrid model was to be developed further, this could be an area that is investigated in more detail to introduce an acceleration profile.

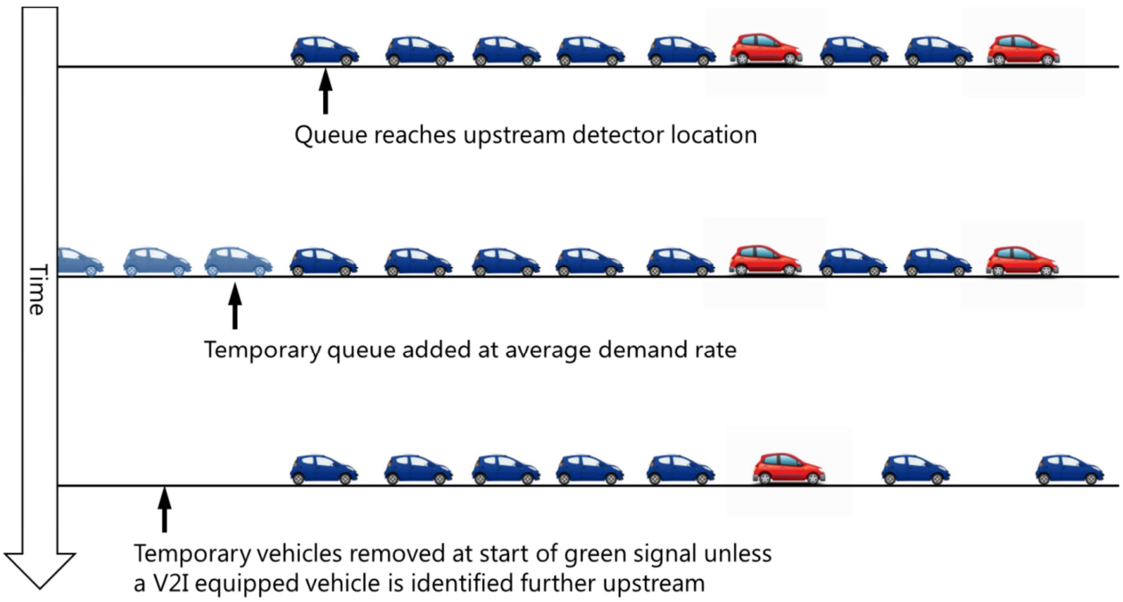

Figure 5.8: Hybrid model methodology for when a queue reaches the conventional upstream detector location.

When a queue reaches the location of the conventional upstream detector and a queue is detected, it cannot be assumed that vehicles further upstream from the detector are V2I equipped unless the entire vehicle fleet is equipped. The model therefore adds 'temporary' vehicles to the back of the queue (whilst the signal state is red) at the average demand rate to ensure that the optimizer judges the queue discharge time appropriately (Figure 5.8). Once the signal state is green, the additional vehicles are removed from the model (unless a V2I equipped vehicle has been identified beyond the location of the upstream detector) as the actual length of the queue is not the important factor in determining when the queue has finished discharging. It cannot be guaranteed that a queue actually exists beyond the upstream detector, or if it does, how

long the queue actually is, unless there is 100% V2I coverage. If a V2I vehicle is present beyond the upstream detector then it can be used to indicate the back of the known queue and used to prevent the optimization process beginning to early. However, there may still be non V2I vehicles queued beyond that. For that reason, the critical gap method or the minimum speed of the last queuing vehicle is used (depending on how many V2I equipped vehicles are present on the link) to identify the point at which the optimization process should begin again.

5.2.3 V2I based classification

The single detector algorithm, first described in Chapter 4 and applied to the MOVA representation in section 5.2, classifies vehicles by length. In reality, the impact of stopping a vehicle will differ according to various characteristics such as powertrain performance, age, vehicle weight and driver behaviour. Most of these factors cannot be derived using conventional detection and control strategies that do utilise a stop penalty value, such as MOVA, instead rely on a user-specified fixed value that is adjusted during initial calibration and validation based on observations.

It is anticipated that smarter detection (e.g. video, radar), and the emergence of V2I will begin to provide accurate data for vehicle type and speed. More accurate and detailed data will enable the performance of each vehicle to be derived more accurately than, for example, using the length based classification applied in this chapter. Here, the effect of using V2I based classification (that assumes knowledge of vehicle type and length) is investigated and compared to the performance of applying V2I detection with no classification.

V2I based classification in this case is applied to the MOVA representation using Equation 5.2. The V2I data is extracted by the Signal Controller from a vehicle record database written to by Vissim every simulation second. The database is

queried and filtered by the Signal Controller to extract the most recent position and speed data for every vehicle on each approach to the junction. Figure 5.9 provides a basic summary of how the data, from conventional detectors and V2I, output from Vissim during a simulation run is used to populate the internal traffic model. Figure 3.11 in Chapter 3 describes the process in more detail.

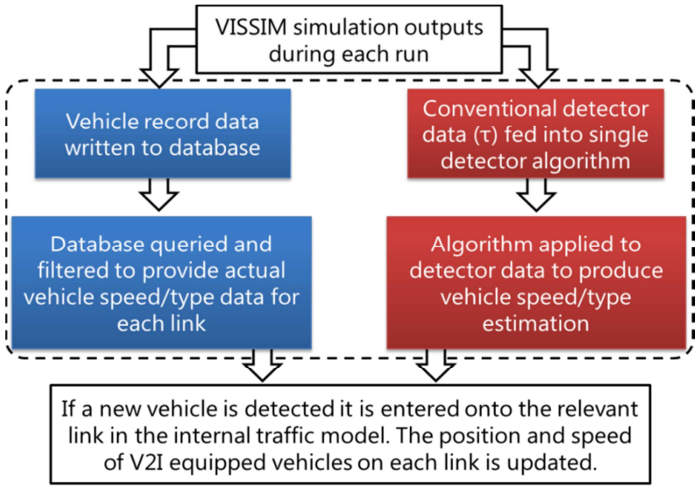

Figure 5.9: Summary of the process used to populate the internal traffic model.

5.2.4 Impact of opposed turning vehicles

The Simulated Environment has been constructed for consistency with the junction layout used in the original Miller (1963) paper. However, the junction layout does not accommodate opposed right turning vehicles as it is primarily used to prove the concept of the Miller algorithm. In this chapter, the Simulated Environment is used to model the each scenario and to test the performance of the hybrid model at different V2I penetration rates. The Simulated Environment is then modified to accommodate storage of a small number of opposed right turning vehicles on each approach and the performance of the hybrid model is tested with different turning proportions and V2I penetration rates.

5.3 Results

5.3.1 Introduction

This section details the outcome of the various model runs used to compare the performance of the single detector classification, V2I and V2I classification models to the MOVA representation described in Chapter 3 for:

- Different proportions of HGV with fixed demand of 600 veh/h (Scenario 1);
- Different demands with a fixed proportion of 5% HGV (Scenario 2);
- Different V2I penetration rates; and
- Different proportions of opposed turning vehicles.

For each scenario the effect of vehicle classification using single detector classification and V2I is demonstrated and, subsequently, the performance of the hybrid model for different penetration rates of V2I equipped vehicles is shown. Finally, the effect of opposed right turning vehicles on the performance of the hybrid model is tested. For the Scenario 1 and Scenario 2 comparisons, the penetration rate of V2I equipped vehicles is assumed to be 100%.

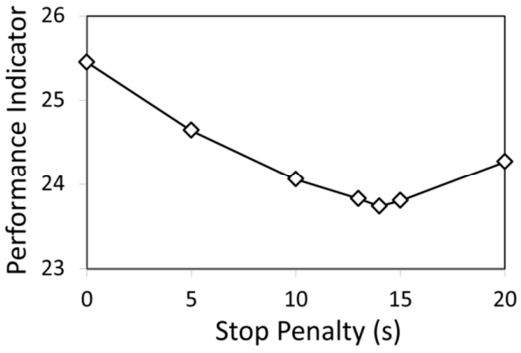

Figure 5.10: Example of process to find the optimal stop penalty

In this section the term HGV refers to the previously defined combination of rigid and articulated HGVs. Optimal stop penalty values have been derived for each scenario by finding the minimum PI value using multiple model runs as shown in Figure 5.10. The introduction of vehicle classification required the optimal stop penalty values for each vehicle type (or class as defined in Vissim) to be derived. The optimal stop penalty is the value that provides the lowest PI value when calculated using Equation 3.8.

The term HGV is used to refer to a combination of rigid and articulated HGVs, as defined previously. However, some specific results for rigid and articulated HGVs are also reported.

A degree of validation that is only possible in a test environment (i.e. identical conditions can be recreated in order to identify optimal parameters) was undertaken for the conventional detection models used as the benchmark for comparison in this work. In reality, for the conventional detection case, some performance deficit would be expected as a result of sub-optimal parameter values being chosen during on-street validation. The documented benefit of V2I could therefore potentially be greater than that demonstrated in this work as V2I would substantially reduce the number of parameters requiring validation. Conversely, it is also the case that this work assumes timely and 'perfect' knowledge of vehicle position, speed and type that, in reality, may be less accurate and/or comprehensive due to, as yet unknown, technological or political constraints.

The majority of results in the following sections and in Chapter 6 are shown with confidence intervals represented by shaded areas around the mean. In all cases, the confidence intervals are shown at 95% and based on the standard deviation. Where one case is being compared to another, the confidence intervals are based on the square root of the sum of the squared standard deviations.

5.3.2 Optimal stop penalty values

Table 5.1 and

Table 5.2 show the optimum stop penalty values for each scenario using a model with conventional detection and the developed hybrid V2I model with 100% penetration rate. In Table 5.1 there is a clear trend, as expected, of increasing stop penalty with percentage of HGV that, for the MOVA representation (labelled as conventional), is broadly in line with the AG45 MOVA manual (Crabtree, M.R. et al., 2012) values for comparable vehicle speed and HGV proportion. The rate of increase of optimal stop penalty value for the hybrid (V2I) model is greater than that for the MOVA representation. It is unclear why that should be the case but an explanation may lie in the V2I method of identifying when to enter the optimization process.

The V2I method for identifying when to enter the optimization process, described in section 5.2.2, requires the last queuing vehicle to have reached a pre-specified minimum speed and for the acceleration of the vehicle to fall within pre-specified bounds. Those identification parameters do not vary between vehicle classes and have been optimized for the base case. It is therefore possible that the parameters are sub-optimal for HGVs and that the increased stop penalty value acts to compensate by encouraging extension of the green signal.

Table 5.2 shows that, whilst the proportion of HGVs and the average free-flow speed remain the same, the optimal stop penalty values also remain similar regardless of the level of demand. There are some discrepancies, most notably at 700 veh/h demand, but no significant deviations.

Table 5.1: Optimal, fixed, stop penalty for Scenario 1 with conventional and V2I detection (seconds).

	0% HGV	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
Conventional	13	17	20	23	27
V2I	11	17	23	29	36

Table 5.2: Optimal, fixed, stop penalty for Scenario 2 with conventional and V2I detection (seconds).

	400 veh/h	500 veh/h	600 veh/h (Base case)	700 veh/h
Conventional	17	17	17	14
V2I	19	17	17	16

The optimal stop penalty values used when implementing vehicle classification are specific to vehicle class. The values have been derived using the same process described for the conventional and V2I detection models. A range of stop penalty values was tested for each vehicle class using a model containing only those vehicles. Table 5.3 shows the derived optimal stop penalty values for each class. The optimum stop penalty value for articulated HGVs is greater than that for rigid HGVs. This reflects the slower acceleration characteristics.

Table 5.3: Optimal stop penalty values for each vehicle class (seconds).

Cars/vans	Rigid HGVs	Articulated HGVs
11	28	44

5.3.3 Benefit of vehicle classification using single detector classification

Having obtained the optimal stop penalty values for each scenario and each vehicle class, the single detector classification algorithm is applied to the MOVA representation to ascertain the benefit provided by vehicle classification within the limits of existing infrastructure. The full tabular results for this and the following section can be found in Appendix B.

Figure 5.11 demonstrates the benefit of using the single detector classification algorithm to apply different stop penalty values to each vehicle class for Scenario 1 and Scenario 2 respectively. The results show that the application of vehicle classification provides a consistent benefit – in terms of delay and stops – of between approximately 5-10% when the proportion of HGVs is 5%. The benefit for articulated HGVs is greatest which reflects the higher stop penalty value applied to that vehicle class.

As the proportion of HGVs increases, the benefit to HGVs quickly reduces. In the case of a 20% proportion of HGVs, the vehicle classification begins to negatively impact on delay and stops for other vehicles. The deterioration in performance of the single detector classification algorithm with increasing HGV proportions, shown in Chapter 4, is likely to influence this result.

The second factor that will affect the performance of vehicle classification in the MOVA representation is the uncertainty of the vehicle position along a link. As described in Chapter 3, vehicles are detected at the upstream location and then traverse the link at a constant speed. It is possible that an HGV can be correctly classified at the upstream detector, the green signal extended by the optimizer to enable it to cross the stop line, but the HGV fail to actually cross the stop line because its speed is less than the average free-flow speed of the link. In that case, the detriment to performance is two-fold as the green has been extended (causing additional delay to opposing traffic) but the expected benefit of saving the HGV a stop has not occurred. The likelihood of this situation will increase as the distance of the upstream detector from the stop line increases.

For this work, HGVs have been assigned a free-flow speed of 80% of the link free-flow speed in an attempt to reduce occurrences of HGVs failing to cross the stop line during a green extension. However, the reduced free-flow speed applies to all HGVs, in the absence of individual vehicle speed data, and

attempts to cater for the majority of HGVs. Consequently, the green will be extended for longer than is necessary on some occasions, causing additional delay to opposing traffic.

It is also worth noting that the single detector classification algorithm was configured to classify into two class bins (short and long) for this application (applying vehicle classification to the MOVA representation). This decision, made to maximise the performance of the algorithm, results in only one 'long vehicle' stop penalty being applied to all HGVs which is likely to have a slight detrimental effect on overall performance of the optimizer as the characteristics of HGVs will vary.

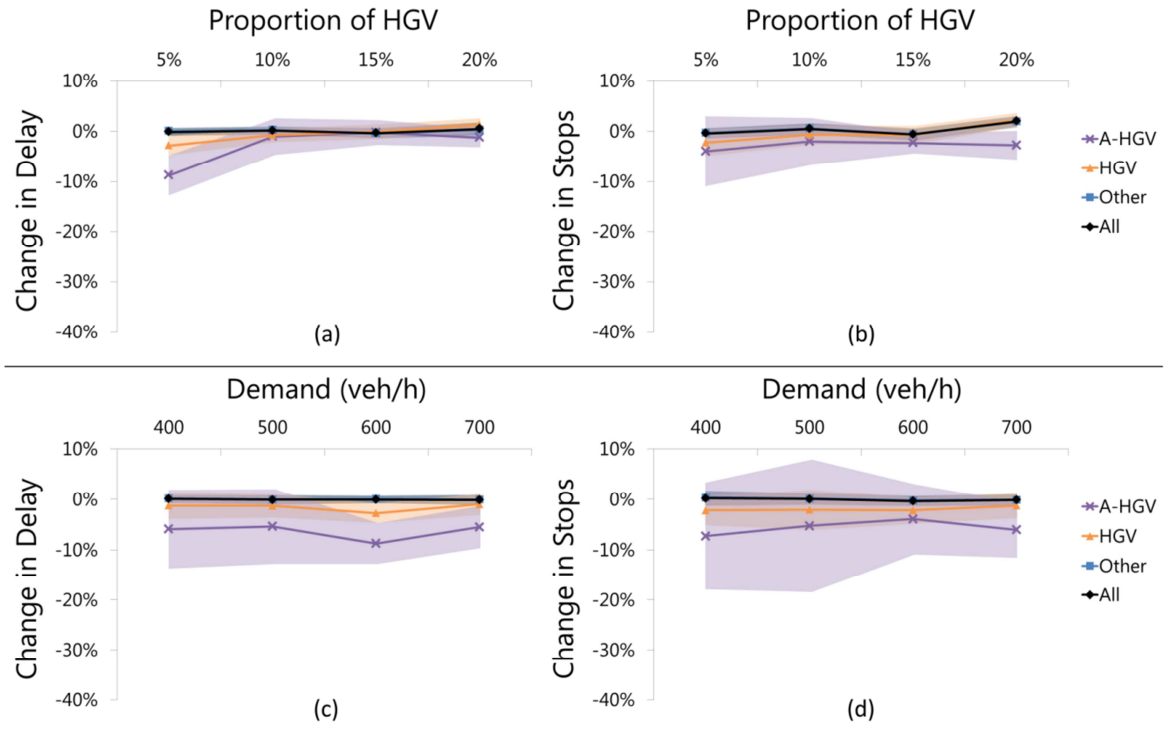

Figure 5.11: Reduction in delay and stops by vehicle class, using single detector classification, compared to the MOVA representation alone for Scenario 1 (a, b) and Scenario 2 (c, d).

5.3.4 Benefit of V2I and vehicle classification using V2I

Having demonstrated the benefit to HGVs of providing vehicle classification in the optimization process, the MOVA representation is now replaced with the

hybrid model which is first tested with a 100% V2I coverage to ascertain the full benefit that can be expected from V2I.

Figure 5.12(a) and Figure 5.12(d) first demonstrate the benefit – in terms of reduction of stops and delay – of using V2I compared to conventional detection (the MOVA representation) with vehicle classification for Scenario 1 and Scenario 2 respectively. The results show that, despite the stochastic nature of micro-simulation, V2I consistently provides an additional overall benefit in the region of 10% – 15% in terms of delay, increasing as the junction approaches saturation. The benefit of V2I is consistent as the proportion of HGVs increases.

Figure 5.12 (b, c) provides a detailed breakdown of the effect of assigning different stop penalty values to each vehicle class, compared to using V2I alone, with an increasing proportion of HGV. The results show the effect on all HGVs (rigid and articulated HGVs combined) as well as on articulated HGVs (A-HGV) separately to demonstrate the benefit given to the slowest accelerating vehicles. As expected, the benefit to HGVs reduces as the proportion of HGVs increases but it is clear that vehicle classification provides a significant benefit to HGVs with very little, and in some cases no, adverse effect on other vehicles. In fact, at the highest proportions of HGVs there is a positive effect on delay for all vehicles. The increase in benefit to HGVs from the previous section can be attributed both to the more accurate V2I classification and the more accurate representation of vehicle trajectories.

It is important to note that, in reality, the vast majority of junctions operate with less than 10% HGVs in the vehicle mix. A benefit of between 10% and 20% in terms of both stop and delays, above and beyond that for V2I alone, would therefore be expected for the slowest accelerating vehicles in most cases, except where the junction is approaching saturation, and between 5% and 10% overall for HGVs combined.

Figure 5.12 (e, f) shows the effect of vehicle classification, compared to V2I alone, with increasing demand. The results indicate that the use of vehicle classification provides a significant benefit in delay and stops of between approximately 5 and 15%, except at 700 veh/h demand where the junction is approaching saturation. Interestingly the largest benefit for HGVs occurs at 500 veh/h demand when it might have been expected to occur at 400 veh/h. However, it should be noted that these results compare to V2I alone. The introduction of V2I alone does not provide the same degree of benefit to every vehicle class compared to the MOVA representation using conventional detection. Some of the benefits to HGVs therefore appear large when compared to V2I alone but a truer representation of the relative benefit is provided by comparing to the consistent benchmark of the MOVA representation (Figure 5.13).

Figure 5.13 (a, b) demonstrates that, when compared to the MOVA representation without vehicle classification, the introduction of V2I with vehicle classification provides consistent benefit in delay for Scenario 1 to all vehicle classes of between approximately 10-20%. The benefit in stops is generally lower at around 10% but is consistent as the proportion of HGVs increases.

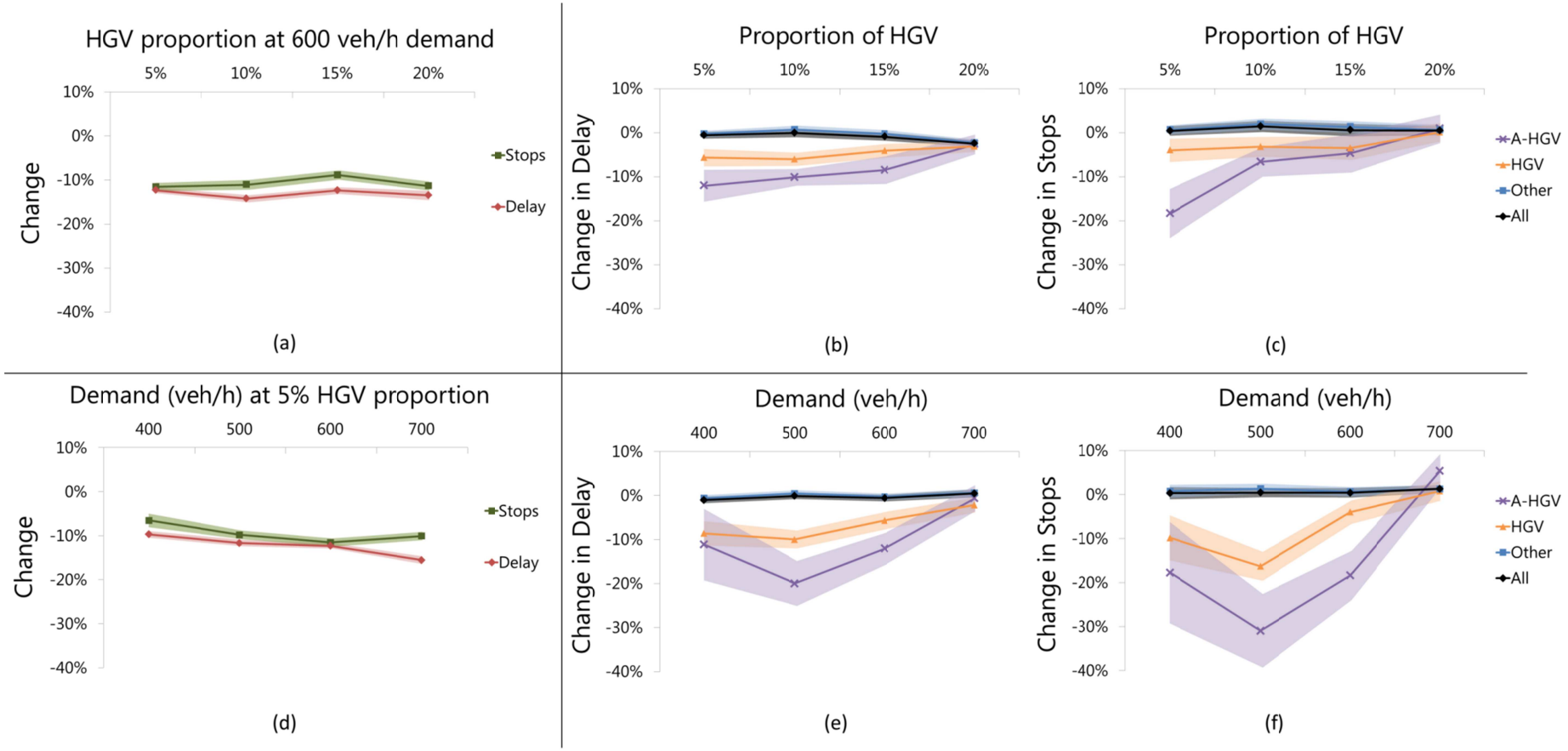

Figure 5.12: Reduction in delay and stops for all vehicles, using V2I alone, compared to conventional detection for Scenario 1 (a) and Scenario 2 (d). Additional reduction in delay and stops by vehicle class, using V2I with vehicle classification, compared to V2I alone for Scenario 1 (b, c) and Scenario 2 (e, f).

The benefit in stops to articulated HGVs is shown to be considerably larger (greater than 20%) at a 5% proportion of HGV than for other cases but the sample of articulated HGVs in that case is small and so it does depend to some degree on whether the HGVs arrive at the signals during a green signal (enabling an extension). The change in stops for the 600 veh/h case does not follow the trend in Scenario 1 or Scenario 2 when compared to other cases. The plotted results for combined HGVs therefore provide a more consistent representation of the benefit in terms of stops.

Figure 5.13 (c, d) shows that the introduction of V2I with vehicle classification provides a greater benefit for HGVs than other vehicle classes, particularly at lower demand where there is a significant benefit to all HGVs of around 25% in delay and stops. At higher demand, the benefit to all vehicles of V2I increases but the relative benefit to HGVs reduces. This is to be expected as the opportunities to provide green extensions over and above the initial queue discharge time decrease.

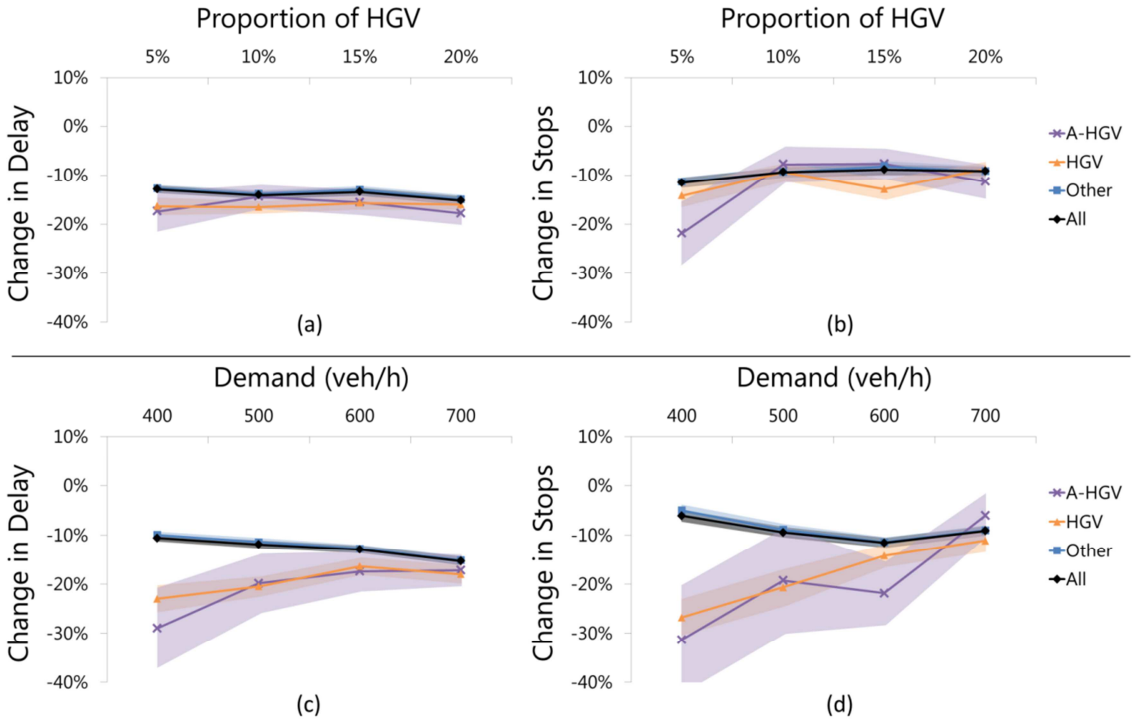

Figure 5.13: Reduction in delay and stops by vehicle class, using V2I with classification, compared to the MOVA representation alone for Scenario 1 (a, b) and Scenario 2 (c, d).

5.3.5 Secondary benefits

More comprehensive data provided by V2I enables the position of individual vehicles, not travelling at a constant speed, to be predicted more accurately for future model time-steps. The accuracy of the predicted vehicle trajectory improves as vehicles accelerating from a queue reach a constant speed. There is therefore a compromise between entering the optimization process early enough during the green to maximize the opportunity for optimal stage end decisions but not so early that the vehicle trajectory predictions are inaccurate due to rapid vehicle acceleration. The decision to enter the optimization process when using V2I is made based on the speed and acceleration of the last queuing vehicle reaching pre-defined thresholds that ensure adequate trajectory prediction accuracy.

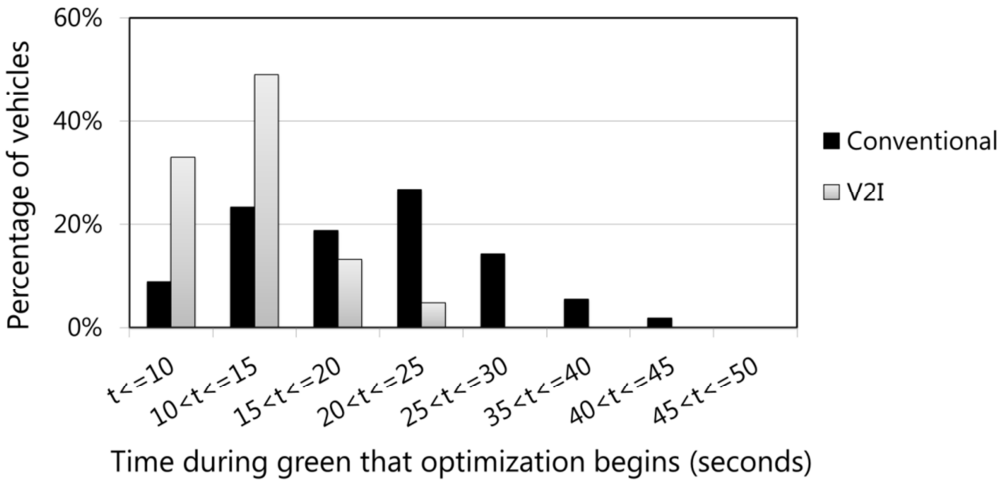

Figure 5.14: Distribution of time during the green that the Miller algorithm begins – 100% V2I coverage.

In the case of 0% V2I penetration rate, the model relies solely on the critical gap method to predict when a queue has discharged. The critical gap method is required to be cautious so as to minimize the instances of the optimization process beginning too early. If the optimization begins before 'non-V2I' vehicles are travelling at an, approximately, constant speed it is likely to result in sub-

optimal decisions being made. As a consequence, the optimization process is entered comparatively later on average than when V2I data is available. In some situations this may lead, for example, to the optimizer missing the opportunity to end a stage at the optimal point, thus reducing efficiency. Figure 5.14 demonstrates that the optimization process is entered consistently earlier with V2I than when using conventional detection.

Figure 5.15 shows a distribution of the time into an amber signal that vehicles cross the stop line. In the conventional detection case it is necessary to be conservative in presenting the amber signal to vehicles to avoid slower moving vehicles either failing to cross the stop line or 'running the red'. V2I detection, specifically knowledge of individual vehicle trajectories, enables the amber signal to be presented at the relevant time based on the individual vehicle speed and, consequently, enables the amber signal to be presented earlier on average. This results in a greater proportion of vehicles crossing the stop line within the amber period (t<=3 seconds).

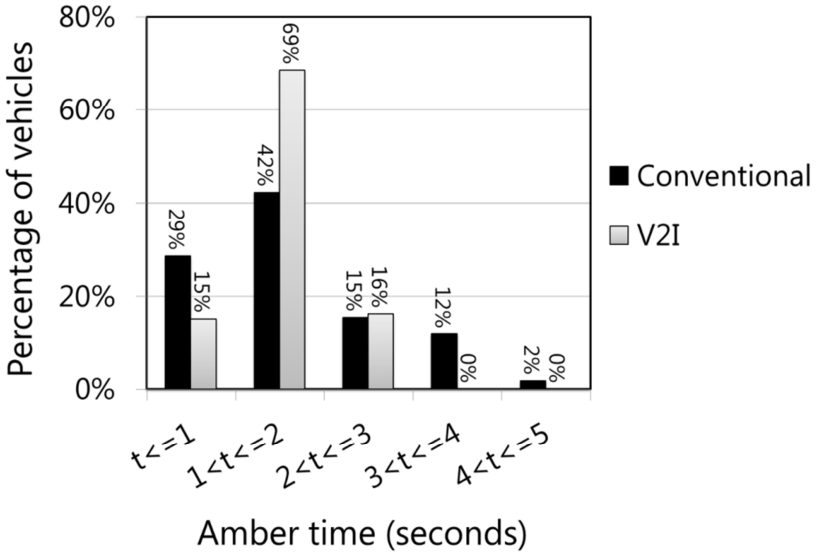

Figure 5.15: Distribution of time into amber signal that vehicles cross the stop line.

In addition, amber times of greater than 3 seconds indicate non-compliance (i.e. vehicles running the red signal). It is clear to see the positive effect on vehicle compliance rates with no vehicles crossing amber after 3 seconds with V2I. Hence, the introduction of V2I not only reduces delays also increases safety by significantly reducing, or in this case eliminating, red running.

5.3.6 V2I penetration rate

In the previous sections, although a hybrid model has been introduced, the V2I penetration rate has been assumed to be 100%. In this section the performance of the hybrid model is tested with different V2I penetration rates to demonstrate the rate at which the total benefit of V2I detection can be realised.

Figure 5.16 demonstrates the performance of the hybrid detection model with an increasing V2I penetration rate. The benefit of V2I in the internal traffic model is primarily accrued from two elements; more accurate queue discharge identification and more efficient stage endings. The profile for higher demand reflects the increased importance of more accurate queue discharge identification as the junction approaches saturation.

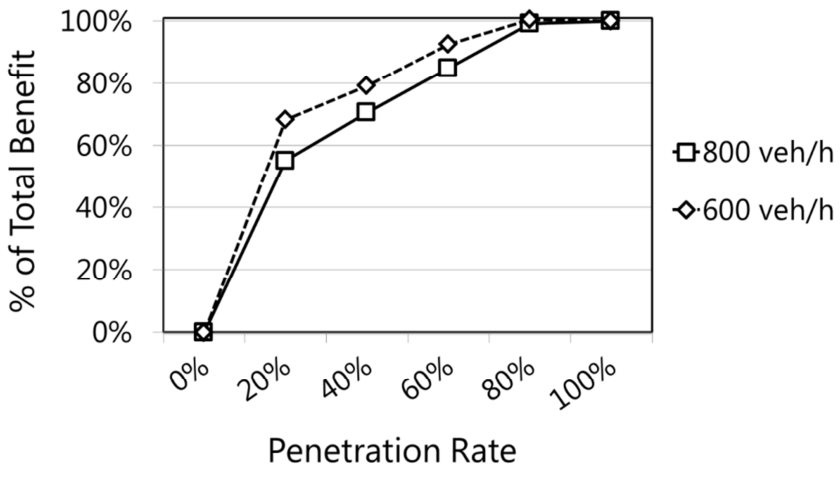

Figure 5.16: Effect of increasing V2I penetration rate with varying demand.

The results suggest that the majority of the benefit provided by more accurate queue discharge identification is accrued with a penetration rate of 20-30% with the remainder of the benefit being primarily provided by improved efficiency of stage endings. The overall rate of benefit realisation follows a similar trend for both presented demand flows, indicating that approximately two thirds of the overall benefit of V2I detection can be realised at approximately one third penetration rate.

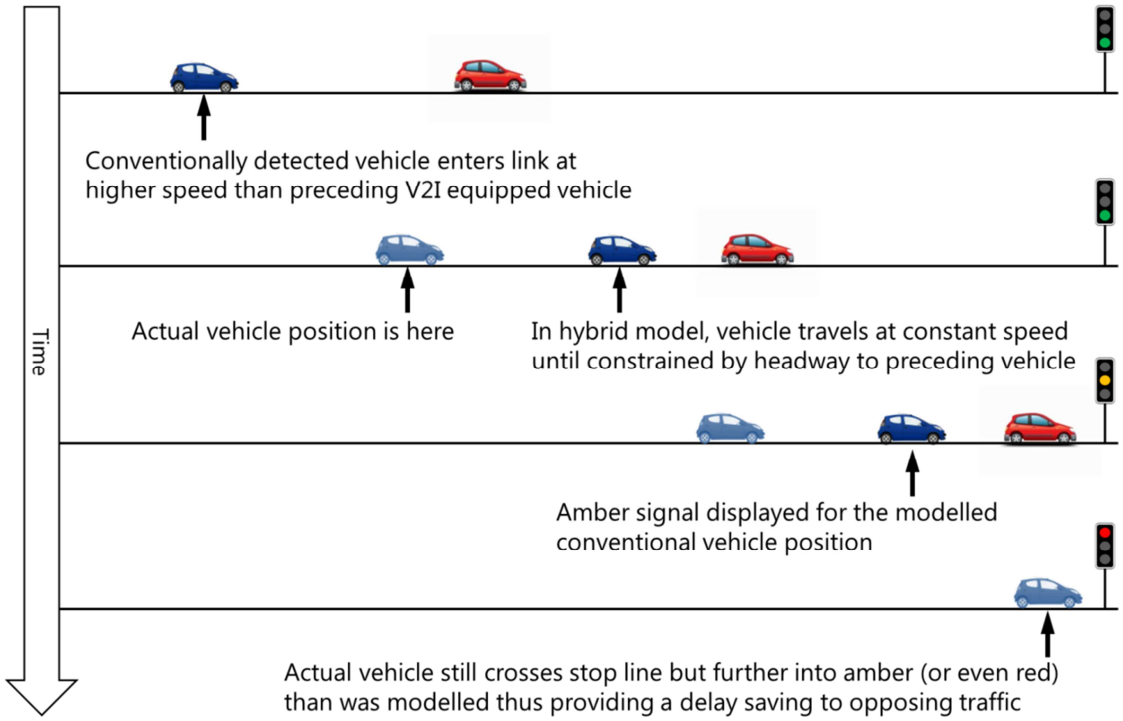

Figure 5.17: Stage ending example of conventionally detected vehicle stage ending.

At 600 veh/h demand, the benefit at 80% V2I penetration rate is slightly more than that for the 100% V2I case. This could simply be explained by the stochastic nature of micro-simulation but could be a result of the way the hybrid model operates. In some situations, conventionally detected vehicles may be presented with an amber signal earlier than a V2I equipped vehicle would due to errors in the modelled vehicle position (Figure 5.17). If the conventionally detected vehicle still crosses the stop line, albeit later during the amber period

or even into the red, then it will experience no additional delay. However, opposing traffic has been saved a small amount of delay as the signal has changed quicker than it would if the green had been extended for the true position of the conventionally detected vehicle.

Clearly, the earlier stage ending is not ideal in terms of road safety and the hybrid model has been configured to ensure that instances of red-running are minimized. However, it does potentially raise questions regarding the balance of safety and performance if the V2I penetration rate reaches 100% of the vehicle fleet.

5.3.7 Turning movements

As described in Chapter 3, the Simulated Environment has been modified to allow a small amount of storage for opposed right turners. The hybrid model has been tested with various turning proportions to understand the effect of opposed right turners on performance. The results for the base case (0% right turning vehicles) vary slightly from those reported in previous sections as a consequence of the minor modifications to the network to accommodate right turning vehicles (Figure 5.18).

The results in Table 5.4 demonstrate that, at any V2I penetration rate, the performance of the optimizer is not significantly affected with a 5% turning proportion. That indicates that the number of vehicles waiting to turn right rarely exceeds the internal storage space. As the turning proportion increases, the performance of the junction begins to deteriorate more quickly due to queued right turning vehicles blocking vehicles travelling ahead. The 15% case is considered a worst case scenario as it would be expected that an early cut-off filter is provided for such significant proportions of turning vehicles or that the junction design would accommodate the movement in a more appropriate manner (see section 3.3.2).

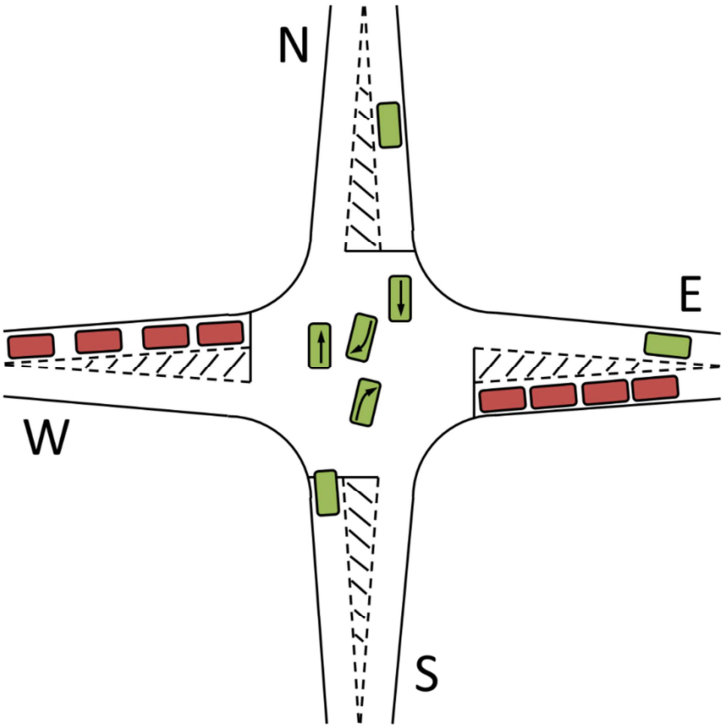

Figure 5.18: Modified Simulated Environment incorporating opposed turning movements with storage space.

The improvements in PI for higher V2I penetration rate compared to 0% are substantial for higher turning proportions. This is a particularly significant result given that the comparison is with the reasonably sophisticated Miller algorithm.

Table 5.4: PI with turning movements for varying V2I penetration rates.

	Proportion of Vehicles Turning Right			
V2I Penetration Rate	**0%**	**5%**	**10%**	**15%**
0%	31.6	34.1	42.6	63.3
33%	30.2	33.1	40.6	59.1
67%	29.1	31.4	37.4	52.4
100%	28.5	30.3	37.9	53.1

An interesting observation is that the 67% V2I penetration rate slightly outperforms the 100% case at higher turning proportions. The likely explanation for this result lies in the queue clearance strategy that is employed before

optimization takes place. At higher turning proportions, the queue on an approach fails to clear more often due to blocking from right turning vehicles (Figure 5.19). At higher V2I rates, the green signal will continue to be extended until the queue clears or the pre-specified maximum is reached. However, in the initial queue clearance period the control strategy is not yet optimizing and, consequently, the longer green extensions may actually result in a sub-optimal stage ending in this circumstance. At lower V2I rates the critical gap method is likely to be employed more often, incorrectly identifying that the queue has fully discharged but, in some circumstances, potentially providing a better result.

The MOVA control strategy attempts to address the problem of exit blocking by monitoring the queue over the X detector from the start of green and, if vehicles have not moved off after a reasonable time, making a decision on whether to terminate the green early. However, if the queue is beyond the downstream detector it is effectively ignored.

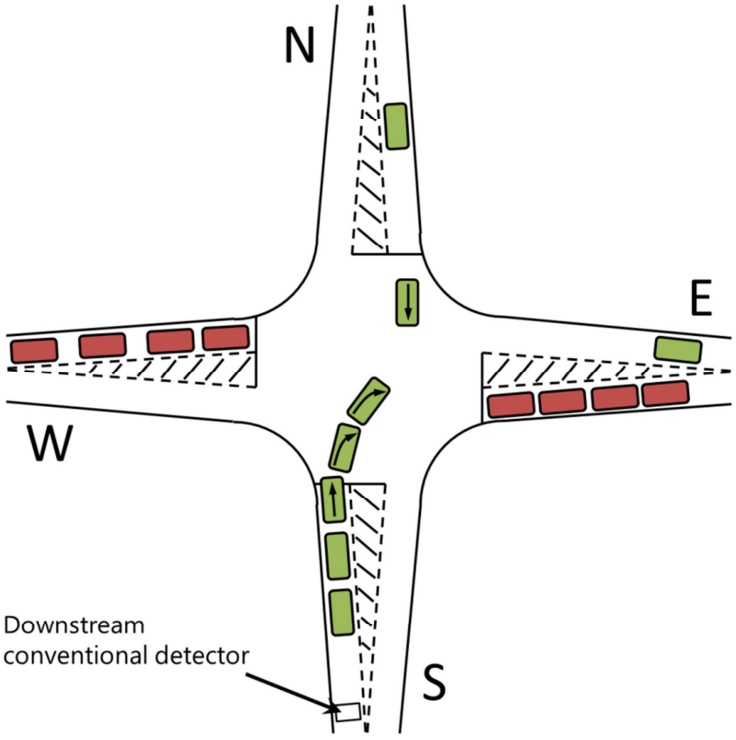

Figure 5.19: Opposed right turners blocking ahead traffic causing a queue downstream of the conventional detector.

5.3.8 Summary

The application of the single detector vehicle classification algorithm, developed in Chapter 4, to the MOVA representation described in Chapter 3 provides a benefit to HGVs of up to 5% in terms of stop and delays. The benefit to the articulated HGV class in Scenario 2 (the 5% HGV proportion cases) is greater (between 5-10% reduction in delay and stops) but the benefit to all HGVs rapidly reduces as the proportion of HGV increases. At the highest proportion of HGVs the introduction of vehicle classification has a negative impact on other vehicle classes.

Introduction of V2I detection without classification – at a 100% penetration rate – provides a significant benefit compared to conventional detection of between 10% and 15% for delay and between 5% and 15% for stops. Increased demand does not have a significant impact on the optimal stop penalty value for either conventional or V2I detection. The benefit in delay provided by V2I increases with demand.

The use of V2I increases the opportunity for optimization and enables the amber signal to be presented earlier, on average, than when using conventional detection. A secondary benefit of improved stage endings is a substantial reduction in vehicles crossing the stop line during the red signal.

The developed hybrid model is capable of delivering a significant proportion of the overall V2I benefit at low penetration rates. For example, it has been shown that approximately 65% of the benefit can be realized with a penetration rate of around 30%.

V2I with classification provides a significant benefit for all HGVs over and above the use of V2I alone. A consistent reduction in delay of around 5% is reported for lower proportions of HGV, reducing as the proportion of HGV increases. The

benefit to articulated HGVs is greater, particularly at lower demand flows. In most cases the introduction of vehicle classification has a negligible effect on non-HGV vehicles but at the highest HGV proportion it provides a noticeable benefit to all vehicles.

Table 5.5 shows that introducing vehicle classification to the MOVA representation improves the overall PI in all but the highest HGV proportion (with the exception of the 10% HGV case). At the highest HGV proportion, the performance of the single detector classification algorithm deteriorates and is likely to be an influencing factor.

The introduction of vehicle classification to the V2I hybrid model improves the overall PI compared to V2I alone except in the case of the highest demand, demonstrating that providing vehicle classification benefits overall performance. At the highest demand it may be the case that providing additional green extensions for HGVs is detrimental to overall performance although it is unclear why it would only affect the V2I case. The ability to benefit HGVs without worsening the overall performance is important when considering road safety and local emissions.

Table 5.5: Percentage improvement in overall PI when implementing vehicle classification.

HGV Proportion	**5% HGV** (Base case)	**10% HGV**	**15% HGV**	**20% HGV**
Conventional	0.4	-0.2	0.5	-1.1
V2I	0.5	0.0	0.9	1.3
Demand	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
Conventional	0.1	0.1	0.4	0.3
V2I	0.7	0.7	0.5	-0.9

5.4 Conclusions

Initially, this chapter considered the effect of applying the single detector classification algorithm, developed in Chapter 4, to the MOVA representation. The length based classification of the algorithm provides a proxy for vehicle acceleration performance and enables the effect of applying different stop penalty values to each vehicle class to be evaluated.

Assumptions were made about the frequency and type of data (e.g. vehicle speed and type) that will be made available through V2I. Importantly, the purpose of incorporating V2I in this case was to quantify the benefit that can be achieved by improving the accuracy of the traffic model of the existing Miller optimization method. To implement V2I, a hybrid model was developed that is capable of combining conventionally detected data with V2I data to realise the benefits of V2I as quickly as possible. The benefit of introducing vehicle classification to enable assignment of optimization weightings by vehicle type was subsequently considered and compared to V2I alone.

Considerable benefit, compared to a MOVA representation, of between 10% and 15% in terms of delay has been demonstrated for all vehicles as a consequence of the more accurate and more comprehensive data provided by V2I. The magnitude of benefit is consistent with the outcomes reported by Feng et al. (2015) for a 100% penetration rate, supporting the findings of this chapter. However, the benefit for HGVs can be substantially increased, by almost 20% in the case of the slowest accelerating vehicles, as a result of incorporating vehicle classification. The additional benefit to HGVs is particularly apparent at lower demand. The results also suggest that providing HGVs with a superior level of service can benefit all vehicles, particularly as the proportion of HGVs increases.

It is important to note that the comparisons made in this chapter are not to a fixed-time, or even the UK standard System-D VA strategy, but rather a reasonably sophisticated and well-established optimization technique.

The vehicle classification section of this paper has focused on providing vehicles with slow acceleration a greater weighting in the optimization process. This differs from the majority of bus priority applications in that the optimizer weighting is based on vehicle performance rather than factors such as vehicle occupancy or policy objectives.

Such a large reduction in stops per vehicle, particularly for HGVs – including up to 30% for articulated HGVs – as a result of introducing V2I technology with vehicle classification would have a considerable positive effect on fuel consumption and emissions. The freight industry would benefit from this but the impact on public health due to reduced emissions is just as important, if not more.

The hybrid detection model tests indicate that a substantial proportion of the benefits demonstrated in this paper can be achieved at relatively low penetration rates (e.g. 65% of the full benefit with 30% penetration rate) and that there is no lower threshold of V2I coverage required before benefits start to materialize. Consequently, the findings of this chapter have important policy and implementation implications given the immediate nature of return on investment.

Chapter 6
A car-following model based optimizer

6.1 Introduction

The previous chapter investigated whether modifying an existing, state-of-the-art, algorithm to explicitly account for vehicle class in the optimization process, could improve the performance of the optimizer. It also described the development of a hybrid model that enables V2I data to be used alongside conventional detection to enhance the performance of the algorithm. The hybrid model is capable of providing a relatively simple spatial representation of vehicles traversing approach links to a junction when the V2I penetration rate is greater than 0%.

The spatial representation of the hybrid model takes into account some of the interaction between vehicles and enables an improved estimation of vehicle trajectories. However, the Miller algorithm incorporated in the MOVA representation still makes use of the data provided by the hybrid model in the same way as the simple vertical queue model presented in Chapter 2. That is, no changes are made to the Miller optimizer itself but rather to the accuracy of the data presented to the optimizer and the identification of when the optimization process should begin.

It is recognised that, although the hybrid model provides some representation of vehicle interaction, it is not a sophisticated car-following model and is based on a set of simple rules (i.e. retaining a minimum headway between vehicles) as opposed to being derived from traffic flow theory. Furthermore, in the previous chapter, the hybrid model is used in conjunction with the Miller (1963)

algorithm, thus necessitating the retention of a pre-specified stop penalty to ensure that the delay caused by a vehicle stop is fully accounted for.

This chapter investigates whether a more detailed spatial model, capable of representing interaction between vehicles, could provide an additional performance enhancement compared to the hybrid model incorporated into the MOVA representation developed in Chapter 3. A popular and more sophisticated car-following model is described here and the methodology used to implement it explained. A new optimization method has been developed to incorporate the more sophisticated car-following model into the optimizer, the development and testing of which is also described.

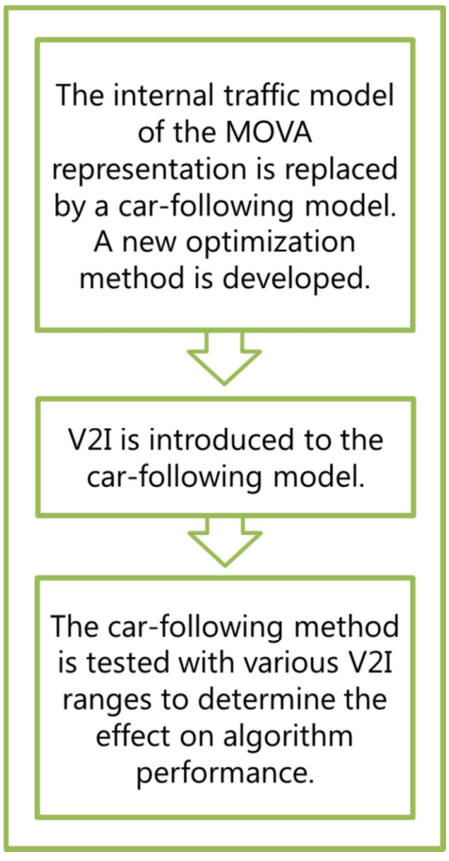

Figure 6.1: Schematic showing Chapter 6 structure.

The car-following traffic model is first implemented with the classification algorithm developed in Chapter 4 and is then developed to incorporate V2I detection with classification. Finally, the developed control strategy is tested with various V2I ranges to demonstrate the effect of providing an increased visibility of vehicle arrivals on optimizer performance.

6.2 Model type

The vertical queue model used in the MOVA representation, described in Chapter 2, assumes that vehicles travel at a constant speed from the point of detection to the signal stop line and are then added to a 'vertical' queue whilst the signal is at red. When the signal becomes green, the queue is assumed to discharge at a constant rate, the saturation flow rate. In the modified Miller method, described in Chapter 3 the saturation flow rate is primarily used to estimate the length of time a queue on an approach, currently receiving a red signal, will take to discharge once the signal becomes green.

The vertical queue model does not represent any interaction between vehicles and the delay calculated in the Miller method is therefore a simplification and possibly, in some cases, an underestimation of the reality. As discussed in Chapter 3, it is not clear whether the Miller algorithm fully accounts for the delay caused by a vehicle stop or whether it only accounts for the initial delay between the start of the green signal and the average queue discharge rate being achieved (Figure 6.2). Consequently, a modification is made that is described in Chapter 3 section 3.2.3 as part of the MOVA representation. The stop penalty value, as implemented in the MOVA control strategy, must be calibrated by a practitioner during a relatively short time period based on observations of vehicle mix, speed and site specific factors that affect driver behaviour. It can be seen that the selection of a stop penalty value is therefore

subject to error that can result in deterioration of performance from the optimal value.

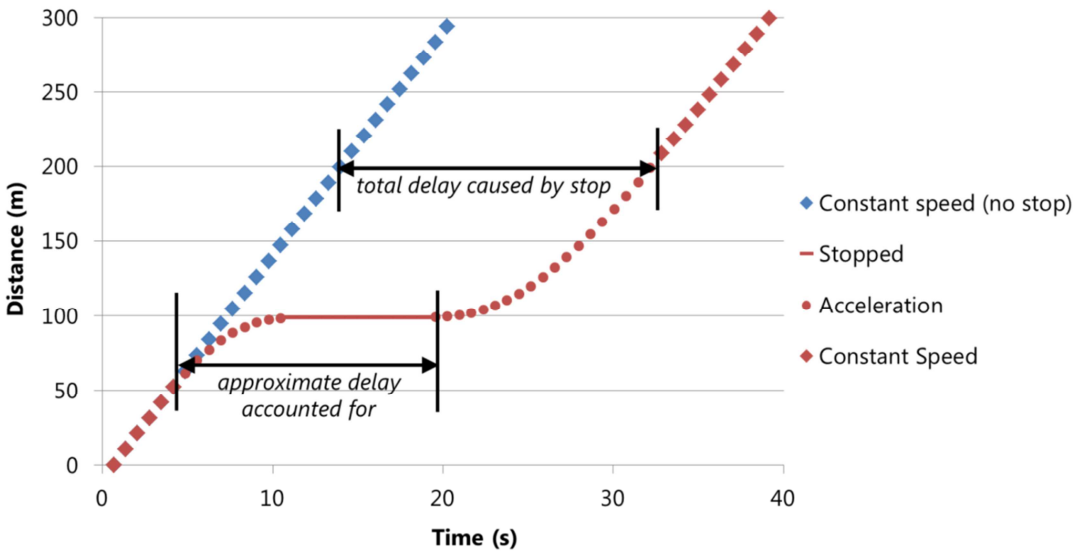

Figure 6.2: Vehicle trajectory example showing the stop penalty purpose.

Conversely, a spatial model such as a car-following model can more accurately represent the interaction between vehicles. The deviation of journey time from an 'ideal' (i.e. unstopped) journey can be recorded to effectively include the delay incurred throughout the stop/start process. When considering emerging technology such as V2I, the increase in the volume and type of information available would suggest the use of a more detailed model to enable more of the information to be utilised. For the existing situation, the use of a model that can more accurately reflect the impact of vehicle stops, thereby removing the need for a stop penalty value, is an attractive proposition. However, the use of a more complex model poses a number of challenges that need to be considered when developing a practical solution, some of which are addressed in this chapter.

6.3 The Gipps model

6.3.1 Choice of model

Any reasonably sophisticated car-following model could realistically have been chosen for this application. The aim of this chapter is not to re-invent or develop a new car-following model but rather to evaluate the effect of applying a more detailed traffic model to the optimization process. The Gipps (1981) model was chosen for this study due to the relative ease with which it could be implemented into the optimizer and the correlation between model parameters and driver behaviour/vehicle characteristics. The correlation between model parameters and real-world properties is important if the car-following method is developed further. As with existing models, practitioners would need to calibrate the model to account for site specific conditions, particularly in the case of conventional detection. The easier it is to relate the model to the real-world, the more likely it is that the model will be accurately calibrated.

The Vissim micro-simulation software package, in which the Simulated Environment is constructed, employs the Wiedemann (1974) car-following model described in Chapter 2. The Wiedemann (1974) model is a psycho-physical model based on a drivers' perception of the difference in speed between their vehicle and the preceding one. The model parameters that affect saturation flow are less relatable to real-world characteristics. For example, in the Wiedemann (1974) model, one of the terms used to calculate the following distance consists of two parameters, the additive part of desired safety distance and a multiplicative part of the safety distance. The correlation between these parameters and real-world vehicle/driver characteristics is unclear in the available literature.

A consequence of using different car-following models is that the model behaviour will differ slightly and model parameters will not directly correlate. Conversely, using the same car-following model in the Simulated Environment and the internal traffic model of the optimizer would not present a reliable indication of the ability of a car-following model to reflect real-world conditions as model parameters could be calibrated to match those of the Simulated Environment. The use of a different model ensures that the optimizer does not have perfect knowledge of the following behaviour and is therefore more representative of a system that is calibrated to a real-world set of parameters such as vehicle performance characteristics.

Ordinarily, in micro-simulation software, the characteristics such as acceleration, deceleration and desired speed are randomly chosen from a pre-specified distribution as the vehicle is entered onto the modelled network. The stochastic nature of micro-simulation is created as the random seed specified for the model is adjusted from one simulation run to the next, resulting in different parameter values being selected from the various distributions.

In this application, certainly in the case of the real-time element of the optimizer algorithm, much of the stochasticity is removed as parameters such as the desired speed, for example, are not known. In future, it may be possible to gather more information regarding vehicle and driver characteristics such as acceleration, deceleration, desired speed and reaction time but such information is subject to various issues regarding data privacy and policy, discussed in Chapter 7. Ultimately, if the performance characteristics of a vehicle such as acceleration and deceleration can be tracked at upstream junctions through a network, the parameters used in the model could be made vehicle specific rather than vehicle type specific.

6.3.2 The Gipps model principles

The car-following model proposed by Gipps (1981)is based on the principle of setting performance limits on the driver and vehicle that are used to calculate a safe speed, that is, a speed that ensures the vehicle can stop safely if the preceding vehicle stops rapidly. The equation used to describe the Gipps model is split into two terms. The first term (Equation 6.1) describes two constraints.

The first constraint is that a vehicle should not exceed its driver's desired speed and the second is that the acceleration should first increase with speed as engine torque increases and then decrease to zero as the vehicle approaches the desired speed.

The second term (Equation 6.2) considers the limitation of braking. If the preceding vehicle brakes at its driver's maximum desirable rate then the following vehicle must be travelling at a far enough distance behind to allow for the driver reaction time and the vehicle stopping distance. A further safety margin is also included that allows for an additional delay in driver reaction time. This allows for some margin of error for the following driver to react to the vehicle ahead.

The vehicle speed in the new time step is calculated by Gipps (1981) as the minimum of the two terms Equation 6.1 and Equation 6.2, described by Equation 6.4.

$$v_n^{\,1}(t+\tau) = v_n(t) + 2.5a_n\tau\left(1-\frac{v_n(t)}{V_n}\right)\sqrt{\left(0.025+\frac{v_n(t)}{V_n}\right)} \qquad (6.1)$$

$$v_n^{\,2}(t+\tau) = b_n\tau + \sqrt{\left(b_n^2\tau^2 - b_n\left[2y - v_n(t)\tau - \frac{v_{n-1}(t)^2}{\hat{b}}\right]\right)} \qquad (6.2)$$

$$y = x_{n-1}(t) - s_{n-1} - x_n(t) \tag{6.3}$$

$$v_n = \min(v_n^{\,1}, v_n^{\,2}) \tag{6.4}$$

where:

a_n = Maximum acceleration that the driver of vehicle *n* wishes to undertake (ms^{-2}).

b_n = The maximum deceleration that the driver of vehicle *n* wishes to undertake (ms^{-2}).

$\hat{b}$ = The maximum deceleration that the driver of vehicle *n-1* wishes to undertake (ms^{-2}).

s_n = The effective length of vehicle *n*. The sum of the vehicle length and the distance to the preceding vehicle at standstill (m).

V_n = The desired speed of vehicle *n* (ms^{-1}).

$x_n(t)$ = The position of the front of vehicle *n* at time *t* (m).

$v_n(t)$ = The speed of vehicle *n* at time *t* (ms^{-1}).

τ = The reaction time (s).

6.3.3 Calibration of parameters

As already discussed, in micro-simulation applications vehicle and driver characteristics are randomly assigned from a distribution of values (based on observed data) as vehicles enter the road network. Assigning characteristics with a certain degree of randomness helps to reproduce the stochasticity observed in reality.

For this application, the car-following model is being used to replicate the behaviour of real vehicles where the exact parameter values are not known. In the optimizer algorithm (discussed later) the car-following model parameters are assigned average values based on evaluation of optimizer performance over multiple model runs. This applies to the real-time and 'cloned run-ahead' elements of the optimizer algorithm.

However, as is explained later in the chapter, random variation could be introduced for the 'cloned run-ahead' element of the optimizer algorithm. Introducing noise in the optimizer calculations would be an interesting further study and could potentially improve the performance of the optimizer but would also increase the computational requirements considerably and thus the number of random seeds (i.e. the number of model runs with different random noise distribution) would need to be carefully considered.

The calibrated parameters in Table 6.1, with the exception of desired speed, have been evaluated in quarter second steps using the PI described in Chapter 3. The desired speed has been evaluated in half second steps. The various parameters were calibrated by evaluating the results from multiple model runs containing only vehicles of the respective classes. An indication of the parameter values could be gained from the speed distributions used in Vissim, derived from ATC data (as described in Chapter 3). However, as previously discussed, the values are not necessarily directly comparable given the difference in car-following models.

Table 6.1: Calibrated Gipps vehicle characteristics.

	Car/Van	Rigid HGV	A-HGV
a_n (ms^{-2})	3.0	1.5	2.0
b_n (ms^{-2})	-3.5	-2.5	-2.5
V_n (ms^{-1})	13.5	13.5	13.5

Table 6.1 shows that the desired speed parameter is identical across all three vehicle classes. This reflects the urban nature of the link and that speeds are constrained by a speed limit. Interestingly, the optimum desired speed value is lower than the free flow speed used in Chapter 5. The maximum acceleration of the rigid HGV class is slightly lower than for the articulated HGV class reflecting the different vehicle power distributions used in Vissim.

The parameter $\hat{b}$ has been set using a method that users can employ in the Aimsun micro-simulation software package, referenced by Ciuffo et al. (2012), which averages the maximum desired deceleration of the leading and following vehicle:

$$\hat{b} = \frac{b_{n-1} + b_n}{2} \tag{6.5}$$

It is recognised that this does not guarantee there will be no intrusions on the leading vehicle. However, in this context it is unlikely to significantly affect the outcome of the optimization. It is possible that an alternative method of estimating $\hat{b}$ could provide superior results. However, it should be noted that, as $\hat{b}$ is an estimation by a driver of the maximum deceleration of the preceding vehicle, it is impossible to derive the value with 100% accuracy.

6.4 Optimizer development

The Simulated Environment has been modified, as shown in Figure 6.3, to relocate the conventional downstream detector to beyond the stop line. The primary reason for relocating the detector is to allow the flow rate to be monitored at the start of the green signal. Monitoring the rate of queue discharge enables adjustments to be made to the model, discussed in the following section, to account for instances of a queue discharging more slowly, or quickly, than predicted. This can particularly present an issue at higher HGV proportions where errors in the single detector classification algorithm, described in Chapter 4, become more common. The consequence of an increase in error is that some HGVs may be incorrectly classified as short vehicles and thus the predicted rate of discharge is higher than the actual rate.

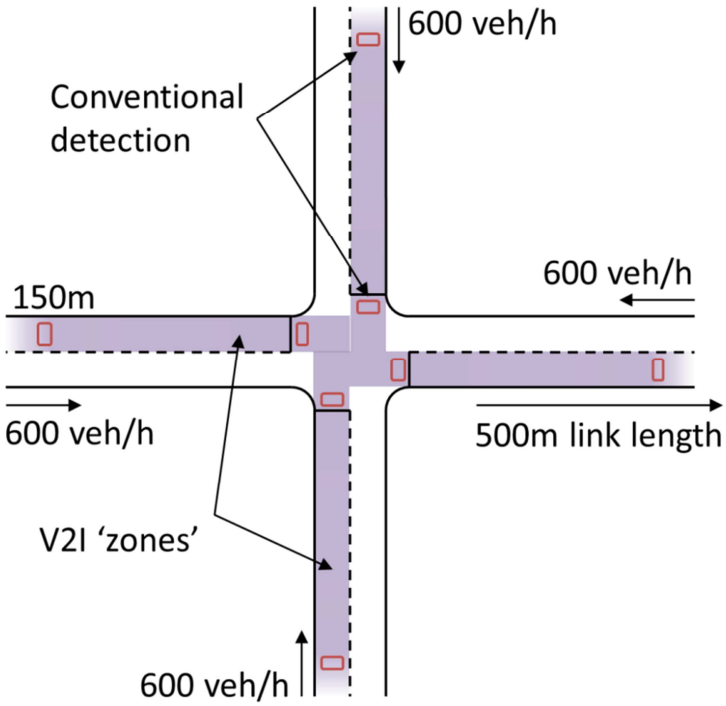

Figure 6.3: The updated simulated layout for the car-following method - conventional and V2I detection overlaid

Relocating the detector is made possible due to the change in model type. The modified Miller method employed in the MOVA representation relies on a critical gap being identified over the 'X' detector to identify the point at which the optimization begins. That method eliminates the requirement to explicitly model the process of the queue discharging but, as a result, prevents any spatial representation of interaction between vehicles.

The relocation of the X detector to beyond the traffic signal stop line requires new infrastructure to be implemented where a site already has a MOVA type layout. In some cases, stop line detection will already be present. This is usually the case where the signals do not automatically revert to green when no demand is present as it ensures a fail-safe for cyclists if the vehicle detection does not register the demand. Where no detection currently exists, the detectors would make use of existing ducting infrastructure (required for the traffic signals themselves) which would significantly reduce the cost of installation compared to detection further upstream.

In the car-following method, the last queuing vehicle is instead tagged in the internal traffic model and monitored to ascertain the point at which the optimization process can be entered. The method of identifying the point at which optimization begins is consistent with that introduced for the hybrid model developed in the previous chapter. The use of the latter method negates the need for a second detector upstream of the stop line to monitor gaps but introduces a requirement to accurately monitor the initial queue discharge rate. Thus, relocating the detector to immediately beyond the stop line satisfies this requirement whilst enabling the physical infrastructure requirements to be kept to two detectors per lane.

6.4.1 Accommodating model error

To implement any traffic model in an on-line optimizer requires that practical issues regarding the response of the model to limited data availability and model error are addressed. The introduction of a car-following model is no different in this regard and some of the techniques used to accommodate practical issues are described in Appendix C.

6.4.2 A car-following model based optimizer

6.4.2.1 Justification

As described in Chapter 3, the Miller method is implemented using a vertical queue model. The algorithm makes assumptions regarding the queue discharge rate and the 'penalty', in terms of overall delay, of a vehicle approaching the stop line being stopped. The queue discharge rate is assumed to be constant from one cycle to the next. Although the discharge rate, selected by a practitioner during observations, implicitly takes into account the vehicle mix, it does so as an average value rather than reflecting the impact that individual vehicle performance can have on queue clearance time from one cycle to the next.

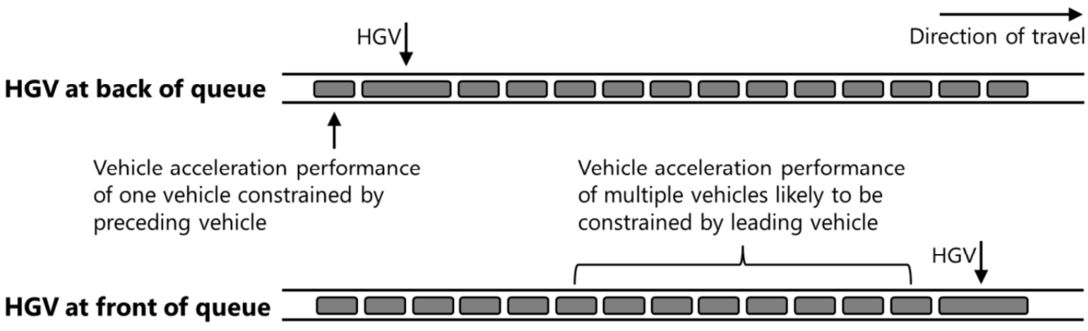

Figure 6.4: Examples of the placement of slow accelerating vehicles in a queue.

As an example, for a vehicle mix with an otherwise low HGV proportion, the discharge rate may be selected at a relatively high value. However, in some cycles the queue may contain at least one HGV with substantially inferior acceleration performance compared to other vehicles. In that situation, the queue clearance time could be significantly underestimated, resulting in the optimization process being entered prematurely and sub-optimal decisions being made. The degree to which the heterogeneous nature of the vehicle mix will impact the performance of the optimizer is also likely to depend on the position within the queue of the vehicle with inferior performance characteristics.

If a vehicle with a slow acceleration profile is leading the discharging queue, all following vehicles will be constrained to the performance of the leading vehicle (Figure 6.4). The 'stop penalty', according to the modified Miller algorithm could therefore effectively be far greater depending on the number of vehicles in the queue and their acceleration characteristics. It is clear that the dynamics of interaction between vehicles results in greater complexity than can be represented by a simple vertical queue model, hence the proposal for a car-following model. The Miller algorithm used in the MOVA representation cannot, using average values for parameters, fully reflect the resulting variability in discharge rate or impact of stops. Consequently, a new optimization method has been developed to accommodate a car-following model.

6.4.2.2 Optimization principles

The developed optimizer operates on the same principle as the Miller method, principally that a rolling horizon scheme is employed, updated each time step with the most recent data available. As in the Miller method, the benefit of extending a green signal is compared to the additional delay caused to waiting

vehicles. Retaining the principle of the Miller method allows for comparison of the vertical queue and car-following traffic models.

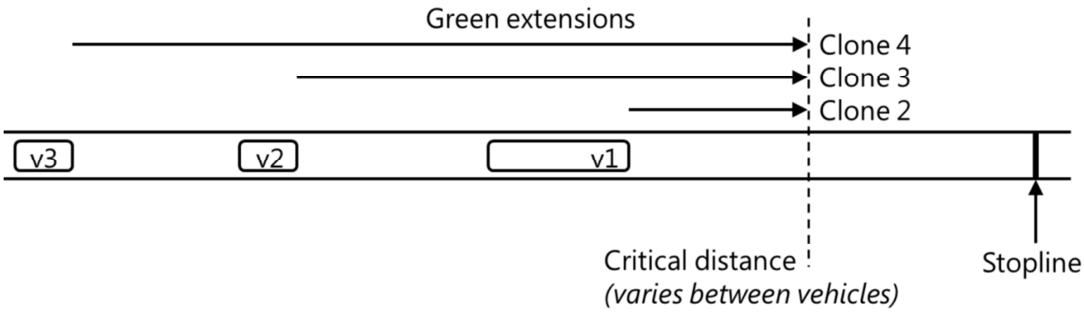

Figure 6.6: Initial green extension time for each signal controller clone.

However, rather than evaluating extensions of *h*, 2*h*, 3*h* etc time steps, the green signal is instead extended until the next vehicle reaches the 'critical distance', the point where it is committed to crossing the stop line even if the green signal were terminated (Figure 6.6). To enable the total delay experienced at the junction to be evaluated for different green extensions, this process is repeated for each vehicle on the modelled links currently at green (i.e. every vehicle that has been detected is tested).

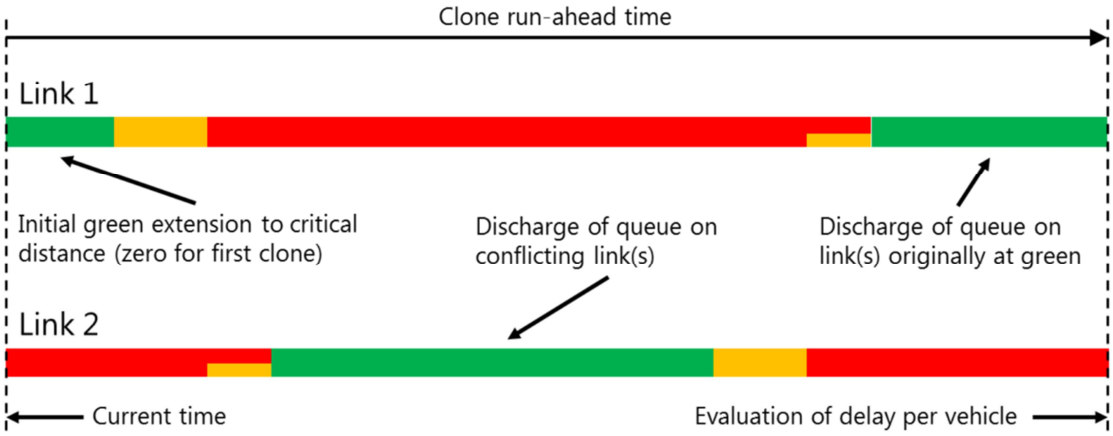

Figure 6.5: Example of one-cycle horizon.

In practise, to evaluate the delay resulting from different green extensions requires that a clone of the signal controller model is created for each vehicle on the link. For each clone, the model of the traffic signal junction is run ahead for the initial green extension time required for the identified vehicle on the approach link to reach the point at which it will cross the stop line. At this point the green signal is terminated, queues on conflicting links are discharged and then, finally, any residual queue on the links that were at green when the signal controller was cloned is discharged to complete evaluation over one cycle (Figure 6.5). The final part of the process ensures that any consequential delays to vehicles in the next cycle as a result of a green extension are accounted for.

The total 'run-ahead' time ensures that the delay, incurred by vehicles that are stopped as a result of terminating the green after the initial green extension, is fully accounted for. The delay for each vehicle is recorded by the optimizer as it crosses a point downstream of the stop line (to ensure the delay has been fully captured) and is retrieved at the end of each cloned signal controller run-ahead along with the total number of vehicles to cross the stop line to enable the average delay per vehicle to be calculated.

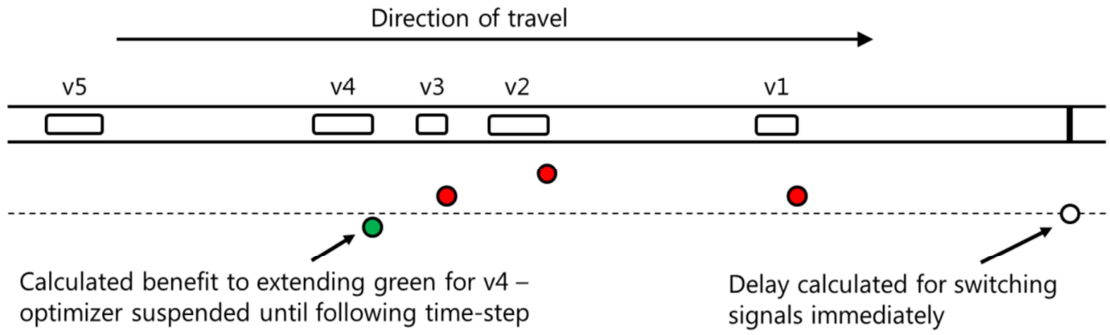

Figure 6.7: Example of an optimizer result for one time-step.

The value of delay is compared to that resulting from the signal switching immediately. If, at any point, it is identified that there is an advantage to

extending the green signal, the optimizer is suspended until the following time step (Figure 6.7).

6.5 Implementing V2I

The introduction of V2I to the car-following model is considered with an assumed 100% coverage. The interaction between leading and following vehicles, represented in the car-following model, makes the implementation of V2I more complex than in a vertical queue model. For example, the headway between modelled vehicles depends on the value of the parameters used in the car-following model. However, there may be cases where, in reality, a vehicle is willing to follow another more closely than the model will allow. Whilst this is not an issue when the model is running in real-time and updating vehicle positions from V2I data, it causes issues when the signal controller model is cloned to run ahead, at which point the car-following model is applied.

Once the car-following model is applied, the model parameters take effect and, for a vehicle following more closely than the parameters will allow, the result is a rapid deceleration, or in some cases a 'jump' backwards, to the point at which the vehicle is following at a headway that satisfies a safe following distance in the car-following model (Figure 6.8). This perturbation will propagate along the link to any vehicles following within a reasonable distance and potentially cause vehicles to discharge at a slower rate than would otherwise be the case.

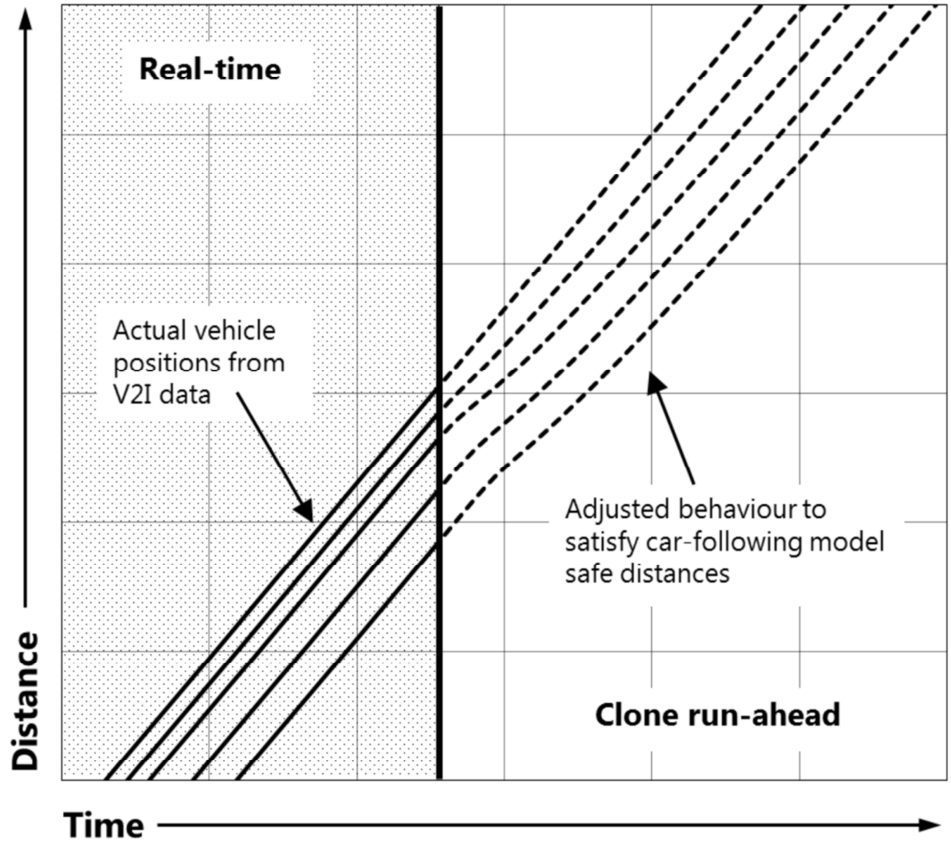

Figure 6.8: Example of the effect of switching from real-time V2I data to car-following model calculated vehicle positions in a cloned signal controller model.

To mitigate this effect, the headway between vehicles is monitored in the real-time instance of the model and, where it falls below the safe following distance dictated by the car-following model, the parameters of the model are modified to reduce the safe distance. To modify the parameters, an iterative process is undertaken whereby the relevant parameters are adjusted until the car-following derived speed (and thus safe distance) matches the actual speed taken from the V2I data, or until the parameter values reach a pre-defined limit of acceptability.

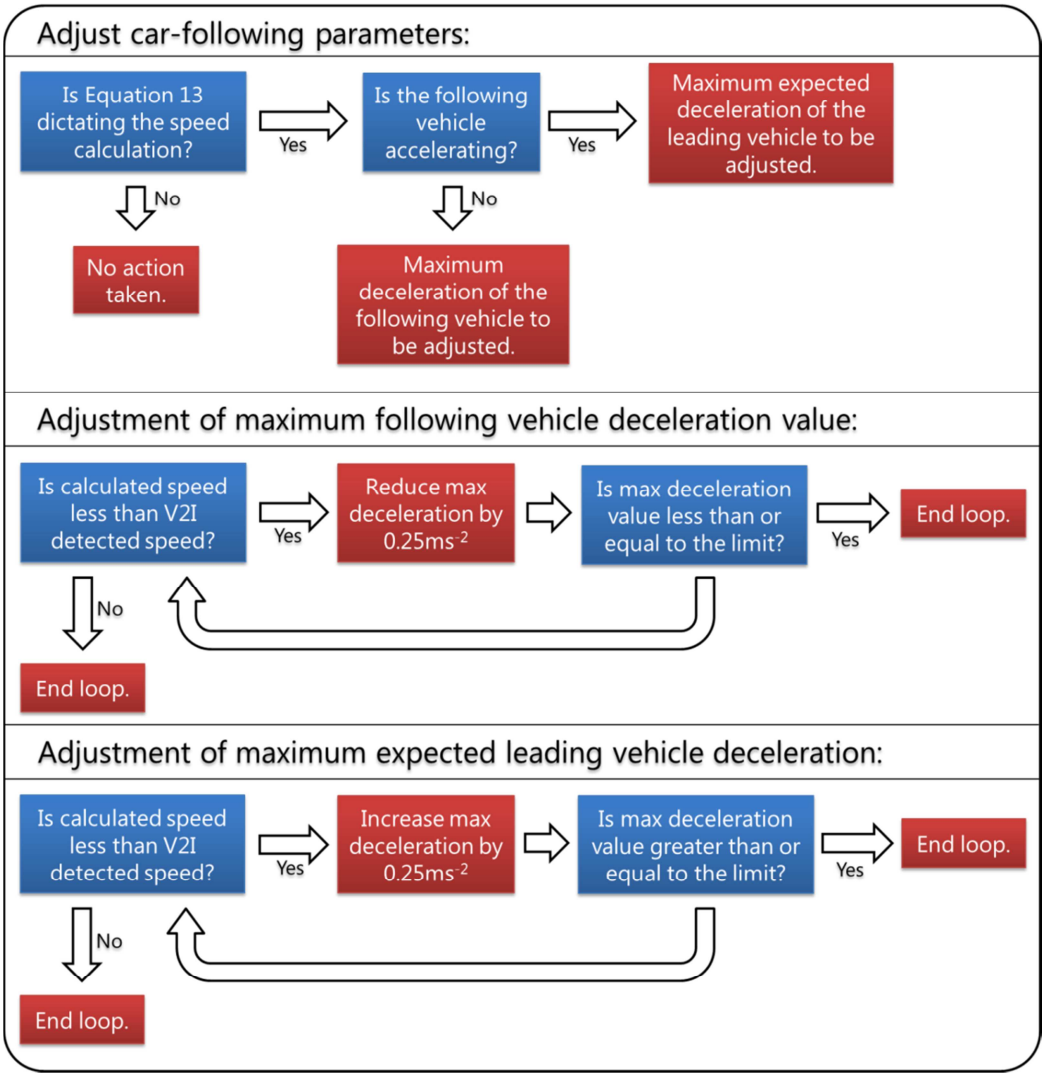

Figure 6.9: Schematic of the iterative process followed to adjust car-following parameters based on actual vehicle speed and position data from V2I.

First, it is determined whether Equation 6.1 or Equation 6.2 is dictating the following vehicle speed. If Equation 6.2 is dictating the vehicle speed then the speed is being constrained by the leading vehicle. The nature of the constraint is determined by whether the following vehicle is decelerating or accelerating. If the following vehicle is decelerating then the maximum deceleration that the driver of the vehicle wishes to undertake b_n is likely to be the dominant factor in the derivation of vehicle speed. If, however, the following vehicle is accelerating then the maximum expected deceleration $\hat{b}$ of the leading vehicle

is likely to become the dominant factor. If parameter b_n is dominant then the value of b_n is reduced in steps of 0.25ms^{-2} down to a limit of -4.5ms^{-2}. If parameter $\hat{b}$ is dominant then the value of $\hat{b}$ is increased in 0.25ms^{-2} steps up to a limit of -1.75ms^{-2}. This process is described in Figure 6.9.

The parameter value limits chosen here are approximately 25% above and below the respective -2.5ms^{-2} and -3.5ms^{-2} maximum deceleration values assigned to the vehicle classes in Table 6.1. The 25% limit was chosen as a reasonable performance deviation from the original parameter values but could be subject to additional investigation if this method were to be developed further. It is perhaps the case that the limit values could be absolute values rather than being relative to the validated vehicle class parameter values.

Finally, if the vehicle speed is very low then the distance between vehicles becomes the dominant factor. In this case, the position of the leading and following vehicle is known and so the modelled effective vehicle length (that is, the sum of vehicle length and safety standstill distance between vehicles) can be adjusted to the actual effective vehicle length.

Once the parameters have been modified, the new values are retained until the vehicle leaves the link unless further interaction with a preceding vehicle triggers subsequent parameter changes.

6.6 Results

6.6.1 Introduction

The results for implementing a car-following model into the traffic signal optimizer, using conventional detection with the single detector classification algorithm and V2I detection with classification, are detailed in this chapter. The

results are compared to the MOVA representation described in Chapter 3 and the hybrid model presented in Chapter 4 for:

- Different proportions of HGV with fixed demand of 600 veh/h (Scenario 1); and
- Different demands with a fixed proportion of 5% HGV (Scenario 2).

Results are also presented for various V2I detection ranges to demonstrate the effect of an increased range.

As in Chapter 4, the term HGV refers to the previously defined combination of rigid and articulated HGVs. However, some specific results for rigid and articulated HGVs are also reported.

Initial testing showed that applying a stop penalty value to the car-following model does not provide a benefit in terms of the PI described in Chapter 3. The reason for this is likely to be that the car-following model method already includes the full delay associated with a vehicle stop and therefore the addition of a stop penalty would be solely for reasons of policy. If more than one junction was to be considered in an optimizer then the introduction of a stop penalty may have more of an influence in determining the offset between the junctions.

6.6.2 Implementation with conventional detection

Results are first presented for the implementation of a car-following model using conventional detection with the single detector classification algorithm presented in Chapter 4. The results are compared to the MOVA representation, described in Chapter 3 to provide a consistent evaluation.

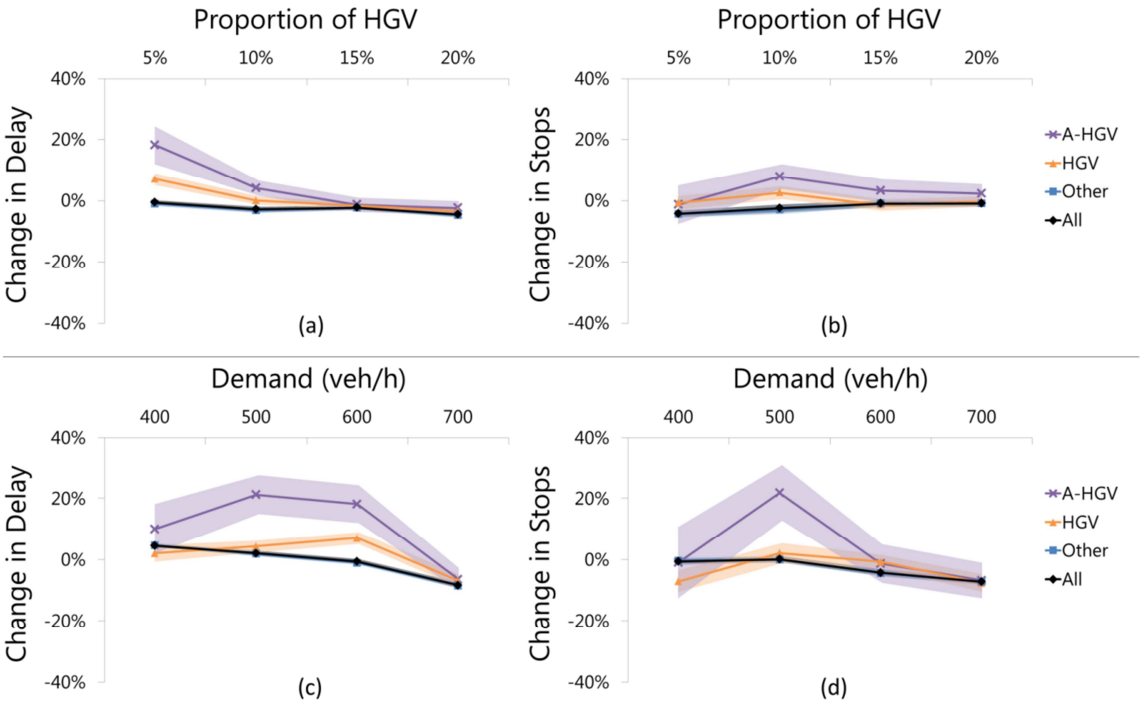

Figure 6.10: Change in delay and stops by vehicle class, using a car-following model with single detector classification, compared to the MOVA representation for Scenario 1 (a, b) and Scenario 2 (c, d).

Figure 6.10 demonstrates the change in delay for each vehicle class compared to the MOVA representation when using a car-following model with conventional detection. The results show that, for Scenario 1 (a, b), whilst there is an overall benefit for delay, there is actually an increase in delay to HGVs at lower HGV proportions. For vehicle stops there are mixed results for HGVs but the general trend is a benefit to all vehicles that decreases as the proportion of HGVs increases. The results for Scenario 2 (c, d) suggest that the car-following model only begins to provide a benefit at higher demand and that, in general any benefit is relatively minor.

The negative impact of the car-following model on delay at lower levels of demand is possibly a result of the increased uncertainty of vehicle positions between the upstream detector and stop line when compared to the MOVA representation. Moving the downstream detector beyond the stop line provides

a benefit when monitoring the rate of queue discharge during a green signal but it increases the reliance on the internal traffic model to accurately estimate the position of vehicles upstream of the stop line. It could also be the case that, at low demand, the vertical queue model provides a good approximation of traffic flow that then deteriorates as the conditions approach saturation.

Another contributing factor, and also the probable reason for the mixed results for HGVs, is the single detector classification algorithm. In Chapter 5, the single detector classification algorithm is implemented into the MOVA representation to allow stop penalty values to be assigned based on vehicle length (as a proxy for class). In that case, the accurate classification of vehicles is most important during the Miller optimization process. During that process, the speed of vehicles over the upstream detector is assumed to be approximately free-flow and approximately constant. The results of Chapter 4 show that the single detector classification performs better in those conditions than when the vehicle speed is reduced, or is fluctuating over the detector (Figure 4.17).

However, when applying the classification algorithm to the car-following model, the performance of the model relies on accurate data during all conditions as it affects interaction between vehicles during initial queue discharge. The performance of the classification algorithm at lower vehicle speeds therefore has a greater impact on the overall performance of the car-following model than on the MOVA representation.

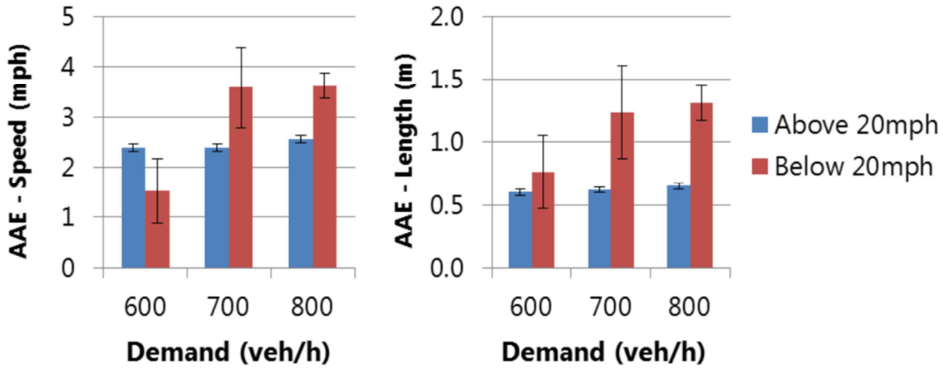

Figure 6.11: Average absolute error for speed and length estimation of single detector classification algorithm.

At higher demand, the relocation of the downstream detector to beyond the stop line is likely to provide a greater benefit to the car-following model as the accurate estimation of the rate of queue discharge becomes an increasingly dominant factor in the control strategy performance.

6.6.3 Implementation with V2I detection

The next results are presented for the implementation of a car-following model using V2I detection. Two comparisons are shown in this section to provide a demonstration of the relative performance of the car-following based optimizer:

- A comparison to the hybrid model developed in Chapter 5 (using V2I but without vehicle classification); and
- A comparison to the hybrid model developed in Chapter 5 with V2I based detection and vehicle classification.

Figure 6.12 demonstrates a much more consistent performance benefit for the car-following model, over and above the hybrid model with V2I detection alone. The increased consistency compared to the use of conventional detection supports the hypothesis provided in the previous section regarding the impact of the single detector classification algorithm performance on the car-following model.

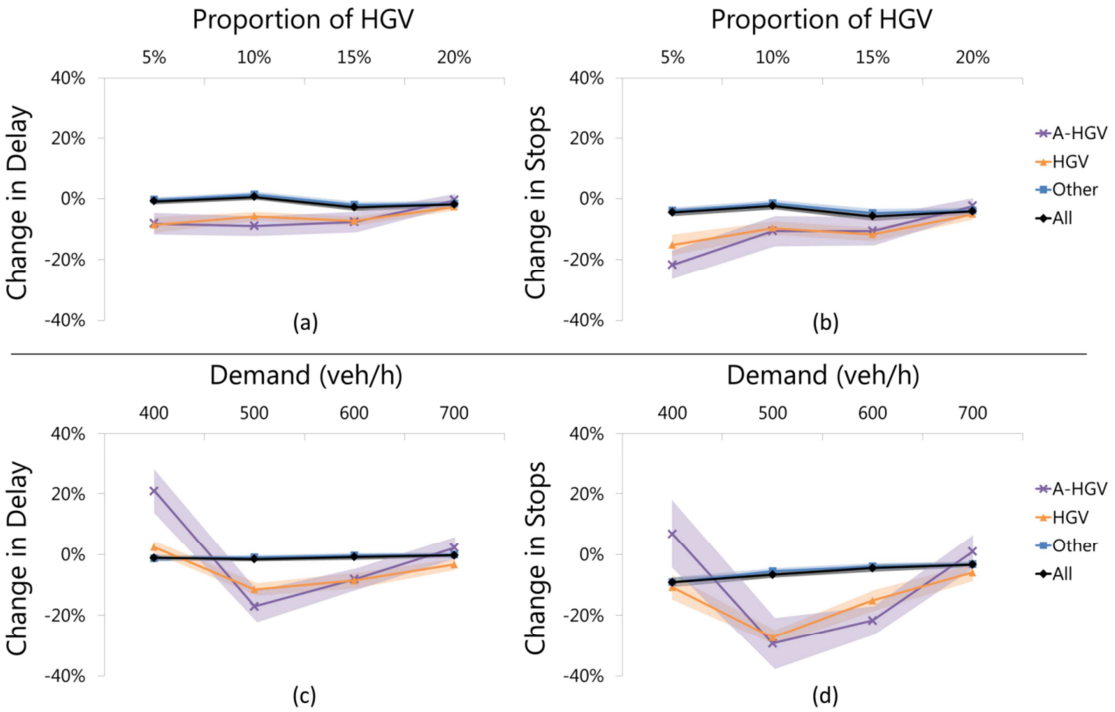

Figure 6.12: Change in delay and stops by vehicle class, using a car-following model with V2I detection, compared to the hybrid model with V2I detection for Scenario 1 (a, b) and Scenario 2 (c, d).

As expected, the benefit to HGVs decreases as the proportion of HGVs increases. However, for Scenario 2 (c, d), although there is an expected trend of reducing benefit to HGVs as the demand increases, there is also an anomaly in the lowest demand case. The increase in articulated HGV delay and stops compared to the hybrid model does not follow the general trend of the results. The full results (Appendix B) show that the hybrid model (without classification) tends to provide the most benefit to non-HGVs. However, in the lowest demand case the benefit of the hybrid model to HGVs is greater than for non-HGVs and greater than the benefit provided by the car-following model. The benefit of the car-following model is actually more equitable across vehicle types in this case.

Aside from the anomaly for HGVs in the lowest demand case, the results for all vehicles provide a clearer demonstration of the performance benefit that the car-following model provides. The graphs for vehicle stops show that there is a consistent benefit for all vehicles of approximately 5% when compared to the hybrid model, increasing at lower demand, whilst the results for delay remain largely consistent with the hybrid model.

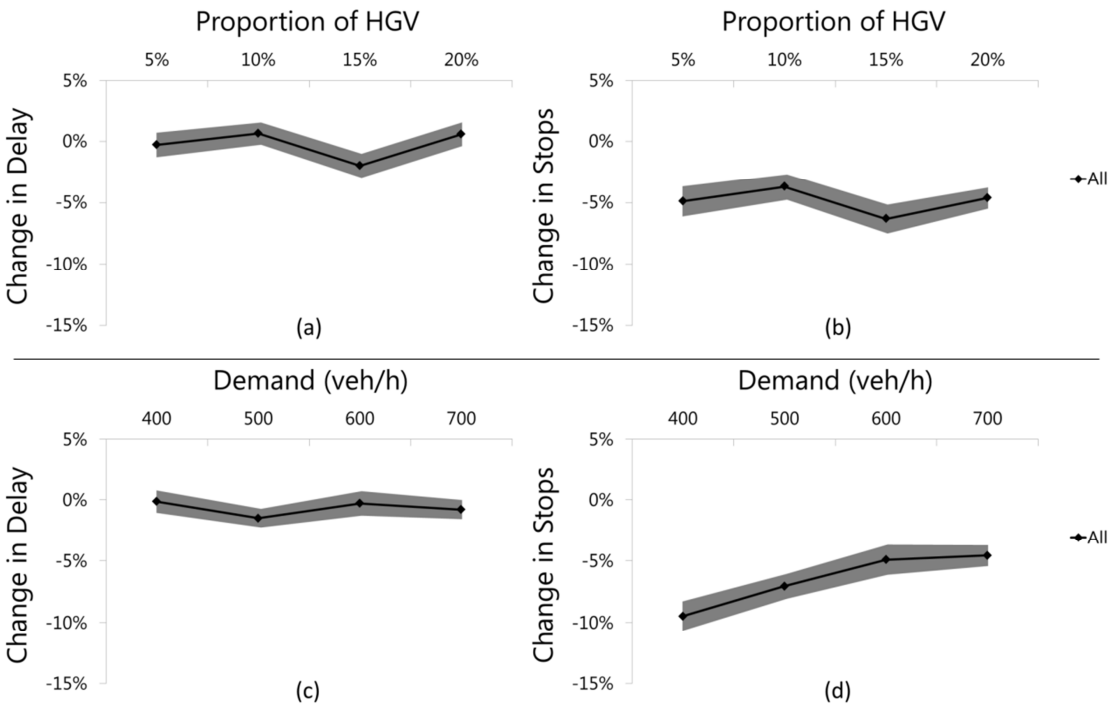

Figure 6.13: Change in delay and stops for all vehicles when using a car-following model with V2I detection compared to the hybrid model with V2I detection (including classification).

To compare the relative performance of the developed hybrid model and the car-following based optimizer, both with V2I detection and vehicle classification, the results are presented in Figure 6.13. The results show that, in terms of delay, the car-following model based optimizer performs similarly to the hybrid model with the differences generally not statistically significant. However, when considering vehicle stops, the car-following model provides a clear benefit compared to the hybrid model. There is a consistent overall benefit, regardless of HGV proportion, of approximately 5% and a more significant benefit at lower

demand of almost 10%. This suggests that the combination of a more comprehensive representation of interaction between vehicles provided by the car-following model and the modified optimizer does indeed provide an overall benefit in performance.

6.6.4 The effect of V2I range

Finally, the Simulated Environment has been modified to enable the V2I range to be varied to understand the effect on the performance of the optimizer. Figure 6.14 shows the normalized PI for various demand cases indicating the optimal range. Normalization has been achieved by taking the maximum and minimum values for each data series and scaling each series between those points.

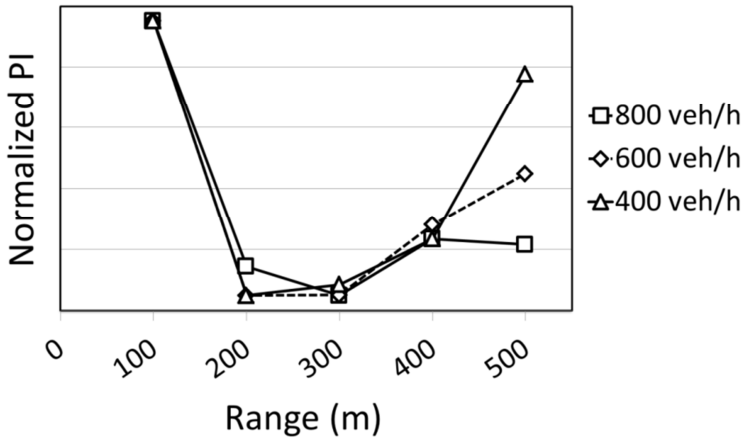

Figure 6.14: Normalized PI for various demand cases for an increasing V2I detection range.

The range of between 200-300 metres equates to a travel time of approximately 15-22 seconds. It could be assumed that an ever increasing detection range would provide diminishing benefits but benefits nonetheless. However, these results suggest that there is no benefit to providing a range in excess of approximately 22 seconds travel time from the junction (based on the average free-flow speed of 13.5ms^{-1}), a finding that is consistent with the 25 seconds derived by Robertson and Bretherton (1974).

The reason for this is likely to be a result of the uncertainty associated with vehicle travel time to the stop line. The V2I data for vehicle position and speed is available to the optimizer continuously and, as is the case for the Miller algorithm, the effect of extending the green signal is considered up to the maximum range of the detection on links currently at green. However, as the distance from the stop line increases, the benefit associated with providing an increased look-ahead to the optimizer eventually becomes outweighed by the uncertainty in vehicle travel time. Thus, there is a point beyond which the increased range is no longer useful.

In Figure 6.14 the optimiser performance worsens significantly beyond approximately 400 metres in the low demand case but less so for the high demand scenario (800 veh/h). As the junction approaches saturation, vehicles on the approach travel more closely together and so there is less variation in speed. Consequently, there is a greater certainty of travel time further from the junction in the 800 veh/h scenario than in the 400 veh/h scenario, hence the less significant worsening in performance.

In this study, the same weighting has been applied to all green extensions during the optimization. However, an increasing penalty could be applied to longer green extensions to account for the increased uncertainty in the estimated arrival time at the junction. This may reduce the extent to which overall performance worsens at longer detection ranges. Full results are available in Appendix B.

6.7 Conclusions

In this chapter a car-following model has been implemented to replace the vertical queue method employed in the MOVA representation described in Chapter 3. The implementation of the car-following model required

development of a new optimizer and a modification to the Simulated Environment to relocate the downstream conventional detector beyond the signal stop line.

The results showed that implementation of the car-following model using conventional detection provides a general overall benefit for delay and stops when compared to the MOVA representation. However, the MOVA representation outperforms the car-following model at lower demand and the results for HGVs are mixed. It is likely that the mixed results are at least partly a consequence of vehicle classification errors when using the single detector classification algorithm presented in Chapter 4. As discussed in section 6.6.2, the way that the classification algorithm is applied to the MOVA representation reduces the impact of classification errors compared to the car-following model.

The addition of V2I detection to the car-following model introduces greater consistency. The performance of the optimizer is first compared to the hybrid model alone, developed in Chapter 5 (both with 100% V2I coverage). The hybrid model makes use of the Miller method and the car-following model uses the optimizer developed in this chapter. The comparison between methods demonstrates that the car-following model provides a benefit for stops, particularly at low levels of demand, of up to almost 10% with little or no impact on delay. The benefit for HGVs is more pronounced, with the exception of the lowest demand case, with a reduction in stops of approximately 30% in one case.

The subsequent comparison of the car-following based optimizer with the hybrid model, this time with V2I-based vehicle classification included, is perhaps the most representative comparison of the two techniques. The results show a consistent and statistically significant benefit provided by the car-following based method in terms of stops of almost 10% in the lowest demand case. The

reduction in stops is over and above the benefit (compared to the MOVA representation) already demonstrated in Chapter 5 by the hybrid model with V2I-based vehicle classification.

The introduction of V2I to the car-following model demonstrated the advantage of modelling the interaction between vehicles more comprehensively but also highlighted the importance of providing accurate data to the traffic model. The car-following model has been shown to provide important benefits compared to the hybrid model but provided mixed results when implemented with conventional detection.

The effect of V2I range on optimizer performance was also tested in this chapter. Normalized PI results demonstrated that the optimal V2I detection range for the Simulated Environment is between 200-300 metres from the signal stop line. At the average free-flow speed, this equates to approximately 15-22 seconds travel time, a finding consistent with Robertson and Bretherton (1974). It is possible that at a junction with more 'friction' on the approach to a junction, such as on-street parking or bus stops, the increased travel time uncertainty would reduce the useful V2I range further. The findings, primarily that there is an optimal range and that the benefit does not simply continue to increase with range, has potential implications on infrastructure requirements. For example, the range of current DSRC communications devices is approximately 1000 metres and so it may be possible to deploy fewer devices across a network.

Chapter 7
Conclusions

This chapter draws together the conclusions from each of the previous chapters and provides an overall summary of the contributions of this research in relation to the objectives set at the beginning of the study. There then follows a discussion of the findings and the possible next steps to further develop the research undertaken in this thesis.

7.1 Main conclusions and contributions

The main objectives of this study were set out in Chapter 1 and then refined in Chapter 2. A Simulated Environment is constructed and described in Chapter 3. The research described in each of the following chapters (Chapter 4, 5 and 6) contributes to achieving the objectives of the study. In this section the conclusions from those chapters are brought together and the contributions discussed. The objectives of this thesis were:

1) To investigate whether a single detector vehicle classification method can be developed, using only existing infrastructure, which is capable of explicitly classifying HGVs. Then, to assess whether the accuracy of such a method is adequate to provide a performance benefit in a traffic signal optimizer in terms of minimizing delay and stops;
2) To develop a control method capable of incorporating a hybrid detection traffic model that can simultaneously make use of conventional and connected vehicle (i.e. V2I) data in order to test whether benefits of V2I can be realised sooner than the, approximately, 30-50% penetration rate described in current literature;

3) Extend the hybrid model to utilise V2I-based vehicle classification data and investigate the effect on optimizer performance;
4) Apply a microscopic traffic model to an on-line responsive traffic signal optimizer to identify whether the more detailed representation of vehicle interaction can provide a benefit to the performance of an optimizer in terms of its ability to minimize stops; and finally
5) Assess the impact of V2I range on the performance of the developed control method.

A summary table is provided in Table 7.1 to provide clarity on how various sections of the study relate to each of the objectives. The conclusions related to each objective are then discussed in greater detail below.

Table 7.1: Objective summary table.

	Relevant sections of study
Objective 1	Two single detector classification algorithms were developed in Chapter 4 with testing of the preferred algorithm in Chapter 5 (modified MOVA representation) and Chapter 6 (car-following model based optimizer).
Objective 2	The development of a hybrid model that is capable of using conventional and V2I-based detection data was developed in Chapter 5.
Objective 3	An extension of the hybrid model to incorporate V2I-based vehicle classification data was developed later in Chapter 5.
Objective 4	A car-following model based optimization method was developed in Chapter 6.
Objective 5	The impact of range on the performance of the car-following model based optimization method was investigated later in Chapter 6.

The **first objective** was initially addressed in Chapter 4 with the development of two algorithms to estimate the speed of vehicles crossing a detector and to classify them by length. The preferred algorithm was then applied to a modified

existing traffic signal optimizer in Chapter 5 and a developed optimizer in Chapter 6 with the findings concluding the aims of the objective.

The second single detector classification algorithm described in Chapter 4 incorporates exponential smoothing and knowledge of the queue length from downstream traffic signals. It was found to consistently outperform the first algorithm that operated with no knowledge of the traffic signal control.

It was concluded that neither algorithm would realistically be a replacement for vehicle classification at a dual loop detector site. However, the preferred algorithm could prove useful to traffic authorities for basic monitoring of HGV proportions on the urban road network at traffic signal sites where single detectors are employed. The 'instant' nature of the speed estimation would also provide a quick response to 'abnormal' drops in vehicle speed that could be used for incident detection purposes.

The primary objective of developing a single detector classification algorithm was to provide accurate vehicle type information to a traffic signal optimizer using existing infrastructure. The MOVA representation method of signal optimization, described in Chapter 3, requires that any queue that exists at the start of a green signal is discharged before optimization begins. This suits the developed algorithm because the speed and length estimation error is lower when vehicles are travelling closer to a free-flow speed and hence the variation in speed between consecutive vehicles is less.

It was demonstrated in Chapter 5 that incorporating the single detector classification algorithm into a modified MOVA representation could indeed provide a benefit to HGVs in terms of stops and delay without adversely affecting other vehicles. However, the results showed that the benefit is not large.

Even so, application of the single detector classification algorithm to the MOVA representation reduced delay and stops for the largest articulated HGVs by 5-10% at a 5% proportion of HGVs with negligible impact on other vehicles. There are over 1000 MOVA controlled junctions in the UK that could potentially be retrofitted with the preferred algorithm to reduce HGV stops and, by extension, reduce associated emissions.

Conversely, the initial results of Chapter 6 showed that the algorithm does not perform well enough for providing vehicle type data into the car-following model. This is because, unlike the MOVA representation, the car-following model requires accurate classification data regardless of vehicle speed over the detector.

The development of a single detector classification algorithm that can be successfully deployed on existing infrastructure is an important contribution and satisfies the first objective of this study. The finding that the algorithm did not perform well when applied to the car-following model in Chapter 6 is also an important contribution as it demonstrates the limitations of using existing infrastructure. It also confirms that the required accuracy of the classification algorithm is, to some extent, dependent on the nature of the application.

The **second objective**, to develop a control method capable of incorporating a hybrid detection traffic model, was addressed in the second part of Chapter 5. Initially, V2I data was implemented with a 100% penetration rate of equipped vehicles to understand the maximum benefit that could be gained by applying V2I to an existing optimization method. The introduction of V2I data resulted in benefit, in terms of stops and delay, of between 10-15% that supported the (10-16%) findings of Feng et al. (2015). The 10-15% reduction is achieved at 100% V2I coverage but, as the review of literature in Chapter 2 showed, the previously developed V2I based optimization algorithms require approximately 30-50% V2I

penetration rate before providing a benefit compared to conventional detection. The significant contribution of this research is to show that, by developing a hybrid detection traffic model, the potential benefit of V2I detection can be realised immediately as the technology begins to penetrate the vehicle fleet. The results suggest that more than 50% of the maximum benefit of V2I data could be achieved with just 20% of the vehicle fleet enabled and that the benefit achieved at 80% penetration rate is similar to the maximum. The hybrid model enables conventional detection to be used alongside V2I data to enhance the performance of an existing optimization strategy and, in doing so, achieves the second objective of this study.

Later in Chapter 5, the **third objective** was addressed by introducing V2I based vehicle classification to the hybrid detection traffic model. The results clearly demonstrated the benefit of using vehicle classification in an optimization method. The initial results from the first objective, applying the single detector classification algorithm to the MOVA representation, were mixed with modest benefits in terms of stops and delay for all HGVs combined of less than 5%. Larger benefits were demonstrated for articulated HGVs of up to between 5-10% but, as the proportion of HGV increased, the benefit quickly reduced as a result of an increase in error from the classification algorithm.

The use of V2I based vehicle classification resulted in larger and more consistent benefits in terms of stops and delay for all HGVs of up to 15% in Scenario 2. Articulated HGVs experienced even greater benefit at lower demand when compared to V2I alone.

The increase in benefit compared to applying vehicle classification with the single detector classification algorithm demonstrates the importance of accurate vehicle classification. It also shows that the ability to track vehicles on the approach to a junction is vital to eliminating the estimation error of

conventional detection systems. For example, when vehicle position and speed is estimated, it can lead to a green signal being extended for a vehicle but, due to the vehicle travelling more slowly than estimated, it fails to cross the stop line. This situation increases delay to opposing traffic whilst not delivering the planned delay saving benefit for the stopped vehicle.

This contribution has important implications for minimizing vehicle emissions at traffic signal junctions. The benefit, particularly for larger HGVs, is shown to be particularly significant with a reduction in stops of more than 25% in one case compared to the MOVA representation without classification. Reducing the number of stops will reduce vehicle emissions and, in doing so, could provide a valuable tool for traffic managers in helping to achieve air quality objectives.

Finally in Chapter 5, the modified MOVA representation was tested with opposed right turning movements that periodically blocked vehicles travelling ahead at the junction. The results showed the modified MOVA representation to be robust to such issues, providing a significant benefit compared to conventional detection, particularly as the proportion of right turning vehicles increases.

In Chapter 6, the internal traffic model of the optimizer was replaced with a car-following model, in this case using the Gipps (1981) model. The aim of using a more sophisticated car-following model, as stated in the **fourth objective**, was to better evaluate the impact of interaction between vehicles on the number of stops. To facilitate the use of the more sophisticated traffic model, a new optimizer was developed. Initially the car-following model was tested using conventional detection. Various modifications were made to the control algorithm to enable the traffic model to be adjusted when discrepancies between the modelled representation and actual traffic conditions were detected.

The results showed that, with conventional detection, the use of a car-following model provided mixed results. The overall performance of the algorithm, compared to the MOVA representation, was improved with higher demand but was less effective with low demand. This is likely due to the relocation of the first upstream detector (50 metres from the stop line in the MOVA representation) to immediately downstream of the stop line. The relocation enabled vehicles to be counted at the beginning of the green to increase the accuracy of the modelled queue discharge rate but reduced the information available during free-flow conditions.

The benefit to HGVs was also mixed with some reduction in stops for HGVs but only at low demand and with a low proportion of HGVs. As already discussed, this highlights the increase in error of the single detector classification algorithm at higher proportions of HGV and when speeds of consecutive vehicles over the detector are increasingly varied (i.e. with higher demand).

Introducing V2I with classification clearly demonstrated the benefit of the car-following model in terms of reduction of stops. The change in delay for all vehicles when using the car-following model compared to the MOVA representation (both methods incorporating V2I based vehicle classification) is negligible at lower proportions of HGV but, as the proportion of HGV increases, the delay is reduced. Vehicle stops are reduced for all proportions of HGV by up to 5% and, in the lower demand case, by up to 10%.

The work in Chapter 6 satisfies the fourth objective of this study and its contribution to current literature is important in establishing that the use of a more sophisticated traffic model does provide a useful performance benefit to a signal optimizer and is feasible to implement in terms of computational power. Issues regarding the practicality of implementation are discussed later.

In the final section of Chapter 6, the effect of V2I detection range on the optimizer performance was investigated. The results showed that, for the Simulated Environment used in this study, the optimum detection range was between 200-300 metres from the junction. This equates to a travel time of approximately 15-22 seconds depending on the approach speed. The finding was consistent with Robertson and Bretherton (1974), that there is no benefit to providing information to the optimizer more than 25 seconds in advance. This element of work satisfies the **fifth and final objective** of the thesis.

7.2 Discussion

The outcomes of this study have shown that explicitly classifying vehicles in a traffic signal optimization strategy can lead to significant benefits in stops and delay for those vehicles whilst having a negligible impact on other vehicles. If a more sophisticated internal traffic model is incorporated then those benefits, particularly in terms of stops, are increased further.

7.2.1 Practical issues

Chapter 6 described the process of incorporating a more sophisticated internal traffic model into an optimizer. As discussed in the previous section, there are certainly some practical issues to overcome if this method is to be developed further. Firstly, there is an issue associated with how capable such models are of accommodating 'unexpected' driver behaviour. The second issue related to the car-following model is how easy it would be for a practitioner to understand and validate model parameters.

For the first issue, it was shown in Chapter 6 that methods can be employed to enable the model to be adjusted without causing significant perturbations. However, those methods (involving adjusting model parameters) can only be

effective in a situation where a significant proportion of the vehicle fleet is V2I enabled. The study suggests that the more sophisticated traffic model may only be feasible to introduce at higher levels of V2I penetration unless video detection, discussed later in this section, could be used as a proxy for V2I.

The second issue is possibly more difficult to address. An important factor that affects the performance of any traffic model is the accuracy of the model parameter values. To maximise the performance of an optimizer, the traffic model must be calibrated to accurately reflect real-world conditions. Existing control strategies such as SCOOT and MOVA incorporate many parameters that must be understood and adjusted by a practitioner based on site specific conditions.

The introduction of V2I detection will greatly reduce the number of parameters to calibrate but will not eliminate the requirement completely. For example, the car-following model parameters for each vehicle type (i.e. acceleration, deceleration) must still be validated with the most suitable values for each vehicle class. Those parameter values are likely to also be dependent to a certain extent on site characteristics such as the approach gradient and external factors such as weather.

One method would be to build a database, through simulation, of various vehicle types and gradients that would enable the traffic model to assign appropriate model parameters based on the junction topology entered by the practitioner. This is something that is already undertaken by micro-simulation software manufacturers (PTV-Vision, 2011), albeit with limited documentation as to the evidence behind the modification of acceleration parameters to accommodate gradient. However, it would be necessary to ensure that such a database was managed carefully and values updated regularly.

Ultimately, to provide the most accurate vehicle specific (rather than class specific) data, the performance of each vehicle in terms of acceleration and deceleration could be recorded prior to arrival at a junction. If that were possible, profiles for the acceleration and deceleration performance could be constructed and refined as vehicles travel through the network. The relevant parameter values could then be applied to each vehicle as it is inserted into the traffic model on the approach to a junction. Using this method would help to take into account driver behaviour as well as vehicle performance. It would also reduce error associated with larger vehicles where HGV performance, for example, may differ significantly depending on whether a vehicle is fully laden or an 'empty runner'. However, the viability of this method would depend largely on whether any data privacy issues surrounding the storage of vehicle specific parameter values could be overcome.

The pursuit of increasingly accurate vehicle performance data will eventually result in diminishing returns in terms of the improvement in performance of the optimizer. However, even if vehicle performance data was perfect, the variability and uncertainty of driver behaviour would still affect the accuracy of the traffic model. Consequently, there is a limit to the usefulness of ever increasing vehicle performance data accuracy whilst vehicles are human operated.

Fully autonomous vehicles, and maybe those with a high level of autonomy, will reduce or remove the influence of driver behaviour on the accuracy of the traffic model. Hypothetically, if the vehicle performance and driver behaviour error could be, in effect, eliminated then the remaining error would reside with the accuracy of the traffic model itself and its ability to represent traffic conditions. In the scenario of full autonomy it may be possible to change the way that traffic signals are optimized to take advantage of trajectory manipulation and/or remove the requirement for physical traffic signals altogether. There are

practicalities to be overcome in this scenario such as how cyclists and unequipped vehicles (e.g. classic cars) would be accounted for but initially it could be introduced in an area restricted to fully autonomous vehicles.

In this study, vehicle stops have been used as a proxy for emissions. The use of a sophisticated emissions model would, in theory, enhance the performance of an optimizer in achieving the explicit objective of minimising emissions. However, that performance enhancement would not be realised unless vehicle specific data that affects emissions (i.e. knowledge of engine type, age, current vehicle weight and driver behaviour) was available.

7.2.2 V2I data quality and availability

It has been assumed throughout the thesis that V2I data will become available and that it will penetrate the vehicle fleet to a point where there is effectively 100% coverage. It may never be the case that 100% of vehicles are V2I equipped but it could be that information regarding the location, speed etc of non-equipped vehicles is derived from those that are equipped. However, it is currently uncertain whether V2I data will indeed become widely available and, if it does, how quickly that transition will occur.

The business case for developing a traffic signal control strategy that incorporates V2I data depends on the availability of the data. If vehicle manufacturers choose not to provide the data then, unless regulation is introduced to force manufacturers to supply it, it is unlikely that a traffic signal control strategy that can make use of it will be developed into a commercial solution. However, there are potential alternative solutions that would enable data to be provided to an optimizer that mimics the data provided by V2I and would therefore allow the development of a V2I capable traffic signal optimizer.

Radar detection is widely used in the traffic signal control industry but its application is generally limited to providing outputs that would be expected from conventional detection. The aim of most above ground detection has been to reduce the requirement for physical works to install ducting and inductive loop detection. However, radar detection is capable of tracking vehicles and providing real-time speed and position data so could therefore potentially be used to mimic V2I data.

Video detection is also capable of tracking vehicles and, with recent advances in machine learning, can also provide classification. The continual development of processing power and camera technology has made the use of more sophisticated above-ground detection technically viable and more financially competitive with conventional detection. All 'above-ground' detection solutions are prone to issues with occlusion and some development is required to determine how a traffic model would accommodate such issues. The range is also limited by the capability (and expense) of the above-ground detection and is likely to be less than that expected from V2I detection. However, an advantage of using this technique is that all vehicles can be tracked and there is therefore effectively 100% penetration rate of V2I –type data (i.e. position, speed, type etc) within the range of the detection. Trials are currently underway to test the capability of radar and video detection to mimic V2I detection (NIC, 2018).

The use of alternative solutions to either supplement or replace V2I data may prove important to overcome barriers to development but it could also help to address another issue. Box and Waterson (2013) test their optimization method with various degrees of random noise applied to the V2I data. This has been done on the basis that GPS devices, even in the most accurate cases, contain a degree of error and that error can increase significantly in 'urban canyons'

where there are a large number of tall buildings. In this thesis the V2I data has been assumed to be more or less perfect and, as such, the results could be viewed as optimistic. However, the nature of information to be supplied by V2I is still not yet entirely clear and it has been assumed in this study that GPS information supplied by vehicles would be supplemented by other on-board sensors to increase the accuracy of the positional data. If the data provided was limited to the accuracy of GPS data then the use of alternative methods such as radar and video detection is likely to provide an enhanced performance compared to V2I, albeit with a shorter range.

In the previous chapter, the issue of data privacy was discussed in the context of collecting performance data from vehicles as they travel through a network. This data could be used to create acceleration/deceleration profiles and provide preferred following headway/driver reaction time to an internal traffic model. Data privacy could make this difficult to achieve unless V2I devices could be anonymised in much the same way as Bluetooth MAC addresses are in journey time monitoring systems. This would enable the combination of vehicle and driver performance to be reflected in the traffic model to improve accuracy.

It is perhaps worth noting that Local Authorities could potentially use interventions such the introduction of Clean Air Zones across the UK to incentivise vehicle manufacturers to provide V2I data. For example, a discount or exemption from the enforcement charge could be provided for vehicles that provide V2I data on the basis that the data will enable emissions for that vehicle to be reduced through improved signal optimization.

7.2.3 Research limitations

The research in this study is only applicable to under-saturated conditions. The MOVA representation described in Chapter 3 reflects the under-saturation algorithm of the MOVA control strategy. When over-saturation is detected,

MOVA switches to an alternative algorithm with the objective function of maximising capacity rather than minimising delay. There is less available literature on the method that is employed by MOVA in the alternative algorithm but Vincent and Peirce (1988) describe it as a redistribution of green time based on the efficiency of flow over the X detector, the detector closest to the stop line.

It should be feasible to apply V2I detection to the alternative algorithm to improve the measurement of flow efficiency and thus improve optimizer performance to provide an increase in capacity during periods of over-saturation. Incorporating V2I data would also be useful to ensure reliable detection of over-saturated conditions and prevent instances of the under-saturated algorithm being used in oversaturated conditions and vice versa.

The focus of this research was on applying V2I detection and vehicle classification to an existing optimization method, although a new optimizer was developed to accommodate the car-following traffic model. The methodology of the study was, in part, based on the idea that a lack of accurate and comprehensive data contributes to estimation error in the internal traffic models of signal optimizers thus limiting their performance. The development of a hybrid model enables the benefits of V2I to be realised as data becomes available rather than using purely conventional data until the penetration rate of V2I technology reaches, say, 30% of the vehicle fleet. However, the emergence of above-ground detection as a technique for mimicking V2I data may present an alternative scenario:

- **Scenario 1:** An existing junction has conventional detection and associated infrastructure. It is therefore not immediately cost effective to implement above-ground detection. V2I data is instead incorporated with the hybrid model as it becomes available; and

- **Scenario 2:** An existing junction has no conventional detection (common where infrastructure such as ducting cannot be installed) making the implementation of above-ground detection cost effective. A new optimization algorithm designed specifically to use V2I-type data could be developed to take advantage of the effectively '100%' vehicle fleet data coverage.

This thesis has focused on isolated intersections to enable the research to be conducted within the timescales of the study. The focus on isolated intersections does not preclude the application of the study outcomes to wider network optimization. The use of V2I detection and vehicle classification in particular would improve the accuracy of any internal traffic model that uses conventional detection. However, the computational requirements of the car-following model implementation in Chapter 6 may deem it inappropriate for network-wide use unless it was within a hierarchical control strategy.

7.3 Suggestions for further research

As discussed in the previous section, the collection of vehicle specific data regarding acceleration and deceleration performance prior to a vehicle arriving at a junction would enable the optimizer to more accurately determine the degree of weighting to be applied to it. In the MOVA representation that would manifest itself in the value of the stop penalty applied to each vehicle. In the car-following method described in Chapter 6 it would enable the model parameter values to be adjusted to more accurately represent vehicle specific behaviour. Research could focus on the methodology of collecting vehicle performance data as a vehicle travels through a network and at what point the collected data provides a benefit compared to a pre-specified value applied to all vehicles of the same class (used in this study).

All work in this study is undertaken in a Simulated Environment that allows repeatable conditions for representative comparisons of optimizer performance. Further research would seek to apply the developed hybrid traffic model and the alternative car-following model solutions to real-world conditions to investigate the practical issues that have been raised in this chapter.

In Chapter 6, a V2I penetration rate of 100% was assumed. Allowing for a situation where the V2I penetration rate is less than 100% introduces additional complexities in using the car-following model. For example, a common scenario will arise where a following vehicle is V2I equipped but the leading vehicle has been detected conventionally. In that scenario, the V2I equipped vehicle effectively provides 'ground truth' data whereas the leading vehicle position is estimated after leaving the upstream detector.

In the case of full V2I coverage, the car-following parameters can be adjusted with the knowledge of the actual position and speed of both vehicles. However, in this situation it is not desirable to change the deceleration parameters of the following vehicle based only on an estimation of the leading vehicle position. There is therefore a decision to make when the V2I equipped following vehicle violates the safe following distance calculated by the car-following model. Is the leading vehicle effectively 'pushed' forward to maintain the safe headway (which assumes the model parameter values are correct) or is it instead assumed that the violation of the safe following distance does indeed reflect reality, in which case the model parameters should be adjusted.

It is certainly not a straightforward exercise to adjust the model in a consistent manner that maintains the integrity of the vehicle trajectories along the link and so a decision was made not to pursue the development of a model that caters for a mix of V2I and conventionally detected vehicles. It may be interesting to pursue this in more detail as part of a further study but it would need to be

considered in the context of the discussions regarding V2I data availability, accuracy and possible supplementary data.

In Chapter 6, the possibility of adding random noise to the car-following model parameter values was discussed. It was identified that adding noise to the cloned run-ahead element of the optimizer could be an interesting further study in that the behaviour of the optimizer may become more or less cautious depending on the amplitude of the random noise. However, the number of variations of parameters would have to be carefully considered as the computational requirements would increase proportionally with the number of random seeds used (i.e. the use of 10 random seeds would increase the computation requirements by one order of magnitude).

Finally, the MOVA representation and the more sophisticated car-following model method used in this study are applicable to under-saturated conditions. To develop a control strategy that is generally applicable, an alternative strategy must be developed for over-saturated conditions. In the hybrid model, conventional detection could continue to be used in the short term but further research would enable alternative V2I based methods of monitoring efficiency of flow to be investigated.

References

Aboudolas, K., Papageorgiou, M. and Kosmatopoulos, E. 2007. Control and optimization methods for traffic signal control in large-scale congested urban road networks. In: *American Control Conference, 2007. ACC'07*: IEEE, pp.3132-3138.

Aboudolas, K., Papageorgiou, M. and Kosmatopoulos, E. 2009. Store-and-forward based methods for the signal control problem in large-scale congested urban road networks. *Transportation Research Part C: Emerging Technologies*. **17**(2), pp.163-174.

Adams, E. 2017. *How Long, Really, Until Self-Driving Cars Hit the Streets?* [Online]. [Last accessed 10 August]. Available from: http://www.thedrive.com/tech/16768/how-long-really-until-self-driving-cars-hit-the-streets

Aghabayk, K., Sarvi, M. and Young, W. 2015. A state-of-the-art review of car-following models with particular considerations of heavy vehicles. *Transport reviews*. **35**(1), pp.82-105.

Aghabayk, K., Sarvi, M., Young, W. and Kautzsch, L. 2013. A novel methodology for evolutionary calibration of Vissim by multi-threading. In: *Australasian Transport Research Forum*, pp.1-15.

Ahmane, M., Abbas-Turki, A., Perronnet, F., Wu, J., El Moudni, A., Buisson, J. and Zeo, R. 2013. Modeling and controlling an isolated urban intersection based on cooperative vehicles. *Transportation Research Part C: Emerging Technologies*. **28**, pp.44-62.

Alessio, F.M., K.; Daalderop, G.; Alexander, P. D.; Schober, F.; Pfliegl, W. 2017. Ready to roll: Why 802.11p beats LTE and 5G for V2x. [Online]. [Accessed 12 August 2018]. Available from: https://www.siemens.com/content/dam/webassetpool/mam/tag-siemens-com/smdb/mobility/road/connected-mobility-solutions/documents/its-g5-ready-to-roll-en.pdf

Amsterdam Group. 2015. *Signal Phase and Time (SPAT)and Map Data (MAP*. Amsterdam Group.

Athol, P. 1965. Interdependence of certain operational characteristics within a moving traffic stream. *Highways Research Record*. **72**, pp.58-87.

Bell, M.C. and Bretherton, R.D. 1986. Ageing of fixed-time traffic signal plans. In: *International Conference on Road Traffic Control*, pp.77-80.

Bodenheimer, R., Brauer, A., Eckhoff, D. and German, R. 2014. Enabling GLOSA for adaptive traffic lights. In: *Vehicular Networking Conference (VNC), 2014 IEEE*: IEEE, pp.167-174.

Bodger, M. 2011. SCOOT MMX (SCOOT Multi-Modal 2010). In: *JCT Symposium*. JCT Consultancy.

Box, S. and Waterson, B. 2012. An automated signalized junction controller that learns strategies from a human expert. *Engineering Applications of Artificial Intelligence*. **25**(1), pp.107-118.

Box, S. and Waterson, B. 2013. An automated signalized junction controller that learns strategies by temporal difference reinforcement learning. *Engineering applications of artificial intelligence*. **26**(1), pp.652-659.

Bretherton, D., Hounsell, N. and Radia, B. 1996. Public transport priority in SCOOT. In: *Intelligent Transportation: Realizing the Future. Abstracts of the Third World Congress on Intelligent Transport SystemsITS America*.

Cai, C., Wong, C.K. and Heydecker, B.G. 2009. Adaptive traffic signal control using approximate dynamic programming. *Transportation Research Part C: Emerging Technologies*. **17**(5), pp.456-474.

Chazan, G. 2018. *Hamburg becomes first German city to ban old diesel vehicles*. [Online]. [Last accessed 12 August 2018]. Available from: https://www.ft.com/content/0eadb66c-5e6f-11e8-9334-2218e7146b04

Cheung, S.Y., Coleri, S., Dundar, B., Ganesh, S., Tan, C.-W. and Varaiya, P. 2005. Traffic measurement and vehicle classification with single magnetic sensor. *Transportation Research Record*. **1917**(1), pp.173-181.

Ciuffo, B., Punzo, V. and Montanino, M. 2012. Thirty years of Gipps' car-following model: Applications, developments, and new features. *Transportation Research Record: Journal of the Transportation Research Board*. (2315), pp.89-99.

Clearview Intelligence. n.d. Product Specification: M100 Wireless vehicle detection system. [Online]. [Accessed 10 August 2018]. Available from: https://www.clearview-intelligence.com/uploads/specifications/CI-PS-44-M100-4pp-FINAL-LR.pdf

Coifman, B. 2001. Improved velocity estimation using single loop detectors. *Transportation Research Part A: Policy and Practice*. **35**(10), pp.863-880.

Coifman, B., Dhoorjaty, S. and Lee, Z.-H. 2003. Estimating median velocity instead of mean velocity at single loop detectors. *Transportation Research Part C: Emerging Technologies*. **11**(3-4), pp.211-222.

Coifman, B. and Kim, S. 2009. Speed estimation and length based vehicle classification from freeway single-loop detectors. *Transportation Research Part C: Emerging Technologies*. **17**(4), pp.349-364.

Collier, B.C., S.; Hallworth, B. 2017. Sprucing up your ITS. In: *JCT Symposium*. JCT Consultancy.

Crabtree, M.R. 2011. *MOVA Traffic Control Manual*. TRL Limited.

Crabtree, M.R., Henderson, I.R., Drury, R.D. and Burtenshaw, G. 2012. *Guide to MOVA Data Set-Up and Use*. TRL Limited.

Crabtree, M.R.K., J.V. 2005. *The safety of MOVA at high speed junctions*.

Cuff, M.M., J. 2018. *ClientEarth wins third battle with government over air quality plan*. [Online]. [Last accessed 11 August]. Available from: https://www.businessgreen.com/bg/news/3027090/clientearth-wins-third-battle-with-government-over-air-quality-plan

Daganzo, C.F. 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*. **28**(4), pp.269-287.

Dailey, D.J. 1999. A statistical algorithm for estimating speed from single loop volume and occupancy measurements. *Transportation Research Part B: Methodological*. **33**(5), pp.313-322.

DEFRA. 2015. *Improving air quality in the UK: Tackling nitrogen dioxide in our towns and cities*. Department for Environment Food & Rural Affairs.

DEFRA. 2017. *Clean Air Zone Framework: Principles for setting up Clean Air Zones in England*. Department for Environment Food & Rural Affairs.

Department for Transport. 1995. *Traffic Advisory Leaflet 4/95: The "SCOOT" Urban Traffic Control System*. Department for Transport.

Department for Transport. 2000. *Traffic Advisory Leaflet 8/00: Bus priority in SCOOT*. Department for Transport.

Department for Transport. 2018. Road traffic (vehicle miles) by vehicle type in Great Britain, annual from 1949. [Online]. [Accessed 11 August 2018]. Available from: www.gov.uk/government/organisations/department-for-

transport/series/road-traffic-statistics

DEVPOST. 2018. *Connected Vehicle Technology Challenge*. [Online]. [Last accessed 12 August 2018]. Available from: https://connectedvehicle.devpost.com/details/understanding-dsrc

Diakaki, C., Papageorgiou, M. and Aboudolas, K. 2002. A multivariable regulator approach to traffic-responsive network-wide signal control. *Control Engineering Practice*. **10**(2), pp.183-195.

Dodsworth, J., Shepherd, S. and Liu, R. 2014. Real-time single detector vehicle classification. *Transportation Research Procedia*. **3**, pp.942-951.

Driver & Vehicle Licensing Agency. 2018. *Get vehicle information from DVLA*. [Online]. [Last accessed 11 August]. Available from: https://www.gov.uk/get-vehicle-information-from-DVLA#more-information

Eckhoff, D., Halmos, B. and German, R. 2013. Potentials and limitations of green light optimal speed advisory systems. In: *Vehicular Networking Conference (VNC), 2013 IEEE*: IEEE, pp.103-110.

ETSI. 2010. *Intelligent Transport Systems (ITS): Communications Architecture*. European Telecommunications Standards Institute.

European Union. 2008. Directive 2008/50/EC of the European Parliament and of the Council. *Official Journal of the European Union*. (L 152).

Feng, Y., Head, K.L., Khoshmagham, S. and Zamanipour, M. 2015. A real-time adaptive signal control in a connected vehicle environment. *Transportation Research Part C: Emerging Technologies*. **55**, pp.460-473.

FirstGroup Plc. 2018. For the technically minded: How bus priority in Swansea works. [Online]. [Accessed 11 August 2018]. Available from: https://www.firstgroup.com/south-west-wales/about-us/technically-minded-how-bus-priority-swansea-works

Fox, K., Chen, H., Montgomery, F., Smith, M. and Jones, S. 1998. Selective Vehicle Priority in the UTMC Environment (UTMC01). [Online]. [Accessed 11 August 2018]. Available from: http://www.its.leeds.ac.uk/projects/spruce/utmc1rev.html#_Toc428247164

Gartner, N.H. 1983. OPAC: A demand-responsive strategy for traffic signal control. In: *62nd Annual Meeting of the Transportation Research Board*. Transportation Research Record, pp.75-81.

Gipps, P.G. 1981. A Behavioral Car-Following Model for Computer-Simulation. *Transportation Research Part B-Methodological*. **15**(2), pp.105-111.

Gissler, A. 2015. Connected vehicle: Succeeding with a disruptive technology. [Online]. [Accessed 10 August 2018]. Available from: https://www.accenture.com/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Dualpub_21/Accenture-digital-Connected-Vehicle.pdf

Hamilton, A., Waterson, B., Cherrett, T., Robinson, A. and Snell, I. 2013. The evolution of urban traffic control: changing policy and technology. *Transportation Planning and Technology*. **36**(1), pp.24-43.

He, X., Liu, H.X. and Liu, X. 2015. Optimal vehicle speed trajectory on a signalized arterial with consideration of queue. *Transportation Research Part C: Emerging Technologies*. **61**, pp.106-120.

He, Y., Du, Y. and Sun, L. 2012. Vehicle classification method based on single-point magnetic sensor. *Procedia-Social and Behavioral Sciences*. **43**, pp.618-627.

Hellinga, B.R. 2002. Improving freeway speed estimates from single-loop detectors. *Journal of transportation engineering*. **128**(1), pp.58-67.

Henry, J.-J., Farges, J.L. and Tuffal, J. 1983. The PRODYN real time traffic algorithm. In: *4th IFAC/IFIP/IFORS conference on control in transportation systems, Baden-Baden*. pp.305-310.

Hertfordshire County Council. 2009. Intelligent Transport System (ITS) Strategy. [Online]. [Accessed 11 August 2018]. Available from: https://www.hertfordshire.gov.uk/media-library/documents/highways/transport-planning/local-transport-plan-live/intelligent-transport-systems-its-inventory.pdf

Heydecker, B. 2004. Objectives, stimulus and feedback in signal control of road traffic. *Journal of Intelligent Transportation Systems*. **8**(2), pp.63-76.

Highways Agency. 2002. *MCE 0108 Issue C: Siting Of Inductive Loops For Vehicle Detecting Equipments At Permanent Road Traffic Signal Installations*. Highways Agency.

Highways Agency. 2004. *The Geometric Layout of Signal-Controlled Junctions and Signalised Roundabouts*.

Highways Agency. 2005. MCH 1542 Issue C: Installation Guide for MOVA. [Online]. MCH 1542. (C). [Accessed 11 August 2018]. Available from:

http://www.ukroads.org/webfiles/MCH1542C.pdf

Hounsell, N., McLeod, F., Landles, J. and Gardner, K. 1996. Bus priority in London: Building on PROMPT. In: *Intelligent Transportation: Realizing the Future. Abstracts of the Third World Congress on Intelligent Transport SystemsITS America*.

Hounsell, N. and Shrestha, B. 2005. AVL based bus priority at traffic signals: A review of architectures and case study. *European Journal of Transport and Infrastructure Research*. **5**(EPFL-ARTICLE-158663), pp.13-29.

Hunt, P., Robertson, D., Bretherton, R. and Winton, R. 1982. SCOOT-A Traffic Responsive Method of Coordinating Signals. *Traffic Engineering & Control*. **23**, p190.

INRIX. 2018. *INRIX Global Traffic Scorecard*. [Online]. [Last accessed 10 August]. Available from: http://inrix.com/scorecard/

Jiang, D. and Delgrossi, L. 2008. IEEE 802.11 p: Towards an international standard for wireless access in vehicular environments. In: *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*: IEEE, pp.2036-2040.

Kamalanathsharma, R.K. and Rakha, H.A. 2016. Leveraging connected vehicle technology and telematics to enhance vehicle fuel efficiency in the vicinity of signalized intersections. *Journal of Intelligent Transportation Systems*. **20**(1), pp.33-44.

Katsaros, K., Kernchen, R., Dianati, M. and Rieck, D. 2011. Performance study of a Green Light Optimized Speed Advisory (GLOSA) application using an integrated cooperative ITS simulation platform. In: *7th International Wireless Communications and Mobile Computing Conference*: IEEE, pp.918-923.

Ki, Y.-K. and Baik, D.-K. 2006. Vehicle-classification algorithm for single-loop detectors using neural networks. *IEEE Transactions on Vehicular Technology*. **55**(6), pp.1704-1711.

Kwon, J., Varaiya, P. and Skabardonis, A. 2003. Estimation of Truck Traffic Volume from Single Loop Detectors with Lane-to-Lane Speed Correlation. *Transportation Research Record: Journal of the Transportation Research Board*. **1856**(-1), pp.106-117.

Lee, J., Park, B. and Yun, I. 2013. Cumulative travel-time responsive real-time intersection control algorithm in the connected vehicle environment. *Journal of Transportation Engineering*. **139**(10), pp.1020-1029.

Leeds City Council. 2016. Leeds Transport Strategy. [Online]. [Accessed 11 August 2018]. Available from: https://www.leeds.gov.uk/docs/Leeds%20Transport%20Strategy.pdf

Leeds City Council. 2018. Best Council Plan: Tackling poverty and reducing inequalities. [Online]. Available from: https://www.leeds.gov.uk/your-council/plans-and-strategies/council-plans

Lighthill, M.J. and Whitham, G.B. 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proc. R. Soc. Lond. A.* **229**(1178), pp.317-345.

Lo, H.K. 1999. A novel traffic signal control formulation. *Transportation Research Part A: Policy and Practice.* **33**(6), pp.433-448.

Mauro, V. and Di Taranto, C. 1989. UTOPIA. In: *Proceedings of the IFAC-IFIP-IFORS Conference on Control, Computers, Communications in Transportation*, pp.242-252.

McCoy, P.T., Balderson, E.A., Hsueh, R.T. and Mohaddes, A.K. 1983. Calibration of TRANSYT platoon dispersion model for passenger cars under low-friction traffic flow conditions. *Transportation Research Record.* **905**, pp.48-52.

Meta, S. and Cinsdikici, M.G. 2010. Vehicle-classification algorithm based on component analysis for single-loop inductive detector. *IEEE Transactions on Vehicular Technology.* **59**(6), pp.2795-2805.

Mikhalkin, B., Payne, H.J. and Isaksen, L. 1972. Estimation of speed from presence detectors. *Highways Research Record.* (388), pp.73-83.

Miller, A.J. 1963. A computer control system for traffic networks. In: *2nd International Symposium on the Theory of Traffic Flow, London*. pp.200-220.

Mirchandani, P. and Head, L. 2001. A real-time traffic signal control system: architecture, algorithms, and analysis. *Transportation Research Part C: Emerging Technologies.* **9**(6), pp.415-432.

Air pollutants: Nitrogen oxides. 2018. [Online database]. National Atmospheric Emissions Inventory.

Neelisetty, S. and Coifman, B. 2004. Improved single loop velocity estimation in the presence of heavy truck traffic. In: *Proc. of the 83rd Annual Meeting of the Transportation Research Board*.

NIC. 2018. *Five shortlisted to develop 'Roads for the Future' and driverless cars.* [Online]. [Last accessed 12 August]. Available from:

https://www.nic.org.uk/news/five-shortlisted-to-develop-roads-for-the-future-and-driverless-cars/

NIS Ltd. n.d. National Traffic Information Service: Datex II Service. [Online]. [Accessed 12 August 2018]. Available from: http://www.trafficengland.com/resources/cms-docs/user-guide.pdf

Oh, S., Ritchie, S. and Oh, C. 2002. Real-time traffic measurement from single loop inductive signatures. *Transportation Research Record: Journal of the Transportation Research Board*. (1804), pp.98-106.

Ordnance Survey. 2018. *Beginner's guide to GPS*. [Online]. [Last accessed 11 August]. Available from: https://www.ordnancesurvey.co.uk/business-and-government/help-and-support/navigation-technology/gps-beginners-guide.html

Papageorgiou, M., Ben-Akiva, M., Bottom, J., Bovy, P.H.L., Hoogendoorn, S.P., Hounsell, N.B., Kotsialos, A. and McDonald, M. 2007. Chapter 11 ITS and Traffic Management. In: Cynthia, B. and Gilbert, L. eds. *Handbooks in Operations Research and Management Science*. Elsevier, pp.715-774.

Peirce, J. and Webb, P. 1990. *MOVA: The 20 site trial*. TRL.

Peterson, A., Bergh, T. and Steen, K. 1986. LHOVRA, A new traffic signal control strategy for isolated junctions. *International Conference on Road Traffic Control*. pp.98-101.

Porche, I. and Lafortune, S. 1997. Dynamic traffic control: decentralized and coordinated methods. In: *Proceedings of Conference on Intelligent Transportation Systems*, pp.930-935.

Porche, I., Sampath, M., Sengupta, R., Chen, Y.-L. and Lafortune, S. 1996. A decentralized scheme for real-time optimization of traffic signals. In: *Proceedings of the 1996 IEEE International Conference on Control Applications.*: IEEE, pp.582-589.

PTV-Vision. 2011. *VISSIM 5.30-05 User Manual*. PTV-Vision.

Pushkar, A., Hall, F.L. and Acha-Daza, J.A. 1994. Estimation of speeds from single-loop freeway flow and occupancy data using cusp catastrophe theory model. *Transportation research record*. **1457**, p149.

Quinn, D., Montgomery, F. and May, A. 1988. Control of congestion in highly saturated networks: Working paper 251–Experimental results and conclusions. Institute for Transport Studies, University of Leeds, UK.

Reid, D.S., B. 2007. LinSig for Signalled Roundabouts. In: *JCT Symposium*. JCT Consultancy.

Robertson, D. 1969. *TRANSYT: A traffic network study tool.* TRRL, Crowthorne, Berkshire, United Kingdom.

Robertson, D., Baker, R. and Lucas, C. 1980. *Coordinating traffic signals to reduce fuel consumption.* TRRL, Crowthorne, Berkshire, United Kingdom.

Robertson, D. and Bretherton, R. 1974. Optimum control of an intersection for any known sequence of vehicle arrivals. In: *Proceedings of the 2nd IFAC/IFIP/IFORS Symposium on Traffic Control and Transportation Systems*.

RTEM. 2013. SLC4-Single loop Classification Firmware. [Online]. [Accessed 11 August 2018]. Available from: https://www.rtem.co.uk/slc4.html

Sen, S. and Head, K.L. 1997. Controlled optimization of phases at an intersection. *Transportation Science*. **31**(1), pp.5-17.

Shepherd, S. 1994. Traffic control in over‐saturated conditions. *Transport Reviews*. **14**(1), pp.13-43.

Siemens Plc. 2015. ST950 Facilities Handbook. [Online]. [Accessed 11 August 2018]. Available from: https://www.siemens.co.uk/traffic/pool/downloads/handbooks/st950/667_hb_46000_001.pdf

Siemens Plc. 2016a. SCOOT User Guide. [Online]. [Accessed 12 August 2018]. Available from: https://www.siemens.co.uk/traffic/pool/downloads/handbooks/utc/666_hf_16940_000.pdf

Siemens Plc. 2016b. SLD4 User Manual. [Online]. [Accessed 10 August 2018]. Available from: https://www.siemens.co.uk/traffic/pool/downloads/handbooks/road_signals/667_hb_45200_000.pdf

Sims, A.G. and Dobinson, K.W. 1980. The Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. *IEEE Transactions on Vehicular Technology*. **29**(2), pp.130-137.

Stahlmann, R., Möller, M., Brauer, A., German, R. and Eckhoff, D. 2016. Technical evaluation of GLOSA systems and results from the field. In: *Vehicular Networking Conference (VNC), 2016 IEEE*: IEEE, pp.1-8.

Steel, P.A., M.; Hodge, A.; Saini, O.; Leigh, A.; Clayton, D.A.; Fulton, L.; Pursey, R.; Ball, J.; Foote, M.; Bounds, D. 2012. *Chameleon Outstation User Handbook*. Peek. p.86. [User Handbook].

Sun, C. and Ritchie, S.G. 1999. Individual vehicle speed estimation using single loop inductive waveforms. *Journal of Transportation Engineering*. **125**(6), pp.531-538.

Tafish, H., Balid, W. and Refai, H.H. 2016. Cost effective Vehicle Classification using a single wireless magnetometer. In: *Wireless Communications and Mobile Computing Conference (IWCMC), 2016 International*: IEEE, pp.194-199.

TomTom. 2018. *TomTom Traffic Index*. [Online]. [Last accessed 10 August]. Available from: https://www.tomtom.com/en_gb/trafficindex/

Transport for London. 2018. *Ultra Low Emission Zone*. [Online]. [Last accessed 12 August 2018]. Available from: https://tfl.gov.uk/modes/driving/ultra-low-emission-zone

TRL Limited. 2018. *MOVA*. [Online]. [Last accessed 12 August]. Available from: https://trlsoftware.co.uk/products/traffic_control/mova

Tucker S., S.I., Fletcher J., Mustard D.,. 2006. *Evaluating the benefits of MIDAS automatic queue protection*. Traffic Engineering & Control.

Van den Berg, M., Hegyi, A., De Schutter, B. and Hellendoorn, J. 2003. A macroscopic traffic flow model for integrated control of freeway and urban traffic networks. In: *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*: IEEE, pp.2774-2779.

Vincent, R.A. and Peirce, J.R. 1988. *MOVA: Traffic responsive, self-optimising signal control for isolated intersections*. TRRL.

Wang, Y. and Nihan, N.L. 2003. Can single-loop detectors do the work of dual-loop detectors? *Journal of Transportation Engineering*. **129**(2), pp.169-176.

Whittingham, M. 2004. *SIETAG handbook - installation, commissioning and maintenance*. Siemens.

Wiedemann, R. 1974. *Simulation des Strassenverkehrsflusses*. Schriftenreihe des Instituts für Verkehrswesen der Universität Karlsruhe.

Witts D. 2013. The missing link. In: *JCT Symposium*. JCT Consultancy.

Yang, K., Guler, S.I. and Menendez, M. 2016. Isolated intersection control for various levels of vehicle technology: Conventional, connected, and automated

vehicles. *Transportation Research Part C: Emerging Technologies*. **72**, pp.109-129.

Ye, Z., Zhang, Y. and Middleton, D. 2006. Unscented Kalman filter method for speed estimation using single loop detector data. *Transportation Research Record: Journal of the Transportation Research Board*. (1968), pp.117-125.

Zhanfeng, J., Chao, C., Coifman, B. and Varaiya, P. 2001. The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors. In: *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE, 2001*, pp.536-541.

Appendix A
Flow-occupancy derivations

A.1 Introduction

This appendix provides a detailed derivation of space-mean speed, using the relationship between speed, flow and occupancy. The space-mean speed approximation is the basis of many single detector speed estimation algorithms. The derivation below is based on that by Kwon et al. (2003).

$$q_i = \frac{n_i}{T_i} \tag{A.1}$$

$$O_i = \frac{\sum_{k \in K_i} t_k}{T_i} \tag{A.2}$$

where:

q_i = Flow in sample time period *i* (vehicles/second).

n_i = Number of vehicles in sample time period *i*.

T_i = Sample *i* interval length (seconds).

O_i = Occupancy in the sample time period *i*.

t_k = 'On-time' of vehicle *k* (seconds).

K_i = Set of all vehicles that pass over the detector in the sample time period.

The space mean speed for a particular sample of vehicles is defined by Equation (A.3) where v_k is the speed of vehicle k. If all vehicle speeds were known (i.e. if all vehicles in a sample were providing real-time tracking data) then this would be straightforward to calculate. However, in this case the vehicle speeds are not known and so instead, an estimation of vehicle length is used (Equation (A.4)) along with the sum of the time vehicles spend over the detector during a sample time period, as shown in the following equations.

$$\overline{v}_i = \frac{\sum_{k \in K_i} v_k}{n_i} \tag{A.3}$$

The vehicle on-time (the duration for which a vehicle is present over a detector) is related to the speed and length of a particular vehicle by

$$L_{tot} = v_k t_k \tag{A.4}$$

where L_{tot} is the sum of the length of the vehicle and the effective length of the detector. So, from Equation (A.2),

$$O_i = \frac{1}{T_i} \sum_{k \in K_i} \frac{L_k}{v_k} \tag{A.5}$$

and substituting Equation (A.1),

$$O_i = q_i \frac{1}{n_i} \sum_{k \in K_i} \frac{L_k}{v_k} \tag{A.6}$$

Equation (A.6) can be re-written as,

$$O_i \approx \frac{q_i \overline{L}_i}{\overline{v}_i} \tag{A.7}$$

or

$$\overline{v}_i \approx \frac{q_i \overline{L}_i}{O_i} \tag{A.8}$$

assuming that vehicle speeds are approximately constant during the sample interval.

Equation (A.9) reiterates the relationship between parameters for reference given the various terminologies in the related literature reviewed here.

$$\overline{L_i} = MEVL = \frac{1}{n_i} \cdot \sum_{k \in K} L_{tot} = \frac{1}{g} \tag{A.9}$$

where:

$\overline{L_i}$ = Space-mean length (metres).

n = Number of vehicles in the specified vehicle sample.

$MEVL$ = Mean effective vehicle length (metres).

g = Reciprocal of MEVL.

For a given sample time period, Equation (A.8) can also be expressed as

$$\overline{v} \approx \frac{n}{T \times O \times g} \tag{A.10}$$

as is the case in Wang and Nihan (2003). However, Athol (1965) referred to the *g*-factor as the *K*-factor.

Appendix B
Tabular results

B.1 Summary

The full tabular results for Chapter 5 and Chapter 6 are included in this appendix for reference. A key is provided below:

- A: MOVA representation;
- B: Modified MOVA representation with single detector vehicle classification;
- C: Hybrid model (with 100% V2I penetration rate);
- D: Hybrid model (100% V2I) with V2I-based classification;
- E: Car-following model based optimizer using conventional detection; and
- F: Car-following model based optimizer using V2I based detection.

The results are presented for delay, stops and PI for all vehicles. Within each category a breakdown of results for all vehicles, all non-HGV vehicles, all HGV vehicles and articulated HGVs is also provided. All results are an average taken from 10 simulation runs with different random seeds.

Finally, the PI results from testing the performance of the car-following based optimizer with different V2I ranges are presented.

B.2 Tabular results for Chapter 5 and Chapter 6

B.2.1 Delay

Table B.1: Average delay (seconds per vehicle) for all vehicles.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	13.6	14.7	15.9	17.0
B	13.5	14.8	15.8	17.0
C	11.9	12.7	13.9	14.4
D	11.8	12.7	13.7	14.7
E	13.5	14.4	15.5	16.3
F	11.9	12.7	13.5	14.5
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	9.4	11.1	13.6	17.6
B	9.4	11.1	13.5	17.6
C	8.5	9.8	11.9	14.9
D	8.4	9.8	11.8	14.9
E	9.8	11.3	13.5	16.1
F	8.4	9.6	11.9	14.8

Table B.2: Average delay (seconds per vehicle) for all non-HGV vehicles.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	13.5	14.6	15.7	16.7
B	13.5	14.7	15.7	16.8
C	11.8	12.5	13.7	14.2
D	11.8	12.6	13.7	14.6
E	13.4	14.2	15.3	16.0
F	11.8	12.7	13.4	14.3
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	9.3	11.1	13.5	17.6
B	9.3	11.1	13.5	17.6
C	8.4	9.8	11.8	14.8
D	8.4	9.8	11.8	14.9
E	9.8	11.3	13.4	16.1
F	8.3	9.7	11.8	14.8

Table B.3: Average delay (seconds per vehicle) for all HGVs.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	14.7	15.7	16.8	17.8
B	14.3	15.6	16.8	18.1
C	13.1	14.0	14.8	15.0
D	12.3	13.1	14.2	15.4
E	15.3	15.6	16.5	17.4
F	12.1	13.2	13.7	15.0
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	10.5	11.9	14.7	18.5
B	10.3	11.7	14.3	18.3
C	8.8	10.5	13.1	15.5
D	8.1	9.4	12.3	15.2
E	10.6	12.2	15.3	17.1
F	9.0	9.3	12.1	15.0

Table B.4: Average delay (seconds per vehicle) for articulated HGVs.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	14.7	15.5	16.4	18.3
B	13.4	15.4	16.4	18.1
C	13.8	14.8	15.1	15.1
D	12.2	13.3	13.8	15.5
E	15.9	16.0	16.1	17.6
F	12.5	13.4	14.0	15.4
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	9.9	11.7	14.7	19.2
B	9.3	11.1	13.4	18.2
C	7.9	11.7	13.8	16.0
D	7.0	9.4	12.2	15.9
E	10.2	13.4	15.9	17.0
F	9.6	9.7	12.5	16.4

B.2.2 Stops

Table B.5: Average stops per vehicle for all vehicles.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	0.60	0.62	0.66	0.67
B	0.60	0.62	0.65	0.68
C	0.53	0.55	0.59	0.61
D	0.53	0.56	0.60	0.61
E	0.57	0.61	0.65	0.68
F	0.51	0.54	0.56	0.58
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	0.41	0.49	0.60	0.74
B	0.41	0.49	0.60	0.74
C	0.39	0.44	0.53	0.66
D	0.39	0.44	0.53	0.67
E	0.41	0.49	0.57	0.69
F	0.35	0.41	0.51	0.64

Table B.6: Average stops per vehicle for all non-HGV vehicles.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	0.60	0.62	0.66	0.68
B	0.60	0.63	0.65	0.69
C	0.53	0.55	0.60	0.61
D	0.53	0.56	0.60	0.61
E	0.57	0.61	0.65	0.68
F	0.51	0.54	0.57	0.59
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	0.41	0.49	0.60	0.74
B	0.41	0.49	0.60	0.74
C	0.39	0.44	0.53	0.67
D	0.39	0.45	0.53	0.67
E	0.41	0.49	0.57	0.69
F	0.35	0.41	0.51	0.64

Table B.7: Average stops per vehicle for all HGVs.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	0.58	0.59	0.65	0.65
B	0.56	0.59	0.64	0.67
C	0.52	0.55	0.58	0.59
D	0.50	0.54	0.56	0.59
E	0.56	0.60	0.63	0.66
F	0.45	0.50	0.52	0.56
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	0.43	0.46	0.58	0.72
B	0.42	0.45	0.56	0.71
C	0.35	0.44	0.52	0.63
D	0.31	0.37	0.50	0.64
E	0.39	0.46	0.56	0.65
F	0.31	0.32	0.45	0.60

Table B.8: Average stops per vehicle for articulated HGVs.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	0.55	0.56	0.57	0.63
B	0.53	0.55	0.56	0.61
C	0.53	0.55	0.55	0.56
D	0.43	0.52	0.53	0.55
E	0.52	0.59	0.58	0.62
F	0.42	0.50	0.50	0.54
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	0.32	0.42	0.55	0.68
B	0.30	0.40	0.53	0.64
C	0.27	0.50	0.53	0.61
D	0.22	0.34	0.43	0.64
E	0.30	0.49	0.52	0.59
F	0.29	0.35	0.42	0.61

B.2.3 Performance Indicators

Table B.9: Average PI in each test case for Scenario 1 and Scenario 2.

Scenario 1	5% HGV (Base case)	10% HGV	15% HGV	20% HGV
A	27.5	30.7	34.6	38.0
B	27.4	30.8	34.4	38.4
C	24.2	27.1	30.8	33.2
D	24.1	27.0	30.5	33.6
E	26.9	30.1	33.9	37.4
F	23.7	26.6	29.3	32.6
Scenario 2	**400 veh/h**	**500 veh/h**	**600 veh/h** (Base case)	**700 veh/h**
A	19.0	22.5	27.5	35.0
B	19.0	22.4	27.4	34.9
C	17.3	20.1	24.2	30.3
D	17.2	19.9	24.1	30.6
E	19.4	22.7	26.9	32.1
F	16.5	19.1	23.7	29.8

B.3 Tabular results for V2I range

Table B.10: Average PI for increasing V2I range.

	100m	200m	300m	400m	500m
400 veh/h	18.5	16.2	16.3	16.7	18.0
600 veh/h	26.3	23.6	23.6	24.3	24.8
800 veh/h	44.2	38.1	37.4	38.8	38.7

Appendix C
Accommodating car-following model practical issues

C.1 Introduction

This appendix is provided to describe in more detail some of the practical issues associated with implementing the car-following model in an on-line optimizer. The following examples describe how the model accommodates discrepancies in the representation of reality once those discrepancies become apparent.

C.2 Discharge rate

As already discussed, relocation of the downstream detector to beyond the stop line allows the discharge rate of vehicles over the stop line to be monitored. The car-following model has been implemented with the ability to 'pause' if it is identified that the model is discharging vehicles more quickly than has been recorded over the detector. The ability to pause the model has been added to provide some protection against incorrect classification of vehicles and to accommodate uncertainty in driver reaction times. For example, an HGV that is incorrectly classified as a car/van may accelerate much more slowly than the model predicts. If this is left unchecked it may result in premature termination of the green signal as the modelled queue will discharge too quickly. To avoid this potential problem, the car-following model on the relevant link can effectively be paused until another vehicle has crossed the downstream detector.

The MOVA strategy does not explicitly model the discharge of the queue and, instead, identifies the end of queue discharge by monitoring gaps over the X detector (approximately 3.5 seconds travel time upstream of the stop line).

Pausing of the model is only permitted during the start of the green signal as it is important to avoid over-reliance on precisely counting vehicles in and out of the link. In many real world situations there are traffic sources and sinks along a link, between detectors, that are either too difficult or too minor to justify additional detection. If the model relied entirely on counting vehicles in and out of the link it would not cope with such perturbations. The pause function is also only employed if the count difference between the modelled and actual discharge is greater than two vehicles to allow for some variation between the model and reality.

C.3 Queues beyond the detector

When a vehicle is detected travelling over the upstream detector, the classification algorithm detailed in Chapter 4 is employed to identify the vehicle class. The vehicle is then entered onto the modelled link with relevant characteristics based on the identified vehicle class. However, once a queue reaches the upstream detector, there is no more data available to inform the model of queuing beyond that detector.

To overcome this issue, when the queue extends close to the upstream detector, 'temporary' vehicles are added to the model at an average demand flow rate. Those vehicles are considered by the optimizer but do not necessarily exist in reality. If a new vehicle is detected (i.e. more space between the actual end of queue and the upstream detector than is modelled) then the temporary vehicles are removed from the model and the process restarted.

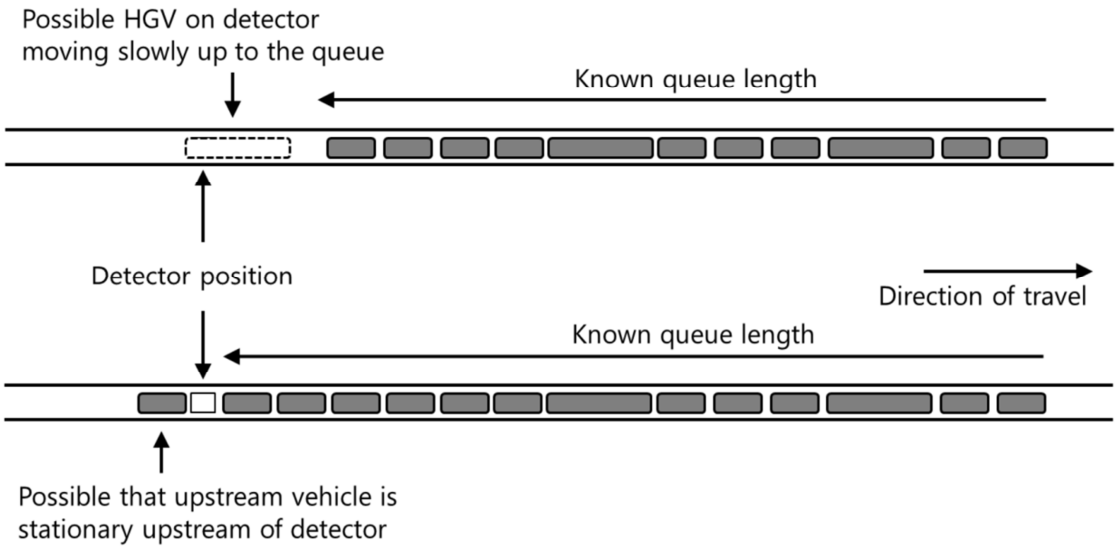

Figure C.1: Examples of queues extending back to an upstream detector.

This process is necessary as it enables the optimizer to take into account the likelihood that the queue has extended beyond the upstream detector and thus consider the likely increase in delay resulting from a further green extension to the opposing links. If this process was not undertaken then the modelled queue lengths would never extend beyond the upstream detector and the optimizer may continue to extend an opposing green signal on the assumption that the queue will still take the same amount of time to discharge whether the signal changes now or at some indeterminate time in the future.

At the start of the following green, the temporary vehicles are removed from the model as they are no longer required in the optimizer.

C.3.1 Queue length error

In some instances (for example, when a queue has discharged more quickly than the model has predicted) there may be a modelled vehicle located at the point a newly detected vehicle is planned to be inserted. In this situation the model must be adjusted to accommodate the new vehicle.

The condition described occurs most commonly in the following two situations, particularly when the link is approaching saturation:

- **Situation 1** – The signal state is red and the modelled queue extends to the upstream detector location (Figure C.2); and
- **Situation 2** – The signal state is green, traffic has begun discharging but the last vehicle in the queue is still stationary or slow moving (Figure C.3).

In **Situation 1**, there are two reasons (in the Simulated Environment) that the modelled queue could be longer than the actual queue. The effective vehicle lengths in the queue may have been overestimated, or, the vehicle at the front of the queue could have incorrectly been modelled as stopping at the red signal whereas in reality it crossed the stop line after the end of the previous effective green (the green signal plus some period of amber). Of course, it could also be a combination of the two. In this case, the vehicle at the front of the queue can simply be removed from the model and the position of the other vehicles adjusted to move one space up in the queue. This process can be repeated until there is sufficient space to accommodate the newly detected vehicle.

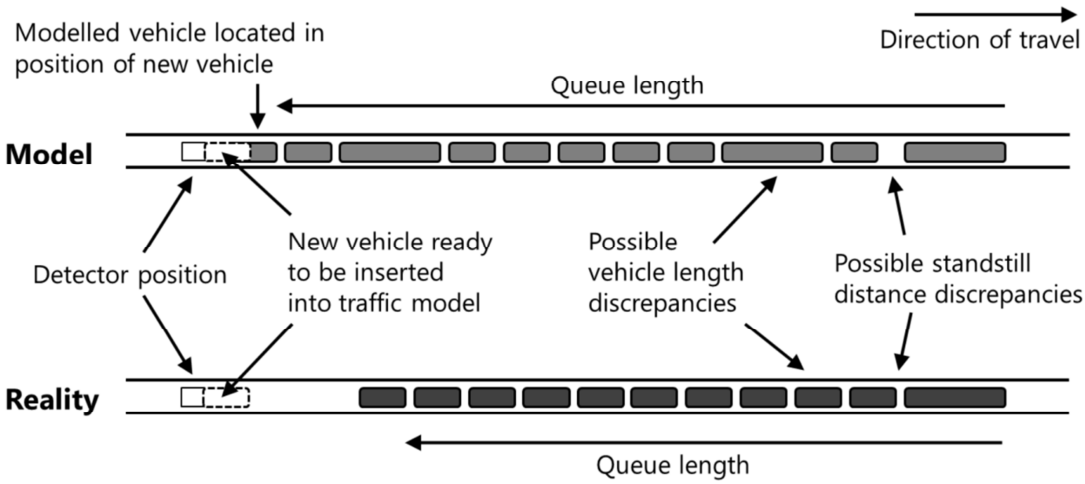

Figure C.2: An example of Situation 1 occurring where the model (top) queue length extends to the upstream detector whereas the reality (bottom) is that the queue is shorter.

An alternative option could be to iteratively reduce the effective length of the vehicles in the queue (and adjust the position of each vehicle to maintain the minimum safety distance at standstill) until there is sufficient space for the newly detected vehicle. In this situation there is an uncertainty as to the cause of the model error and therefore neither method of model adjustment carries greater validity. For this work the first method has been implemented.

If this method were to be developed further and implemented in a real world scenario, a third reason for overestimation of the queue could occur. Junction layouts that include a traffic 'sink' between the upstream and downstream detector (i.e. traffic can leave the link before it crosses the stop line) will result in consistent overestimation of queue lengths if the proportion of traffic leaving the link is significant. In that case, an additional detector could be installed at the location of the sink to enable the model to be adjusted appropriately. This is something that is already applied in MOVA using IN-SINK and X-SINK detectors for vehicles leaving the link after the upstream and downstream detector respectively. It is also incorporated into SCOOT through the use of 'subtractive' links (Siemens Plc, 2016a).

In **Situation 2**, the queue is already in the process of discharging and so it is assumed that the vehicles in the modelled queue have not discharged as quickly as in reality. Rather than removing any vehicles from the link model, the model for the specific link (not the whole junction) is run ahead until there is sufficient space on the link for the newly detected vehicle to be inserted.

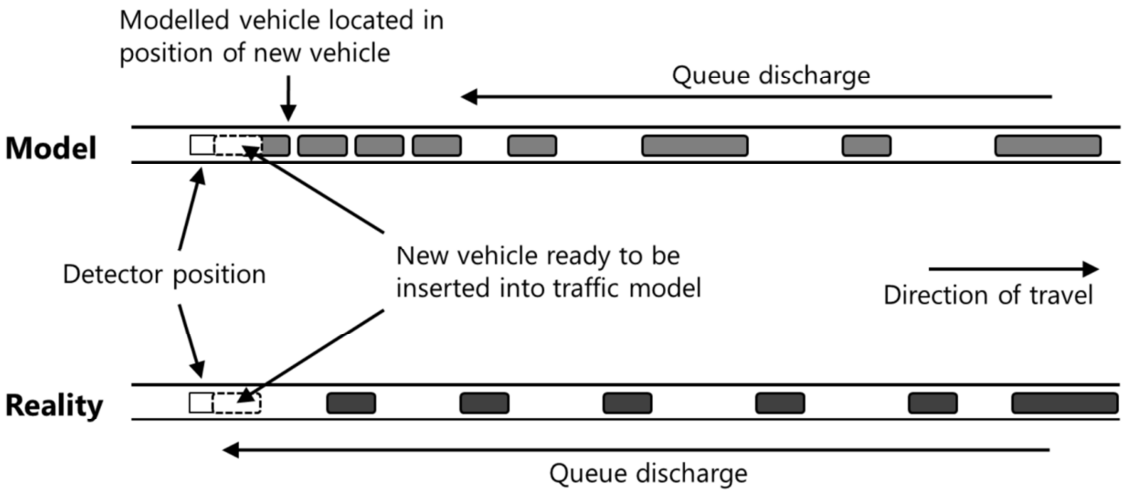

Figure C.3: An example of Situation 2 occurring where the model (top) discharges less quickly than in reality (bottom).

Appendix D
Software architecture and source code

D.1 Introduction

An outline of the software architecture used to create the Simulated Environment is described by the class diagram shown in Figure D.1. The detailed source code is available on the CD included with this thesis.

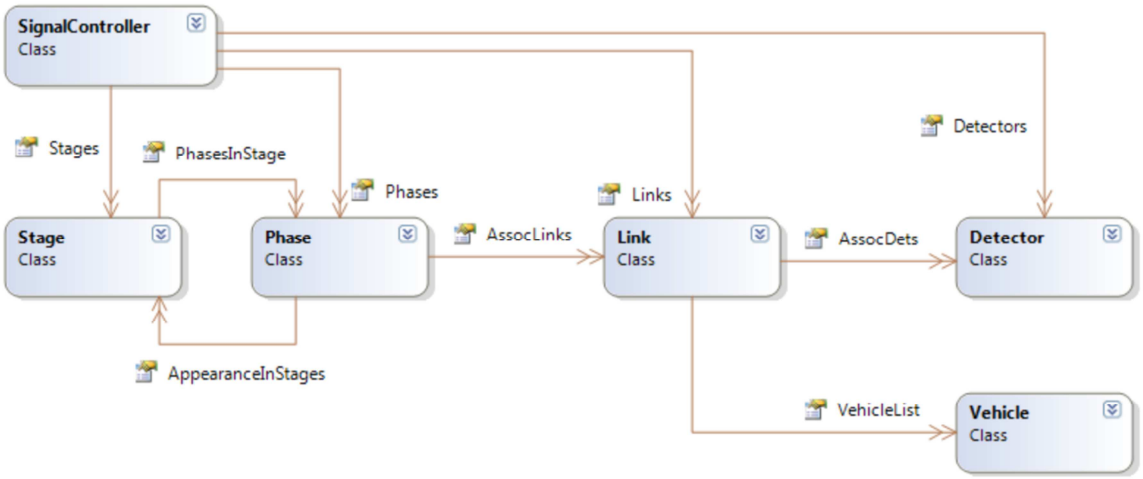

Figure D.1: Architecture for Simulated Environment Signal Controller software.