Modelling and Optimisation of Dynamic Motorway Traffic

Thesis submitted to University College London for the degree of Doctor of Philosophy

by

Ying Li

Department of Civil, Environmental & Geomatic Engineering

Centre for Transport Studies

University College London

December 2015

Declaration

I, Ying Li, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

_____

Date:

_____

Abstract

Ramp metering, variable speed limits, and hard shoulder running control strategies have been used for managing motorway traffic congestion. This thesis presents a modelling and optimisation framework for all these control strategies. The optimal control problems that aim to minimise the travel delay on motorways are formulated based upon a macroscopic cell transmission model with piecewise linear fundamental diagram. With the piecewise linear nature of the traffic model, the optimal control problems are formulated as linear programming (LP) and are solved by the IBM CPLEX solver. The performance of different control strategies are tested on real scenarios on the M25 Motorway in England, where improvements were observed with proper implementation.

With considering of the uncertainties in traffic demand and characteristics, this thesis also presents a robust modelling and optimisation framework for dynamic motorway traffic. The proposed robust optimisation aims to minimise both mean and variance of travel delays under a range of uncertain scenarios. The robust optimisation is formulated as a minimax problem and solved by a two stage solution procedure. The performances of the robust ramp metering are illustrated through working examples with traffic data collected from the M25 Motorway. Experiments reveal that the deterministic optimal control would outperform slightly the robust control in terms of minimising average delays, while the robust controller gives a more reliable performance when uncertainty is taken into account. This thesis contributes to the development and validation of dynamic simulation, and deterministic and robust optimisation.

Acknowledgements

First and foremost, I would like to present my deep gratitude to my principal supervisor Andy Chow for his incredible support, insightful instructions and detailed comments and suggestions on the thesis. I am forever grateful for his help. I would also like to thank my secondary supervisor Professor B.G. Heydecker for his time to provide sound advice and valuable criticism for completing the thesis. I am also grateful to both examiners Professor Richard Gibbens and Dr Ke Han for their useful comments and suggestions on the thesis.

Many staff at the Centre for Transport Studies have assisted me during my stay at University College London. I am particularly grateful to Catherine Holloway, Taku Fujiyama, Nicola Christie, and Helena Titheridge. Special thanks to Angela Cooper (in UCL Centre for Languages & International Education) who supports me on the thesis writing.

I would like to convey my warm and profound thanks to my dearest friends in the United Kingdom and China for their patience and invaluable support throughout my study period. I am grateful to my friends and fellow colleagues RuNing Ye, ChienPang Liu, Nuo Duan, Simrn Gill, Shuai Li, Rui Sha, Fang Xu, Aris Pavlides, Matthew Tsang, HuaiDong Wang, SiDi Sun, ZiJia Wang, Kun Liu, MengYang Qin, and Li Zhao for standing beside me along this path and for all the time, joy and sorrow we shared together.

I am tremendously indebted to my family members, especially my parents XiaoJuan Zhang and ZhongXian Li, and my master for all their support, care and encouragement. They are the sources of confidence and strength for me to overcome difficulties in study and life.

Finally, I would also like to acknowledge Highways England for providing the MIDAS traffic data, and IBM Academic Initiative for providing the software and license for using the CPLEX Optimisation Studio. The contents do not reflect the official views or policies of Highways Agency, and other organisations.

Contents

List of Figures

List of Tables

Abbreviations

LWR	**L**ighthill-**W**hitham-**R**ichards
GHR	**G**azis-**H**erman-**R**othery
CTM	**C**ell **T**ransmission **M**odel
MIDAS	**M**otorway **I**ncident **D**etection and **A**utomatic **S**ignalling
IQR	**I**nter-**Q**uantile **R**ange
VDT	**V**ehicle **D**istance **T**ravelled
VHT	**V**ehicle **H**ours **T**ravelled
TTD	**T**otal **T**ravel **D**istance
TTT	**T**otal **T**ravel **T**ime
TSD	**T**otal **S**ystem **D**elay
LP	**L**inear **P**rogramming
MILP	**M**ixed **I**nteger **L**inear **P**rogramming
CFL	**C**ourant **F**riedrichs **L**ewy
RM	**R**amp **M**etering
VSL	**V**ariable **S**peed **L**imits
HSR	**H**ard **S**houlder **R**unning
RMVSL	**R**amp **M**etering with **T**wo **V**ariable **S**peed **L**imits

Symbols

Q	capacity of main road	[veh/hr]
v	free flow speed of main road	[km/hr]
w	wave speed of main road	[km/hr]
$\mathring{\rho}$	critical density of main road	[veh/km]
$\bar{\rho}$	jam density of main road	[veh/km]
t	time index	
k	time step counter	
i	segment index	
Δx	cell length	[km]
Δt	time step	[hr]
I	number of cells	
K	number of simulation time steps	
f	cell outflow of main road	[veh/hr]
ρ	cell density of main road	[veh/km]
o	cell occupancy of main road	
$\hat{\rho}$	measured density of main road	[veh/km]
s	exit flow from main road	[veh/hr]
λ	the traffic demand that wants to enter the system	[veh/hr]
r	the actual demand that enters the system	[veh/hr]
ρ_b	the average density in quantile regression	[veh/km]
f_b	the largest non-outlier flow in trained regression	[veh/hr]
$Q1$	the 25^{th} percentile flow in trained regression	[veh/hr]
$Q3$	the 75^{th} percentile flow in trained regression	[veh/hr]
$\bar{r}$	the ramp capacity	[veh/hr]

l	queue length on the ramp	[veh]
$\bar{l}$	the maximum queue length on the ramp	[veh]
β	fraction of traffic that leaves main road (split ratio)	
d	system delay	[veh-hr]
Φ	fundamental diagram	
η	the balance between the main road delay and ramp delay	
ζ	the balance between safety and efficiency	
$\mathbf{c}$	optimal control policy	
$\mathbf{c_r}$	optimal ramp metering policy	
$\mathbf{c_v}$	optimal VSL control policy	
$\mathbf{c_h}$	optimal HSR control policy	
$\mathbf{c_{rhv}}$	optimal integrated control policy	
$\mathbf{c}^*$	optimal robust ramp metering policy	
α	decision variables for VSL	
μ	decision variables for HSR	
ϵ	mean absolute percentage error in density	
P	relative delay reduction	
P_m	relative main road delay reduction	
P_t	relative total system delay reduction	
$\bar{L}_v$	effective vehicle length	[metre]
L_d	loop detector length	[metre]
K_r	the regulation parameter in ALINEA	
$\widehat{o}$	pre-set expected value for downstream occupancy in ALINEA	

Chapter 1

Introduction

1.1 Background and Motivation

The performance of transport infrastructure is closely linked to the associated society's prosperity, economic growth and quality of life. Sustained growth in demand in virtually all modes of transport places tremendous pressure on the infrastructure. This ever-increasing demand for travel results in considerable congestion and economic loss in many cities around the world. Traffic congestion has significant impacts on many related important issues including energy consumption, public health, safety, environment, and security. The grand challenge for the world's major cities in the 21st century is to make economic growth and sustainability compatible.

Traffic volume has grown rapidly over the last few decades in the United Kingdom [3]. Traffic congestion is a major bottleneck for economic and social development. It is one of the key challenges for major cities around the world, especially for the United

Kingdom (UK). The UK Eddington study [28] states that the monetary cost due to road congestion will reach £22 billion (at 2002 prices) per annum for all road users by 2025, in which 13 per cent of road traffic will be subjected to stop-and-go travel conditions. In a report published in 2009, the UK Department for Transport (DfT) also suggests that traffic congestion across the English road network as a whole will increase from 2003 levels by 27 per cent by 2025, and 54 per cent (67 per cent in the London road network) by 2035. This represents an average increase in time spent travelling of 9 per cent (6 seconds) for each kilometre travelled [3].

Continuous construction of new roads will not be a sustainable solution to traffic congestion due to the increasingly constrained financial, physical, and environmental conditions. Consequently, governments, businesses, and research teams around the world want to explore alternative ways to effectively utilise and manage existing road infrastructure. In fact, adding new infrastructure may induce more demand for travel and hence is not an effective option. Therefore, a sustainable solution for mitigating congestion calls for the effective management of the existing infrastructure through appropriate control methods. This study investigates the modelling and optimisation framework for dynamic motorway traffic.

1.2 Review of Traffic Models

A traffic model is a mathematical representation that is used to describe and estimate the behaviour of traffic flow. The importance of a traffic model for effective transport management is highlighted in Kotsialos and Papageorgiou [52]. Various traffic models have been developed in the past decades. Traffic models can be broadly categorised into

microscopic and macroscopic paradigms according to the level of detail. Microscopic models simulate the traffic behaviour of individual vehicles, while macroscopic models represent traffic at an aggregated level. Details of different models are presented below.

1.2.1 Microscopic models

Microscopic models simulate the traffic behaviour of each vehicle and their laws are drawn from cognitive studies, artificial intelligence, and measurements with the use of in-vehicle devices. Some established examples of microscopic models include AIMSUN [16], DRACULA [63], Paramics [97] and VISSIM [87]. The car-following models are typical microscopic models with the consideration of interaction between adjacent vehicles. In a car-following representation [49], the motion of a group of vehicles by supporting that the motion of the foremost vehicle is known and then applying the model to calculate that each successive vehicle. The general form of car-following models conforms to that of stimulus-response. This can be expressed in broad terms as:

$$Response(later) = Sensitivity \times Stimulus(now) \tag{1.1}$$

where the stimulus is defined by the velocity difference between adjacent vehicles, and the response refers to the current action (e.g. acceleration and braking) of the following vehicle [47], and the sensitivity is a parameter specifying how the following drivers react to this action.

Different forms of sensitivity give rise to different car-following models. Various car-following models are prepared in the literature, including Gazis-Herman-Rothery (GHR)

models, Safety distance or Collision avoidance models, Linear models, and Psychophysical or Action point models. The simplest stimulus-response relationship of GHR model [18] is as follows:

$$a_n(t+T) = \mu_1\big(v_{n-1}(t) - v_n(t)\big) \tag{1.2}$$

The response $[a_n(t+T)]$ depends linearly in the stimulus $[v_{n-1}(t) - v_n(t)]$. In particular, the sensitivity is supposed to independent of both speed and spacing, and is the same for acceleration as it is for braking. The response $[a_n(t+T)]$ is the acceleration of the n^{th} vehicle at time $t+T$, and T is the duration of the lag between stimulus and response, which due to the driver's behaviour (such as perception, interpretation, evaluation and action) or the vehicle's mechanical lag. The notation μ_1 is the sensitivity coefficient which is a constant. The stimulus $[v_{n-1}(t) - v_n(t)]$ is the relative speed of the n^{th} vehicle to its front vehicle $n-1$ at the time t.

The stimulus-response type car-following model has been modified by many researchers since 1950s [12]. Gazis *et al.* [30] propose a further modification of the GHR model such that the distance between adjacent vehicles is considered in the sensitivity of the model. Edie [29] modifies the GHR model by considering that the speed of the leading vehicle will also influence the following vehicle. The general form of the GHR model can be written as:

$$a_n(t+T) = \mu_2 \frac{v_n(t+T)^l}{s_n(t)^m}\big(v_{n-1}(t) - v_n(t)\big) \tag{1.3}$$

where the response $[a_n(t+T)]$ is the acceleration of the n^{th} vehicle at time $t+T$. The stimulus $[v_{n-1}(t) - v_n(t)]$ is same as the stimulus used in Equation 1.2. The sensitivity $[\mu_2 \frac{v_n(t+T)^l}{s_n(t)^m}]$ considers the speed of the vehicle and the distance between adjacent vehicles, where $s_n(t)$ is the spacing in front of vehicle n at time t, and $v_n(t+T)$ is the

speed of the n^{th} vehicle at time $t + T$, and μ_2 is the sensitivity coefficient. Different combinations of m and l have been investigated by some researchers (e.g. May and Harmut [69]; Ozaki [76]). May and Harmut [69] consider a range of possibilities of m and l, and they found that the best fit for motorway data was achieved with values $m = 0.8$, $l = 2.8$, and for traffic in tunnels with values $m = 0.6$, $l = 2.1$. However, it is difficult to choose a uniform value of m and l to fit all traffic conditions.

The original formulation of safety distance or collision avoidance models dates to 1959 [51]. The collision avoidance model considers the following vehicle keeping the minimum safe following distance. Gipps [32] develops the model by considering several mitigating factors. Since then, the collision avoidance model is widely used in simulation.

The other kind of microscopic model is Cellular automata models [104] which use discretised time and space. A set of local rules are defined to describe the relationship between the centre cell and its neighbouring cells in the isotropic case. The rules of Cellular automata models can be modified intuitively and flexibly by considering different traffic conditions. Cellular automata models can be divided into two categories: one is the basic one-dimensional Cellular automata models suitable for motorway traffic such as rule-184 model [103] and NS model [71], and the other two-dimensional Cellular automata models (e.g. BML4 model [11]) are suitable for urban network traffic. Cellular automata models retain non-linear behaviour and some physical characteristics of traffic flow, and those models are easier to simulate on the computer.

1.2.2 Macroscopic models

In contrast to microscopic models, macroscopic models represent traffic dynamics in terms of aggregated quantities: flow, density and mean speed of traffic. The relationship between density and flow is known as the fundamental diagram (see Figure 1.1). The fundamental diagram plays an important role in modelling the dynamics of traffic propagation.

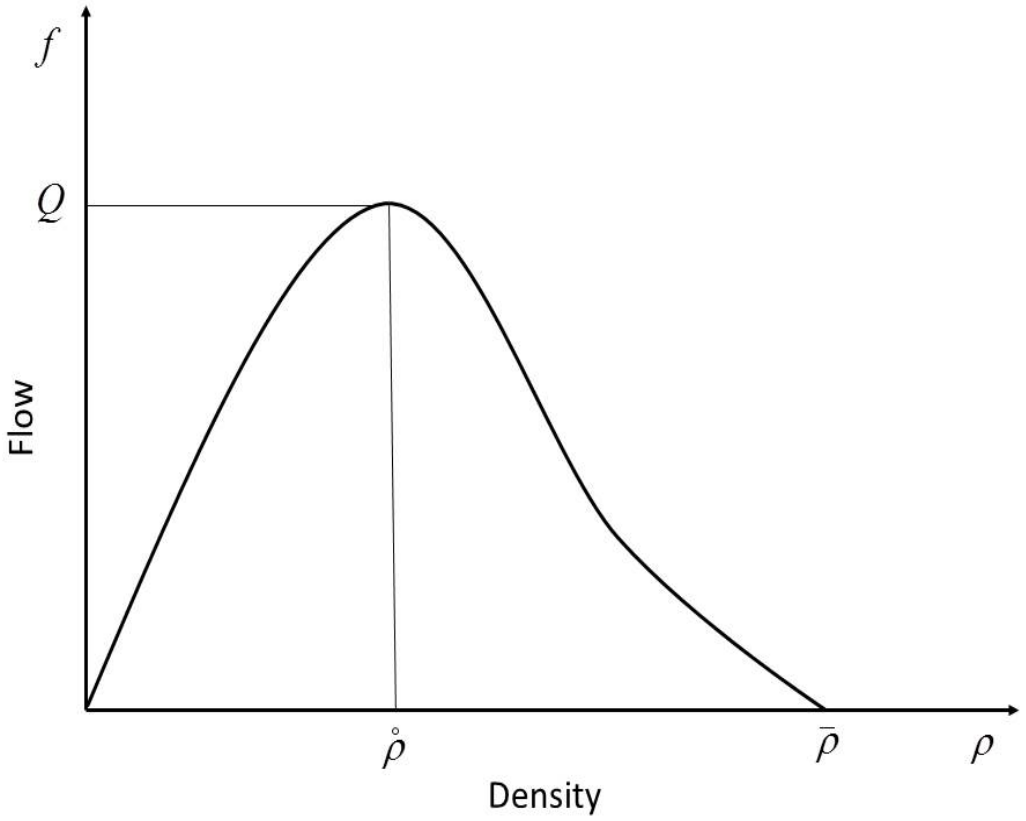

FIGURE 1.1: Fundamental diagram

The fundamental diagram should be able to represent traffic in all possible states: free flow, at capacity and congested. In general, the function of fundamental diagram Φ is assumed to be concave and is defined for density $\rho \in [0,\ \bar{\rho}]$, where $\bar{\rho}$ is the jam density

which is the maximum density of traffic that can be achieved at the corresponding location. The maximum value of flow observed on the diagram is regarded as the capacity at that location, which is Q as shown in Figure 1.1. The density at which the capacity flow is observed is called the critical density $\mathring{\rho}$, in which Φ is increasing for all $\rho < \mathring{\rho}$ and Φ is decreasing for $\rho > \mathring{\rho}$. The data with a density value higher than critical density are classified as congested data, which are characterised by the fact that an increase in density will induce a decrease in flow which is because of the speed reduction due to the onset of congestion at high densities. Data points associated with a density lower than critical density are regarded as free flow. Traffic in a free flow state generally proceeds with a relatively steady speed known as the free flow speed. As a result, the flow increases as the density of traffic increases.

A number of functional forms for fundamental diagram have been proposed in the literature to model this fundamental relationship between traffic flow and density. Some well-known ones include Greenshields [36], Greenberg [35], and Edie [29]. Moreover, Carey and Bowers [14] present a comprehensive review of different fundamental diagrams adopted for traffic flow modelling. Newell [73] proposes the use of a piecewise linear fundamental diagram (triangular shape) in his seminal paper on traffic flow theory, which is then adopted by Daganzo [23] in his cell transmission model formulation. As pointed out by Papageorgiou *et al.* [80]; Lo [64]; Kotsialos and Papageorgiou [52] and many others, a macroscopic traffic model should incorporate the fundamental diagram of traffic flow in order to produce plausible estimations. Based upon the fundamental diagram of traffic, macroscopic traffic models can further be distinguished into two approaches: first order models and higher order models.

Lighthill and Whitham [61], and Richards [90] propose a pioneering first order macroscopic model of traffic flow which is called a kinematic wave model. The kinematic wave model is also known as the Lighthill-Whitham-Richards (LWR) model. They applied the method of fluid dynamics to the traffic flow by making an analogy with traffic flow and continuous fluid. The model considers two spatiotemporal variables $f(x,t)$ and $\rho(x,t)$, which represent the flow and density at location x along a road section at time t respectively.

Consider a stretch of motorway, the number of vehicles flowing from x to $x+\Delta x$ during time interval $[t, t+\Delta t]$. At time t the traffic density on this section is $\rho(x,t)$, and at time $t+\Delta t$ the traffic density on this section is $\rho(x,t)-\Delta\rho$. The negative sign of the $\Delta\rho$ represents the congested traffic condition because that a increase in traffic density will induce a decrease in flow under heavy traffic. The traffic flows into the section at a rate of $f(x,t)$, and flows out of the section at a rate of $f(x,t)+\Delta f$. Suppose no entries and exists in this section, then the rate of change in the number of vehicles should equal to the net flow into this section as follows:

$$\Big(\rho(x,t)-\big(\rho(x,t)-\Delta\rho\big)\Big)\Delta x = \Big(f(x,t)-\big(f(x,t)+\Delta f\big)\Big)\Delta t \tag{1.4}$$

Simplifying the Equation 1.4 gives,

$$\Delta\rho\Delta x = -\Delta f\Delta t \tag{1.5}$$

that is,

$$\frac{\Delta\rho}{\Delta t}+\frac{\Delta f}{\Delta x}=0 \tag{1.6}$$

Then to take limit,

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial f(x,t)}{\partial x} = 0 \tag{1.7}$$

The conservation law (Equation 1.7) presented above assumes the traffic on a homogeneous motorway without entering and exiting traffic. However, in reality, the motorway always has on-ramps and off-ramps, we extend the conservation law to model such inhomogeneitoes. The revised conservation law (Equation 1.8) describes the number of vehicles in the road section during a time interval equals to the total vehicles that have entered the road section by that time minus the total vehicles that have exited the section by the same time.

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial f(x,t)}{\partial x} = r(x,t) - s(x,t) \tag{1.8}$$

for all x and t, where $r(x,t)$ and $s(x,t)$ are the exogenous inflow (e.g., on-ramp or main road inflow) and outflow (e.g., off-ramp or main road outflow) at (x,t) respectively.

Following Equation 1.8, the LWR model assumes that the traffic flow $f(x,t)$ adjusts instantaneously to the associated traffic density $\rho(x,t)$ through a predefined fundamental diagram function Φ as follows:

$$f(x,t) = \Phi\big(\rho(x,t)\big) \tag{1.9}$$

As a first order model, the speed $v(x,t) = f(x,t)/\rho(x,t)$ is assumed to adjust to the associated traffic density instantaneously (see further discussion in Papageorgiou [77]). The LWR model has been one of the most widely accepted models due to its plausibility and simplicity. However, it is not easy to solve the LWR model analytically. A number

of numerical schemes have been proposed for solving the LWR model. Daganzo [23] develops the cell transmission model (CTM), which discretises the LWR model over time and space. The cell transmission model divides the road network into a collection of sub-sections or cells whose length is equal or greater than the distance travelled by a single vehicle on the free flow speed in one time step. Lebacque [57] demonstrates that CTM is an application of the Godunov scheme [33], which is an established discretisation method for solving partial differential equations with discontinuous solutions. The cell transmission model (CTM) proposed by Daganzo [23] remains one of the most efficient discretisation schemes due to its simplicity and credibility, which will be discussed in details in Section 2.2.

A major criticism of the LWR model is its implicit assumption of unrealistically high acceleration and deceleration of vehicles through Equation (1.9) and its incapability of capturing a complex phenomenon such as capacity drop and stop-and-go wave [105]. Further discussion on LWR model can be found in Nagel and Nelson [70]. In an attempt to remedy the deficiency of first order models, Payne [88] develops a second order model, which considers the transient dynamics of acceleration and deceleration, and drivers reaction time during the state transition. The second order models consider the dynamic of speed. Therefore, a change in the downstream density influences the speed of the upstream after a period of time.

1.2.3 Discussion

Microscopic models have been adopted in many practical applications, where they are able to describe fine details of traffic flow in the real world. The primary advantage of

microscopic models is that the individual vehicle motion can be modelled. Nevertheless, calibrating microscopic models can be extremely expensive, as they require a huge amount of data and manpower in order to collect and capture the fine details of traffic flow to be modelled, and it hinders them from system wide applications and optimisation. Moreover, some information such as drivers' reaction times and decision rules on lane-changing behaviour are latent and may never be observable.

In contrast to microscopic models, macroscopic models are simpler due to the fewer model parameters involved. Macroscopic models represent traffic dynamics in terms of aggregated quantities: flow, density and mean speed of traffic. In general, the required data for macroscopic models can be readily obtained from standard surveillance infrastructure such as loop detectors, cameras and other kinds of fixed sensors. Moreover, macroscopic models are much more computationally efficient than their microscopic counterparts, while actually not sacrificing too much accuracy considering the fact that the amount of data required to fully calibrate microscopic models are often not obtainable in practice. This makes macroscopic models a feasible candidate for large-scale applications in the real world.

Compared with the first order macroscopic models, the second order macroscopic models can reproduce more complex traffic phenomena because the dynamic speed is adopted in the second order models. However, Daganzo [25] identifies several deficiencies of second order models including the possibility of violation of causality and negative flows. Moreover, the calibration of second order models involve determination of additional parameters such as acceleration/deceleration rates, and drivers' reaction time. Considering desirable mathematical properties and computational efficiency, this study adopts the first order macroscopic model (CTM) to represent traffic dynamics in the optimisation

framework.

1.3 Current Practice of Traffic Control

Over the past decades, significant advances have been made by introducing appropriate traffic control strategies to make the best use of existing motorways [2]. This section reviews those control strategies that are used for motorways such as ramp metering, variable speed limits (VSL), and hard shoulder running (HSR).

1.3.1 Ramp metering

Ramp metering has been proved to be effective in reducing traffic congestion and travel time. It has been used in the United Kingdom, United States, Germany, France and other parts of the world (see Bellemans *et al.* [6]; Haj-Salem *et al.* [39]; Harbord [41]; Zhang and Levinson [110]). Researchers have developed various ramp metering strategies. Details can be found in Papageorgiou and Kotsialos [78] which provides a comprehensive review on ramp metering.

Wattleworth [102] proposes a fixed-time ramp metering strategy. However, it leads either to overload of the main road flow or underutilisation of the motorway due to the dynamic variations of traffic conditions, which were not taken into account. To accommodate the temporal variation of traffic, traffic responsive ramp metering strategies were developed. Traffic responsive ramp metering are based on real time measurements from loop detectors or other kinds of sensors installed in the motorway. Ramp metering can further be classified into local and coordinated ramp metering. Figure 1.2 shows the

typical configuration of traffic responsive ramp metering system. Loop detectors (the rectangle in Figure 1.2) are installed in the on-ramp and main road of the motorway. The control device gathers data (e.g. traffic flow, speed and occupancy) from loop detectors and processes the information. Then the traffic lights can be set according to the current traffic condition.

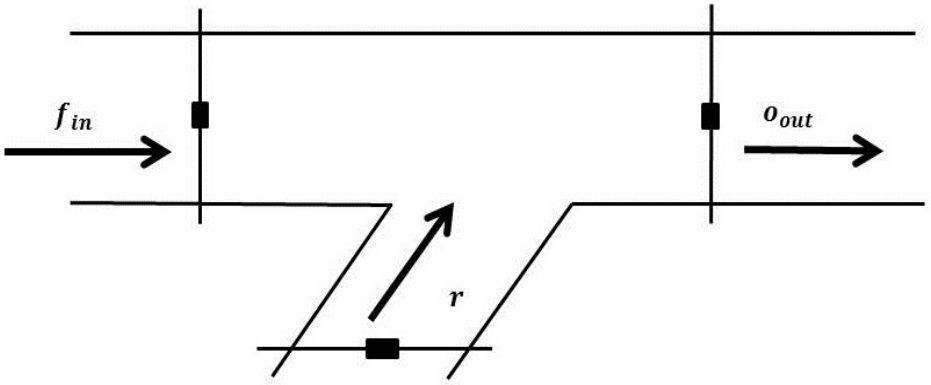

FIGURE 1.2: Configuration of responsive ramp metering system

Local ramp metering strategies make use of measurements from the vicinity of a single ramp. Most prominent examples of local ramp metering strategies are the demand-capacity (DC) and the occupancy (OCC) strategies. The control rule of the demand-capacity strategy [68] is shown as follows:

$$r(k) = \begin{cases} Q - f_{in}(k-1) & \text{if} \quad o_{out}(k) < \mathring{o} \\ r_{min} & \text{else} \end{cases} \tag{1.10}$$

where $r(k)$ is the ramp inflow over simulation time step k, Q is the downstream capacity of the on-ramp, $o_{out}(k)$ and $\mathring{o}$ are the occupancy measurement over simulation time

step k and the critical occupancy at the downstream of the on-ramp respectively. The notation $f_{in}(k-1)$ is the upstream motorway flow measurement of the on-ramp over simulation time step $k-1$, and r_{min} is the pre-set minimum value of the on-ramp inflow. Figure 1.3 shows the demand-capacity strategy is an open loop control strategy which is sensitive to various disturbances [78].

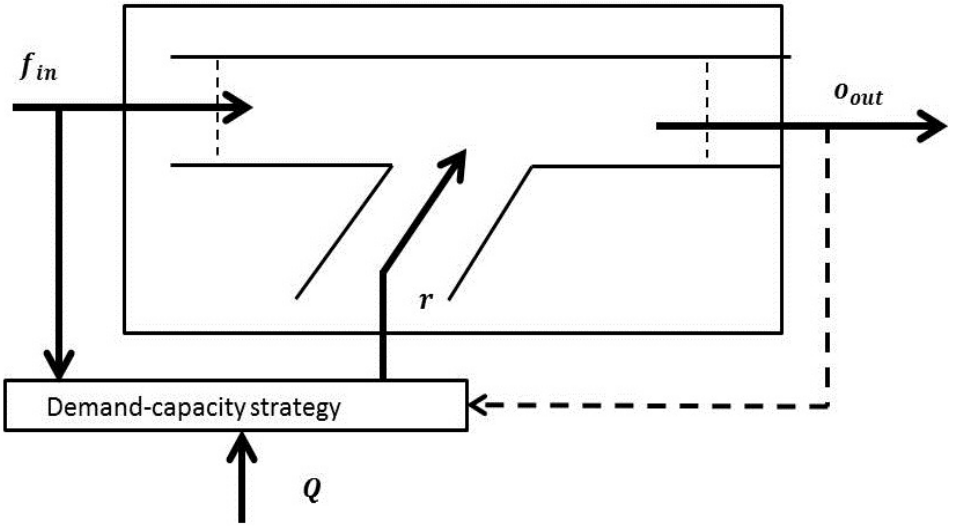

FIGURE 1.3: Demand-capacity ramp metering strategy

(Source: Papageorgiou 2000 [78])

The occupancy strategy (e.g. ALINEA) is similar to the demand-capacity strategy, but it replies on downstream occupancy measurements of the on-ramp. The ALINEA [82] strategy and its variations are feedback control schemes (as shown in Figure 1.4) targeting a set-point for the downstream occupancy. ALINEA regulates the ramp inflow $r(k)$ according to the downstream measurement of main road occupancy $o_{out}(k-1)$ of the on-ramp as follows:

$$r(k) = r(k-\Delta c) + K_r\big(\widehat{o} - o_{out}(k-1)\big) \tag{1.11}$$

where $r(k)$ is the ramp inflow over simulation time step k, K_r is the control gain, $o_{out}(k)$ is the occupancy measurement over simulation time step k, Δc is the control period governing the frequency of updating the ramp metering strategy, $r(k - \Delta c)$ is the ramp inflow over time step $k - \Delta c$, and $\widehat{o}$ is a pre-set target occupancy value at the downstream of the on-ramp, which is typically set as the associated critical occupancy [83]. If the measured downstream occupancy is less than (greater than) the required downstream occupancy, the new ramp metering rate is increased (decreased) on the basis of the last ramp metering rate. This ramp metering strategy measures only the downstream occupancy, so it requires fewer measurements than any other local ramp metering strategies.

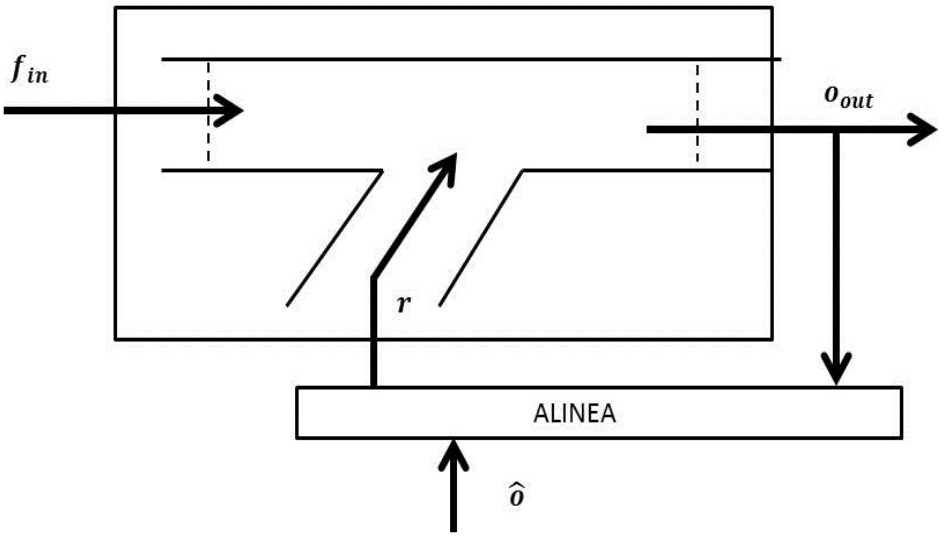

FIGURE 1.4: ALINEA ramp metering strategy

(Source: Papageorgiou 2000 [78])

Kerner [50] develops an alternative ramp metering strategy (ANCONA). In three-phase traffic theory, the traffic is classified into free-flow (F), synchronised flow (S), and wide moving jam (J). ANCONA will not restrict the on-ramp inflow when bottleneck under

free-flow (F) traffic, therefore the on-ramp inflow $r(k)$ is as follows:

$$r(k) = \lambda(k) \tag{1.12}$$

where $\lambda(k)$ is the traffic demand that wants to enter the network, and $r(k)$ is the actual demand that enters the network.

When synchronised flow (S) is measured at the bottleneck, ANCONA starts to control the on-ramp inflow based on main road speed measurements. The detected average speed v_{dec} and predefined predefined congestion speed v_{con} are used to define the traffic condition. If $v_{dec} \leq v_{con}$, ANCONA reduces the on-ramp inflow as follows:

$$r(k) = \lambda_1(k) < \lambda(k) \tag{1.13}$$

When the reduced on-ramp inflow applied for a period of time, the traffic condition will return to free-flow (F) indicated by $v_{dec} > v_{con}$. ANCONA increases on-ramp inflow by introducing a greater on-ramp inflow $\lambda_2(k)$ which is $\lambda_2(k) > \lambda_1(k)$.

Local ramp metering strategies operate the motorway traffic depending on the local measurements, therefore the controllers performance independently at different locations. In order to achieve greater efficiency, coordinated ramp metering strategies are developed to coordinate local ramp metering actions for the global motorway system rather than a single junction. Coordinated ramp metering strategies make use of measurements over a region of motorway to control all metered ramps included therein [86].

A number of coordinated ramp metering approaches have been proposed in the literature. Kotsialos and Papageorgiou [53] and Papamichail *et al.* [86] present a model

predictive framework for coordinated ramp metering based on a second order model (METANET), which considers capacity drop and transient behaviour of traffic congestion. Hegyi *et al.* [43] solve this problem based on METANET model. Gomes and Horowitz [34] propose a deterministic ramp metering optimiser based on a cell transmission model (CTM) [23] and formulate it as a linear programming. Kurzhanskiy and Varaiya [54] presente a CTM-based performance evaluation tool which is called TOPL [19] (TOPL is mainly a macroscopic performance evaluation tool which does not involve optimisation). Therefore, optimal ramp metering strategies need to be developed for motorway networks.

There are also some other rule based coordinated control systems such as HERO [85] (an extension of ALINEA) and others (see Zhao *et al.* [111] and Zhang and Levinson [108]). Rule based ramp metering control strategies are useful to a certain situation on the road. Most of the control systems are based upon a traffic model and the importance of a reliable model of traffic to effective transport management has been highlighted in Kotsialos and Papageorgiou [52]. This study adopts a model based optimiser for deriving coordinated ramp metering strategy.

1.3.2 Variable speed limits

Variable speed control schemes adjust the speed limit according to the current traffic condition with variable message signs (see Figure 1.5). The objective is to improve mobility through managing the formulation of congestion and smoothing traffic flows, as well as safety through reducing the variance in speed. Variable speed limits (VSL) can be regarded as main road metering, which reduce the incoming flow on the main

road. Moreover, reducing speed as traffic density approaching critical value also helps to prevent breakdown as empirical study shows that the journey times remain constant under VSL even with increased traffic volume and reduced speed [41]. Several simulation studies present the influence of VSL on traffic. Hadiuzzaman and Qiu [38] report VSL is an effective control method during congested periods. Hegyi *et al.* [44] test VSL on a hypothetical network based on a second-order traffic model (METANET), which shows total travel time decreased as VSL eliminate the effect of a shock wave. In addition to mobility, it is also revealed that VSL have a positive impact on safety and mobility (Lee *et al.* [58]). Empirical study has also been conducted in the Netherlands by van den Hoogen and Smulders [99], who analyse the result of VSL on the A2 Motorway in the Netherlands, which indicates VSL are useful to control traffic flow and unsafe driving behaviour.

FIGURE 1.5: Variable speed limits in operation

(Source: UK Highways Agency, Feb 14, 2007)

In this study, we develop a mathematical tool for minimising the total travel delay of a motorway system through determining a set of optimal VSL. The formulation extends [34] by considering main road variable speed control and formulating the corresponding optimisation problem as a mixed integer linear programming (MILP). This study also includes analyses on the sensitivity of the optimal control solutions with respect to different assumptions of changes in traffic characteristics under reduced speed limits.

1.3.3 Hard shoulder running

The hard shoulder is originally designed for managing incidents on motorways. Hard shoulder running (HSR) is a strategy which increases the motorway capacity by opening the hard shoulder to road users as a traffic lane during peak periods or accidents. This control can be operated temporarily through the utilisation of the variable message signs mounted on overhead gantries (see Figure 1.6). In practical applications, HSR is only used when the speed limit is reduced to at least 60 mph [4] with the consideration of safety.

Hard shoulder running has been used in a number of European motorways. A pilot scheme involving the HSR operates in Birmingham since 2006 on the M42 Motorway over a 17 km stretch between Junctions 3A and 7. The signs and signals on the managed M42 Motorway inform drivers of the speed limit in operation and the availability of hard shoulder lane. The results show that it is a cost-effective way to increase throughput along congested road sections, and an additional 15 per cent reduction in travel time is observed [100]. Cohen [21] presents the effect of using HSR on the capacity and speed

FIGURE 1.6: Operation of hard shoulder running on the M42 Motorway

(Source: ITS International 2009)

based on the traffic data collected from the A86 Motorway in Paris. Geistefeldt [31] summarises the effect of temporary hard shoulder running in the German federal state of Hesse. In addition, Samoili *et al.* [91] investigate the lane flow distribution based on the data collected from a Swiss motorway section. As discussed above, significant advances have been made by introducing HSR to make the best use of the motorway over recent decades. However, little research has been conducted in mathematical analysis of HSR. This study develops a mathematical tool for minimising the total travel delay of a motorway system through HSR.

1.4 Research Contributions

This thesis develops a mathematical tool for optimising the performance of motorway through modelling and regulating the traffic flowing on it. Unlike usual practice that relies on computational expensive microscopic simulations, the tool developed herein adopts a macroscopic representation of traffic flow which is called cell transmission model (CTM). We demonstrate how one can use traffic data collected from standard loop detection systems to develop and calibrate a CTM-based model of a specific road network. The CTM traffic model is parsimonious and piecewise linear which allows us to formulate the corresponding system optimisation as a linear programming (LP) problem or Mixed Integer Linear Programming (MILP). It is known that LP problems can be solved by established solution algorithms such as SIMPLEX or interior-point methods for the global optimal solution, and Mixed Integer Linear Programming (MILP) can be solved by using a branch-and-bound algorithm for the global optimal solution. The concept is illustrated through a real case study of UK M25 Motorway. To the best of our knowledge, this is one of the few systemic studies on motorway modelling and optimisation using real scenario data. By considering the demand and supply uncertainty, the robust ramp metering is developed. The optimal solution calculated can provide useful insights and guidance on how we should manage traffic flow on motorway in order to maximise the corresponding efficiency. The main contributions of this thesis are as follows:

1. it contributes to the development and validation quick and reliable traffic models and algorithms to process raw traffic flow data collected from motorways;

2. it contributes to the development of effective optimisation algorithms to derive optimal ramp metering based on the sensor measurements and model estimations;

3. it contributes to the sensitivity of optimal ramp control solutions with respect to different scenario configuration and model assumptions. We provide a numerical examples and results using our optimal model regarding the effectiveness of ramp metering with respect to network configuration and separation of ramps. Moreover, the sensitivity analysis of the weighting parameter in the objective function is presented;

4. it contributes to the development of the optimal control formulation and solution for variable speed limits and hard shoulder running operations. We adopt a Mixed Integer Linear Programming (MILP) approach and demonstrate the optimal speed control and hard shoulder solutions over a range of scenarios including a real case of M25 Motorway in the UK;

5. it contributes to the analyses of variables speed limits operations with different assumptions on fundamental diagram transformation. It is interesting to observe that the effectiveness of variable speed limits depends heavily on changes in capacity rather than changes in shape of the fundamental diagrams;

6. it contributes to the motorway of traffic management with the consideration of uncertainties in traffic demand and characteristics. We do not only apply existing optimisation methods (e.g. [64]; [114]; [34]), which seek optimal control strategies where the travel demands and fundamental diagrams are considered to be deterministic. The paper extends these existing optimisation formulations to incorporate the uncertainties in demand flows and the set-valued fundamental diagrams under congested situation. We present how the uncertainties of the traffic variables can be specified by using likelihood region (Ω) and how the robust optimisation can be solved an iterative two-stage solution algorithm. The robust

optimisation model for dynamic motorway traffic control is novel and has never been seen in the literature.

1.5 Thesis Outline

This thesis is organised as follows:

Chapter 2 reviews cell transmission model, and demonstrates the use of traffic data collected from standard loop detection systems for calibrating the fundamental diagram.

Chapter 3 presents an optimisation framework for motorways through ramp metering maximising. The concept is illustrated through a case study of the M25 Motorway in the UK.

Chapter 4 presents an optimisation framework with variable speed limits, hard shoulder running, ramp metering with variable speed limits, and integrated control strategies. The concept is also illustrated through a case study of the M25 Motorway in the UK.

Chapter 5 presents a robust ramp metering optimisation framework for a motorway system with uncertainties in traffic demand and characteristics. The uncertainties are quantified through set-valued functions.

Chapter 6 provides some concluding remarks and a future research plan.

Chapter 2

Modelling of Motorway Traffic

2.1 Introduction

In this chapter, we discuss modelling of dynamic road traffic flow. The chapter is organised as follows: Section 2.2 begins with an introduction of cell transmission model (CTM); Section 2.3 discusses the processing of the raw traffic data collected from standard loop detection system. Section 2.4 presents the model calibration procedure with loop detector data. Section 2.5 provides some concluding remarks.

2.2 Cell Transmission Model

The cell transmission model is a finite difference approximation of the first order LWR model proposed by Daganzo [23]. Under the cell transmission formulation, the road section is discretised into a collection of sub-sections or 'cells' as shown in Figure 2.1.

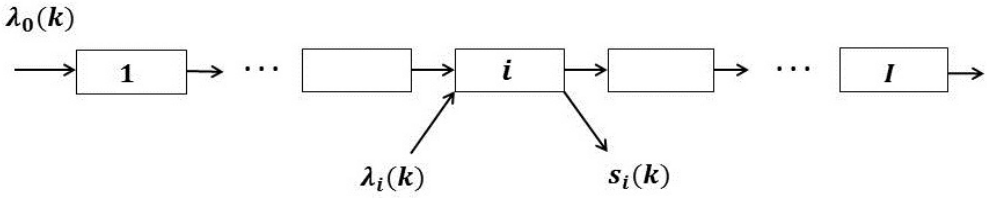

FIGURE 2.1: Discretisation of a road section

The cells are numbered from the upstream 1 to the downstream I in the figure. The first cell (main source) is connected to a main road demand $\lambda_0(k)$. Moreover, each cell i can also be associated with an external incoming flow $\lambda_i(k)$ (e.g. an on-ramp) and an external outgoing flow $s_i(k)$ (e.g. an off-ramp) at each simulation time step k. In the cell transmission formulation, traffic dynamics are characterised by flow and density. The evolution of traffic flow and density is governed by the principles of flow conservation and propagation.

Define $f_i(k)$ the traffic outflow from cell i during time step k, and hence $f_{i-1}(k)$ (outflow from upstream cell $i-1$) will be the inflow to cell i during the same time step k. The density in cell i at the following time step $k+1$ can then be updated by the conservation equation as follows:

$$\rho_i(k+1) = \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - f_i(k) + r_i(k) - s_i(k)\Big) \tag{2.1}$$

where Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. $f_i(k)$ is the traffic outflow from cell i during time step k, and $\rho_i(k)$ is the density in the cell i during time step k. $r_i(k)$ and $s_i(k)$ represent the actual on-ramp inflow and

off-ramp outgoing flow during time step k respectively.

It is noted that Equation (2.1) indeed is a discretised version of Equation (1.8). The time step size Δt is set such that $\Delta t \leq \min_i \frac{\Delta x_i}{v_i}$, where $\min_i \frac{\Delta x_i}{v_i}$ refers to the smallest ratio of cell length to the associated free flow speed along the section. The above condition is known as the Courant-Friedrichs-Lewy (CFL) condition [22], which is used to ensure the numerical stability and non-negativity of traffic quantities by constraining the traffic to not travel further than the length of the cell in one simulation time step.

By assuming a piecewise linear fundamental diagram (see Figure 2.2), the cell transmission rule calculates the outflow from cell i within time step k, which can be updated under the given cell density as follows:

$$f_i(k) = \min\{v_i \rho_i(k),\ \ Q_i,\ \ Q_{i+1},\ \ w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big)\ \ \} \tag{2.2}$$

where $f_i(k)$ is the traffic outflow from cell i during time step k, and $\rho_i(k)$ is the density in the cell i during time step k. The notations v_i, w_{i+1}, and $\bar{\rho}_{i+1}$ are free flow speed, shockwave speed and jam density respectively. The notation Q_i denotes the capacity flow at cell i which corresponds to the maximum flow that can leave cell i, and Q_{i+1} represents the capacity flow at cell $i+1$ which corresponds here to the maximum flow that can enter cell $i+1$. The inclusion of both capacity flows at adjacent cells is due to the consideration of a non homogeneous section where different locations hold different capacities.

Equation (2.2) can be regarded as a piecewise linear approximation of Equation (1.9). When there is no congestion, the traffic moves from one cell to the next cell at free

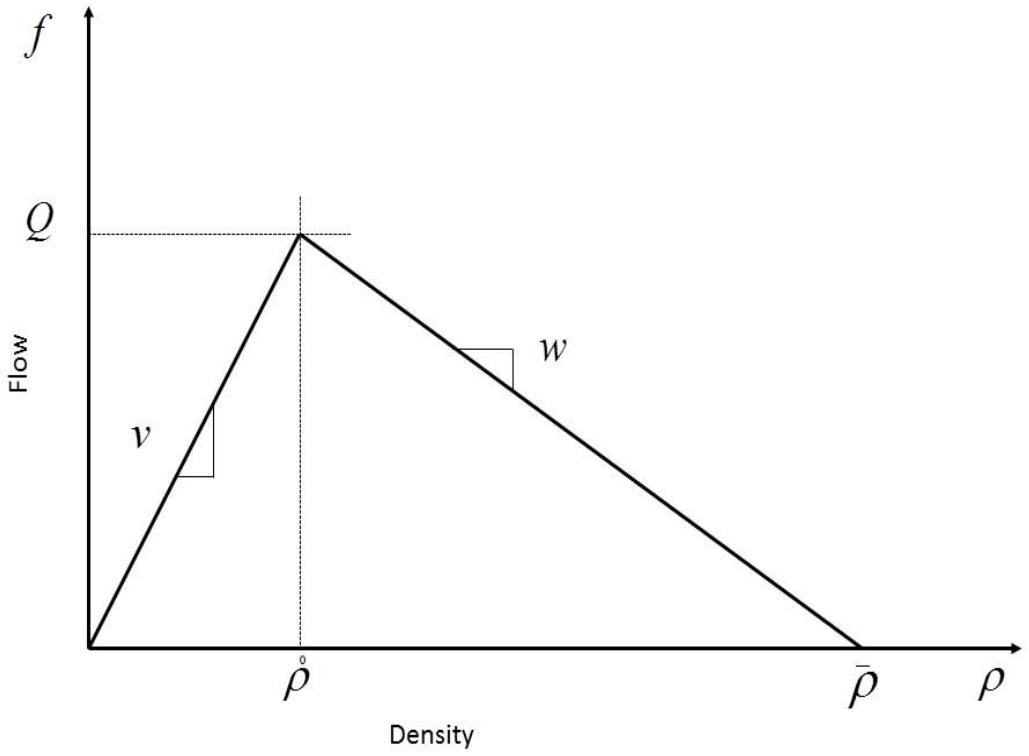

FIGURE 2.2: Triangle shaped fundamental diagram

flow speed v_i. The quantity $\left(v_i \rho_i(k) \right)$ represents the total traffic flow travels from cell i to cell $i+1$. The notation w_{i+1} denotes the backward shockwave speed specified by the fundamental diagram at the downstream cell $i+1$, and $\bar{\rho}_{i+1}$ is the jam density at cell $i+1$. The quantity $\Big(w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big) \Big)$ specifies the available space for incoming traffic at the downstream cell $i+1$ during time step k. The above formulation covers both congested and uncongested regimes. It is known that the minimum operator in Equation (2.2) can be formulated as Linear Programming (Lo [64]; Ziliaskopoulos [114]; Lo [65]; Gomes and Horowitz [34], and others). It will be discussed in Chapter 3.

The form of outflow Equation (2.2) is due to the assumption of the piecewise linear fundamental diagram shown in Figure 2.2. This linearity assumption facilitates the

development of several efficient algorithms for cell based estimation and optimisation (see examples: Daganzo [23]; Gomes and Horowitz [34]; Kurzhanskiy and Varaiya [54]). However the outflow function can be generalised as:

$$f_i(k) = \min\{S_i(k),\ R_{i+1}(k)\} \tag{2.3}$$

where $S_i(k)$ is the send flow function, which represents the traffic demand advancing from upstream i to downstream $i+1$. $R_{i+1}(k)$ is regarded as the receive flow function, which means the available space in the downstream cell $i+1$. This general form of outflow allows the different forms of the fundamental diagram used in the CTM including the 'inversed-λ' shaped fundamental diagram, which also exhibits a capacity drop (capacity drop is known as a reduction in discharge flow after queue formation observed at downstream of an active bottleneck [17]) as non-linear ones like Greenshields [36]; Underwood [98]; Edie [29].

2.3 Traffic Data

In this study, data for calibrating the traffic model are obtained through the Motorway Incident Detection and Automatic Signalling (MIDAS) double-loop detection system. This section describes how to process the raw MIDAS data. The MIDAS dataset contains traffic flow, occupancy (e.g. percentage of time that the loop detector is occupied in each time interval), and speed [48], which are measured directly from double-loop detectors. The MIDAS system consists of inductive loop detectors at 500-metre intervals installed on all lanes. The data are processed and stored in 1-min intervals in csv files. You can

request the access to the Highways Agency MIDAS data via the Mott MacDonald help desk by providing a brief description of your research project and some contact details.

2.3.1 Flow

Flow (f) is defined as the number of vehicles passing a point in a given time period. The MIDAS system classifies flow data into the following two kinds:

- Flow over different categories: f_c denotes the average flow of category c. The traffic flow data are categorised under MIDAS into four categories according to the detected vehicle length:

  1. Category 1: Vehicles with a length less than 5.2 metres (e.g. Car or small van);
  2. Category 2: Vehicles with a length between 5.2 metres and 6.6 metres (e.g. Large van);
  3. Category 3: Vehicles with a length more than 6.6 metres and less than 11.6 metres (e.g. Rigid HGV);
  4. Category 4: Vehicles with a length greater than or equal to 11.6 metres (e.g. Articulated HGV).

- Flow on different lanes: f_l denotes the average flow on lane l, and the maximum number of lanes is 7. Then the flow of the given road section can be calculated from lane flow or category flow by using the following formula:

$$f = \sum_{l=1}^{L} f_l = \sum_{c=1}^{4} f_c \tag{2.4}$$

where L refers to the number of lanes at the corresponding road section, and $c = 1, 2, 3, 4$ refers to the vehicle category as discussed above.

2.3.2 Occupancy

Occupancy (o) is defined as the proportion of time during which a single point on a road is covered by vehicles where $0 \leq o \leq 1$, where 0 means there is no vehicle occupied the loop detector, and 1 means the loop detector is fully occupied during the interval. Then the mean occupancy of the given road section can be calculated as:

$$o = \sum_{l=1}^{L} o_l \Big/ L \tag{2.5}$$

where o_l refers to the occupancy on the lane l of the corresponding road section, and L is the number of lanes at the corresponding road section.

2.3.3 Speed

Speed ($\overline{v}$) is defined as the distance of a vehicle travelling on the road per unit time. With the double-loop configuration, the speed of each vehicle can be measured directly as the spacing between two loop detectors (which is 5 metres) divided by the time gap between the vehicles' detection at two loop detectors [41]. Then the mean speed of the given road section can be calculated as:

$$\overline{v} = \sum_{l=1}^{L} v_l f_l \Big/ \sum_{l=1}^{L} f_l \tag{2.6}$$

where v_l refers to the speed on lane l of the corresponding road section, f_l denotes the average flow on lane l, and L is the number of lanes at the corresponding road section.

2.3.4 Density

Density (ρ) reflects the proximity of vehicles on the road. It is defined as the number of vehicles occupying a given length of lane or road section at a specified time instant. Like all other fixed loop detection information systems, MIDAS can only observe temporal occupancy but not spatial density, which will be required for calibrating the fundamental diagram. Following Papageorgiou and Vigos [79] and others, density can be calculated from the occupancy by using the following formula (see Appendix A):

$$\rho = \frac{o}{\overline{L}_v + L_d} \tag{2.7}$$

where o is the measured occupancy ($0 \leq o \leq 1$) from the corresponding detector, and L_d is the length of the detector (which is taken as 2 metres), and $\overline{L}_v$ is an estimation of average vehicle length (in metres) passing the corresponding location.

The length of the corresponding vehicle can be estimated by multiplying its speed with the average duration of the 'on' time (i.e. the period when the loop is occupied by a vehicle) of the loops. The average vehicle length at each 1-min interval is estimated as:

$$\overline{L}_v = \sum_{c=1}^{4} f_c \overline{L}_{vc} \Big/ \sum_{c=1}^{4} f_c \tag{2.8}$$

where $c = 1, 2, 3, 4$ refers to the vehicle category under MIDAS as discussed in Section 2.3.1; f_c is flow of vehicles in category c measured during the corresponding time interval;

$\overline{L}_{vc}$ is the average vehicle length (in metres) in category c where $\overline{L}_{v1} = 5.2$ metres, $\overline{L}_{v2} = 5.9$ metres, $\overline{L}_{v3} = 9.1$ metres, and $\overline{L}_{v4} = 11.6$ metres are taken in this study [5].

2.4 Model Calibration

Traffic herein is described in terms of the basic quantities such as flow, speed and density, which are introduced in the previous Section 2.3, which underlies the fundamental diagram. Measurements from loop detectors allow fundamental diagram to be estimated for the corresponding road section. In this section, we will look at how to calibrate piecewise linear fundamental diagram in the CTM with the loop detector data. Calibrating CTM will be equivalent to determining the underlying fundamental diagram. For the piecewise linear fundamental diagram in the CTM, we can divide and calibrate the fundamental diagram in three components: free flow line, capacity, and congested line. The associated critical density and jam density can be derived accordingly after obtaining the three lines.

As an illustration, Figure 2.3 shows a scatter plot of flow-density data collected at detector station (4936A) on the M25 Motorway (direction: clockwise) in the United Kingdom. The data are collected over 2 weekdays (2 September 2014 (Tuesday) and 3 September 2014 (Wednesday)) through MIDAS [1] loop detection system in the United Kingdom. To reduce noise in the dataset, the data are processed into 5-min averages (MIDAS stores 1-min data). The detector station (4936A), which consists of 6 lanes, is located at the downstream of the on-ramp at Junction 14 and the upstream of the off-ramp at Junction 15. Junction 14 connects to London Heathrow (LHR) Airport and Junction 15 is a busy interchange with the M4 Motorway in the west of the Greater London Area.

This study adopts a calibration procedure developed by Dervisoglu *et al.* [26]. Details of calibration procedure is presented below.

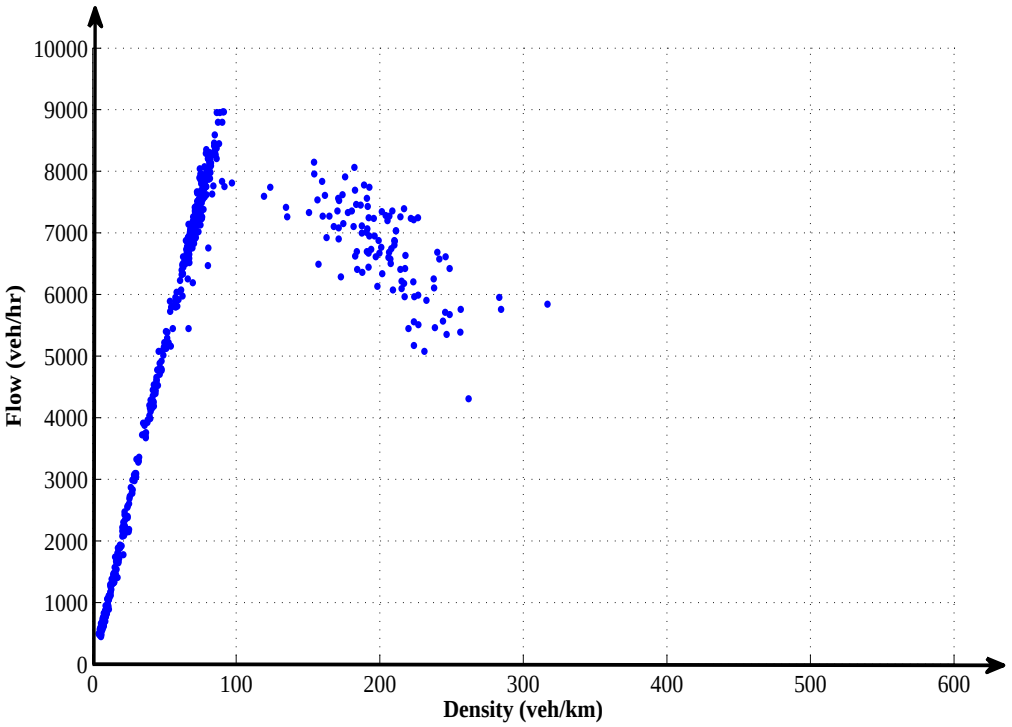

FIGURE 2.3: Empirical scatter plot of traffic flow and density

(Detector Station: 4936A on M25-clockwise [2 and 3 September 2014])

2.4.1 Step 1: Free flow line

For each detector station, a speed threshold for extracting data points in the free flow state from the dataset should be defined. The speed threshold can be a fixed value (e.g. 70 mile/hr, as adopted in Bickel *et al.* [10] and Dervisoglu *et al.* [26]). Alternatively, the speed threshold can also be set as a percentile (e.g. 85^{th} percentile) of all measured speeds. From a statistical perspective, the latter is a more robust classification, as it reduces the impact of potential outliers due to extreme conditions such

as weather, incidents or detector failure. Moreover, the corresponding speed threshold is site-dependent, which is preferable as different locations can have different free flow speeds due to different speed limits.

With the defined speed threshold, a data point is regarded as free flow if its associated speed is higher than that of the speed threshold. The free flow part (left-hand side) of the fundamental diagram typically exhibits a linear relationship with little variability from the empirical findings [26]. Then the extracted flow density data are fitted by using a standard linear regression method, which was regarded as the free flow line. The slope of the free flow line gives the free flow speed (v) of traffic at that location. It often requires this free flow line to pass though the origin of the flow-density plane reckoning that the flow value should be zero when the density is zero (see Appendix B for a short note on the constrained regression). The regression line for the free flow speed based on the collected data is shown in Figure 2.4.

2.4.2 Step 2: Capacity

The capacity (Q) is taken as the maximum observed flow values at the detector station observed over a period of time. The capacity line based on the collected data is shown in Figure 2.5. Given capacity, the corresponding critical density ($\mathring{\rho}$) is determined as:

$$\mathring{\rho} = Q / v \tag{2.9}$$

where free flow speed (v) is determined in Step 1. Recent empirical findings (Brilon *et al.* [13]; Dervisoglu *et al.* [26] and Chow *et al.* [20]) suggest that the capacity is

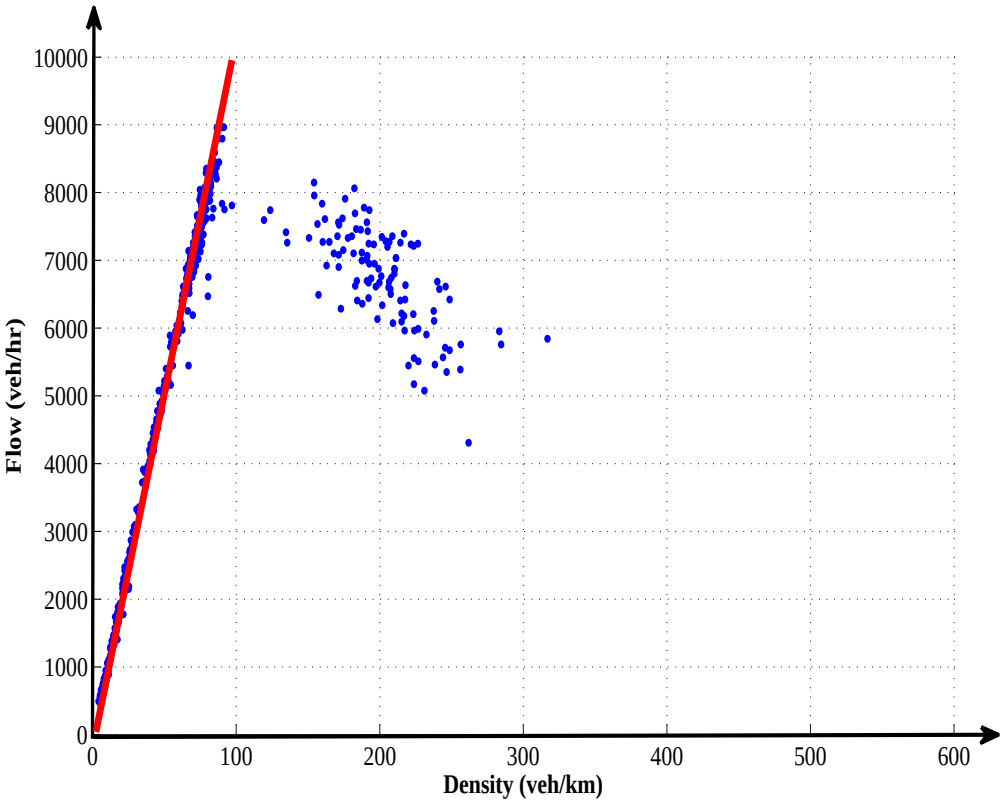

FIGURE 2.4: Estimation of free flow speed

(Detector Station: 4936A on M25-clockwise [2 and 3 September 2014])

a random variable which may vary depending on various external conditions such as weather, composition of traffic, and approaching traffic states. This stochastic variation of capacity will be discussed in Chapter 5.

2.4.3 Step 3: Congested line

After extracting the free flow data following Step 1, any remaining data points are regarded as congested data if its density is higher than the critical density determined in Step 2. Similar to capacity, the congested data often exhibits a high degree of variability, which is due to the heterogeneity in vehicular speeds and drivers' behaviour under congested conditions (Ngoduy and Liu [74]; Ngoduy [75]). The variability in congested

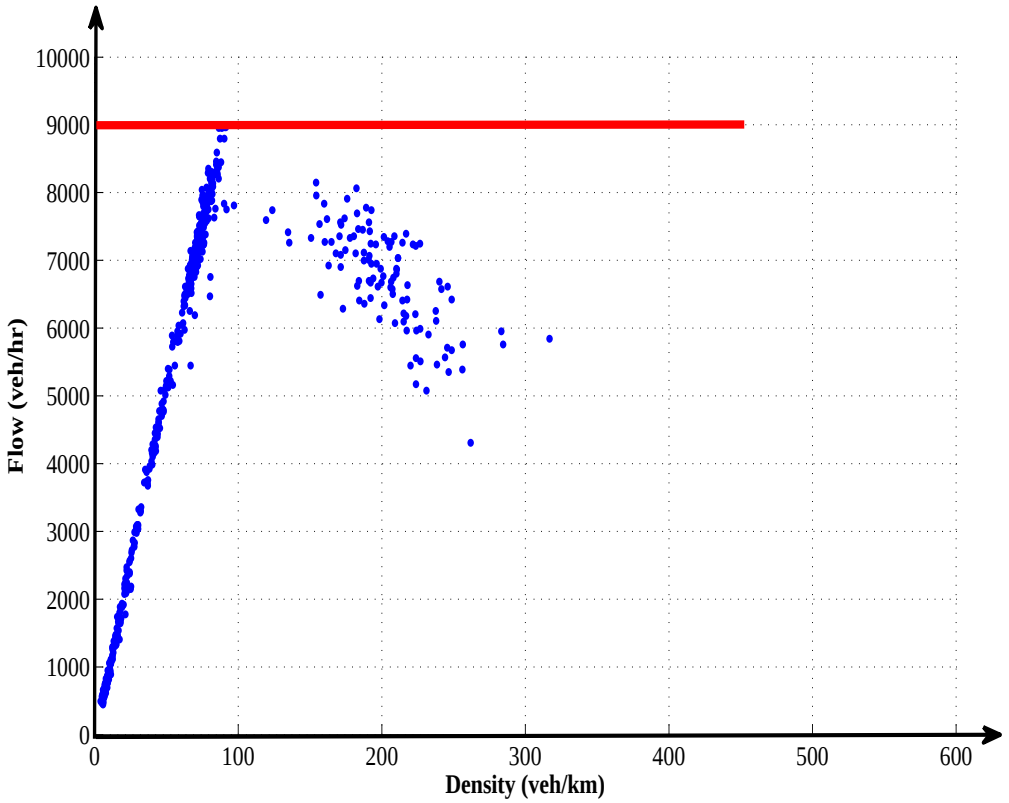

FIGURE 2.5: Estimation of capacity

(Detector Station: 4936A on M25-clockwise [2 and 3 September 2014])

data means fitting them with a function challenging. Further discussion of the variability of the congested data will be presented in Chapter 5. The congested data can be fitted by either linear or non-linear flow-density functions.

Greenshields *et al.* [37]; Greenberg [35]; Underwood [98] and many others fit the congested data with a nonlinear flow-density function. Heydecker and Addison [46] recently conduct an empirical study with the United Kingdom M25 Motorway data where they analyse and investigate the goodness-of-fit of a number of models. Interestingly, it is found that Underwood's model can indeed produce a better fitting result than the traditional Greenberg's and Edie's models for congested traffic data. Moreover, it is found

that the convex region in Underwood's model helps to reproduce complex dynamic traffic phenomenon like the stop-and-go waves (Heydecker and Addison [46]; Carey and Bowers [14]). Furthermore, Underwood's model does not have a jam density that will drive the speed to be zero. However, this can easily be remedied by truncating the flow-density function to a user-defined jam density. In fact, unlike urban networks, it is extremely rare to observe a data point that is sustainably associated with a zero speed on motorways. As reported in Heydecker and Addison's empirical study, there is no significant effect on traffic state estimation even when the model does not have a finite jam density.

If the congested flow-density data are fitted by a linear function, then the slope of the regression line will give the shockwave speed (w), which represents the average speed of backward propagation of congestion. Jam density is the intersection of congested line and density (horizontal) axis. The congested line can be derived by applying an ordinary least square estimation on the congested dataset. To construct a closed fundamental diagram, the congested line should be constrained such that it will pass through the capacity point ($\mathring{\rho},\ Q$). Should this constraint be removed, the corresponding fundamental diagram will become 'reverse-λ' shaped with a discontinuity at ($\mathring{\rho},\ Q$). This discontinuity can be regarded as a capacity drop.

As an alternative to ordinary regression, Dervisoglu *et al.* [26] propose a trained regression approach to fit congested data. Trained regression is more robust than ordinary regression in statistical science because it reduces the influence of outlier data. Different from the ordinary regression, which looks at the expected values of the regressors, trained regression considers the quantiles of the regressors and hence reduces the influence of outlying data due to slow moving heavy vehicles, unstable stop-and-go motions,

and hardware failure.

Following the trained regression developed by Dervisoglu *et al.* [26], the congested data along the density (horizontal) axis can be divided into a series of non-overlapping bins ($\mathbf{B} = \{B_1,\ B_2\ ...\ B_i\}$), in which each bin B_i contains a certain number of data points (e.g. N_i data points in B_i). Horizontally, each bin is represented by a 'BinDensity', which is the average of all the density values in the bin. Vertically, each bin is represented by a 'BinFlow', which is the largest non-outlier flow value among the flow values in the bin. Given the density-flow data points in the bin $B_i = \{(\rho_{i1}, f_{i1}),\ (\rho_{i2}, f_{i2})\ \ \ldots\ \ (\rho_{iN_i}, f_{iN_i})\}$. Then, the 'BinDensity' and 'BinFlow' can be determined for the bin B_i as:

$$\rho_{bi} = \sum_{n=1}^{N_i} \rho_{in} \Big/ N_i \tag{2.10}$$

$$f_{bi} = \max_{f_{in}} \{f_{in} | B_i, f_{in} \in f_{in} < Q3_i + 1.5IQR_i\} \tag{2.11}$$

where $Q3_i$ is the 75^{th} percentile flow in the bin B_i. The notation IQR_i represents the inter-quantile range, which is defined as the difference between 25^{th} percentile and 75^{th} percentile flow in the bin B_i with the number of data points (N_i).

For example, given 10 density-flow data points in the bin B_j. The 10 density values and the corresponding 10 flow values as shown in Table 2.1. Then the mean of those 10 density values is 63.9 as follows:

$$\text{BinDensity} = \frac{\sum_{n=1}^{10} \rho_{jn}}{10} = 63.9 \tag{2.12}$$

Table 2.1: Density-flow data points (ρ_{jn}, f_{jn}) in the bin B_j

n	1	2	3	4	5	6	7	8	9	10
ρ_{jn}	62	65	82	70	72	57	64	60	61	46
f_{jn}	6840	7240	8820	7320	7440	6180	7080	6600	6360	4920

The 'BinFlow' is set to be 7440 as shown in Equation 2.13, which is the 'maximum non-outlier' within the bin B_j.

$$\begin{aligned} f_{bj} &= \max_{f_{jn}}\{f_{jn}|B_j, f_{jn} \in f_{jn} < Q3_j + 1.5IQR_j\} \\ &= \max_{f_{jn}}\{f_{jn}|B_j, f_{jn} \in f_{jn} < 7320 + 1.5 \times 960\} \\ &= \max_{f_{jn}}\{f_{jn}|B_j, f_{jn} \in f_{jn} < 8760\} \\ &= 7440 \end{aligned} \tag{2.13}$$

Moreover, Figure 2.6 shows how to define the 'BinDensity' and 'BinFlow' and calibrate the wave speed based on the collected data. The density-flow data points (measurements) as shown in Figure 2.6 in a series of blue points. The left two vertical solid lines represent the lower and upper bound of the density of the bin (B_1 is shown in the figure). The red points between two vertical solid lines represent the data in the bin (B_1), specifically the 'BinDensity' and 'BinFlow' on the horizontal and vertical axis respectively.

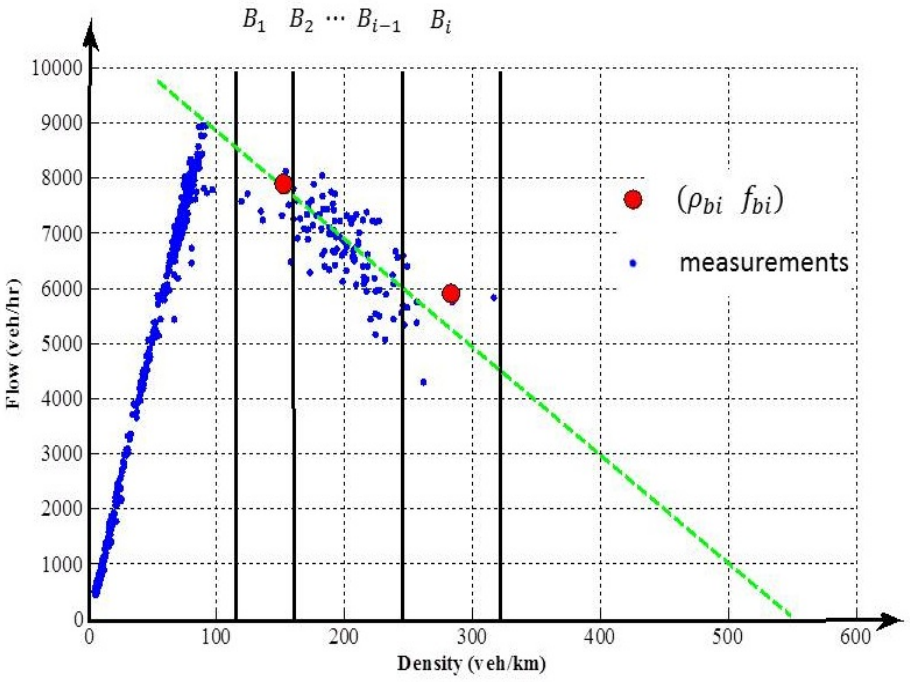

FIGURE 2.6: Trained regression

2.4.4 Review of procedure

Figure 2.7 shows a flow-density scatter plot at detector station (4936A) on the orbital M25 Motorway in the UK on 2 and 3 September 2014. The data are collected over 2 weekdays (consists of 6 lanes) which is the same dataset used in Figure 2.3. The data are first classified into 'free flow' and 'congested' portions based on a percentile-based classification. A data point is recognised as 'free flow' if its speed is higher than the 85^{th} percentile of all measured speeds. Figure 2.7 shows the flow-density data in free flow exhibiting a strong linear relationship which can be fitted by standard linear regression. A constraint is added for the free flow data regression line, which needs to pass through the origin. The slope of the free flow line gives the free flow speed v, which is 100 km/hr. The capacity Q is determined as the maximum flow values observed, which is 9000 veh/hr. Finally, congested line pass through the capacity point $(\mathring{\rho},\ Q)$,

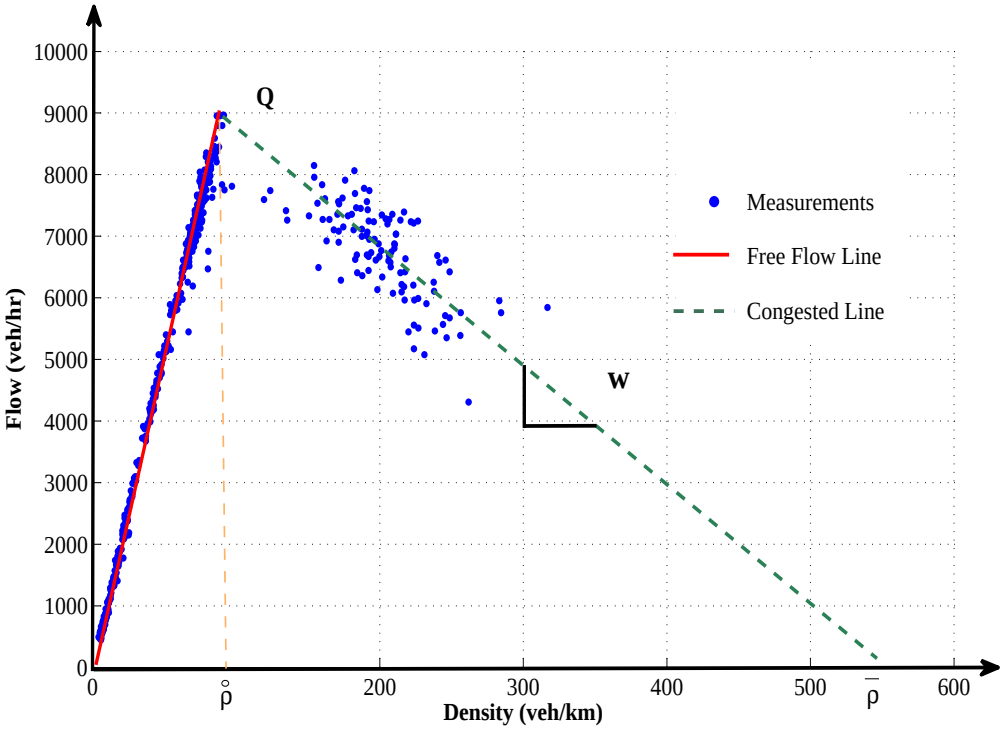

FIGURE 2.7: Calibrated fundamental diagram for CTM

(Detector Station: 4936A on M25-clockwise [2 and 3 September 2014])

and the slope of the congested line gives the wave speed -19 km/hr.

The free flow data can generally be fitted well with the linear regression function. The free flow line (the red solid line in Figure 2.7) tends to overestimate the flow values as they approach the capacity. It is due to the speed reduction as the traffic state approaches congestion, while this speed reduction cannot be captured by the linear regression function. The consequence of this will be an underestimate of the value of the critical density and hence the congestion will be predicted to onset earlier than it actually does. To improve the estimation, one can adopt a modified fundamental diagram which includes a non-linear region near the capacity (see for example, Skabardonis and Geroliminis [93]).

The green dash line in Figure 2.7 shows the congested line. Data with density values

higher than the critical density are classified as congested, which are characterised by the fact that an increase in density will induce a decrease in flow. The intersection of congested line and horizontal line (density axis) shows the jam density 540 veh/km. The road section (detector station 4936A) contains 6 lanes. If we assume the average legnth of the vehicle is 10 metres, then the jam density should be $(1000 \times 6)/10 = 600$ vehicles. The calibrated jam density (540 veh/km) closes to the theoretical jam density (600 veh/km). The congested data often exhibit a high degree of variability due to a number of reasons such as heterogeneity in vehicular speeds and drivers' behaviour under congested condition (Ngoduy and Liu [74]; Ngoduy [75]). Some recent studies propose the use of the set-valued function for regressing the congested data (e.g. Kurzhanskiy and Varaiya [55]) while the discussion will be presented in Chapter 5.

2.5 Summary

This chapter consists of three parts. In the first part, we introduce the first order model CTM in detail. In the second part, how to process the raw traffic data collected from the MIDAS dataset is introduced, especially how to calculate the density based on the measured occupancy. Then in the third part, how CTM can be calibrated with real traffic data (MIDAS) is presented. The calibration procedure presented here consists of three main steps: determining the free flow line, capacity, and congested line. The free flow line can be determined with flow-density data under free flow which exhibits a linear relationship. Capacity is taken as the maximum flow value observed over a period of time, and trainted regression is used for the congested part. The modelling and optimisation framework for ramp metering to be presented in following chapter.

Chapter 3

Optimisation of Ramp Metering

3.1 Introduction

We start with the optimisation model of motorway traffic through ramp metering strategy. Ramp metering is a motorway control method to avoid onset of congestion through limiting the access of ramp inflows into the main road of the motorway. The optimisation model of ramp metering is developed based upon cell transmission model (CTM). With the piecewise linear structure of CTM, the corresponding motorway traffic optimisation problem can be formulated as a linear programming (LP) problem. It is known that LP problem can be solved by established solution algorithms such as SIMPLEX or interior-point methods for the global optimal solution. The commercial software (CPLEX) is adopted in this study to solve the LP problem within reasonable computational time. The concept is illustrated through a case study of the United Kingdom M25 Motorway. The optimal solution provides useful insights and guidances on how to manage motorway traffic in order to maximise the corresponding efficiency.

This chapter is organised as follows: Section 3.2 introduces the performance indicators that are used as objective function for optimal ramp metering. Section 3.3 presents the formulation that is used to derive the optimal ramp metering for maximising the efficiency of the motorway traffic. Section 3.4 explores the sensitivity analysis of the distance between on-ramp and off-ramp. Section 3.5 explores the sensitivity analysis of the balance parameter η between the main road delay and ramp delay. Section 3.6 presents the implementation of CTM and the optimal ramp metering will be illustrated through a case study of the United Kingdom M25 Motorway. Section 3.7 provides some concluding remarks.

3.2 Objective Function

A standard optimisation formulation consists of three components: objective function, constraints, and decision variables. Formulating an optimisation problem for motorway operations first requires defining a sensible performance indicator as the objective function. Some typical performance indicators include VDT (vehicle-distance-travelled), VHT (vehicle-hours-travelled), and Delay (see Kurzhanskiy and Varaiya [54]).

1. VDT (unit: [veh-km]) is defined as the sum of the products of the vehicle with the associated distance travelled. The VDT is a measure of the throughput of a road section (or cell) during a particular time period. The higher the VDT is, the more productive the system is, as that implies more traffic can be served in a given time. Given the traffic flow $f_i(k)$ at a road section i of length Δx_i during the time interval k of length Δt, the associated VDT is calculated as:

$$\mathrm{VDT}_i(k) = f_i(k)\Delta x_i \Delta t \tag{3.1}$$

TTD (total-travel-distance) refers to the sum of VDT over a road section during a particular time period. Following the definition of the VDT above, TTD is calculated as:

$$\begin{aligned} \mathrm{TTD} &= \sum_{i=1}^{I}\sum_{k=1}^{K} \mathrm{VDT}_i(k) \\ &= \sum_{i=1}^{I}\sum_{k=1}^{K} f_i(k)\Delta x_i \Delta t \end{aligned} \tag{3.2}$$

2. VHT (unit: [veh-hr]) is defined as the sum of the products of the vehicle with the associated travel time. The VHT is a measure of the efficiency of a road section during a particular time period. The lower VHT is, the more efficient the system is, as that implies traffic can be served in less time. Given the traffic density $\rho_i(k)$ at cell i of length Δx_i during time interval k of length Δt, the associated VHT is calculated as:

$$\mathrm{VHT}_i(k) = \rho_i(k)\Delta x_i \Delta t \tag{3.3}$$

TTT (total-travel-time) refers to the sum of VHT over a road section during a particular time period. Following the definition of the VHT above, TTT is calculated as:

$$\begin{aligned} \text{TTT} &= \sum_{i=1}^{I}\sum_{k=1}^{K} \text{VHT}_i(k) \\ &= \sum_{i=1}^{I}\sum_{k=1}^{K} \rho_i(k)\Delta x_i \Delta t \end{aligned} \tag{3.4}$$

It is noted that the ratio of VDT to VHT gives the mean speed of the traffic in the road section.

3. Delay (unit: [veh-hr]) is one of the most effective ways to evaluate congestion level, which can be derived from VDT and VHT as:

$$\begin{aligned} d_i(k) &= \text{VHT}_i(k) - \frac{\text{VDT}_i(k)}{v_i} \\ &= \rho_i(k)\Delta x_i \Delta t - \frac{f_i(k)\Delta x_i \Delta t}{v_i} \end{aligned} \tag{3.5}$$

where v_i is the free-flow speed of cell i.

If the delay function is adopted as the ramp metering objective fucntion, the throughput will be increased. However, long queues may be created at some on-ramps due to the objevtive function just considers the delay on the main road. The equity of ramp metering is considerd in many literatures along with ramp metering efficiency. Zhang and Levinson [109] adopt the weighted travel time as the objective function to balance the efficiency and equity of ramp metering. Zhang and Levinson [108] suggest that the leaset equitable one is the most efficient ramp control. Levinson and Zhang [59] evaluate the data of eight weeks collected from the case study with and without ramp metering for several representative motorways in Twin cities. By considering various performance measures, they found

that the ramp delay need to be considered even at the expense of overall motorway efficiency when the objective balance efficiency and equity of ramp meters. In this study, the total system delay (TSD) is adopted as the objective function for the ramp metering optimisation problem due to the ramp delay includs in the objective function.

4. TSD (total-system-delay) is defined as the sum of total main road delay and ramp delay in this study. For a road section during a particular time period, it is calculated as:

$$\text{TSD} = \sum_{i=1}^{I}\sum_{k=1}^{K} d_i(k) + \eta \sum_{j=1}^{J}\sum_{k=1}^{K} l_j(k)\Delta t \tag{3.6}$$

where the first term $\sum_{i=1}^{I}\sum_{k=1}^{K} d_i(k)$ represents the total main road delay over the whole road section (number of cells equals to I) during a particular time period (number of simulation time steps equals to K). The second term $\sum_{j=1}^{J}\sum_{k=1}^{K} l_j(k)\Delta t$ represents the total ramp delay on all on-ramps (number of on-ramps equals to J) during a particular time period. The parameter η adjusts the balance between the main road delay and ramp delay. The sensitivity ananlysis of parameter η is presented in Section 3.5.

3.3 Ramp Metering Formulation

Ramp metering is one of the most widely used control methods, which aims to improve the throughput on the main road through limiting access of traffic from on-ramps. This section presents a model-based approach, which enhances the efficiency of motorway operations through modelling and regulating the traffic flowing on it. There are some important applications of the first-order macroscopic model (CTM) which were presented in the literature for traffic modelling and management. Lo [64] presents a novel traffic signal control formulation, which developed through a mixed integer programming technique. Ziliaskopoulos [114] presents a system optimum dynamic traffic assignment problem as a linear programming problem based on CTM. Gomes and Horowitz [34] show that the linearity of CTM enables formulating optimal ramp control problems as a linear programming (LP) problem. In the CTM formulation, traffic dynamics are characterised by flow and density in each cell at each time step. The evolution of traffic flow and density is governed by the principles of flow conservation and propagation. It is convenient (e.g. Gomes and Horowitz [34]) to specify the exit flow $s_i(k)$ through a split ratio β_i, where $0 \leq \beta_i \leq 1$, to represent the proportion of traffic leaving the system through the sink link i during time step k as shown in Figure 3.1.

Following this specification, the relationship between exit flow and outflow can be defined as follows:

$$s_i(k) = \beta_i\big(s_i(k) + f_i(k)\big) \tag{3.7}$$

where $s_i(k)$ is the exit flow from cell i, and $f_i(k)$ denotes the outflow from cell i which is the flow remaining in the system after deducting $s_i(k)$. The notation $\rho_i(k)$ represents the density in cell i at time step k, and β_i is the split ratio of cell i.

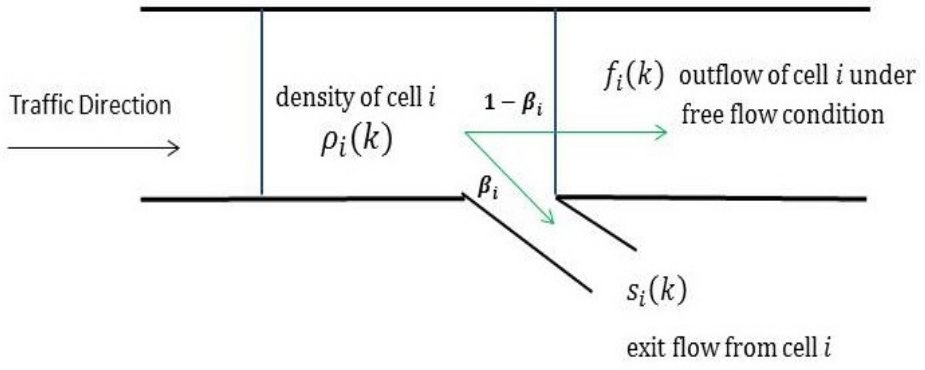

FIGURE 3.1: Specification of outflow ratio

Equation 3.7 can be rearranged to obtain an expression of the exit flow as follows:

$$s_i(k) = f_i(k)\big(\beta_i/\bar{\beta}_i\big) \tag{3.8}$$

where $\bar{\beta}_i = 1 - \beta_i \geq 0$. Then $s_i(k) + f_i(k)$ can be rewrote as follows:

$$\begin{aligned} s_i(k) + f_i(k) &= f_i(k)\big(\beta_i/\bar{\beta}_i\big) + f_i(k) && (3.9)\\ &= f_i(k)\big((\beta_i + 1 - \beta_i)/\bar{\beta}_i\big) && (3.10)\\ &= f_i(k)/\bar{\beta}_i && (3.11) \end{aligned}$$

Substituting Equation 3.11 into Equation 2.1 gives:

$$\rho_i(k+1) = \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - f_i(k)/\bar{\beta}_i + r_j(k)\Big) \tag{3.12}$$

where Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. The notations $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i

respectively, and $r_j(k)$ is the actual ramp inflow that want enter the main road.

Given the cell density, the CTM calculates the outflow from cell i within time step k by assuming a piecewise linear fundamental diagram as shown in Equation 2.2. If we assume the exit flow is located at the end of the cell i, the Equation 2.2 can be revised as follows:

$$f_i(k) = \min\{\rho_i(k)v_i\bar{\beta}_i,\ \ Q_i,\ \ Q_{i+1},\ \ w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big)\ \ \} \tag{3.13}$$

where notations v_i, w_{i+1}, and $\bar{\rho}_{i+1}$ denote free flow speed, shockwave speed and jam density respectively. The notation Q_i is the capacity flow at cell i which corresponds to the maximum flow that can leave cell i; Q_{i+1} is the capacity flow at cell $i+1$ which corresponds here to the maximum flow that can enter cell $i+1$.

The minimum operator in Equation 3.13 can be reformulated as the following linear programming (see Lo [64]; Gomes and Horowitz [34]):

$$\begin{aligned} &\min \quad \big(- f_i(k) \ \big) \\ &\text{subject to:} \\ &f_i(k) \quad \leq \quad \rho_i(k)v_i\bar{\beta}_i \\ &f_i(k) \quad \leq \quad Q_i \\ &f_i(k) \quad \leq \quad Q_{i+1} \\ &f_i(k) \quad \leq \quad w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big) \end{aligned} \tag{3.14}$$

The above set of expressions (Equation 3.14) form the core of the optimisation problem. This linearity facilitates the development of several efficient algorithms for cell based

estimation and optimisation (see examples, Ziliaskopoulos [114]; Sun *et al.* [96]; Gomes and Horowitz [34]; Sumalee *et al.* [95]).

In original CTM model, the traffic flow is determined by the available space of upstream and downstream. However, the linearisation of the original non-linear flow propagation formulation caused the holding back problem [27] in the earliest analytical formulation [114]. Because it does not require the solution to fall on the fundamental diagram. Therefore, the vehicles will be held at some upstream cells even there is enough space in the downstream cells. Gomes and Horowitz [34] introduce total-travel-distance (TTD) to the objective function to eliminate the holding-back problem. Therefore, the delay is adpted as the objective function because TTD is included. The motorway traffic optimisation without ramp metering can be formulated as follows:

$$\min_{\mathbf{c}} \; Z = \sum_{i=1}^{I} \sum_{k=1}^{K} \left(\rho_i(k) \Delta x_i \Delta t - \frac{f_i(k) \Delta x_i \Delta t}{v_i} \right) \tag{3.15}$$

subject to:

$$\rho_i(k+1) = \rho_i(k) + \frac{\Delta t}{\Delta x_i} \left(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k) \right) \tag{3.16}$$

$$f_i(k) \leq \rho_i(k) v_i \bar{\beta}_i \tag{3.17}$$

$$f_i(k) \leq Q_i \tag{3.18}$$

$$f_i(k) \leq Q_{i+1} \tag{3.19}$$

$$f_i(k) \leq w_{i+1} \left(\bar{\rho}_{i+1} - \rho_{i+1}(k) \right) \tag{3.20}$$

$$r_j(k) = \lambda_j(k) \tag{3.21}$$

where Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. The notations v_i, w_{i+1}, and $\bar{\rho}_{i+1}$ are free flow speed, shockwave speed and jam density respectively. The notation Q_i is the capacity flow at cell i which corresponds to the maximum flow that can leave cell i, and Q_{i+1} is the capacity flow at cell $i+1$ which corresponds here to the maximum flow that can enter cell $i+1$. $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. $\lambda_j(k)$ denotes the traffic demand that wants to enter the system through on-ramp j during time step k, and $r_j(k)$ is the actual demand that enters the system.

The constraint set (3.16 - 3.20) equivalents to the CTM as shown by Lo [64]; Gomes and Horowitz [34] and others. The holding back problem [114] is addressed here due to the $[-f_i(k)]$ term in the objective function which will maximise the outflow from each cell i over time step k. As a consequent, at least one of the constraints (3.17 – 3.20) must be binding. Constraints (3.17) and (3.18) can be regarded as the (demand) limitations on outflow under free flow condition, while constraints (3.19) and (3.20) can be regarded as the (supply) limitations on outflow under congested condition. Constraint (3.16) is the conservation equation to update the density in cell i for next time step $k+1$. Moreover, the constraint (3.21) shows the ramp inflow equals to the actual ramp demand, which means there is no ramp control is applied on the road.

The formulation (3.15 - 3.21) is generic and applicable to the general case. Additional constraints may be added for specific applications. For ramp metering, this study adopts total-system-delay (TSD) as the objective function because ramp delay is considered. However, it is noted that the choice of objective function is flexible where different objective functions can be used for different applications. The motorway traffic optimisation with ramp metering can be formulated as follows:

$$\min_{\mathbf{c_r}} \; Z = \sum_{i=1}^{I}\sum_{k=1}^{K}\Big(\rho_i(k)\Delta x_i\Delta t - \frac{f_i(k)\Delta x_i\Delta t}{v_i}\Big) + \eta\sum_{j=1}^{J}\sum_{k=1}^{K} l_j(k)\Delta t \tag{3.22}$$

subject to:

$$\begin{aligned}
\rho_i(k+1) &= \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k)\Big) && (3.23)\\
f_i(k) &\leq \rho_i(k) v_i \bar{\beta}_i && (3.24)\\
f_i(k) &\leq Q_i && (3.25)\\
f_i(k) &\leq Q_{i+1} && (3.26)\\
f_i(k) &\leq w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big) && (3.27)\\
l_j(k+1) &= l_j(k) + \big(\lambda_j(k) - r_j(k)\big)\Delta t && (3.28)\\
l_j(k) &\leq \bar{l}_j && (3.29)\\
r_j(k) &\leq \bar{r}_j && (3.30)\\
r_j(k) &\geq 0 && (3.31)\\
r_j(k) &\leq \frac{l_j(k)}{\Delta t} + \lambda_j(k) && (3.32)\\
r_j(k) &\leq \big(\bar{\rho}_j - \rho_j(k)\big)\frac{\Delta x_j}{\Delta t} && (3.33)
\end{aligned}$$

where Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. The notations v_i, w_{i+1}, and $\bar{\rho}_{i+1}$ are free flow speed, shockwave speed and jam density respectively. Q_i is the capacity flow at cell i which corresponds to the maximum flow that can leave cell i; Q_{i+1} is the capacity flow at cell $i+1$ which corresponds here to the maximum flow that can enter cell $i+1$. $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. $\lambda_j(k)$ denotes the traffic demand that wants to

enter the system through on-ramp j during time step k, and $r_j(k)$ is the actual demand that enters the system. The notation $\bar{l}_j$ refers to the maximum ramp queue length at on-ramp j, and the notation $\bar{r}_j$ refers to ramp capacity at on-ramp j.

The optimisation problem seeks the optimal control policy $\mathbf{c_r}$ to be implemented over time $k = 1, 2 \dots K$ and cells $i = 1, 2 \dots I$ that minimises the total system delay Z in the system. The first term $\sum_{i=1}^{I} \sum_{k=1}^{K} \left(\rho_i(k) \Delta x_i \Delta t - \frac{f_i(k) \Delta x_i \Delta t}{v_i} \right)$ is the main road delay in the objective function (Equation 3.22), and the second term $\sum_{j=1}^{J} \sum_{k=1}^{K} l_j(k) \Delta t$ is delays on the boundary links. One will get a trivial solution: $r_j(k) = 0, \forall j, k$, which simply prohibits any traffic from on-ramps entering the system if this second term $\sum_{j=1}^{J} \sum_{k=1}^{K} l_j(k) \Delta t$ is omitted. The notation η is a parameter that adjusts the balance between the main road delay and ramp delay (boundary queues). The sensitivity analysis is presented in Section 3.5. In this study, the value of η is set to be 1 indicating all road sections are equally weighted.

For the ramp queue length, Equation (3.28) is used to capture the evolution of queues $l_j(k)$ on on-ramps $j = 1, 2 \dots J$, where J is the total number of on-ramps. Moreover, one may add a upper bound $\bar{l}_j$ (Equation (3.29)) for some on-ramps to specify the maximum queue length of the on-ramps such that an unacceptable long queue on the on-ramp will not be obtained as an optimisation result.

For ramp inflow, Equations (3.30) and (3.31) are additional constraints on the control variable $r_j(k)$ to ensure its upper bound and non-negativity respectively. Equations (3.32) and (3.33) are constraints on ramp demand and main road space respectively. Note that $\lambda_j(k)$ does not necessarily equal to $r_j(k)$ due to various reasons including gridlock which traffic cannot be freely flowing into the system but subject to downstream

traffic condition. The difference between $r_j(k)$ and $\lambda_j(k)$ can also be due to various control strategies such as ramp metering (Gomes and Horowitz [34]) and variable speed limits (Smulders [94]).

With the formulation presented above, the size of the optimal ramp metering problem basically depends on the number of cells I and time steps K considered. Constraint set (3.23 - 3.27) each gives a total of $I \times K$ constraints. Moreover, the on-ramp inflow and on-ramp queue constraint set (3.28 - 3.33) each gives an additional set of $J \times K$ constraints with J on-ramps considered. Regarding the number of decision variables, it varies under different scenarios.

The above LP problem can be solved by a number of established algorithms such as SIMPLEX or interior-point methods for global optimal solutions (Vanderbei [101]). Nevertheless, it is known that there is no guarantee that the optimal solution can be found in polynomial time while the complexity of a LP problem basically depends on the number of decision variables and constraints involved.

3.4 Effect of Off-ramp Position

This section explores the sensitivity of the performance of the optimal ramp metering with respect to the distance between on-ramp and off-ramp. A corridor consisting of 25 cells with one on-ramp fixed at cell 20 and one off-ramp. The length of each cell (Δx_i) is 500 metres, and the simulation time step (Δt) is 20 seconds. To test the effect of distance between on-ramp and off-ramp, the location of the off-ramp varies from cell 2 to cell 18. The main road and ramp demands are 3000 veh/hr and 1200 veh/hr respectively. In

addition, the on-ramp capacity and maximum on-ramp queue are set to be 1500 veh/hr and 60 vehicles respectively. The split ratio at the off-ramp is set to be 0.1. The same free-flow speed (100 km/hr), capacity (3600 veh/hr) and jam density (240 veh/km) are set for all cells. The simulation horizon is one hour, and the cool down period is 30 minutes. In order to compare the performance of different scenarios, the relative delay reduction P is defined as:

$$P = \frac{D_n - D_c}{D_n} \ 100\% \tag{3.34}$$

where D_n and D_c are total system delay (or main road delay) under no control case and control case respectively.

Figure 3.2 shows the effectiveness of ramp metering with respect to the distance between on-ramp and off-ramp. The horizontal axis is the distance between on-ramp and off-ramp in terms of the number of cells between them. The left and right vertical axes are the relative total delay (main road + ramp) reduction and relative main road delay reduction respectively. Figure 3.2 shows that the relative total delay reduction (solid line) reduces as the distance between the ramps increases and becomes insignificant when the ramps are more than five cells apart. Moreover, the relative total delay reduction not reduce to zero (0.005) when the distance between the on-ramp and off-ramp is 18 cells. This finding indeed supports the argument made in [78] which suggests that a major benefit of ramp metering is due to the reduction of spillover of congestion at the associated on-ramp to the upstream junction.

In this example, when the distance between the on-ramp and off-ramp pair is less than

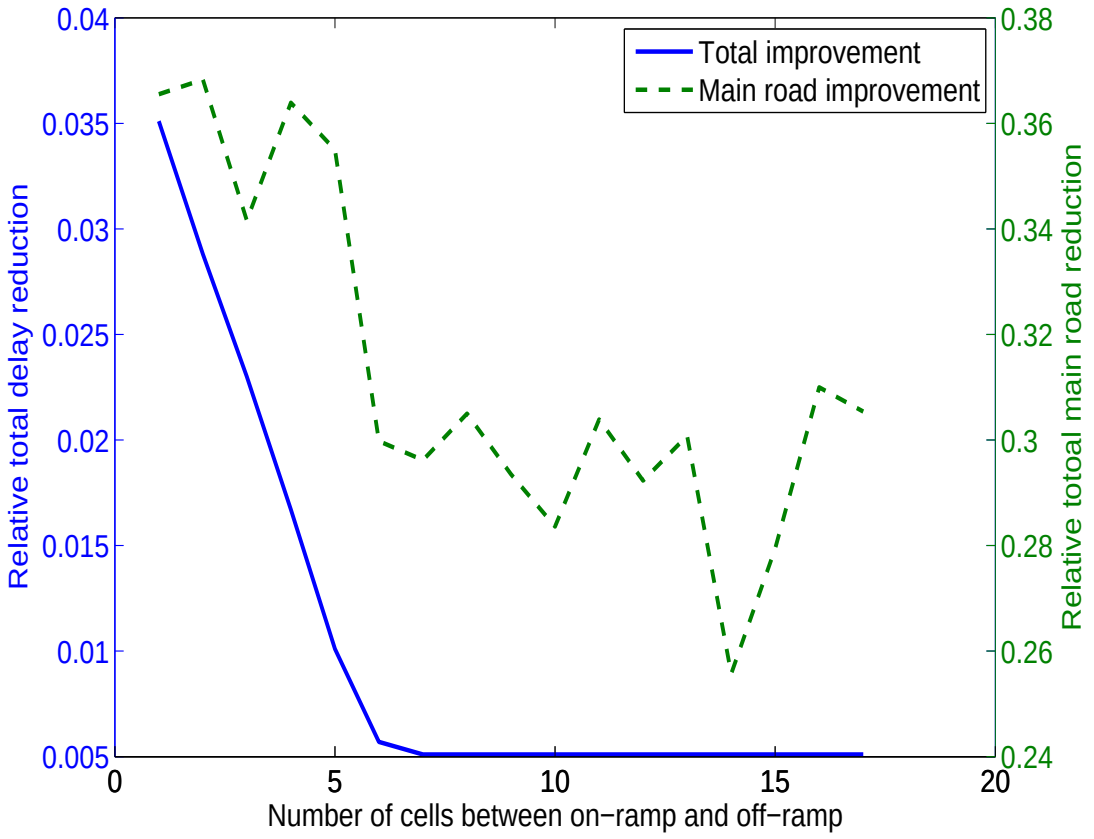

FIGURE 3.2: Performance of ramp metering

five cells, the congestion formed at the on-ramp at cell 20 could spill over to the off-ramp at its upstream. Metering the on-ramp at cell 20 reduces the degree of spillover effect, and hence the discharge of main road traffic through the off-ramp. Nevertheless, when the two ramps are too far apart, for example beyond five cells, the congestion formed at the on-ramp at cell 20 indeed cannot reach the location of the off-ramp even under no-control case. As a result, metering the on-ramp does not have an impact on facilitating the main road traffic discharge.

Without the benefit of reducing spillover congestion, it makes no difference from a system perspective whether queuing up the traffic on the main road or on the ramp (see Papageorgiou and Kotsialos [78]). Hence, no reduction in the total system delay

will be observed under such cases as revealed in Figure 3.2. The relative main road delay reduction (dotted line) has a similar trend to the relative total delay reduction, but with some fluctuations over the trend. A possible reason is that the optimiser aims to minimise the total system delay regardless of whether the reduction comes from the main road or on-ramps. This can induce some instability in main road (and ramp) delay calculations.

3.5 Sensitivity Analysis of Balance Parameter η

This section explores the sensitivity of the parameter η of the optimal ramp metering. A corridor consisting of 25 cells with one on-ramp at cell 21 and one off-ramp at cell 19. The length of each cell (Δx_i) is 500 metres, and the simulation time step (Δt) is 15 seconds. The on-ramp capacity and maximum on-ramp queue are set to be 1500 veh/hr and 60 vehicles respectively. The split ratio at the off-ramp is set to be 0.15. The same free-flow speed (100 km/hr), capacity (3600 veh/hr) and jam density (240 veh/km) are set for all cells. The simulation horizon is 4 hours with the cool down period of 30 minutes.

Table 3.1: Delays of ramp metering with η higher than one (veh-hr)

η	1.0	1.11	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19
Ramp	166.65	163.31	163.29	163.26	163.25	163.20	163.11	162.86	0.16	0.00
Main	538.20	541.55	541.56	541.59	541.61	541.66	541.79	542.06	733.47	733.66
Total	704.85	704.85	704.85	704.86	704.86	704.87	704.88	704.92	733.63	733.66

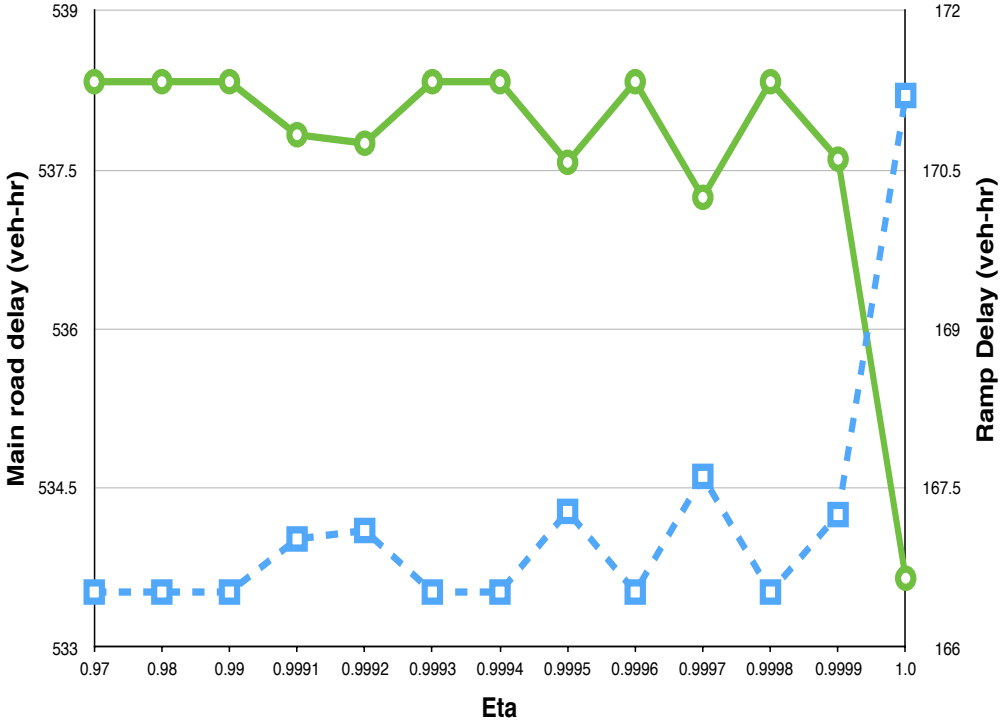

FIGURE 3.3: Delays of ramp metering with η lower than one

The optimal ramp metering control strategies are calculated with different value of η, then the ramp metering control strategies are simulated on the same CTM simulation platform, and the total system delay is calculated with the η of one. The balance parameter η of main road delay and ramp delay is sensitive around one. Figure 3.3 and Table 3.1 show the delays of ramp metering with the value of η lower than one and higher than one respectively.

Figure 3.3 shows the main road delay and ramp delay of ramp metering. The horizontal axis is the value of η between 0.97 and 1.0. The delays of η lower than 0.97 are not shown in the figure due to the delays are constant for $\eta \leq 0.97$. That means there is no more space to store more vehicles to relieve the main road congestion even the lower η gives more weight on main road delay. The left and right vertical axes are the main road

delay and ramp delay respectively. Figure 3.3 shows that the main road delay (green solid line with circles) reduces as the increases of η and becomes insignificant when the η approach to one. On the contrary, the ramp delay (blue dotted line with squares) has the opposite trend as shown in the figure. However the total system delay of ramp metering is constant at the value of 704.85 veh-hr for $\eta \leq 1$. Compared with no control case (733.66 veh-hr), there is 28.8 veh-hr reduction on total system delay.

Table 3.1 shows the delays including ramp delay, main road delay and total system delay with η equal and higher than one. The delays are nearly constant for $1 \leq \eta \leq 1.17$. That means η is not sensitive for the region $\eta \in [1,\ 1.17]$. However, the ramp delay approach to zero for $\eta > 1.17$. The higher value of η means the system puts more weight on ramp delay than on main road delay. It is important to note that the ramp delay equals to zero under the extreme case ($\eta = 1.19$ in this case study). It is not worthwhile to wait the vehicles on the ramp for the large η. Therefore, the vehicle is served when it arrives the ramp. For η lower than one value (1.19 in this case), the total system delay nearly no differece under different value of η, and the ramp delay reduces as the value of η increases and becomes constant when η lower than one value. However, ramp delay equals to zero for η higher than the particular value, then there is no metering on ramps.

3.6 Working Example

The London's orbital M25 Motorway is one of the busiest roads in the United Kingdom, which is used by 250,000 vehicles per day [92]. It is closely monitored and managed by the Highways England. Therefore, the section of the London's orbital M25 Motorway is selected to illustrate overall calibration performance and test the optimal ramp metering control strategy. Moreover, the clockwise direction is selected as it contains data of better quality. The typical traffic at 18:00 on a Thursday is shown in Figure 3.4, which is the screenshot of Google Map. The colour scale (green to red) represents the level of speed. Figure 3.4 shows the congestion generally happened between Junctions 12 and 15 as shown in red colour (slow traffic).

The map of the M25 Motorway between Junctions 10 and 16 is shown in Figure 3.5, and the spatial-temperal traffic pattern on Thursday (4 September 2014) between Junctions 10 and 16 in clockwise direction is shown in Figure 3.6. The colour scale represents the level of traffic density at the corresponding time and location, and the layout of the road section is shown on the left of the figure. Figure 3.6 shows the heavy traffic is observed around Junction 14. Therefore, we select the section of 12.5-km between Junctions 12 and 16 as a test site for the CTM simulation and optimal ramp metering. The section covers two major interchanges: Junction 14 connected with Heathrow Airport; Junction 15 connected with the M4 Motorway to West England and Central London. The on-ramps are located at cell 8 (Junction 13), cell 16 (Junction 14), cell 23 (Junction 15a) and cell 24 (Junction 15b) respectively. The off-ramps are located at cell 4 (Junction 13), cell 10 (Junction 14) and cell 18 (Junction 15) respectively. The list of main road detectors (with MIDAS index) and associated ramp detectors are shown in Table 3.2.

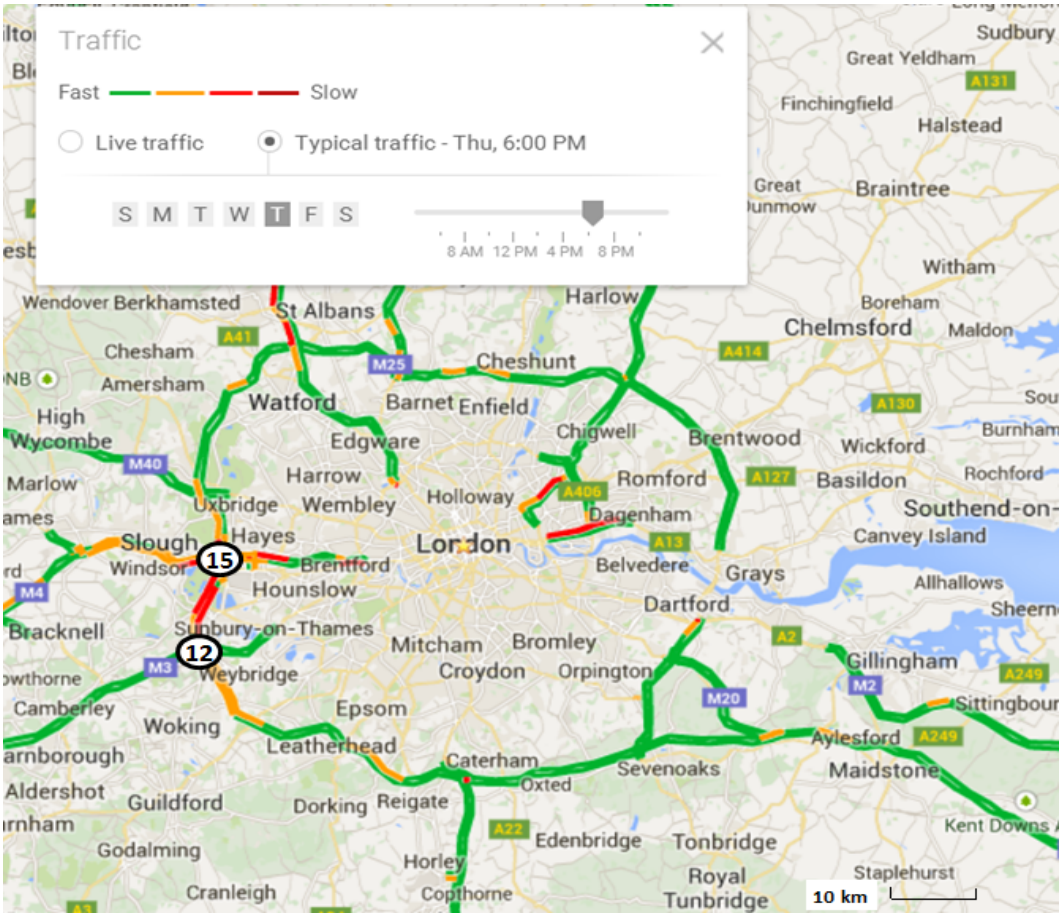

FIGURE 3.4: UK M25 Traffic Speed

(Source: Google Map)

FIGURE 3.5: UK M25 Motorway map - section between Junctions 10 and 16

Source: Highways England

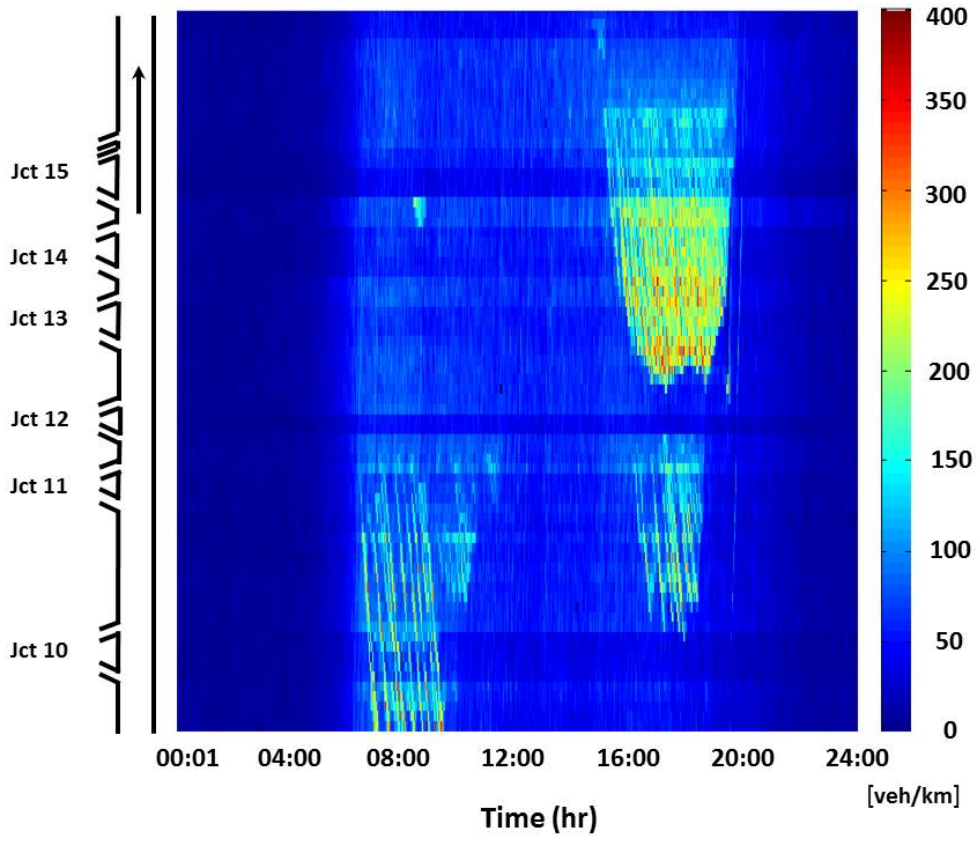

FIGURE 3.6: Observed density count plot - section between Junctions 10 and 16

Table 3.2: The list of MIDAS detector stations between Junctions 12 and 16 on the UK M25 Motorway in clockwise direction

Cell Number	MIDAS ID (Main)	Lanes	Remarks	MIDAS ID (Ramp)	Lanes
1	4866A	5			
2	4871A	5			
3	4876A	5			
4	4879A	5	Off-ramp at Jct 13	4883J	2
5	4883A	4			
6	4887A	4			
7	4892A	4			
8	4898A	5	On-ramp at Jct 13	4892K	2
9	4903A	5			
10	4909A	5	Off-ramp at Jct 14	4912J	2
11	4912A	4			
12	4916A	4			
13	4919A	4			
14	4923A	4			
15	4927A	4			
16	4932A	6	On-ramp at Jct 14	4926K	2
17	4936A	6			
18	4941A	6	Off-ramp at Jct 15	4945J	3
19	4945A	3			
20	4949A	3			
21	4955A	3			
22	4959A	3			
23	4963A	3	On-ramp at Jct 15a	4959K*	2
24	4968A	4	On-ramp at Jct 15b	4963K	1
25	4972A	4			

The 'Lanes' column refers to the number of lanes at the associated detector station. The 'Remarks' column shows the location of the ramps, where 'Jct' means 'Junction'. '*' refers to part of the traffic flow that enter the system through the specially on-ramp.

Following Daganzo [23] and Daganzo [24], the motorway stretch is divided into a series of cells where the length of all cells Δx_i is 500 metres, which is the standard MIDAS detector spacing. The motorway stretch contains 25 detector stations, which are configured such that the centre of upstream and downstream boundaries of each cell will coincide with the location of the associated detector station. The on-ramps and off-ramps are located in the beginning and the end of the cell respectively.

The time step size Δt is set such that $\Delta t \leq \min_i \frac{\Delta x_i}{v_i}$, where $\min_i \frac{\Delta x_i}{v_i}$ refers to the smallest ratio of cell length to the associated free flow speed along the section. The above condition is known as the Courant-Friedrichs-Lewy (CFL) condition [56]. This condition is used to ensure the numerical stability by constraining the traffic flow not to travel further than the length of the cell in one simulation time step. Consequently, the simulation time step Δt is set to 15-sec instead of 1-min as it stored in the dataset.

3.6.1 Without ramp metering

Each cell is characterised by a piecewise linear fundamental diagram which is calibrated by the measurements at the associated detector. The detected flow of the upstream of the first cell and each on-ramp are regarded as the input (demand) of the CTM model. Moreover, each cell has an initial density according to the detected density. If the vehicle cannot flow to the second cell when it arrives, the vehicle will wait at the first cell. We assume the first cell has a enough space to queue all waited vehicles, and the waiting time of the vehicle is counted in the total system delay. That means there is a point queue at the first cell.

A cross-validation is adopted to evaluate the estimation accuracy. The main road data collected on days 2 and 3 September 2014 are used to derive the fundamental diagram, while the on-ramp and main road demand on the day 4 September 2014 (Thursday) are used to construct the boundary conditions. The simulated traffic density, which is 15-sec resolution, is first aggregated into 5-min, and the measured density is also aggregated into 5-min. Then the simulated density $\rho_i(k)$ is compared with the measured density $\hat{\rho}_i(k)$ at each cell by using the mean absolute percentage error in density is defined as:

$$\epsilon = \frac{1}{IK} \sum_{i=1}^{I} \sum_{k=1}^{K} \left| \frac{\hat{\rho}_i(k) - \rho_i(k)}{\hat{\rho}_i(k)} \right| \tag{3.35}$$

where K and I are the number of simulation time steps and cells respectively.

Figure 3.7 shows the density contour plots in which the colour scale represents the level of traffic density at the corresponding time and location. The lower one in Figure 3.7 is the measured density calculated from measured occupancy, while the upper one is modelled density produced by CTM simulation. The mean absolute percentage error ϵ obtained from the CTM modelling conducted in this exercise is 11.5%. The part of the error in density is due to the error associated with conversion of the measured occupancy to density with Equation (2.7), in which the effective vehicle length $\bar{L}_v$ may be over-estimated. With the piecewise linear fundamental diagram, CTM cannot capture fine details of the nonlinear traffic behaviours such as capacity drop, stop-and-go wave, and acceleration-deceleration patterns. Nevertheless, the model can reproduce the general pattern of the traffic congestion (associated with correct location and time) with simple mathematical structure.

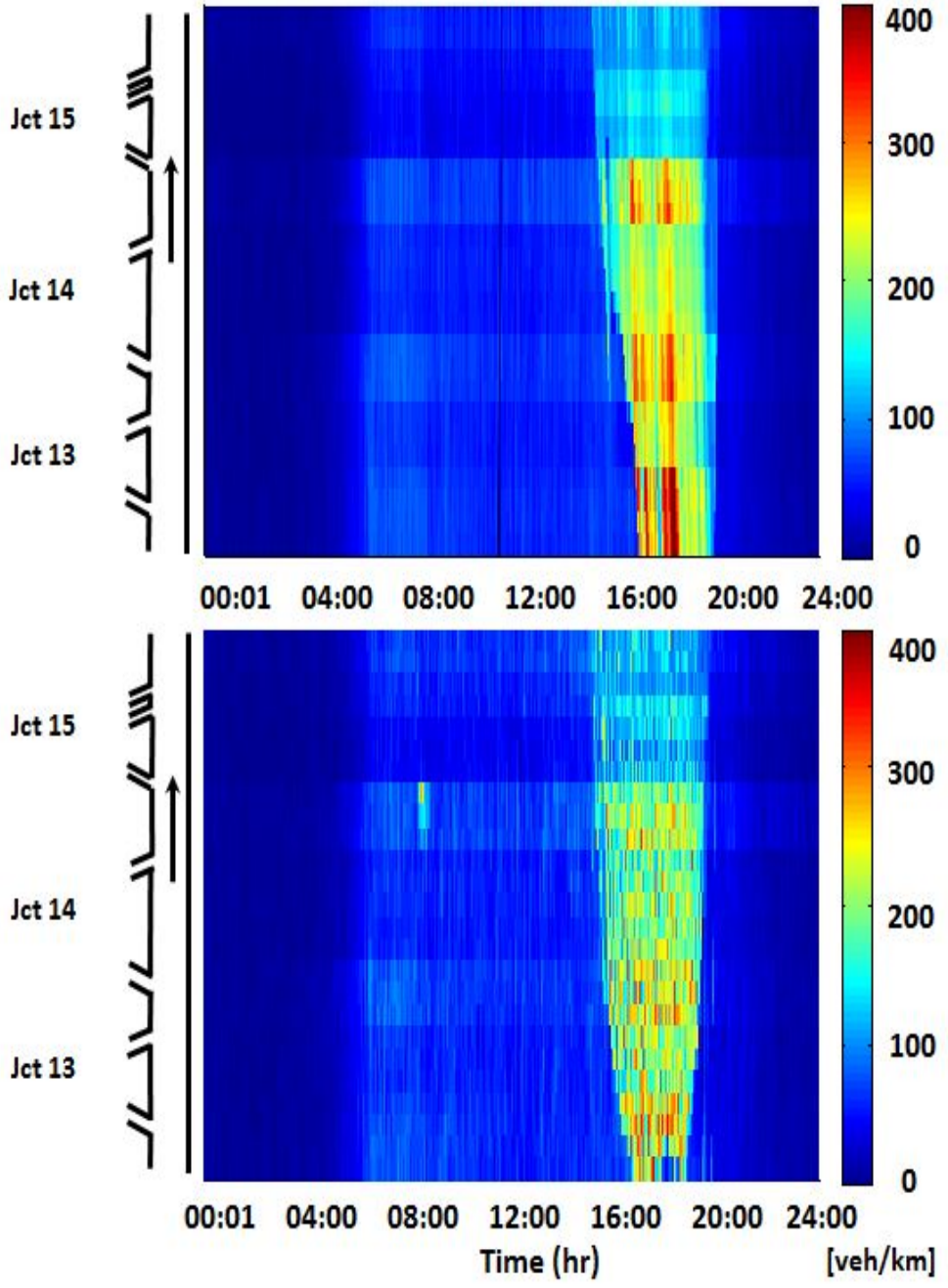

FIGURE 3.7: Modelling result between Junctions 12 and 16 over one day

upper: modelled; lower: observed

3.6.2 With ramp metering

The section of the motorway has been calibrated in Section 3.6.1 by using two days data (2 and 3 September 2014), then this section tests the optimal ramp metering on the calibrated section of the motorway with the data collected on 4 September 2014 (Thursday). The optimisation model is applied to manage afternoon peak hour traffic [14:00 - 21:00] at the congested region. The size of the simulation time step is set to be 15-sec, which gives the total number of time steps $K = 1680$ for a 7 hours [14:00 - 21:00] planning horizon. The optimal control problem is implemented and solved by IBM ILOG CPLEX Optimisation Studio V12.5 running on a desktop computer with Intel Core i5-2400 3.1GHz Processor, 4GB RAM, and Windows 7 64-bit operating system. It takes about four minutes to solve.

To illustrate some fundamental features of the optimal solutions, the optimal ramp metering policy that minimises the total delay along the section of motorway is considered. The problem consists of a total of 6,720 decision variables (ramp inflows, $r_j(k)$, in which 4 (on-ramps) $\times$ 1,680 (time-steps) = 6,720). It can be seen that there is a huge reduction in main road congestion and the associated main road delay reduces from 28,345 veh-hr to 21,829 veh-hr (see Table 3.3) corresponding to 23 % relative main road delay reduction calculated by Equation (3.34). Nevertheless, the reduction in main road delay is made at the expense of the additional queues induced on the on-ramps as shown in Figure 3.9 at Junction 15. It is noted that the size of the ramp queues can be seen reaching almost 500 vehicles at Junction 15. This implies the controller allows nearly 500 vehicles to spill over to London Heathrow Airport from the M25 Motorway, and this is certainly not acceptable in reality.

To produce applicable results, the maximum ramp queue constraint (Equation 3.29) needs to be considered. The maximum allowable queue length $\bar{l}_j$ at all on-ramps is set to be 30 (veh) or 60 (veh), which means the situation where an optimal metering is applied at the on-ramps and the ramp queues are not allowed to exceed 30 or 60 vehicles. In this case, a modest reduction in main road delay (see Table 3.3) is observed due to the consideration of ramp queues. Figure 3.8 compares the main road density with and without the optimal metering control, and the maximum queue length is set to be 60 vehicles at all on-ramps. The colour scale represents the level of traffic density at the corresponding time and location. Figure 3.10 and 3.11 show the main road and ramp delay profiles under different scenarios. Each point on the time series represents the total system delay (unit: [veh-hr]) at the corresponding 15-sec simulation time interval. The total system delay over the entire horizon can be derived by summing up all these 15-sec interval delays. For the 30 maximum ramp queue ramp metering case, the main road delay is 27,031 veh-hr, and the associated ramps' delay is 484.00 veh-hr with metering which gives a total system delay of 27,516 veh-hr, which is smaller than the original 28,345 veh-hr (2.9 %). Nevertheless, this metering policy is a more acceptable scheme as the ramp queues are bounded below a reasonable maximum ramp queue as shown in Figure 3.9.

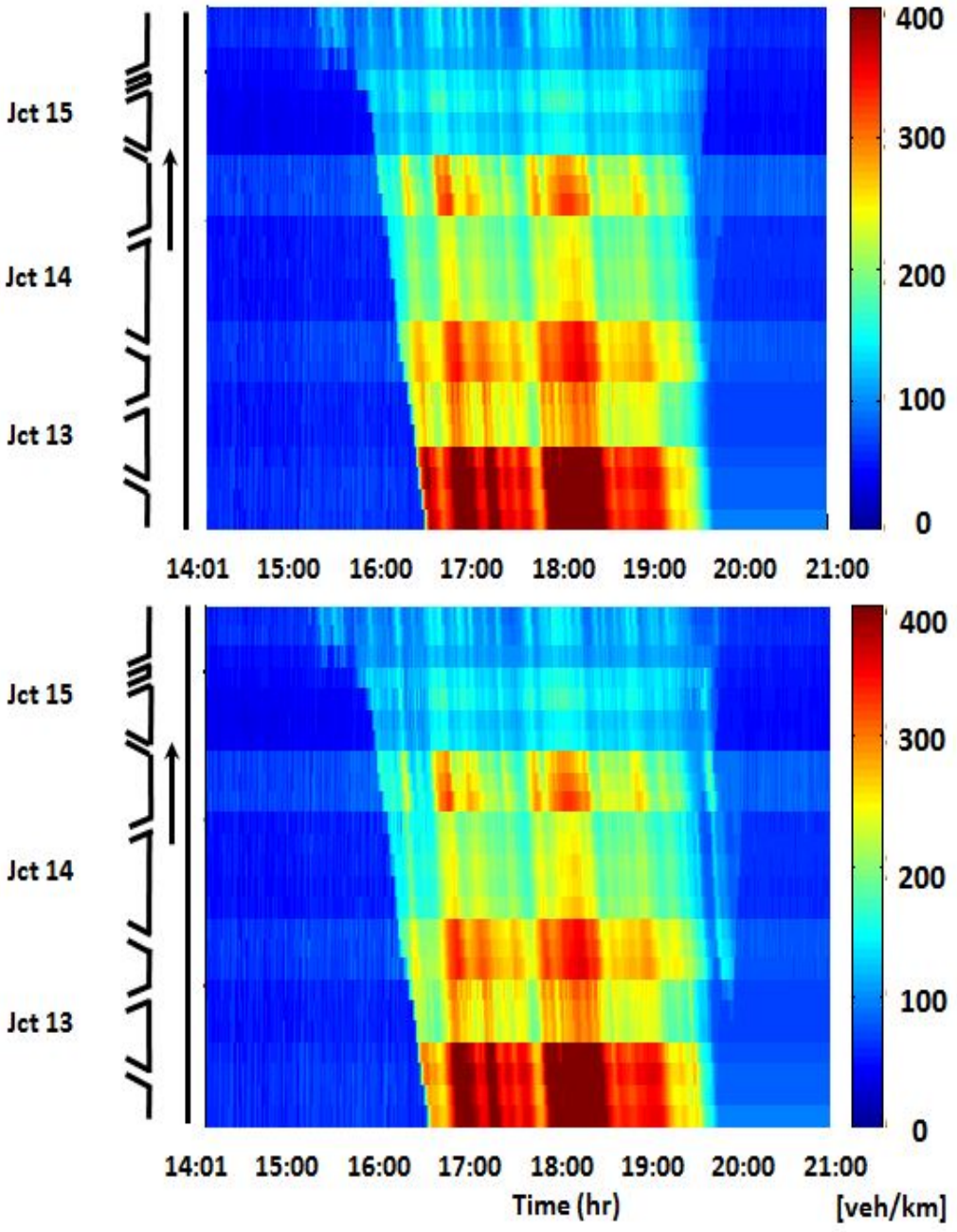

FIGURE 3.8: Comparison of main road densities (Three Junctions)

upper - no control; lower - metered (60 veh)

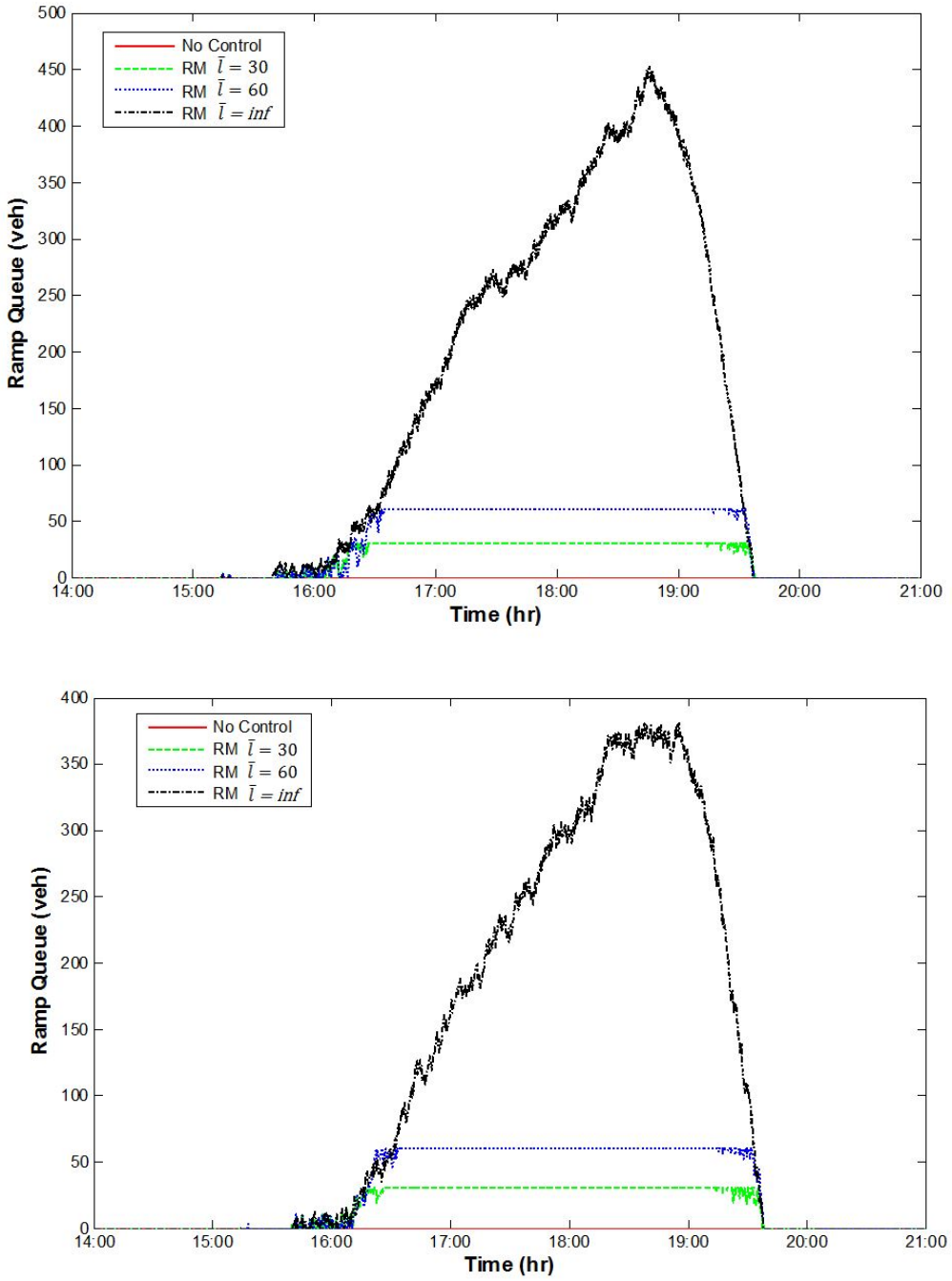

FIGURE 3.9: Ramp queues under metering at Junction 15

upper-Junction 15a; lower-Junction 15b

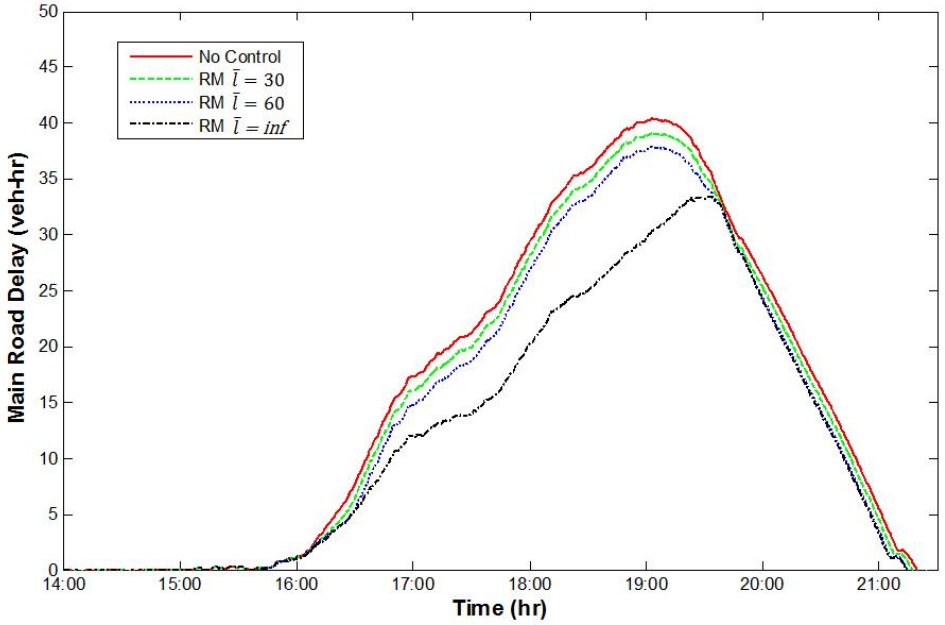

FIGURE 3.10: Comparison of main road delays

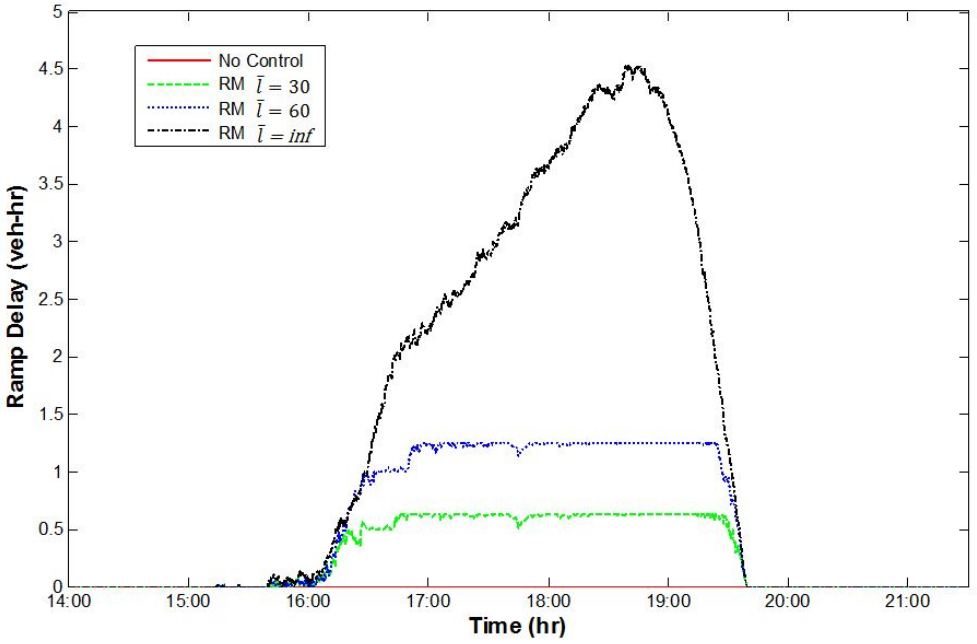

FIGURE 3.11: Comparison of ramp delays

Table 3.3: Delays under different ramp metering strategies

Delay [veh-hr]	**Main**	P_m	**Ramp**	**Total**	P_t
No Control	28,345		0.00	28,345	
RM ($\bar{l} = 30$)	27,031	4.6 %	484.00	27,516	2.9 %
RM ($\bar{l} = 60$)	25,802	9.0 %	938.37	26,740	5.7 %
RM ($\bar{l} =$ **inf**)	21,829	23 %	2,372.0	24,201	14.6 %

3.7 Summary

This chapter presents a mathematical framework that seeks the optimal ramp metering strategy. Cell transmission model is calibrated with traffic data and implemented to model a section of motorway. The validation result reveals the mean absolute percentage error is 11.5 %. With the piecewise linear structure of CTM, the optimal ramp metering problem can be formulated as a LP, which can be solved by a range of established solvers for the global optimal solution. This LP formulation is applied to a scenario of the M25 Motorway where an optimal ramp metering strategy is derived that minimises the total system delay over a fixed space-time horizon. It is shown that optimal solutions are obtainable through CPLEX in a reasonable computational time. We note that the application to motorway traffic is only an illustration and the methodology is indeed generally applicable for other transport networks. We also conduct a sensitivity analysis on the effect of ramp separation on the effectiveness of the ramp metering.

Chapter 4

Optimisation of Variable Speed Limits and Hard Shoulder Running

4.1 Introduction

This chapter discusses the derivation of variable speed limits and hard shoulder running strategies. Variable speed limits (VSL) aim to reduce congestion through homogenising traffic flow by managing their speed. It is shown that VSL have a positive impact on safety and mobility [58]. Hard shoulder running (HSR) increases road capacity by providing an extra lane to road users at specific times, and HSR needs to be applied with VSL because of safety reasons [4]. A pilot scheme involving HSR operates on the M42 Motorway around Birmingham. The results show that it is an effective way to increase

throughput along congested road sections and an additional 15 per cent reduction in travel time is observed [100]. This chapter extends the optimisation formulation to VSL and HSR. The challenge associated with the optimisation formulation is how to capture the transformation of a fundamental diagram under the control.

This chapter is organised as follows: Section 4.2 presents how to implement VSL on the motorway. Section 4.3 describes HSR formulation and the integrated control strategy. Section 4.4 presents the implementation of optimal VSL and HSR through a case study of the United Kingdom M25 Motorway. Section 4.5 provides some concluding remarks.

4.2 Variable Speed Limits

4.2.1 Changes in fundamental diagram under VSL

Variable speed limits affect the traffic on motorways by adjusting the speed limits. The challenge associated with the optimisation for VSL is how to capture the transformation of the fundamental diagram under different speed limits. As discussed in Papageorgiou *et al.* [84], Carlson *et al.* [15], and shown empirically in Heydecker and Addison [46], the fundamental diagram at a specific location will be changed if the speed limit applied at that location changes. Smulders [94] finds that when the speed limit (e.g. 60 mph or 50 mph) is used, the average free flow speed of traffic will be reduced while the capacity will be increased slightly. The slight increase in capacity is due to the shorter headways between adjacent vehicles with lower speed limit.

Currently VSL are in operation on the M25 Motorway where there are four distinct speed limits: 70 mph, 60 mph, 50 mph, and 40 mph. The 70 mph is the normal value, while 60 mph and 50 mph are used for congestion management, and 40 mph is used for serious congestion or incident. In addition to the traffic measurement, MIDAS also records the operating time of speed limits on the motorway. With such information, the relationship between VSL and the shape of the fundamental diagram can be explored.

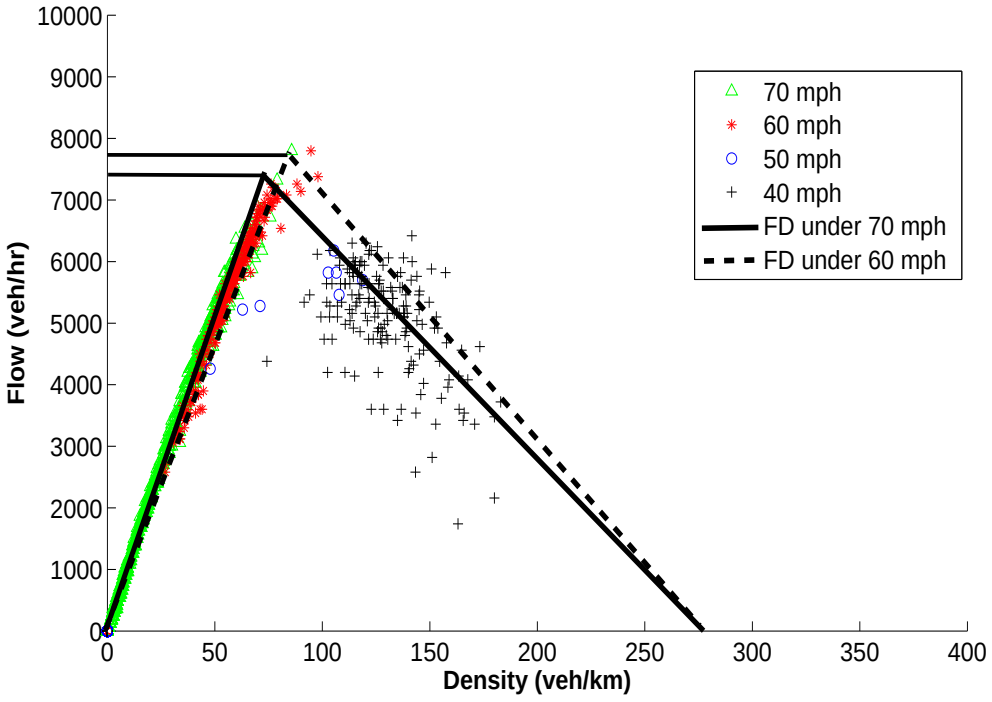

FIGURE 4.1: Changes in fundamental diagram under VSL with real data

Detector Station: 4936A, M25 (clockwise), 26 September 2012

The parameters used in the VSL optimisation are obtained by fitting the fundamental diagram with the traffic data under different speed limits. Figure 4.1 presents an empirical scatter plot of flow-density data collected from a loop detector station (4936A) on the M25 Motorway (clockwise). The detector station, which consists of 6 lanes, and is located at the downstream of the on-ramp at Junction 14 and the upstream of the

off-ramp at Junction 15. The data are classified according to the speed limits (70 mph, 60 mph, 50 mph and 40 mph) in effect when they were collected. The solid and dash lines in Figure 4.1 reveal the fundamental diagrams under 70 mph speed limit and 50 mph speed limit respectively. It is found that when the speed limit (e.g. 60 mph or 50 mph) applied on the motorway for congestion management, the average free flow speed of traffic will be reduced while the capacity will be increased slightly [46, 94]. The empirical observation here generally supports the assumption of the transformation of fundamental diagram under VSL.

4.2.2 Optimisation of VSL

This section extends the optimisation formulation to VSL. The challenge associated with the optimisation for VSL lies on capturing the transformation of the fundamental diagram during the speed transition process. Carlson *et al.* [15] adopt a 'scaling parameter' α to model such a transformation of the fundamental diagram. Each value of α ($0 \leq \alpha \leq 1$) represents one particular speed limit and one particular fundamental diagram generated from a family of exponential speed-density functions under a second-order METANET modelling framework [81]. The objective of the variable speed control problem is to seek the value of α over time and space such that the total delay in the network within a predefined time horizon is minimised. Carlson's formulation has been producing a number of interesting insights on how one should deploy a variable speed control policy. Nevertheless, Carlson *et al.* [15]'s optimal control formulation is non-linear which has a non-linear objective function and a set of nonlinear constraints. As a

result, global optimality cannot be guaranteed. Global optimality may not be an important issue for practical applications. However, it is certainly a desirable property for a theoretical study as the global optimal solution provides a convincing and indisputable benchmark for comparing different implementation plans.

4.2.2.1 Variable speed limits with two speed limits

This study adopts a CTM based on mixed integer linear programming (MILP) formulation for solving the optimal variable speed control problem. It starts with considering only two admissible speed limits. It is similar to the case in the UK where there is a nominal speed limit on motorways at 70 mph, and is reduced to 50 mph for congestion management (Highways Agency [2]). In the model, a set of binary (0 - 1) decision variables $\alpha_i^1(k)$ are introduced to represent the choice between the nominal and reduced speed limits, where $\alpha_i^1(k) = 1$ implies a reduced speed limit which is applied at cell i during time step k; $\alpha_i^1(k) = 0$ means otherwise. The solution of the problem reveals the optimal deployment of the corresponding speed control strategy over time and space. With the binary variable $\alpha_i^1(k)$, constraints (3.17 - 3.20), which are the constraints on the outflow in CTM, are replaced for all i and k as constraints (4.3 - 4.10). Additional constraints on the density conservation equation (Equation 3.16) and ramp inflow (Equation 3.21) are introduced in Chapter 3 . With the objective function (Equation 3.15), the VSL optimisation problem can be formulated as follows:

$$\min_{\mathbf{c_{v2}}} \; Z = \sum_{i=1}^{I} \sum_{k=1}^{K} \left(\rho_i(k) \Delta x_i \Delta t - \frac{f_i(k) \Delta x_i \Delta t}{v_i} \right) \tag{4.1}$$

subject to:

$$\rho_i(k+1) = \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k)\Big) \tag{4.2}$$

$$f_i(k) \leq \rho_i(k) v_i^a \bar{\beta}_i + \alpha_i^1(k) M \tag{4.3}$$

$$f_i(k) \leq Q_i^a + \alpha_i^1(k) M \tag{4.4}$$

$$f_i(k) \leq Q_{i+1}^a + \alpha_{i+1}^1(k) M \tag{4.5}$$

$$f_i(k) \leq w_{i+1}^a\big(\bar{\rho}_{i+1}^a - \rho_{i+1}(k)\big) + \alpha_{i+1}^1(k) M \tag{4.6}$$

$$f_i(k) \leq \rho_i(k) v_i^b \bar{\beta}_i + \big(1 - \alpha_i^1(k)\big) M \tag{4.7}$$

$$f_i(k) \leq Q_i^b + \big(1 - \alpha_i^1(k)\big) M \tag{4.8}$$

$$f_i(k) \leq Q_{i+1}^b + \big(1 - \alpha_{i+1}^1(k)\big) M \tag{4.9}$$

$$f_i(k) \leq w_{i+1}^b\big(\bar{\rho}_{i+1}^b - \rho_{i+1}(k)\big) + \big(1 - \alpha_{i+1}^1(k)\big) M \tag{4.10}$$

$$r_j(k) = \lambda_j(k) \tag{4.11}$$

The optimal VSL policy $\boldsymbol{c_{v2}}$ is to be implemented over all cells and time steps that minimise the total system delay Z in the system. The notation Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. $\lambda_j(k)$ denotes the traffic demand that wants to enter the system through on-ramp j during time step k, and $r_j(k)$ is the actual demand that enters the system. Constraint (4.2) is the conservation equation to update the density in cell i for next time step $k+1$. Moreover, the constraint (4.11) shows the ramp inflow equals to the actual ramp demand, which means there is no ramp control is applied on the road.

The notation M represents a very large number where it is set to be 99,999; The notations

(v_i^a, Q_i^a, w_i^a and $\bar{\rho}_i^a$) and (v_i^b, Q_i^b, w_i^b and $\bar{\rho}_i^b$) are free flow speed, capacity, the shock wave speed and jam density under normal and reduced speed limits respectively. Note that when $\alpha_i^1(k) = 0$, it will disable the constraints (4.7) and (4.8). While with $\alpha_i^1(k) = 1$, the fundamental diagram at cell i will be transformed through switching off constraints (4.3) and (4.4), and switching on constraints (4.7) and (4.8). Considering safety reasons, the VSL should not be changed too frequently because it will confuse drivers. The variable speed limit interval is set to satisfy this constraint to ensure the VSL cannot change in a period of time.

There are indeed a number of different assumptions on how the fundamental diagram can be affected by VSL (see, for example, Hegyi [42]; Carlson *et al.* [15]; Heydecker and Addison [46]), while it is fair to say there is no conclusion of which specification is correct. The MILP formulation here relaxes such restrictions through the 'Big-M' binary constraint set, which allows arbitrary fundamental diagrams to be used under different speed limits.

4.2.2.2 Variable speed limits with three speed limits

The formulation can further be extended to cover more choices of speed limit with additional binary variables and associated constraints. If one more speed limit is considered in the study, at least two sets of binary (0 - 1) decision variables need to be introduced, chosen from three kinds of fundamental diagrams (FD^a, FD^b and FD^c). The notation FD^a (v_i^a, Q_i^a, w_i^a and $\bar{\rho}_i^a$) represnts the fundamental diagram including all relavent parameters under normal speed limit (e.g. 70 mph). The notations FD^b (v_i^b, Q_i^b, w_i^b and $\bar{\rho}_i^b$) and FD^c (v_i^c, Q_i^c, w_i^c and $\bar{\rho}_i^c$) denote fundamental diagrams under higher speed

limit (e.g. 60 mph) and lower speed limit (e.g. 50 mph) respectively. For example, constraint set (3.17 - 3.20) can be replaced by the following constraint set (4.14 - 4.25) to cover three different speed limits with a second binary variable $\alpha_i^2(k)$ as follows:

$$\min_{\mathbf{c_{v3}}} \; Z = \sum_{i=1}^{I} \sum_{k=1}^{K} \Big(\rho_i(k) \Delta x_i \Delta t - \frac{f_i(k) \Delta x_i \Delta t}{v_i} \Big) \tag{4.12}$$

subject to:

$$\begin{aligned}
\rho_i(k+1) &= \rho_i(k) + \frac{\Delta t}{\Delta x_i} \Big(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k) \Big) && (4.13)\\
f_i(k) &\leq \rho_i(k) v_i^a \bar{\beta}_i + \big(\alpha_i^1(k) + \alpha_i^2(k) \big) M && (4.14)\\
f_i(k) &\leq Q_i^a + \big(\alpha_i^1(k) + \alpha_i^2(k) \big) M && (4.15)\\
f_i(k) &\leq Q_{i+1}^a + \big(\alpha_{i+1}^1(k) + \alpha_{i+1}^2(k) \big) M && (4.16)\\
f_i(k) &\leq w_{i+1}^a \big(\bar{\rho}_{i+1}^a - \rho_{i+1}(k) \big) + \big(\alpha_{i+1}^1(k) + \alpha_{i+1}^2(k) \big) M && (4.17)\\
f_i(k) &\leq \rho_i(k) v_i^b \bar{\beta}_i + \big(1 - \alpha_i^2(k) \big) M && (4.18)\\
f_i(k) &\leq Q_i^b + \big(1 - \alpha_i^2(k) \big) M && (4.19)\\
f_i(k) &\leq Q_{i+1}^b + \big(1 - \alpha_{i+1}^2(k) \big) M && (4.20)\\
f_i(k) &\leq w_{i+1}^b \big(\bar{\rho}_{i+1}^b - \rho_{i+1}(k) \big) + \big(1 - \alpha_{i+1}^2(k) \big) M && (4.21)\\
f_i(k) &\leq \rho_i(k) v_i^c \bar{\beta}_i + \big(1 - \alpha_i^1(k) \big) M && (4.22)\\
f_i(k) &\leq Q_i^c + \big(1 - \alpha_i^1(k) \big) M && (4.23)\\
f_i(k) &\leq Q_{i+1}^c + \big(1 - \alpha_{i+1}^1(k) \big) M && (4.24)\\
f_i(k) &\leq w_{i+1}^c \big(\bar{\rho}_{i+1}^c - \rho_{i+1}(k) \big) + \big(1 - \alpha_{i+1}^1(k) \big) M && (4.25)\\
r_j(k) &= \lambda_j(k) && (4.26)\\
1 &\geq \alpha_i^1(k) + \alpha_i^2(k) && (4.27)
\end{aligned}$$

The optimal VSL policy $\boldsymbol{c_{v3}}$ is to be implemented over all cells and time steps that minimise the total system delay Z in the system. The notation Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. The notations $\lambda_j(k)$ and $r_j(k)$ denote the traffic demand that wants to enter the system and the actual demand that enters the system respectively. The constraint (4.26) shows the ramp inflow equals to the actual ramp demand, which means there is no ramp control is applied on the road. Moreover, constraint (4.13) is the conservation equation to update the density in cell i for following time step $k+1$.

The notation M represents a very large number where it is set to be 99,999. The notation (v_i^a, Q_i^a, w_i^a and $\bar{\rho}_i^a$) represents free flow speed, capacity, the shock wave speed and jam density under normal speed limit (70 mph). The notations (v_i^b, Q_i^b, w_i^b and $\bar{\rho}_i^b$) and (v_i^c, Q_i^c, w_i^c and $\bar{\rho}_i^c$) are free flow speed, capacity, the shock wave speed and jam density under higher and lower speed limits respectively. Note that $\alpha_i^1(k)$ and $\alpha_i^2(k)$ represent two set of binary variables, so there are four combinations ([0,0], [0,1], [1,0], and [1,1]). Under constraint 4.27, only three combinations ([0,0], [0,1], [1,0]) works. The first case is that $\alpha^1 = 0$ and $\alpha^2 = 0$, constraints (4.18 - 4.25) will be disabled. Therefore, only constraint (4.14 - 4.17) works. Constraint (4.18 - 4.21) works under $\alpha^1 = 0$ and $\alpha^2 = 1$. The last case is that $\alpha^1 = 1$ and $\alpha^2 = 0$, which means only constraint (4.22 - 4.25) works. Unlike the studies of Papageorgiou *et al.* [84] and Carlson *et al.* [15], one can capture different kinds of transformation by setting appropriate values of parameters.

4.2.2.3 Computational Complexity of VSL

The 'Big-M' formulations in constraint set (4.3 - 4.10) and constraint set (4.14 - 4.25) enable the arbitrary transformation of the fundamental diagram under VSL. The more constraints are introduced for three speed limits than two speed limits. It is known that MILP can induce the 'curse of dimensionality' problem due to the combinatorial nature of the problem (see Luenberger and Ye [67]). For example, consider three speed limits case where there are the two binary variables $[\alpha_i^1(k), \alpha_i^2(k)]$; it implies there can be four (2×2) combinations of them: $\Big([0,0] \quad [0,1] \quad [1,0] \quad [1,1]\Big)$ and hence a larger solution space and computational complexity. To analyse further the computational complexity, suppose that T_n is the control period (typically $5-10$ minutes for variable speed control purposes) that specifies how often the speed limit is being updated. Then further consider R_n to be the number of these control periods in the optimisation problem. As an illustration, if the control period T_n is set to be 5 minutes long, and seeking an optimal speed control strategy over a one hour (60 minutes) time horizon, then R_n will be 60 (minutes)/5 (minutes) = 12. Finally, the number of feasible VSL is defined as V_n. Given these quantities, the total number of possible solutions C_n of the optimisation problem is determined as:

$$C_n = \big((V_n)^{T_n}\big)^{R_n} = (V_n)^{T_n R_n} \tag{4.28}$$

The total number of possible solutions C_n grows exponentially with respect to T_n (corresponding to how often the control is updated) and R_n (corresponding to the length of the optimisation planning horizon). This exponential growth rate of solution space is a

typical feature of MILP problems, which implies one has to consider the problem formulation (e.g. number of decision variables to involve) carefully as it can have a significant impact on the corresponding computational time. With the increase in number of the control interval and control region, the case number of three speed limits increases faster than two speed limits, because the case number increases exponentially. For example, we assume the same control period ($T_2 = T_3 = 5$ min) and the number of control periods ($R_2 = R_3 = 120/5 = 24$ control horizon is set as 2 hours). Then the total number of possible solusions are $C_2 = 2^{5*24} = 2^{120} = 1.3292e{+}36$ and $C_3 = 3^{5*24} = 3^{120} = 1.7970e{+}57$ for two and three speed limits respectively. Because of the computation complexity of the three variable speed limits, the ramp metering with two varibale speed limits is illustrated in the following section.

4.2.2.4 Ramp metering with two variable speed limits

In real practice, the variable speed limit control is always applied with ramp metering. By considering the ramp metering, which is introduced in Chapter 3, the ramp metering with two variable speed limits (RMVSL) can be formulated as follows:

$$\min_{\mathbf{c_{rv}}} \; Z = \sum_{i=1}^{I}\sum_{k=1}^{K}\left(\rho_i(k)\Delta x_i \Delta t - \frac{f_i(k)\Delta x_i \Delta t}{v_i}\right) + \eta \sum_{j=1}^{J}\sum_{k=1}^{K} l_j(k)\Delta t \tag{4.29}$$

subject to:

$$\rho_i(k+1) = \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k)\Big) \tag{4.30}$$

$$f_i(k) \leq \rho_i(k) v_i^a \bar{\beta}_i + \alpha_i^1(k) M \tag{4.31}$$

$$f_i(k) \leq Q_i^a + \alpha_i^1(k) M \tag{4.32}$$

$$f_i(k) \leq Q_{i+1}^a + \alpha_{i+1}^1(k) M \tag{4.33}$$

$$f_i(k) \leq w_{i+1}^a\big(\bar{\rho}_{i+1}^a - \rho_{i+1}(k)\big) + \alpha_{i+1}^1(k) M \tag{4.34}$$

$$f_i(k) \leq \rho_i(k) v_i^b \bar{\beta}_i + \big(1 - \alpha_i^1(k)\big) M \tag{4.35}$$

$$f_i(k) \leq Q_i^b + \big(1 - \alpha_i^1(k)\big) M \tag{4.36}$$

$$f_i(k) \leq Q_{i+1}^b + \big(1 - \alpha_{i+1}^1(k)\big) M \tag{4.37}$$

$$f_i(k) \leq w_{i+1}^b\big(\bar{\rho}_{i+1}^b - \rho_{i+1}(k)\big) + \big(1 - \alpha_{i+1}^1(k)\big) M \tag{4.38}$$

$$l_j(k+1) = l_j(k) + \big(\lambda_j(k) - r_j(k)\big)\Delta t \tag{4.39}$$

$$l_j(k) \leq \bar{l}_j \tag{4.40}$$

$$r_j(k) \leq \lambda_j(k) \tag{4.41}$$

$$r_j(k) \geq 0 \tag{4.42}$$

$$r_j(k) \leq \frac{l_j(k)}{\Delta t} + \lambda_j(k) \tag{4.43}$$

$$r_j(k) \leq \big(\bar{\rho}_j - \rho_j(k)\big)\frac{\Delta x_j}{\Delta t} \tag{4.44}$$

The optimal VSL policy $\boldsymbol{c_{rv}}$ is to be implemented over all cells and time steps that minimise the total system delay Z in the system. The notation η is a parameter that adjusts the balance between the main road delay and ramp delay (boundary queues). In this study, the value of η is set to be 1 indicating all road sections are equally weighted. Δt denotes the length of simulation time step, and Δx_i represents the length of cell

i. The notations $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. Constraint (4.30) is the conservation equation to update the density in cell i for next time step $k+1$.

The notation $\lambda_j(k)$ denotes the traffic demand that wants to enter the system through on-ramp j during time step k, and $r_j(k)$ is the actual demand that enters the system. Equation (4.39) is used to capture the evolution of queues $l_j(k)$. Moreover, one may add a upper bound $\bar{l}_j$ (Equation (4.40)) for some on-ramps to specify the maximum queue length of the on-ramps such that an unacceptable long queue on the on-ramp will not be obtained as an optimisation result. For ramp inflow, Equations (4.41) and (4.42) are additional constraints on the control variable $r_j(k)$ to ensure its upper bound and non-negativity respectively. Equations (4.43) and (4.44) are constraints on ramp demand and main road space respectively.

4.2.3 The effect of fundamental diagram specifications under VSL

It is found that capacity increases slightly when reduced speed limits are applied, while free flow speed reduces under reduced speed limits. This section explores the effect of different assumptions on the transformation of the fundamental diagram under different speed limits. A hypothetical two-lane motorway corridor of 7 km is adopted here which consists of 14 cells with a bottleneck at cell 8. The length of each cell is taken as 500 metres, and the simulation time is set to be 60 min with an extra 15 min cool down period. Two kinds of main road demand are tested in this section. The Demand 1 is set as 3000 veh/hr for the whole simulation horizon, and Demand 2 is set as the trapezoid-shaped with the highest demand of 3400 veh/hr. All cells are assumed to have a common

capacity (3600 veh/hr) and jam density (240 veh/km) except cell 8 where it takes a lower capacity (2800 veh/hr) and jam density (180 veh/km) there. The nominal speed limit is 100 km/hr. The total delays with no control are 103.45 veh-hr and 111.41 veh-hr under Demand 1 (steady demand) and Demand 2 (time-varying demand) respectively.

Two possible transformations of fundamental diagrams (FD' and FD^+) are considered under reduced speed limits (see Figure 4.2). Both FD' and FD^+ consider a reduced free flow speed and a slightly increased capacity as suggested by empirical observations (Heydecker and Addison [46]). The FD' transformation (long red dash lines) assumes the jam density will remain the same, while the FD^+ transformation (dotted lines) assumes the shockwave speed (w) remains the same.

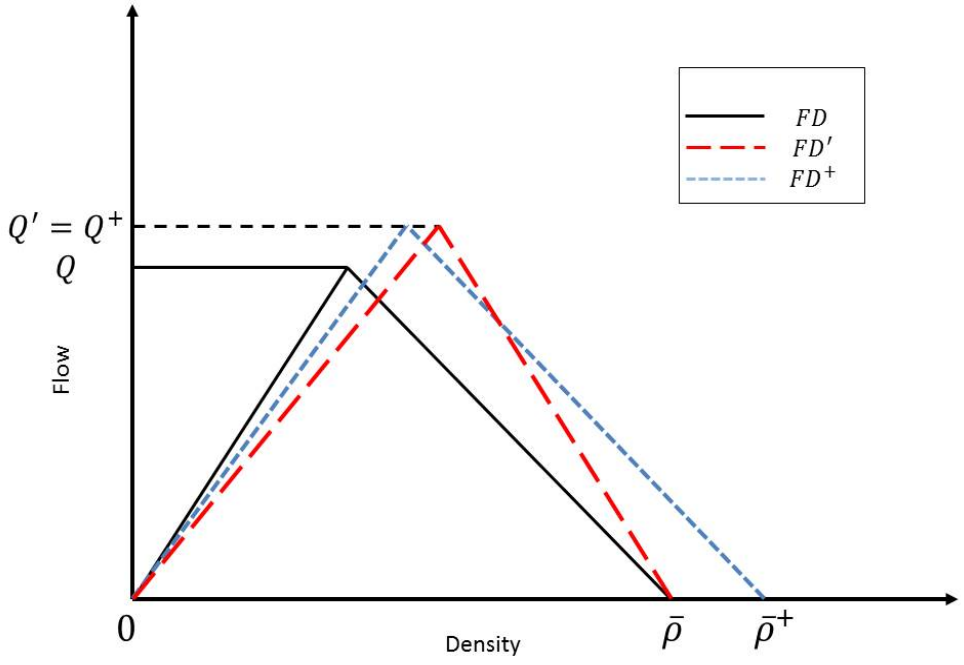

FIGURE 4.2: Changes in fundamental diagram under VSL

Figure 4.3 compares the total delay reduction gained from the optimal VSL with the two transformations FD' and FD^+ under different settings. Scenarios are considered with

6 different capacity settings where '1.00Q' represents a case where there is no change in capacity after reducing the speed limit, '1.01Q' refers to the situation where the capacity flow will be increased by 1% at the location where the speed limit is reduced, and so on. Also considered here are two sets of binary speed limit settings in which 'VSL90' means the alternative (reduced) speed limit is 90 km/hr, while 'VSL80' means the alternative (reduced) speed limit is 80 km/hr. The uneven dash and solid lines with markers represent the percentage of reduced delay under Demand 1 when 'VSL90' and 'VSL80' are applied respectively. The solid and even dash lines represent the percentage of reduced delay under Demand 2 when 'VSL90' and 'VSL80' are applied respectively. Figure 4.3 shows a linear relationship between the capacity improvement under reduced speed limit and the delay reduction. The interesting observation here is that it appears different assumptions of fundamental diagram transformations do not have a significant effect on the eventual performance of the variable speed control.

Table 4.1 further shows the performance of VSL with different spatial and temporal granularity of control where the transformation fundamental diagram (FD') is adopted. In the table, the control interval represents how frequent the speed limit is updated. This control interval represents the temporal granularity of the variable speed control policy. The control region represents the number of cells with the same speed limit. This control region represents the spatial granularity. The numbers in the table are the total system delay (unit: [veh-hr]) under the optimal variable speed control derived with corresponding combination of control region and interval settings. For example, the number '76.94' (veh-hr) in Table 4.1 is the total system delay under the steady demand (Demand 1) optimal speed control with which the speed limit varies every minute and every cell (500 metres), and so forth. It is observed that better performance

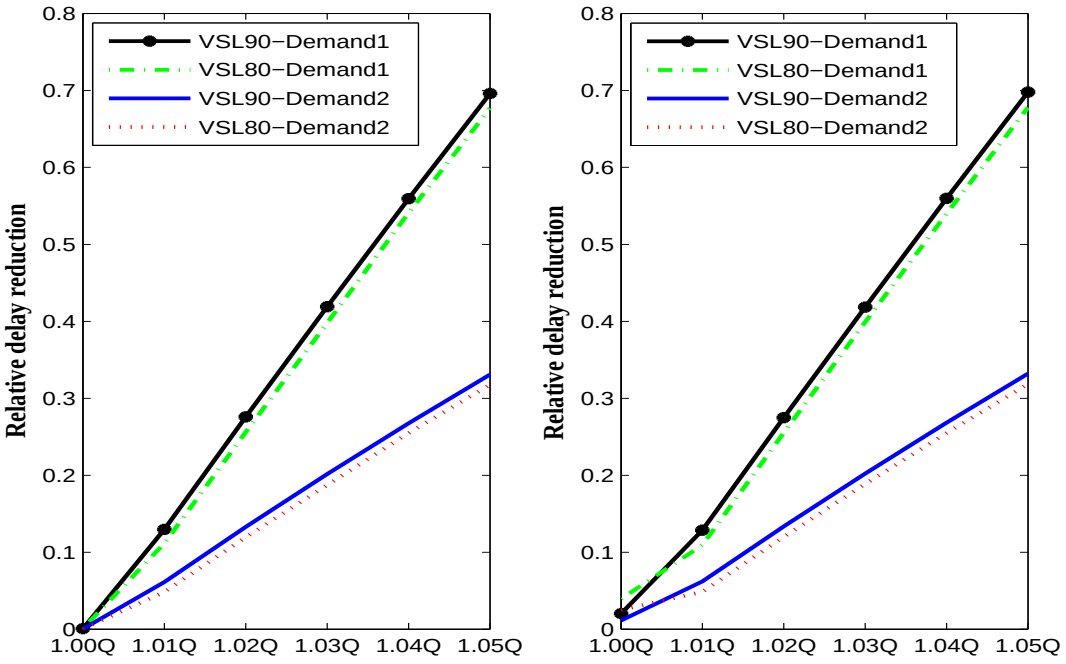

FIGURE 4.3: Comparison of different transformations of fundamental diagram under VSL

(left–FD'; right–FD^{+})

in terms of delay minimisation can be achieved with finer control (i.e. control derived with smaller values of control interval and region). However, one should note that this will have to come at the expense of computational effort as discussed previously. It is also interesting to highlight that the benefit of using a finer control interval indeed depends on the temporal variability of the demand. For example, Table 4.1 reveals that using a finer control interval (1-min) virtually does not bring any additional benefit over the coarser ones (30-min) when the demand profile is steady.

Table 4.1: Performance of VSL with different space-time granularity

Delay	**Control Interval [min]**						
[veh-hr]	**Control region**	**1**	**5**	**10**	**15**	**20**	**30**
Demand 1	**1 Cell**	76.94	76.94	76.94	76.94	76.94	76.94
	2 Cell	80.34	80.34	80.34	80.34	80.34	80.34
	3 Cell	83.68	83.73	83.73	83.73	83.73	83.73
	4 Cell	91.89	92.03	94.44	98.16	99.48	100.6
	5 Cell	92.37	94.35	94.87	98.78	101.1	103.5
Demand 2	**1 Cell**	98.06	98.11	98.34	98.11	98.54	98.54
	2 Cell	100.6	100.8	101.2	100.8	101.6	101.6
	3 Cell	103.0	103.2	104.0	103.4	104.7	104.7
	4 Cell	108.3	108.6	109.4	111.5	109.9	111.5
	5 Cell	108.6	108.9	109.9	111.5	109.9	111.5

4.3 Hard Shoulder Running

4.3.1 Changes in fundamental diagram under HSR

Hard shoulder running (HSR) increases road capacity by providing an extra lane to road users. In real practice, HSR is always applied with reduced speed for safety reasons [4], and the speed limit set to less than or equal to 50 mph. For simplicity, the following discussion of HSR always operates with a speed limit control at 50 mph. The average free flow speed of traffic will be reduced while the capacity will be increased slightly, which is discussed in Section 4.2.1. The slight increase in capacity is attributed to the shorter headways between vehicles under a lower speed limit. Therefore, the free flow speed is reduced when the HSR (including the effect of VSL) applied on the motorway. Moreover the capacity and the jam density increase significantly because an extra lane will be used.

Figure 4.4 illustrates schematically the impact of HSR on the fundamental diagram. The dash and solid lines represent the fundamental diagrams with and without HSR respectively. The capacity and jam density are higher than no control case because an extra lane (hard shoulder lane) opens for road users. The notations (v^*, Q^*, w^* and $\bar{\rho}^*$) and (v, Q, w and $\bar{\rho}$) represent the free flow speed, capacity, wave speed and jam density with and without HSR respectively.

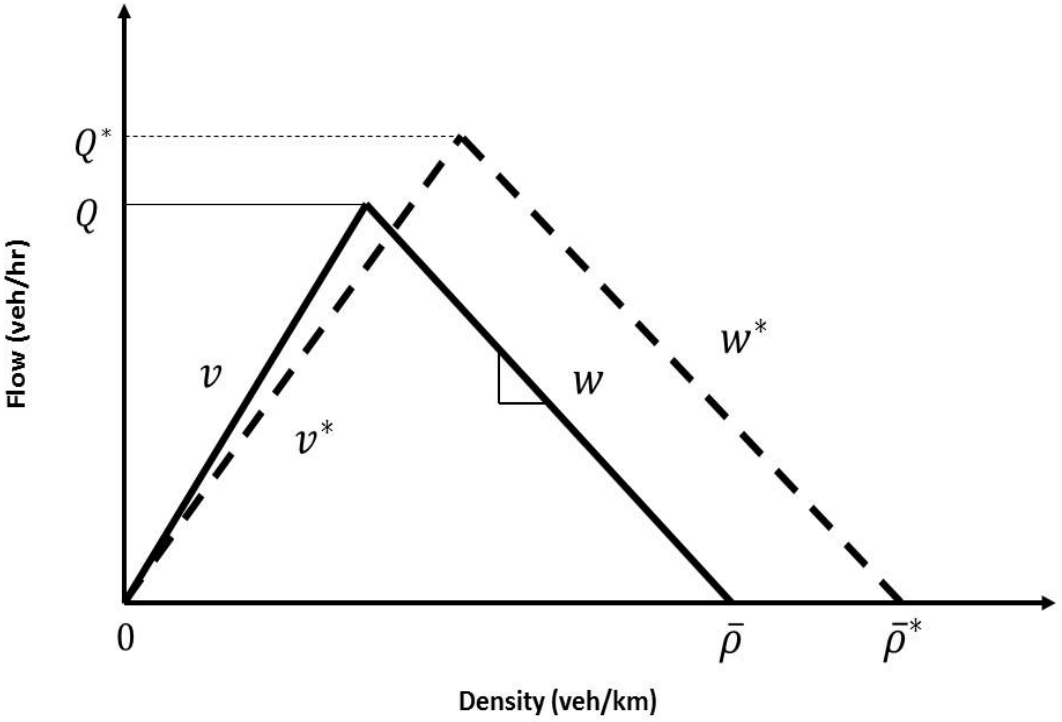

FIGURE 4.4: Changes in fundamental diagram under HSR

4.3.2 Optimisation of HSR

Similar to VSL control, this study adopts a CTM based mixed integer linear programming (MILP) formulation for solving the HSR control policy. A set of 0 - 1 binary decision variables $\mu_i(k)$ is introduced as indicators to choose between the fundamental diagram with and without HSR. The solution of the problem reveals the optimal deployment of the corresponding HSR control strategy over time and space, which means the best time and location to open the hard shoulder as an extra lane with VSL. With the binary variable $\mu_i(k)$, constraints (3.17 - 3.20), which are the constraints on the outflow in CTM, are replaced for all i and k as constraints (4.47 - 4.54):

$$\min_{\mathbf{c_h}} \; Z = \sum_{i=1}^{I}\sum_{k=1}^{K}\Big(\rho_i(k)\Delta x_i\Delta t - \frac{f_i(k)\Delta x_i\Delta t}{v_i}\Big) + \zeta\sum_{i=1}^{I}\sum_{k=1}^{K}\mu_i(k) \tag{4.45}$$

subject to:

$$\begin{aligned}
\rho_i(k+1) &= \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k)\Big) && (4.46)\\
f_i(k) &\leq \rho_i(k)v_i\bar{\beta}_i + \mu_i(k)M && (4.47)\\
f_i(k) &\leq Q_i + \mu_i(k)M && (4.48)\\
f_i(k) &\leq Q_{i+1} + \mu_{i+1}(k)M && (4.49)\\
f_i(k) &\leq w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big) + \mu_{i+1}(k)M && (4.50)\\
f_i(k) &\leq \rho_i(k)v_i^*\bar{\beta}_i + \big(1-\mu_i(k)\big)M && (4.51)\\
f_i(k) &\leq Q_i^* + \big(1-\mu_i(k)\big)M && (4.52)\\
f_i(k) &\leq Q_{i+1}^* + \big(1-\mu_{i+1}(k)\big)M && (4.53)\\
f_i(k) &\leq w_{i+1}^*\big(\bar{\rho}_{i+1}^* - \rho_{i+1}(k)\big) + \big(1-\mu_{i+1}(k)\big)M && (4.54)\\
r_j(k) &= \lambda_j(k) && (4.55)
\end{aligned}$$

The optimal HSR policy $\boldsymbol{c_h}$ is to be implemented over all cells and time steps that minimise the total system delay Z in the system. The notation Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. $\lambda_j(k)$ denotes the traffic demand that wants to enter the system through on-ramp j during time step k, and $r_j(k)$ is the actual demand that enters the system. The constraint (4.55) shows the ramp inflow equals to the actual ramp demand, which means there is no ramp control is applied on the road. Moreover, constraint (4.46) is the conservation equation to update the density in cell i for next time step $k+1$.

The notation M represents a very large number where it is set to be 99,999; The notation (v_i, Q_i, w_i and $\bar{\rho}_i$) and (v_i^*, Q_i^*, w_i^* and $\bar{\rho}_i^*$) are free flow speed, capacity, the shock wave speed and jam density without and with hard shoulder running respectively. The notation $\mu_i(k) = 1$ implies the HSR being opened at cell i during time step k; $\mu_i(k) = 0$ means otherwise. If $\mu = 0$, the constraints (4.51 - 4.54) will be disabled due to the big number M.

In order to maximise the capacity along the motorway, the hard shoulder could be run over the whole optimisation period. However, this will have to come at the expense of safety. The objective function is adjusted according to this, with which the hard shoulder optimisation problem. The last term in the objective function is the sum of all HSR control variables, which increases with the number of HSR in operation. The parameter ζ is the balance between the consideration of safety and capacity. A larger ζ indicates that the more weight on safety. For the extreme case ($\zeta = 0$), all available hard shoulder lanes will be open all the time as that will provide maximum physical capacity when no consideration is given to safety or incident management. The optimal HSR control policy $\mathbf{c_h}$ is to be implemented over all cells and time steps that minimised the total system delay Z.

4.3.3 Optimisation of integrated control strategy

The integration of three control strategies including ramp metering, variable speed limits and hard shoulder running is also considered in this study. The optimisation problem for integrated control strategy is also formulated as a mixed integer linear programming (MILP) problem. The integrated control optimisation problem can be formulated by

considering all relevant ramp metering and HSR constraints as follows:

$$\min_{\mathbf{c_{rvh}}} Z = \sum_{i=1}^{I}\sum_{k=1}^{K}\Big(\rho_i(k)\Delta x_i\Delta t - \frac{f_i(k)\Delta x_i\Delta t}{v_i}\Big) + \eta\sum_{j=1}^{J}\sum_{k=1}^{K} l_j(k)\Delta t + \zeta\sum_{i=1}^{I}\sum_{k=1}^{K}\mu_i(k) \tag{4.56}$$

subject to:

$$\begin{aligned}
\rho_i(k+1) &= \rho_i(k) + \frac{\Delta t}{\Delta x_i}\Big(f_{i-1}(k) - \frac{f_i(k)}{\bar{\beta}_i} + r_j(k)\Big) && (4.57)\\
f_i(k) &\le \rho_i(k)v_i\bar{\beta}_i + \mu_i(k)M && (4.58)\\
f_i(k) &\le Q_i + \mu_i(k)M && (4.59)\\
f_i(k) &\le Q_{i+1} + \mu_{i+1}(k)M && (4.60)\\
f_i(k) &\le w_{i+1}\big(\bar{\rho}_{i+1} - \rho_{i+1}(k)\big) + \mu_{i+1}(k)M && (4.61)\\
f_i(k) &\le \rho_i(k)v_i^*\bar{\beta}_i + \big(1-\mu_i(k)\big)M && (4.62)\\
f_i(k) &\le Q_i^* + \big(1-\mu_i(k)\big)M && (4.63)\\
f_i(k) &\le Q_{i+1}^* + \big(1-\mu_{i+1}(k)\big)M && (4.64)\\
f_i(k) &\le w_{i+1}^*\big(\bar{\rho}_{i+1}^* - \rho_{i+1}(k)\big) + \big(1-\mu_{i+1}(k)\big)M && (4.65)\\
l_j(k+1) &= l_j(k) + \big(\lambda_j(k) - r_j(k)\big)\Delta t && (4.66)\\
l_j(k) &\le \bar{l}_j && (4.67)\\
r_j(k) &\le \bar{r}_j && (4.68)\\
r_j(k) &\ge 0 && (4.69)\\
r_j(k) &\le \frac{l_j(k)}{\Delta t} + \lambda_j(k) && (4.70)\\
r_j(k) &\le \big(\bar{\rho}_j - \rho_j(k)\big)\frac{\Delta x}{\Delta t} + \mu_j(k)M && (4.71)\\
r_j(k) &\le \big(\bar{\rho}_j^* - \rho_j(k)\big)\frac{\Delta x}{\Delta t} + \big(1-\mu_j(k)\big)M && (4.72)
\end{aligned}$$

The optimal policy $\mathbf{c_{rvh}}$ is to be implemented over all cells and time steps that minimise the total system delay Z. The notation Δt denotes the length of simulation time step, and Δx_i represents the length of cell i. $\rho_i(k)$ and $f_i(k)$ are the density in cell i and the outflow from cell i respectively. The notation M represents a very large number where it is set to be 99,999; The notation (v_i, Q_i, w_i and $\bar{\rho}_i$) and (v_i^*, Q_i^*, w_i^* and $\bar{\rho}_i^*$) are free flow speed, capacity, the shock wave speed and jam density without and with hard shoulder running respectively. Constraint (4.57) is the conservation equation to update the density in cell i for next time step $k+1$.

By considering the ramp metering, additional constraints on the ramp queue length (constraints 4.66 - 4.67) should add to the HSR formulations. The notation $\bar{l}_j$ is defined as the maximum queue length on the ramps such that an unacceptable long queue on the on-ramp will not be obtained as an optimisation result. Constraint (4.67) is the conservation equation on the queue length, where $\lambda_j(k)$ denotes the traffic demand that wants to enter the system through on-ramp j during time step k, and $r_j(k)$ is the actual demand that enters the system.

Moreover, constraints on the ramp inflow also need to be considered in the integrated control strategy. The limitations on the maximum value, non-negativity and ramp demand (constraints 4.68 - 4.70) remain the same in the ramp metering control. However, the constraint on main road space (constraint 3.33) should be adapted due to the jam density which could be changed when HSR is applied on the motorway (constraints 4.71 and 4.72). The notation $\bar{r}_j$ refers to ramp capacity at on-ramp j.

4.4 Working Example

The optimal control models are now applied to a case study with traffic data collected from a 10-km section (between Junctions 13 and 15) of the orbital M25 Motorway (direction: clockwise) in London, England. The section covers three on-ramps (one at Junction 14 and two at Junction 15) and two off-ramps (one at Junction 14 and one at Junction 15). The motorway stretch contains 20 detector stations with an average spacing of 500 metres. The data were collected on 3 October 2012 (Wednesday). The on-ramps are located at cell 10 (Junction 14), 17 (Junction 15a) and 18 (Junction 15b). The off-ramps are located at cell 5 (Junction 14) and 15 (Junction 15). The length of the simulation time step is set to be 15 seconds, which gives the total number of time steps $K = 720$ over a three-hour evening peak period [17:00 - 20:00]. The optimal control problems of VSL and HSR are implemented and solved by IBM ILOG CPLEX running on the same desktop computer described previously.

4.4.1 Variable speed limits

The effectiveness of VSL for congestion management is discussed in this section. The normal speed limit is 70 mph on motorways in United Kingdom, while a reduced speed limit of 50 mph is considered as an alternative. Based upon empirical observations, it can be assumed that the capacity is increased slightly by 1 per cent when a lower speed limit (50 mph) is used.

Following real operations, the cells 4 to 19 are specified as the feasible VSL control region, and the speed limit is updated every 5 minutes $\big($equals to 20 (time steps) $= 5 \times 60/15\big)$.

The problem hence consists of 504 VSL control variables $\left(14 \text{ (cells)} \times \frac{720 \text{ (3 hr)}}{20 \text{ (5 min)}} = 504\right)$, and takes about 30 minutes to solve. Figure 4.5 shows the density contours with (right) and without (left) VSL, the total system delay is reduced with VSL from 2145 veh-hr to 1913 veh-hr (see Table 4.3). The main road delay reduction is not as great as the ramp metering case (VSL: 1913 veh-hr versus RM:1757 veh-hr), while there is no extra ramp delay induced with VSL. Considering the overall total system delay, VSL indeed are able to produce a better performance than the ramp metering control (VSL: 1913 veh-hr versus Ramp metering: 2106 veh-hr).

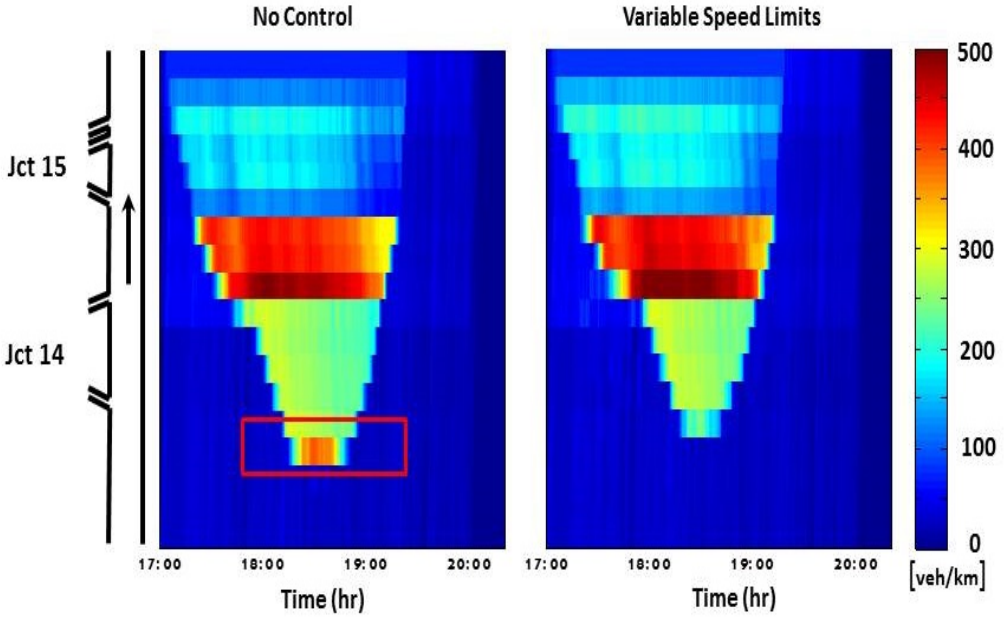

FIGURE 4.5: Main road densities with and without VSL

left - no control; right - VSL

To gain further insight, Figure 4.6 depicts graphically the optimal VSL strategy in which the white grids represent the location (cells) and time (VSL control intervals) where the 50 mph speed limit is used. In general, a lower speed limit will be adopted at congested

regions to gain a slightly higher capacity. Moreover, it is expected that the gain in discharge flow outweighs the reduction in speed as suggested by the overall reduction in total delay.

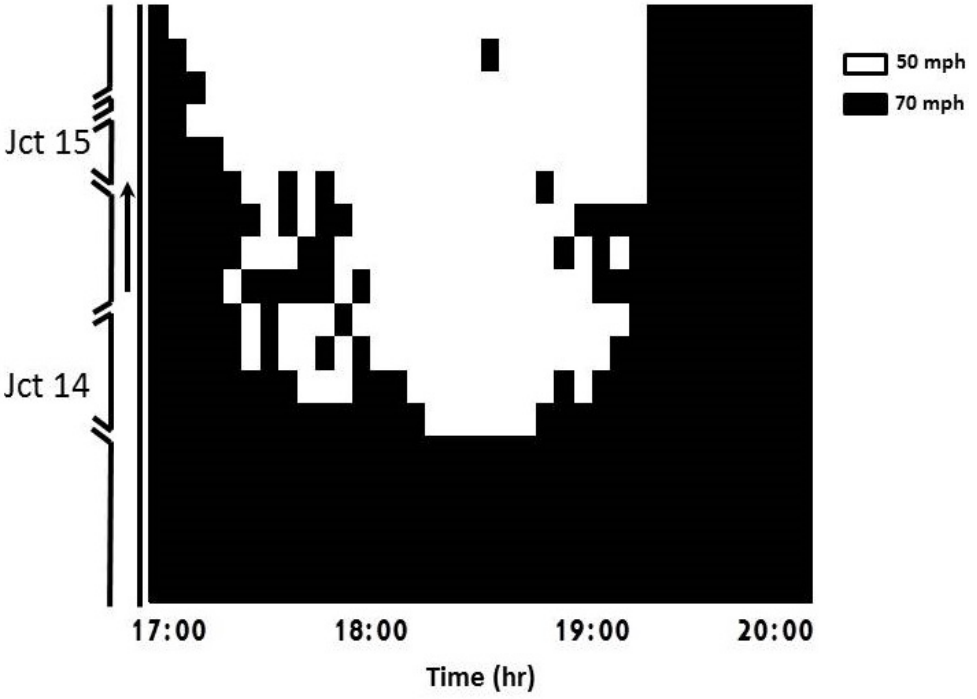

FIGURE 4.6: VSL strategies

4.4.2 Hard shoulder running

An assumption here is that the lane with an extra hard shoulder will give an additional 700 vehicles per hour capacity to the corresponding road section under 50 mph speed limit. In addition, the trade-off parameter ζ between efficiency and safety is set to be 0.3 veh-hr. Moreover, cells 14 to 19 are specified as the feasible HSR control region in which HSR can be applied after the existing road configuration has been checked. Then another assumption here is that the HSR control interval is 5 minutes (equals to 20 time

steps), so that HSR strategies can be updated only every 5 minutes. Consequently, the problem consists of 216 HSR control variables (6 (cells) $\times \frac{720\ (3\ \text{hr})}{20\ (5\ \text{min})} = 216$), and takes about 10 minutes to solve. Figure 4.7 compares the density contours with (right) and without (left) HSR. The layout of the road section is shown on the left of the plots. The main road delay reduces significantly from 214,5 veh-hr (no control case) to 595.2 veh-hr (HSR case).

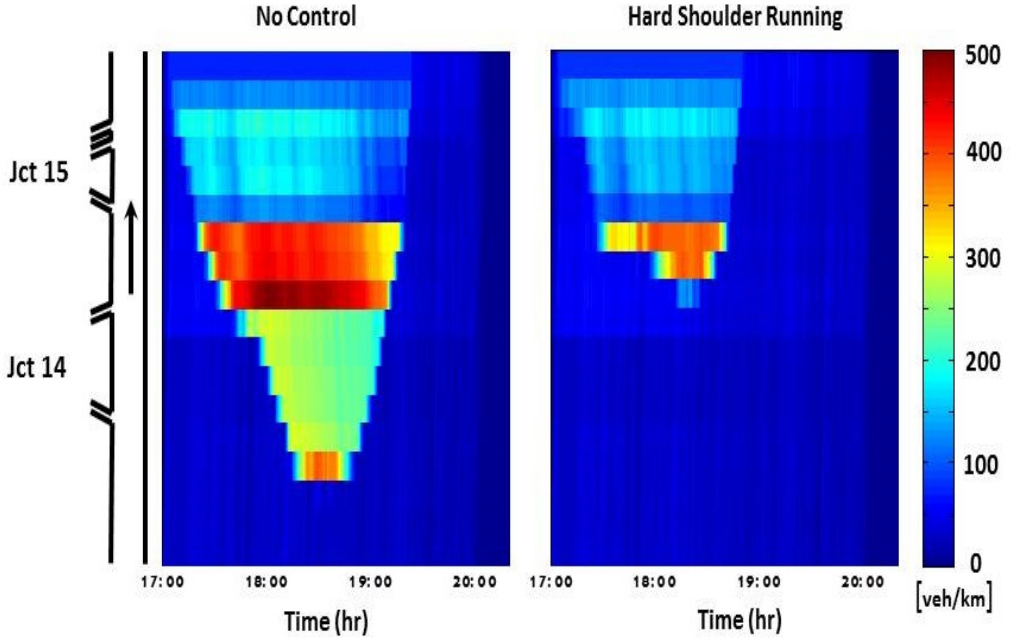

FIGURE 4.7: Main road densities with and without HSR

left - no control; right - HSR

The optimal solution shows that the hard shoulder is opened only at cell 18 during the period from 17:00 to 18:50, with $\zeta = 0.3$. To provide further insight into the sensitivity of ζ on optimal hard shoulder operations, Table 4.2 summarises the performances of HSR under different ζ. Capacity will be given a higher priority with a lower ζ adopted and hence more hard shoulder lanes will be utilised. An extreme case is when ζ is set to be zero, then hard shoulder lanes will be opened at all cells (cells 14 through 19) during

the entire study period. Table 4.2 shows that cell 18 is the first location where the hard shoulder will be used, and it is followed by cells 14 through 19. The sequence of hard shoulder opening generally follows the sequence of the onset of congestion over space.

Table 4.2: Sensitivity analysis of ζ on HSR operations

		Hard Shoulder Running by cell					
ζ	**Total Delay veh-hr**	**14**	**15**	**16**	**17**	**18**	**19**
$\geq$0.20	574.7	N/A	N/A	N/A	N/A	17:00-18:50	N/A
0.19	571.7	17:30-18:40	N/A	N/A	N/A	17:00-18:50	N/A
0.15	569.4	17:30-18:45	17:30-18:35	N/A	N/A	17:00-18:50	N/A
0.14	569.4	17:30-18:40	17:30-18:40	N/A	N/A	17:00-18:50	N/A
0.13	567.4	17:30-18:45	17:30-18:40	17:30-18:35	N/A	17:00-18:50	N/A
0.12	565.5	17:30-18:40	17:30-18:40	17:30-18:40	17:25-18:35	17:00-18:50	N/A
0.10	565.4	17:30-18:40	17:30-18:40	17:30-18:40	17:25-18:40	17:00-18:50	N/A
0.09	565.3	17:30-18:40	17:30-18:40	17:25-18:40	17:25-18:40	17:00-18:50	N/A
0.07	565.2	17:30-18:40	17:30-18:40	17:25-18:45	17:25-18:40	17:00-18:50	N/A
0.05	565.2	17:30-18:45	17:30-18:40	17:25-18:40	17:25-18:40	17:00-18:50	N/A
0.01	565.2	17:30-18:45	17:25-18:45	17:25-18:40	17:25-18:40	17:00-18:50	N/A
0.00	565.2	17:00-20:00	17:00-20:00	17:00-20:00	17:00-20:00	17:00-20:00	17:00-20:00

Note: N/A = Not applicable as hard shoulder is not running in this cell.

4.4.3 Integrated control strategy

To identify the performance of the integrated control, we assume the hard shoulder lane gives an additional (about 700 veh/hr) capacity to the corresponding road section. The trade-off parameter between efficiency and safety is also set to be 0.3 veh-hr. Moreover, cells 14 to 19 are specified as the feasible HSR and VSL (50 mph) control region.

Figure 4.8 compares the no-control case (left) with the integrated control case (right) where the three on-ramps are metered with maximum queue length (60 vehicles). The colour scale represents the level of traffic density at the corresponding time and location. The layout of the road section is shown on the left of the plots. Figure 4.8 shows the integrated control strategy relieves traffic congestion. This reveals that the integrated control strategy is effective in reducing the main road delay. Moreover, the hard shoulder with 50 mph speed limit is only applied at cell 18 (see Table 4.2).

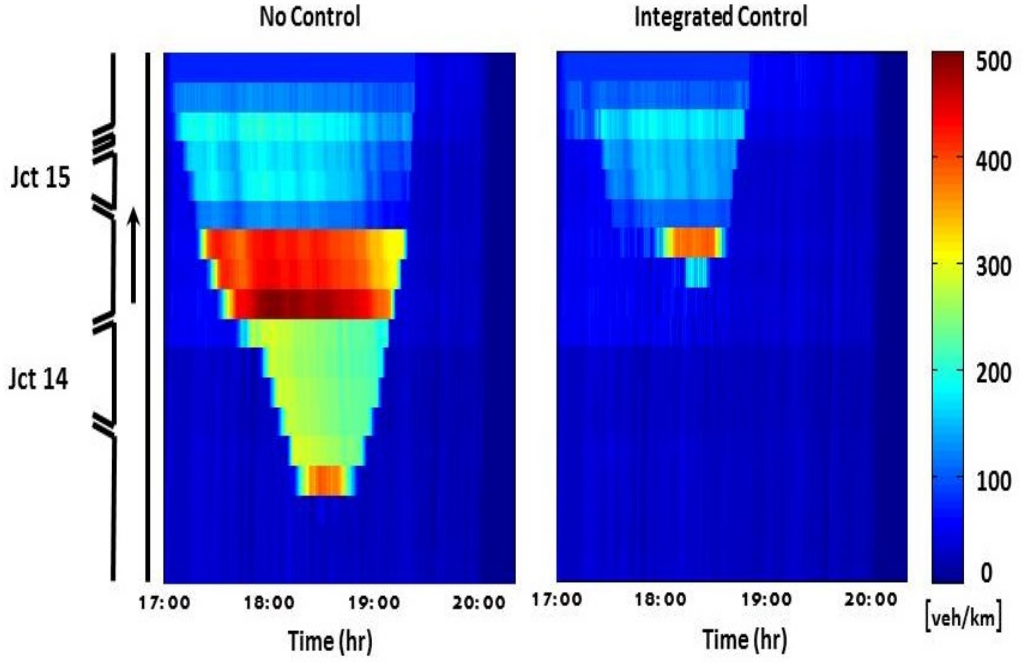

FIGURE 4.8: Main road densities with and without integrated control

left - no control; right - integrated control

The performance of the integrated strategy is better than separated controls as expected because the integrated control obtains the most of benefit from each control strategy. The P_m and P_t in Table 4.3 are the relative reduction on main road delay and total system delay respectively as shown in Equation (3.34). By comparing the numerical results in Table 4.3, it is found that the relative reduction on total system delay of RMVSL

(12%) is approximately equal to the sum of the relative reduction on total system delay under VSL (11%) and ramp metering (1.8%). The small difference (0.8%) comes from the different ramp delay. The ramp delay of RMVSL (321.9 veh-hr) is lower than the ramp delay of ramp metering (349.0 veh-hr). The integrated control strategy shows a similar trend. Variable speed limits and HSR influence the character of the motorway such as free flow speed and jam density, and ramp metering changes the congested location (main road or ramp) and reduces the spill over. Therefore, VSL and HSR induce more reduction on delay compared with ramp metering.

Table 4.3: Delays under different control strategies

Delay [veh-hr]	**Main**	P_m	**Ramp**	**Total**	P_t
NO	2,145		0.00	2,145	
RM	1,757	18 %	349.0	2,106	1.8 %
VSL	1,913	11 %	0.00	1,913	11 %
HSR	595.2	72 %	0.00	595.2	72 %
RMVSL	1,564	27 %	321.9	1,886	12 %
Integrated	359.5	83 %	198.7	558.2	74 %

4.5 Summary

This chapter presents a optimisation framework that seeks an optimal VSL, HSR, RMVSL, and integrated control strategies for motorway traffic management. The optimisation models are formulated based upon CTM. With the linearity of CTM, the optimal control problems are formulated as a MILP problem and solved by CPLEX. One may argue that the computational complexity associated with the MILP formulation, in particular the issue of the 'curse of dimensionality' as the solution space increases. Nevertheless, it should be noted that there are a number of efficient algorithms which exist for solving MILP problems effectively. In particular, the speed control problem presented here can be solved by using a branch-and-bound algorithm which is known to be readily parallelized for the highly effective parallel computation. This chapter also presents various analyses on the sensitivity of the effectiveness of the control strategies with respect to different model settings and assumptions.

The optimisation control strategies are applied to a case study on the UK M25 Motorway. The optimal control strategies are derived by minimising the total system delay, and the solutions are obtained in a reasonable computational time. The results show that all control strategies can effectively reduce total system delay. Hard shoulder running appears very effective approach in reducing total system delay by providing an extra physical lane to motorways. Nevertheless, engineers must be careful in balancing the trade-off between mobility and safety, as HSR removes the buffer intended for incident management. Moreover, an extra lane could imply an increase in travel demand, which has not been considered in this chapter.

Chapter 5

Robust Optimisation of Ramp Metering

5.1 Introduction

It is known that the traffic condition becomes unstable following the onset of congestion [72]. Such traffic flow variability is a growing concern because of its implication on travel reliability in particular during peak periods. This chapter presents a robust ramp metering optimisation framework that incorporates the set-valued fundamental diagram. In addition to the fundamental diagram, we also consider the uncertainty on the demand side due to various measurement or estimation errors [45]. The robust optimisation is formulated as a minimax problem based upon CTM, and solved by a two stage solution procedure.

This chapter is organised as follows: Section 5.2 introduces the basic characteristics of

traffic flow and the sources of uncertainty in traffic state estimation. Section 5.3 presents a robust optimisation framework, which aims to minimise and stabilise travel delay over a range of uncertain scenarios. The performances of different control policies are illustrated and compared through working examples in Section 5.4. Finally, Section 5.5 provides some concluding remarks.

5.2 Characteristic of Road Traffic Flow

5.2.1 Set-valued fundamental diagram

In a traffic model, traffic characteristics are typically represented by a flow-density function, which is known as a fundamental diagram. Fundamental diagrams can be derived by using standard loop detector data (MIDAS in England), which includes measurements of flow (in vehicles per hour), density (in vehicles per kilometre), and speed (in kilometres per hour). Figure 5.1 shows the flow-density scatter plot of data collected at a detector station (4955A) on the M25 Motorway (clockwise) in the United Kingdom. The detector station consisting of three lanes is located at the upstream of Junction 15 connecting with the London Heathrow (LHR) Airport. The data are collected over 5 weekdays from 24 September 2012 (Monday) to 28 September 2012 (Friday) and processed into 5-min averages. The fundamental diagram can be obtained by the three step calibration procedure (free flow line, capacity, and congested line) as described in Chapter 2.

The main challenge of calibrating the fundamental diagram lies in representing the congested data which exhibits a high level of uncertainty due to the underlying complicated

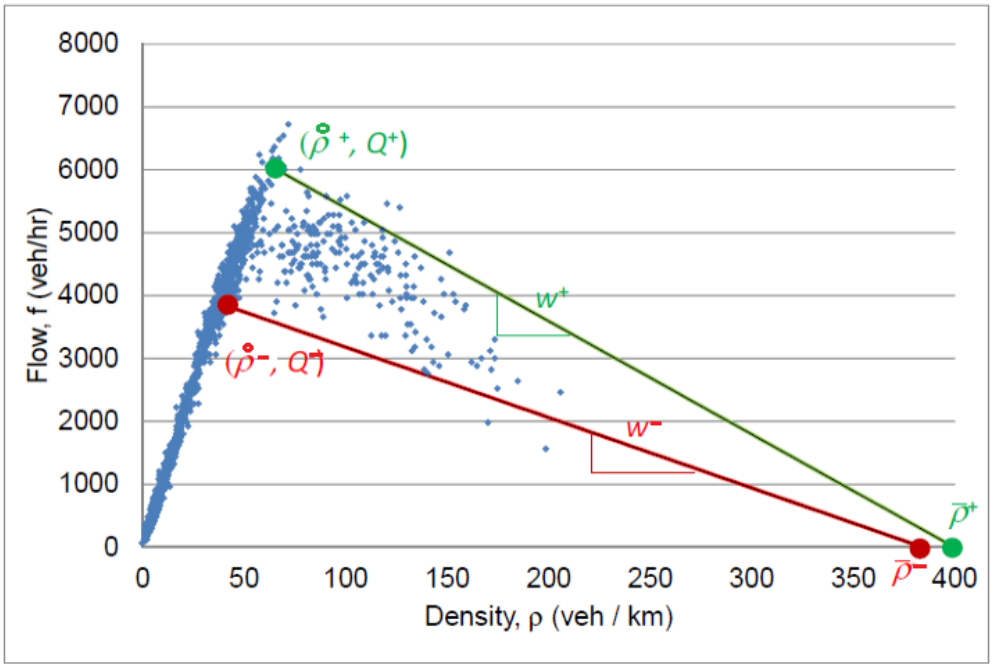

FIGURE 5.1: Set-valued fundamental diagram

(Detector Station: 4955A, M25 (clockwise))

traffic dynamics (e.g. capacity drop). This highlights the difficulty of modelling dynamic traffic with standard specifications of fundamental diagram to model traffic dynamics especially under congested conditions. A number of studies have been conducted to analyse and incorporate the uncertainty in congested traffic. Recently some studies have been proposed using a stochastic traffic modelling framework to address this variability. Estimating the probability distributions associated with these uncertain quantities is shown to be difficult (Brilon *et al.* [13]; Chow *et al.* [20]). Brilon *et al.* [13] present the stochastic concept of capacity based on a series of studies of German motorways. Ngoduy [75] also presents a stochastic fundamental diagram based on multi-class first-order traffic model. The simulation result shows that the wide scattering in the flow-density relationship is due to the random variations in driving behaviour. Sumalee *et al.* [95] propose a stochastic version of CTM called SCTM. Nevertheless, their model requires

pre-definition of various traffic state transitions and associated probabilities which are not easy to determine and implement in practice. Moreover, the computational effort of incorporating explicitly these probability distributions into optimisation framework has also been shown to be demanding [112].

As an alternative, a set-valued fundamental diagram modelling approach is proposed by Kurzhanskiy and Varaiya [55]. Following Kurzhanskiy and Varaiya [55], a value of density in the congested region is associated with a range of flow values which bounded a predefined interval without the need to specify the underlying probability distribution function. This gives rise to an interval estimation of traffic flow in a congested situation and it has been shown this interval estimation provides an important new insight to traffic state estimation.

However, Kurzhanskiy and Varaiya [55] do not describe the calibration methods for the set-valued fundamental diagram. The trained regression proposed in Chapter 2 is extended for the set-valued fundamental diagram. The congested data along the density (horizontal) axis is partitioned into a series of non-overlapping bins containing a certain number (e.g. 10 in this study) data points. Given the 10 density and flow pairs in the bin: $(f_1, \rho_1), (f_2, \rho_2), ..., (f_{10}, \rho_{10})$. Horizontally, each bin is represented by a 'BinDensity', which is the average of the 10 density values in the bin. The formulation is shown as follows:

$$\text{BinDensity} = \sum_{n=1}^{10} \rho_i \Big/ 10 \tag{5.1}$$

Vertically, each bin is represented by a range of flow which is bounded by the largest and the smallest non-outlying flow values in the bin. Under the given flow values, the

largest (BinFlow$^+$) and smallest (BinFlow$^-$) non-outlying values are looked for as:

$$\text{BinFlow}^+ = \max_{f_i}\{f_i|B, f_i \in f_i < Q3 + 1.5IQR\} \tag{5.2}$$

$$\text{BinFlow}^- = \min_{f_i}\{f_i|B, f_i \in f_i > Q1 + 1.5IQR\} \tag{5.3}$$

where $Q1$, $Q3$, and IQR are the 25^{th} percentile flow, 75^{th} percentile flow, and inter-quantile range (i.e. difference between 25^{th} percentile and 75^{th} percentile flow values) in the bin respectively. Two least-square estimations are performed on the [BinDensity - BinFlow$^+$] and [BinDensity - BinFlow$^-$] pairs to obtain the upper and lower bounds of the congested line. The corresponding capacity points $(\mathring{\rho}^+, Q^+)$ and $(\mathring{\rho}^-, Q^-)$ are determined as the intersections of the upper congested line and lower congested line with the free flow line. It is noted that here the constrained least estimation is not adopted because the constraint requiring the congested line passing through a predefined capacity point induces bias in the estimation.

Figure 5.1 shows the result of this set-valued (or interval) estimation at detector station (4955A) at Junction 15 on the M25 Motorway. As shown in Figure 5.1, the capacity is then regarded as an uncertain variable lying within a given range $[Q^-, Q^+]$. Likewise, the jam density is also an uncertain variable lying in $[\bar{\rho}^-, \bar{\rho}^+]$. The set-valued fundamental diagram at other detector stations can be determined by using the same method, and can be incorporated into the CTM simulation framework. With the set-valued fundamental diagram determined, interval estimates of flows $[f^-(k), f^+(k)]$ and densities $[\rho^-(k), \rho^+(k)]$ can be derived from repeated CTM runs with the lower bound and upper bound of the fundamental diagram parameters at each cell [55]. Therefore, the interval estimates will be obtained rather than point estimates as discussed previously.

5.2.2 Demand uncertainty

In addition to the flow-density uncertainty, we also observe the demand uncertainty which is due to the randomness in the estimation of traffic inflows to the motorway. This can be due to transient variations, day-to-day variations in drivers' behaviour, incidents, and other special events (Heydecker [45]; Lo and Chow [66]; Yin [106]). The uncertainty on the demand side can also be associated with the health and quality of the on-site sensors. For example, it is not uncommon to find several percent error in flow measurement due to the configuration of detectors and its alignment with traffic flow. Moreover, for real time control purposes some short-term (e.g. 5 minutes) demand flow prediction will be required, which will induce an additional prediction error [66]. Furthermore, one would expect to see a severe underestimation of demand flows should any associated detector(s) break down or malfunction unexpectedly, which is also common in real life operations.

Figure 5.2 shows a scatter plot of 15-min on-ramp flow (Detector Station: 4959K) against the 15-min main road discharge flow at detector station (4955A) at Junction 15. The data are collected during the afternoon peak [17:30 - 19:00] over 20 weekdays from 10 September 2012 to 5 October 2012 through MIDAS. It is noted that the entire time period is congested and this implies the main road traffic is flowing at capacity flow. The scatter plot (Figure 5.2) shows a high degree of demand variability in which the on-ramp flow varies over the range of [1,500 veh/hr - 2,500 veh/hr], while the discharge flow varies over [4,000 veh/hr - 6,000 veh/hr]. The observed capacity variation in Figure 5.2 is consistent with the one observed from the fundamental diagram in Figure 5.1. This demand variability should be captured in the design of control strategy.

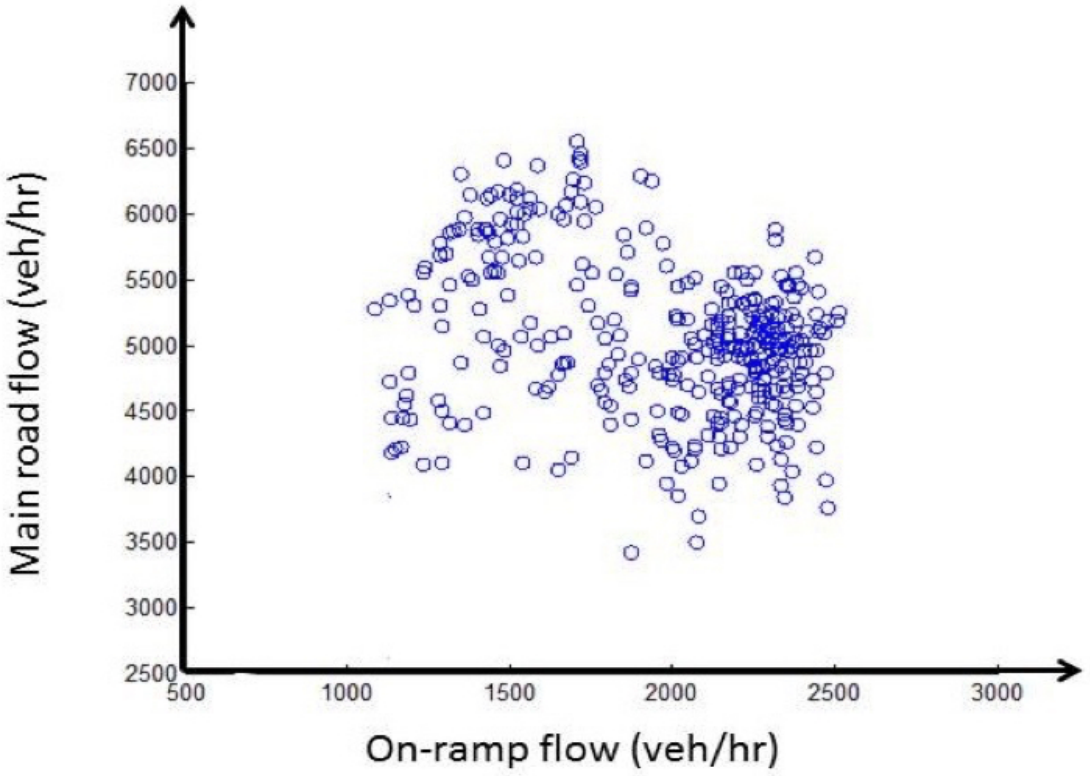

FIGURE 5.2: Demand variability
(Detector Station: 4955A, M25 (clockwise))

5.3 Robust Optimisation of Ramp Metering

5.3.1 Review of robust optimisation

A number of studies have been done to incorporate the stochasticity in traffic dynamics into the modelling framework. It can be broadly categorised into stochastic programming and robust optimisation. The difference between this stochastic programming and robust

optimisation is that the stochastic programming considers the probability distribution of the uncertainties instead of simply using upper and lower bounds. The objective of this stochastic programming is then to minimise the total 'expected' delay instead of the delay under the worst case scenario.

There have been a number of studies attempting to capture this stochastic variation in congested traffic. Sumalee *et al.* [95] propose a stochastic version of CTM. Nevertheless, [95]'s model requires pre-definition of various traffic state transitions (e.g. from free-flow to congestion, and vice-versa) and associated probabilities which are not easy to determine and implement in practice. Such stochastic programming can come up with a more efficient solution through taking explicitly the distribution of uncertainties into account. Nevertheless, the computational process can be intractable as in principle the solution algorithm will have to search through all possible realisations of the uncertain variables. The computational effort hence can increase exponentially with additional variables, and this phenomenon is known as the curse of dimensionality [89]. As discussed in [113], this stochastic approach is theoretically elegant. However, it requires enumerating the set of potential states.

Bertsimas *et al.* [9] present a comprehensive review of different robust optimisation in theory and applications. Some researchers adopt the robust optimisation to incorporate the uncertainty in traffic dynamics. Liu *et al.* [62] adopt the distributionally robust optimisation to solve the on-line signal control problem. Han *et al.* [40] reformulate the LWR-Emission problem as a mixed integer linear program (MILP) by using robust optimisation, which minimise the expected vehicle delays and vehicle emission through road network. Considering the practicality, we adopt an alternative approach proposed

by Kurzhanskiy and Varaiya [54] and Kurzhanskiy and Varaiya [55]. In [55]'s formulation, the fundamental diagram is considered to be uncertain in capacity, critical and jam densities while the density state equation remains the same as the deterministic case (Equation 2.1). The uncertainty in capacities, critical and jam densities is specified by an interval. Kurzhanskiy and Varaiya [55] show that interval estimates of traffic densities can be constructed by simulations with the upper and the lower bounds of the set-valued fundamental diagram. Kurzhanskiy and Varaiya [55] also prove that their stochastic version of CTM is monotonic and it can be reduced to the conventional deterministic form with zero uncertainty considered. Next section illustrates how to define the likelihood region for uncertainty in traffic dynamics.

5.3.2 Likelihood region

Figure 5.3 shows the configuration of on-ramp. The λ_j and r_j denote the traffic demand that wants to enter the system and the actual demand that enters the system respectively, and the road section is characterised by the fundamental diagram $\boldsymbol{\Phi}$ consisting of parameters (capacity: $\boldsymbol{Q}$, wave speed: $\boldsymbol{w}$, free flow speed: $\boldsymbol{v}$ and jam density: $\bar{\boldsymbol{\rho}}$). Let $\boldsymbol{\lambda} = \lambda_j(k)$ be the collection of all demand flows from the on-ramps j at time step k, and $\boldsymbol{\Phi}$ denotes the set of fundamental diagrams of all cells $i \in I$. If we assume $\boldsymbol{w}$ and $\boldsymbol{v}$ are constant values, the capacity $\boldsymbol{Q}$ represents the fundamental diagram $\boldsymbol{\Phi}$ because the jam density is the function of the capacity. Unlike Zhong *et al.* [112], we do not need to specify the probability distributions of these demand flows and fundamental diagrams. We only need to specify the upper and lower bounds for $\boldsymbol{\lambda}$ (demand) and $\boldsymbol{\Phi}$ (supply) within the likelihood set $\boldsymbol{\Omega}$. The corresponding size and geometry of $\boldsymbol{\Omega}$ can be defined based on on-site measurements and engineers' judgement.

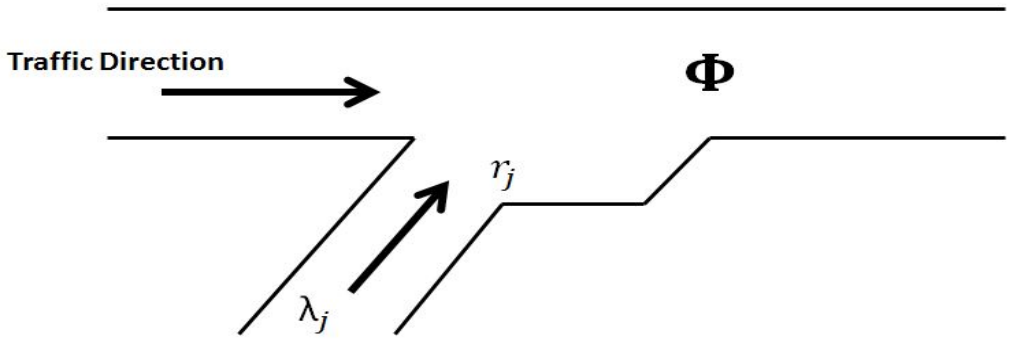

FIGURE 5.3: Configuration of on-ramp

Then we define a likelihood region $\mathbf{\Omega}$ as the union of all of these regions of uncertainty in fundamental diagrams and demand flows. This likelihood region $\mathbf{\Omega}$ can take various shapes such as box, polyhedral, or ellipsoidal [8]. Different geometries reflect different degree of correlation among the fundamental diagrams and demand flows over different locations. Moreover, the size of this likelihood region $\mathbf{\Omega}$ reflects the degree of uncertainty that we wish to take into account in the optimisation. A larger $\mathbf{\Omega}$ means the engineer is more concerned about robustness than efficiency. A 'point' (or single-valued) $\mathbf{\Omega}$ implies no uncertainty is considered and the optimisation will be reduced to the conventional deterministic one with no consideration given to robustness.

5.3.2.1 Box likelihood region

A box-constrained $\mathbf{\Omega}$ implies that the random variables $(\boldsymbol{\lambda}, \mathbf{\Phi})$ vary independently in the intervals as follows:

$$\Omega_{\lambda_j} = [\lambda_j^{min}, \quad \lambda_j^{max}] \tag{5.4}$$

for all λ_j on the on-ramps $j = 1, 2, ..., J$, where λ_j^{min} and λ_j^{max} represent the minimum and maximum possible values of λ_j respectively, and

$$\Omega_{\Phi_i} = [\Phi_i^{min}, \quad \Phi_i^{max}] \tag{5.5}$$

for all fundamental diagrams Φ_i on cells $i = 1, 2, ..., I$, where Φ_i^{min} and Φ_i^{max} represent the two fundamental diagrams that will give the minimum and maximum possible values of flow, respectively, based on a given value of density. The collection of all intervals shown in Equation 5.4 and 5.5 gives a box-constrained likelihood set:

$$\boldsymbol{\Omega_b} = \Omega_{\lambda_1} \times \Omega_{\lambda_2} \times \cdots \times \Omega_{\lambda_J} \times \Omega_{\Phi_1} \times \Omega_{\Phi_2} \times \cdots \times \Omega_{\Phi_I} \tag{5.6}$$

Adopting the box-constrained set $\boldsymbol{\Omega_b}$ in robust optimisation will lead to the worst scenario in which one will have the highest demand flows from all on-ramps and the lowest discharge flows everywhere along the main road at all times, which marks of the red solid circle in Figure 5.4. The box likelihood region is shown as the red solid rectangle in Figure 5.5 with empirical data. The data used in Figure 5.5 is described in Section 5.2.2 as shown in Figure 5.2. Nevertheless, such a scenario will be too conservative for engineering design purposes [8]. Moreover, a number of empirical studies (see Brilon *et al.* [13], Chow *et al.* [20]) reveal the correlation between demand flows and capacity through looking at the stochastic traffic breakdown events. However, this demand-capacity interaction cannot be captured by the box-constrained $\boldsymbol{\Omega_b}$ because the correlation among demand flows is not considered. Therefore, the other shaped uncertainty region need to be considered.

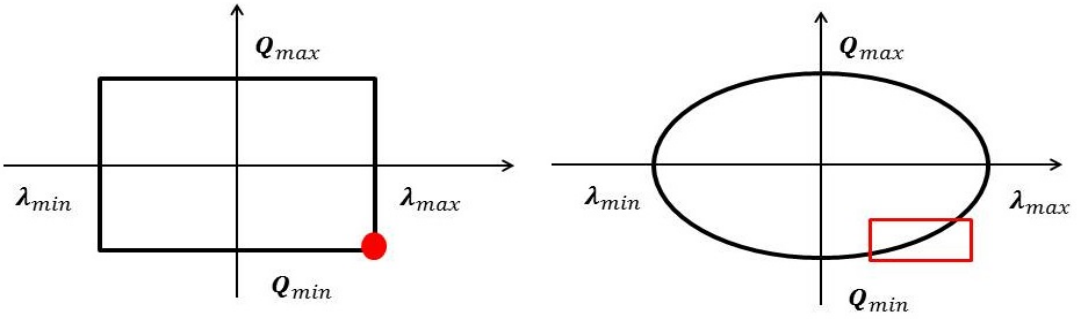

FIGURE 5.4: The geometry of $\mathbf{\Omega}$:

left: Box-constrained; right: Ellipsoidal-constrained

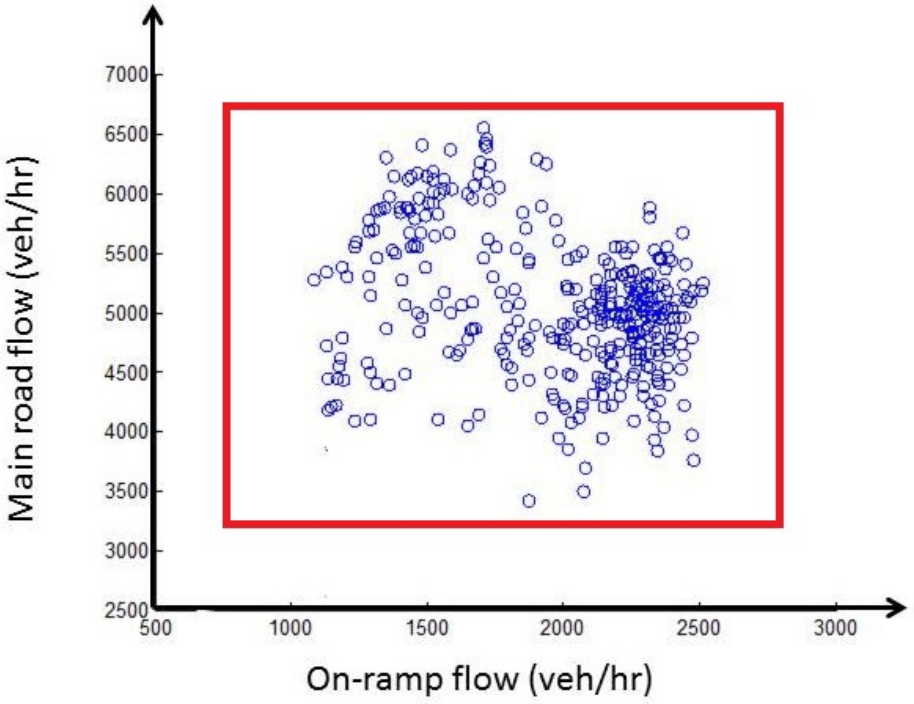

FIGURE 5.5: Uncertainty set with empirical data in box likelihood region
(Detector Station: 4955A, M25 (clockwise))

5.3.2.2 Polyhedral likelihood region

The uncertainty set could be defined as a polyhedral region. A polyhedral set $\mathbf{\Omega}$ implies that the random variables $(\boldsymbol{\lambda}, \boldsymbol{\Phi})$ within a finite number of linear constraints. For example, we may include the inequality constraint as follows:

$$\Omega_{\lambda_j} \leqslant \lambda_j^{max} \tag{5.7}$$

$$\Omega_{\Phi_i} \leqslant \Phi_i^{max} \tag{5.8}$$

where λ_j^{max} represents the maximum possible values of λ_j, and Φ_i^{max} represents the fundamental diagram that will give the maximum possible values of flow based on a given value of density.

5.3.2.3 Ellipsoidal likelihood region

To derive a more practical control strategy, we use an ellipsoidal likelihood set. Defined $\boldsymbol{\varphi} = (\boldsymbol{\lambda}, \boldsymbol{\Phi})$, where $\boldsymbol{\varphi} \in \mathbb{R}^{I+J}$, is the combination of all random variables $\boldsymbol{\lambda}$ and $\boldsymbol{\Phi}$. The ellipsoidal likelihood region Ω_e is defined as a subset of $\boldsymbol{\varphi}$ that satisfies

$$\sum_{s=1}^{I+J} \left(\frac{\varphi_s^{max} - \varphi_s^{min}}{2} \right)^{-2} \left(\varphi_s - \varphi_s^0 \right)^2 \leq \theta^2, \tag{5.9}$$

where $\varphi_s, s = 1, 2, \ldots, I+J$, is an element in $\boldsymbol{\varphi}$, which can refer to λ_j or Φ_i; φ_s^0 is the expected (or nominal) value of φ_s (which, for example, can be the mean demand flow from on-ramp j or the average capacity flow at a location i on the main road);

φ_s^{max} and φ_s^{min} represent the corresponding maximum and minimum possible values of φ_s respectively. The parameter $\theta \in [0, 1]$ reflects the degree of uncertainty taken into account in optimisation [106]. This parameter θ can be regarded as a trade off between efficiency and robustness. The larger θ is, the more preference is given to consideration of robustness. On the other hand, $\theta = 0$ reduces the robust optimisation into conventional deterministic optimisation in which no uncertainty is considered. In ellipsoidal-constrained set, the worst case is marked by the red rectangle in Figure 5.4.

It is noted that the empirical data fits the ellipsoidal region very well as shown in Figure 5.6. The data used in Figure 5.6 is described in Section 5.2.2 as shown in Figure 5.2. Moreover the polyhedral region is more difficult to define than ellipsoidal region. Therefore, the ellipsoidal likelihood set is adopted in this study and introduced in the next section.

5.3.3 Robust ramp metering formulation

The deterministic optimisation described in Chapter 3 is extended to a robust formulation that incorporates uncertainty in both demand and set-valued fundamental diagram. Ben-Tal and Nemirovski [7] analyse the some generic convex optimisation problems with ellipsoidal uncertainty set. The robust optimisation can be formulated as a minimax problem (Vanderbei [101]; Yin [106]; Li [60]), which not only minimises the travel delay, but also minimises the associated delay variability. There is no standard algorithm to solve the minimax optimisation problem. It is conventional to decompose and solve the problem. Given the uncertainty range on demand and set-valued fundamental diagram, the generally idea of the robust optimisation is to first seek the combination of demand

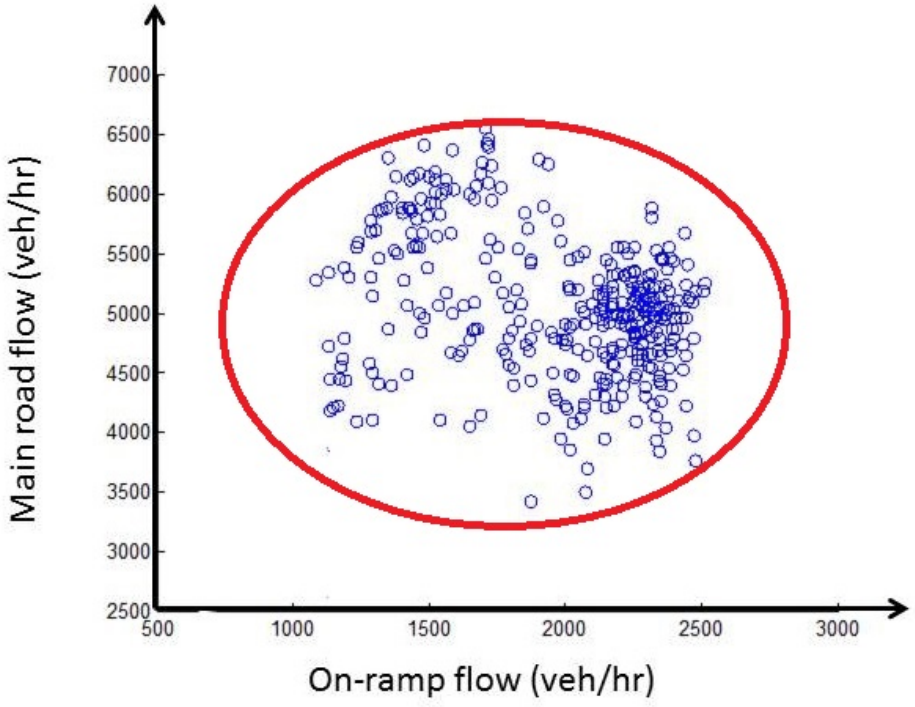

FIGURE 5.6: Uncertainty set with empirical data in polyhedral likelihood region (Detector Station: 4955A, M25 (clockwise))

and fundamental diagram that would gives the highest travel delay. Then, given this set of demand and fundamental diagram, we seek the corresponding optimal ramp metering that will give the lowest travel delay under this worst case scenario.

Given Ω, the robust optimisation can be now formulated as a minimax problem as:

$$\min_{\boldsymbol{c}^*} \max_{\boldsymbol{\lambda},\boldsymbol{\Phi}} Z\big(\boldsymbol{c}, \boldsymbol{\lambda}, \boldsymbol{\Phi}\big) \tag{5.10}$$

subject to:

$$\boldsymbol{\lambda}, \boldsymbol{\Phi} \in \boldsymbol{\Omega} \tag{5.11}$$

This optimisation problem is also subject to traffic dynamics constraints (3.23 - 3.33) as

specified by CTM in addition to constraint (5.11). There is no straightforward solution algorithm for solving the above optimisation problem. It is conventional to decompose and solve the minimax problem as follows:

Step 0. Set the iteration counter n = 1;

Step 1. Given a set of demand flows $\boldsymbol{\lambda}$ and fundamental diagram $\boldsymbol{\Phi}$, a control policy $\boldsymbol{c}^*$ to minimise the total delay Z is sought;

Step 2. Given the control policy $\boldsymbol{c}^*$ determined in Step 1, a set of demand flows $\boldsymbol{\lambda}$ and fundamental diagrams $\boldsymbol{\Phi}$ are sought within the likelihood set $\boldsymbol{\Omega}$ such that the total system delay Z is maximised;

Step 3. If $\varrho = \frac{||\mathbf{c}_n^* - \mathbf{c}_{n-1}^*||}{||\mathbf{c}_n^*||} < \delta$ then stop, where δ is a predetermined error tolerance. Otherwise, let n = n + 1, go to Step 1.

Step 1 is essentially the same optimisation problem as described in Chapter 3. Step 2 involves maximisation of the linear total delay function Z over a constraint set $\boldsymbol{\Omega}$. It can be verified that ellipsoidal constrained set are convex (quadratically-constrained) in $\boldsymbol{\lambda}$ and $\boldsymbol{\Phi}$. The likelihood set constraint can hence be augmented into the objective function through a Lagrangian multiplier. Consequently, the maximisation problem in Step 2 become a convex (indeed quadratic) optimisation subject to linear constraints which can be solved by standard gradient search algorithms. The two-step procedure (Step 1 and Step 2) above will be run iteratively and it can be shown (see Yin and Lawphongpanich [107]; Vanderbei [101], and others). To measure the convergence of the two-stage solution algorithm, we define $\varrho = \frac{||\mathbf{c}_n^* - \mathbf{c}_{n-1}^*||}{||\mathbf{c}_n^*||}$ where $||.||$ is the Euclidean norm, $\mathbf{c}_n^*$ is the vector containing all control variables (i.e. the on-ramp flows) computed at

iteration n in the two-stage algorithm. The convergence proof relies on the assumption that the sub-problems can be solved globally. The working examples shows the iteration always takes less than 10 iterations for the minimax problem to converge (see Figure 5.8). Then the process will converge to an optimal control policy $\boldsymbol{c}^*$ which minimises the total system delay Z under the worst scenario realised in $\boldsymbol{\Omega}$ if global optimal solutions can be found in both sub-problems.

5.4 Working Example

We select a 10-km section of the orbital M25 Motorway as described in Chapter 4. Each cell is characterised by a piecewise linear fundamental diagram which is calibrated by the measurements reported from the associated detector. The fundamental diagram is derived from applying successive linear regressions on the free flow, capacity and congested data as described in Chapter 2. The deterministic optimisation seeks an optimal ramp metering strategy that minimises the total delay along the section in which fundamental diagrams are considered to be single valued and there is no uncertainty in the measured demand. The size of the simulation time step, Δt, is set to be 15-sec, which gives the total number of time steps $K = 720$ for the 3 hours [17:00 - 20:00] planning horizon. The length Δx_i of each cell is 500 metres, which is the same as the detector spacing. The maximum allowable queue length $\bar{l}_j$ on all on-ramps is 60 vehicles. The problem now consists of altogether 2,160 decision variables (ramp inflows, $r_j(k)$, in which 3 (on-ramps) $\times$ 720 (time-steps) = 2,160). The optimisation is implemented and solved by IBM ILOG CPLEX Optimisation Studio V12.5 running on the same desktop computer described previously. The optimisation problem takes a computational time

of about four minutes to solve.

Figure 5.7 compares the density contours under the 'no control' and 'metered' scenarios. The layout of the road section is shown on the left of the figure. The optimisation reduces the main road delay from original 2,145 veh-hr (with zero ramp delays) to 1,776 veh-hr with the ramp queues bounded below a reasonable $\bar{l}_j = 60$ (veh). Moreover, the associated ramps' delay is 332.6 veh-hr with ramp metering, which gives a total system delay (i.e. main road + ramps) of 2,109 veh-hr, which is smaller than the original 2,145 veh-hr. The reduction of a congestion spillover is around the off-ramp (cell 5) at Junction 14 (marked by the red rectangle in the figure). This indicates an overall system-wide benefit from using the on-ramp control policy at cell 10 as it facilitates the discharge of traffic at its upstream off-ramp cell 5.

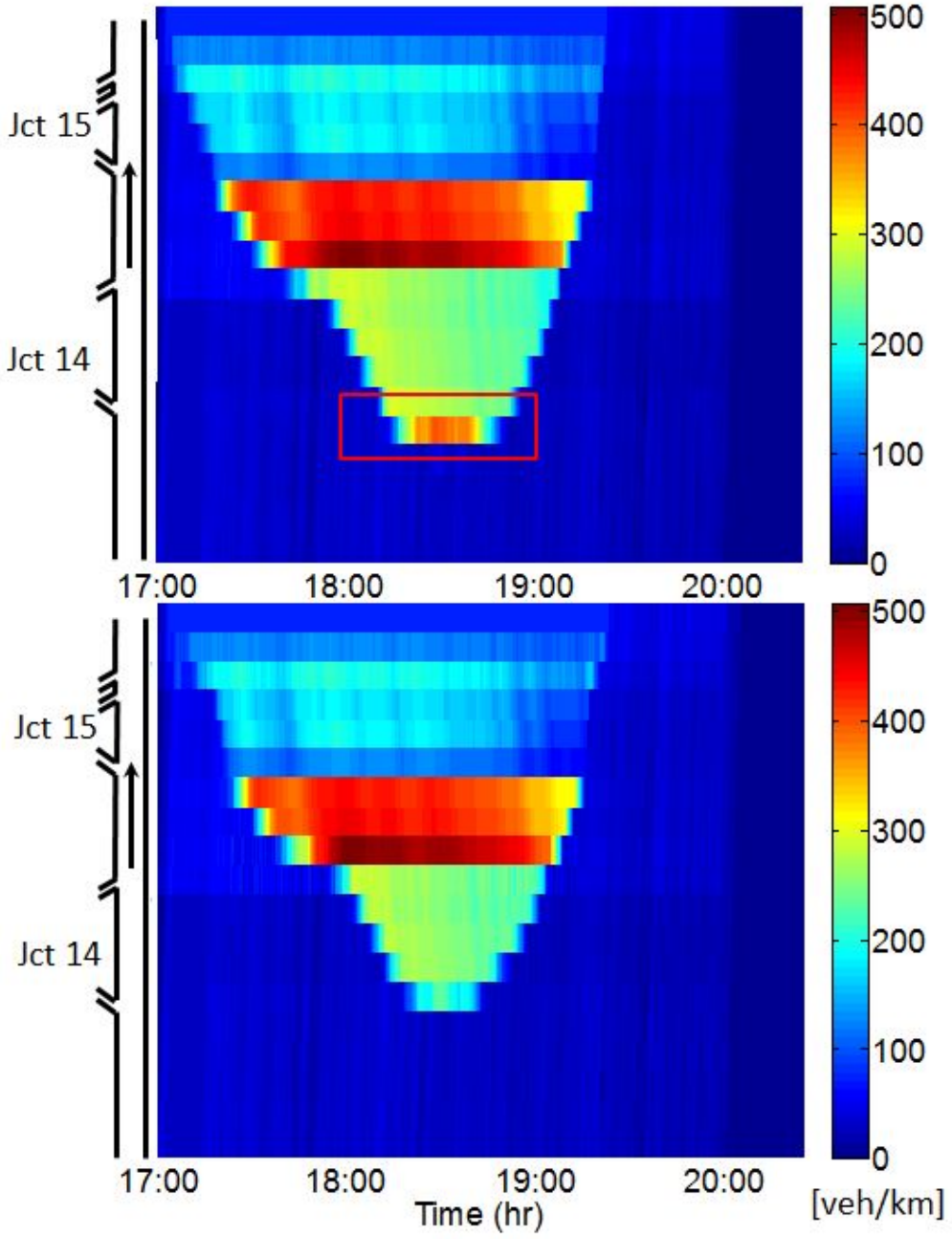

FIGURE 5.7: Comparison of main road densities (Two Junctions):

upper - no control; lower - metered

Now the results of the robust optimisation are presented in which we consider uncertainties which exist in on-ramp flows and fundamental diagrams. The formulation introduced in Section 5.3 is now used to seek a robust control policy that minimises the total system delay and associated variability under a range of scenarios. The network setting is the same as the one presented above while capturing the effect of uncertainties. The uncertainty of ±5% associated with demand flow $\boldsymbol{\lambda}$ measured at each on-ramps as follows:

$$\lambda_j^{min}(k) = 0.95\hat{\lambda}_j(k); \quad \lambda_j^{max}(k) = 1.05\hat{\lambda}_j(k) \tag{5.12}$$

where the notations $\hat{\lambda}_j(k)$ denote the measured value of $\lambda_j(k)$ at time step k as reported from the detectors.

The fundamental diagram $\boldsymbol{\Phi}$ is specified by the free flow speed v, capacity Q, and jam density $\bar{\rho}$, with which other parameters such as shockwave speed w and critical density $\mathring{\rho}$ can be deterministic. To incorporate the set value of the fundamental diagram, the capacity and jam density are assumed with a ±5% estimation error along the main road.

$$Q_i^{min} = 0.95\hat{Q}_i \quad Q_i^{max} = 1.05\hat{Q}_i \tag{5.13}$$

$$\bar{\rho}_i^{min} = 0.95\hat{\bar{\rho}}_i \quad \bar{\rho}_i^{max} = 1.05\hat{\bar{\rho}}_i \tag{5.14}$$

where the notation $\hat{Q}_i$ and $\hat{\bar{\rho}}_i$ denote the calibrated capacity and the jam density.

Given the bounds on $\boldsymbol{\lambda}$ and $\boldsymbol{Q}$, the ellipsoidal likelihood set $\boldsymbol{\Omega_e}$ was then constructed for the robust optimisation problem following Equation (5.9) where the design parameter θ is taken as one. A robust ramp metering policy is derived by solving the minimax optimisation problem with the two-stage algorithm. Figure 5.8 shows the convergence

of the minimax problem. The horizontal (x) axis refers to the number of the iterations and the vertical axis refers the total delay (objective function). The blue dotted line with circles shows the total delay of the maximisation, and the green lines with triangles represent the total delay of minimisation. It shows the minimax problem converge after 5 iterations for this study.

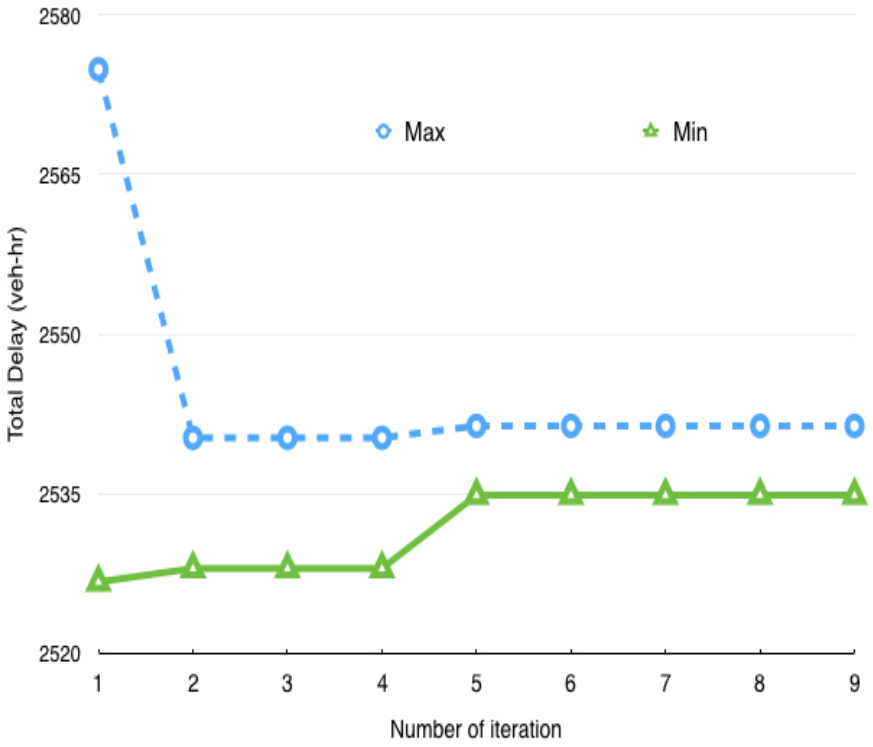

FIGURE 5.8: Convergence of minimax problem

The performance of different control strategies are compared over 11 levels of normal demand values as shown in Figure 5.9. The horizontal (x) axis refers to the demand level on which '100%' is the situation where the simulating with all demands are measured by detectors. '95%' refers to a situation in which all demands are scaled down by 5%, '105%' is the situation where all demands are scaled up by 5%, and so on. A demand

multiplier less than '100%' represents the situation in which the predicted demand is being overestimated (i.e. the actual demand is lower than the design demand) and vice versa. Under each scenario, 300 combinations of demand were randomly generated within the associated Ω_e. The ramp metering strategies are then simulated over these $11 \times 300 = 3300$ scenarios and the results. In Figure 5.9, each box represents the statistis of total network (main road + ramps) delays. The middle line in each box represents the median delay under the corresponding demand level, while the height of the box is the interquartile range of the delays.

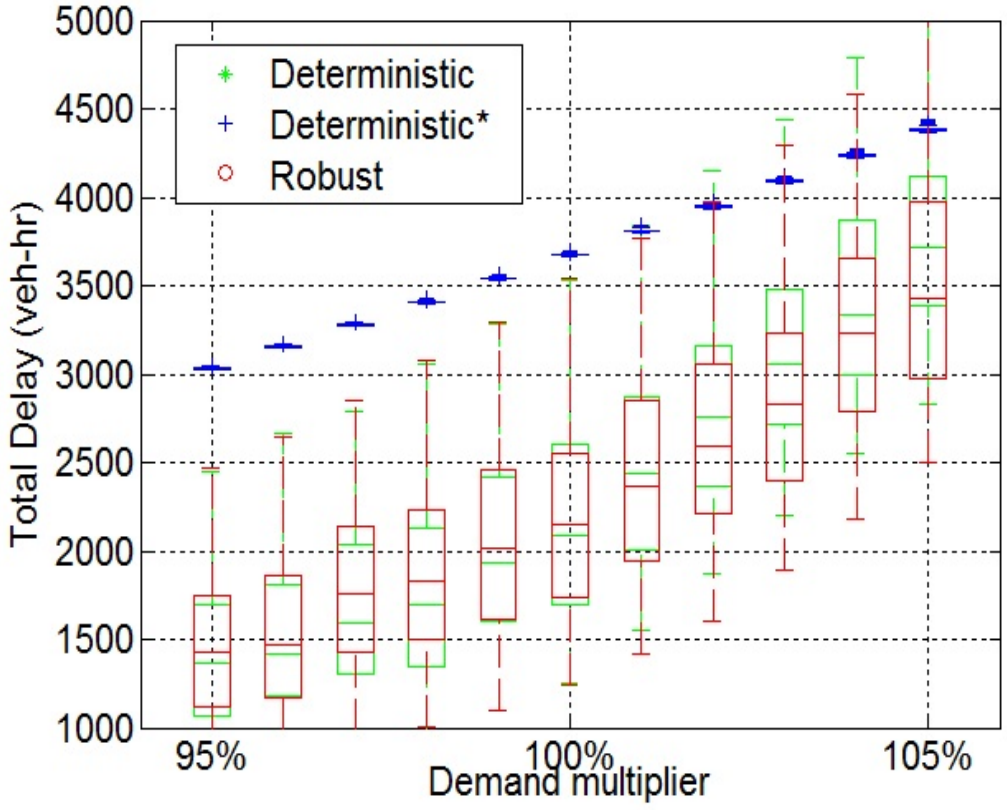

FIGURE 5.9: Comparison of total delay

Figure 5.9 compares the performances of three control strategies including deterministic, deterministic*, and robust. The deterministic control is running under normal demand flows and capacity, while the deterministic* is the control strategy by running a deterministic optimisation with all demand flows multiplied by 1.05, and all capacities multiplied by 0.95. The deterministic* control represents the traditional 'safety factor' based engineering approach in which relevant design parameters are scaled up or down without running the robust optimisation. Note that the maximum ramp queue constraint needs to be disabled when calculating this control strategy as the optimisation will become infeasible with increased demands and decreased capacities due to insufficient space in the system for storing all queues.

Figuer 5.9 reveals that the robust control generally outperforms the deterministic one in particular when the demand is appropriately estimated. Moreover, the ramp delays under robust ramp metering control are lower than those under deterministic control, because an extra buffer on ramps is introduced when the robust ramp metering is calculated. This reveals the advantage of robust control for incorporating potential uncertainties in the overall system and protecting ramps in motorway. The Deterministic* strategy is shown to be very robust as it shows little variations as revealed by the box sizes compared with the other two strategies due to the large buffer considered during their deviation. However, as aforementioned, this Deterministic* control will be too conservative and counter-productive.

Figure 5.10 highlights the cases (100 % - 105 %) when robust optimisation is more effective. Figure 5.10 shows the total delay under deterministic control (upper one) and robust control (lower one). The horizontal (x) axis refers to the total delay and the vertical axis refers to the number of scenarios. Figure 5.10 reveals the robust control

holds a lower mean total delay (around 2500 veh-hr) compared with the deterministic control (around 2800 veh-hr). It is noted that the high total delay (e.g. 3500 veh-hr) of robust control less than deterministic control.

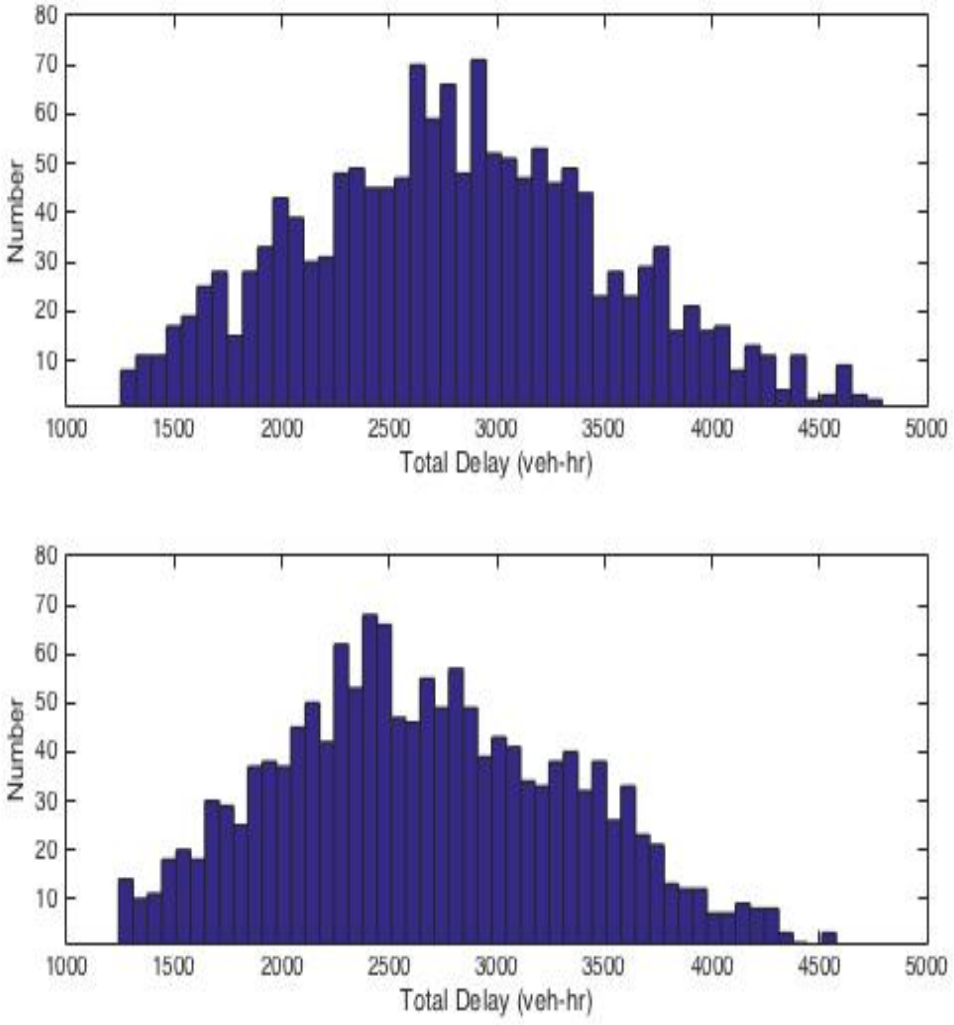

FIGURE 5.10: Comparison of total delay for the cases (100 % - 105 %)

Figure 5.11 summarises all data points shown in Figure 5.9. Each box in Figure 5.11 contains statistics of 300 (scenarios) $\times$ 11 (demand levels) = 3,300 delay measures. It can be seen that the deterministic control gives a slightly better average performance (with a median delay of 2,243 veh-hr) than the robust control (which gives a median

delay of 2,292 veh-hr). Nevertheless, the robust control gives a lower interquartile range of delays ([1,682, 2,903] (veh-hr)) than the deterministic control ([1,574, 3,012] (veh-hr)). Moreover, the maximum observed delay of robust control is also lower than that obtained from deterministic control as shown in Figure 5.9. For the Deterministic* control, it gives the lowest interquartile range of delay and highest median total delay among three control strategies. However, this Deterministic* control will be too conservative and counterproductive. As shown, the Deterministic* control performs worse than the other two control strategies by shutting down ramps for an unnecessarily long time. This suggests that the robust control is useful in stabilising the system when the actual demand is higher than the expected value.

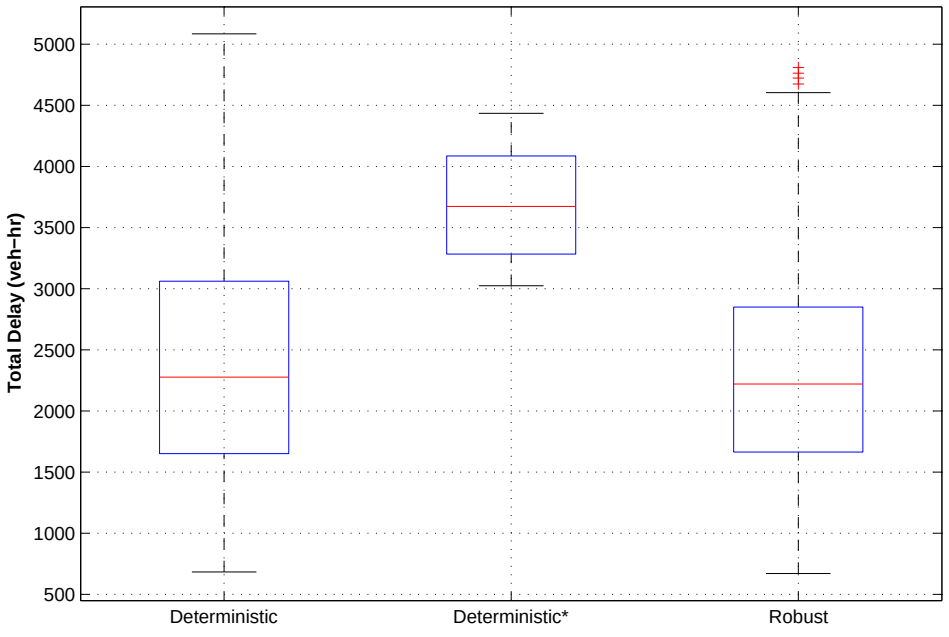

FIGURE 5.11: Overall statistics

5.5 Summary

This chapter presents a robust modelling and optimisation framework for motorway traffic where uncertainties in demand and capacity are explicitly considered. The optimiser aims to minimise both average and variance of total system delay over a range of uncertain scenarios. The optimisation problem is formulated based upon CTM. With the piecewise linear nature of CTM, the optimisation is formulated as a LP. The performance of the optimiser is demonstrated through a set of working examples with traffic data collected on the M25 Motorway in the UK. A deterministic optimisation was first demonstrated which derives an optimal ramp metering policy that minimises the total system delay over a fixed time horizon. The optimal ramp metering problem is solved by using CPLEX and the optimal solution is tractable.

With uncertainty considered, robust optimisation is formulated as a minimax problem and solved in two stages. The two-stage solution procedure first solves a control policy that minimises the system delay given a set of demand and capacity settings. Then with the control policy derived, the minimax solver seeks another set of demand and capacity within a predefined likelihood region such that the system delay is maximised. The two stages are run iteratively and the solution procedure converges in an optimal control policy which minimises both system delay and the associated variability. A series of simulation experiments reveal that the deterministic optimal control outperforms slightly the robust control in terms of minimising average delay over a range of scenarios, while the system performance is shown to be stabler with the robust control strategy implemented.

Chapter 6

Concluding Remarks

6.1 Summary

This thesis investigates a mathematical tool for optimising the performance of motorway through modelling and regulating the traffic flowing on it.

In Chapter 2, the thesis starts with an introduction of cell transmission model (CTM), which describes the relationship and evolution of traffic flow and density over space and time. The model is parsimonious and accurate in predicting traffic dynamics with simple mathematical structure. Then how to process raw MIDAS data is presented. Following this, the calibration of the fundamental diagram for dynamic traffic modelling is presented, which is the core components for building effectiev traffic control system in practice. The calibration procedure presented here consists of three main steps: determining the free-flow line, capacity, and congested line iteratively. The free flow line can be determined as flow-density data under free flow, which exhibits a linear

relationship. Capacity is taken as the maximum flow value observed over a period of time. The main challenge lies on the congested part, which exhibits a high level of variability.

In Chapter 3, the thesis presents a linear modelling and optimisation framework for ramp metering control on motorway traffic. We first present how CTM can be implemented to model real world motorway. The accuracy of the model estimates is validated against real traffic data collected from the M25 Motorway in the UK. The results show CTM reproduces the general pattern of the traffic congestion associated with correct time and location. The validation results presented reveal the mean absolute percentage error ϵ obtained from the CTM modelling conducted in this exercise is 11.5% beyond in many cases. With the piecewise linear structure of CTM, we further develop the motorway ramp metering optimisation problem as a Linear Programming (LP) problem. The LP optimisation problem can be solved by a range of established algorithms and computer solvers for the global optimal solution. This LP formulation is applied to a scenario of UK M25 Motorway where we derive an optimal ramp metering strategy that minimise the total system delay over a fixed space-time horizon. It is shown that optimal solutions are obtainable through CPLEX computer solver in a reasonable computational time. The optimal solution provides useful insights and guidance on how we should manage motorway traffic to maximise its efficiency.

In Chapter 4, a mathematical optimisation framework is presented for four motorway control strategies including VSL, HSR, RMVSL, and integrated control strategies. The optimisation models are formulated based upon the CTM which is known to be able to produce reasonable traffic estimation with a parsimonious mathematical structure. With the linearity of CTM, the optimisation can be formulated as MILP problems. The

optimal control problems are solved by the CPLEX solver and the eventual solution provides useful insights on how we should manage the motorway systems for effective outcomes. The various analyses on sensitivity of the effectiveness of the control strategies with respect to different model settings and assumptions also presented in this Chapter. The optimisation models are applied to a real case scenario on UK M25 Motorway. The solutions are obtainable in a reasonable computational time. The results show that those control strategies can effectively reduce total system delay, especially the integrated control strategy due to it benefits from each control strategy. By comparing the numerical results (see Table 4.3), it is found that the reduction of total system delay of integrated control (74%) is higher than each separate control. The ramp delay of integrated control (198.7 veh-hr) is lower than the ramp delay of ramp metering control (349.0 veh-hr). Hard shoulder running appears to be the effective approach in reducing total system delay by providing an extra physical lane to motorways. Nevertheless, engineers will have to be careful in balancing the trade-off between mobility and safety as HSR reduced the buffer intended for incident management. One should note that the application to this particular motorway is only for illustration purpose and the methodology is generally applicable for other corridors or networks.

In Chapter 5, it presents a robust optimisation framework for motorway where uncertainty in demand and traffic state estimation is explicitly considered. A set-valued fundamental diagram formulation is applied to formulate a robust optimisation framework of motorway traffic. The application of set-valued fundamental diagram opens up new research directions in motorway traffic management with the consideration of uncertainties associated with traffic dynamics. The optimiser aims to minimise and stabilise total system delay. With uncertainty considered, the robust optimisation is formulated

as a minimax problem and solved in two stages. The two-stage solution procedure first solves for a control policy that minimises the total system delay with a set of given demand and fundamental diagram settings. Then with the control policy derived, the minimax solver seeks another set of demand and fundamental diagram settings within a predefined likelihood region such that the system delay is maximised. The two stages are run iteratively and it is shown that the solution procedure converges to an optimal control policy which minimises both total system delay and the associated variability. The results show that the robust control policy derived from the set-valued fundamental diagram helps to minimise both means and variances of travel delays over a range of scenarios with stochastic demand and traffic characteristics. A series of simulation experiments reveals that the system performance is shown to be much stabler with the robust control strategy implemented.

6.2 Future Work

This section identifies several limitations of the work presented in this thesis and suggests possible future research directions.

In this thesis, the modelling and optimisation framework presented here is for an offline planning purpose. A rolling horizon optimisation framework will be studied for real time applications. In order to use a rolling optimisation framework, the forecast for additional periods in the future, and optimisation for several horizons needs to be finished in a reasonable time.

The robust optimisation framework presented here is for ramp metering only. The VSL

and HSR control strategies can also be extended to a stochastic framework, which takes into account the uncertainties exist in demand and fundamental diagram. Because of the combinatorial nature of MILP, solving deterministic VSL (30 min) and HSR (10 min) problems take significantly longer than solving the ramp metering problem (4 min). Therefore, developing more efficient solution algorithms will be a crucial step before robust VSL and HSR formulations can be developed.

Moreover, as a first-order traffic model, CTM assumes arbitrarily large acceleration and deceleration of vehicles with which flow dynamics can be adjusted instantaneously according to the associated traffic density with the predefined fundamental diagram. By ignoring the transient state transition, the current model cannot capture complex but important dynamic behaviour such as capacity drop and stop-and-go traffic pattern [46]. Such transient behaviour of congestion is important when one designs more sophisticated control strategies such as VSL (Papageorgiou *et al.* [84]; Carlson *et al.* [15]). Development of an efficient and plausible modelling and optimisation framework for capturing and managing the nonlinear traffic dynamics will be another important extension of the current study.

Appendix A

Calculate Density from the Occupancy Measurements

Assume that the average length of the vehicles is $\overline{L}_v$, and the length of the detector is L_d. Suppose the speed of vehicles is v , then the occupancy can be estimated as:

$$o = \frac{n(\overline{L}_v + L_d)}{v\Delta t} \tag{A.1}$$

where Δt is sampling time interval (i.e. the frequency of occupancy is measured from the detector), n is the number of vehicles passing the detector in Δt. The quantity $\frac{(\overline{L}_v + L_d)}{v}$ represents the average time that a vehicle will spend on the detector.

Rearranging Equation A.1, the occupancy can be expressed as:

$$o = (\overline{L}_v + L_d)\frac{f}{v} \tag{A.2}$$

as the flow rate $f = \frac{n}{\Delta t}$ by definition.

The fundamental relationship ($f = \rho v$) holds, then the occupancy can be expressed as:

$$o = (\overline{L}_v + L_d)\rho \tag{A.3}$$

Rearranging Equation A.3, the density can be expressed as:

$$\rho = \frac{o}{\overline{L}_v + L_d} \tag{A.4}$$

Appendix B

Constrained Regression

We aim to derive a linear regression model $y' = a_0 + a_1 x$ from a set of data points: $(x_i, y_i) \quad i = 1, 2, ..., n$, where x_i and y_i are dependent variables whose value is found by observation. In this case, x_i is density and y_i is flow. The y' is the estimate of the true value y_i conditional on x_i.

The linear regression model can be constructed by passing through a given point (x_0, y_0) as:

$$\frac{y_i' - y_0}{x_i - x_0} = a_1 \tag{B.1}$$

rearrange as

$$y_i' = a_1(x_i - x_0) + y_0 \tag{B.2}$$

where a_1 is the only parameter to be determined from data. The regression line is constrained by the given point (x_0, y_0). The a_1 is determined by the least-square estimation that aims to minimise the total squared error between the observed y_i and estimated y_i'

$$E = \sum_{i=1}^{n} \big(y_i' - y_i\big)^2 = \sum_{i=1}^{n} \big(a_1(x_i - x_0) + y_0 - y_i\big)^2 \tag{B.3}$$

where n is number of data points. The model parameter a_1 can be determined such that E is minimised, in which we can solve a_1 from $\frac{\partial E}{\partial a_1} = 0$

Hence,

$$\begin{aligned} \sum_{i=1}^{n} \big(a_1(x_i - x_0) + y_0 - y_i\big)\big(x_i - x_0\big) &= 0 \\ \Rightarrow a_1 = \frac{\sum_{i=1}^{n}(y_i - y_0)(x_i - x_0)}{\sum_{i=1}^{n}(x_i - x_0)^2} & \end{aligned} \tag{B.4}$$

References

[1] MIDAS n.d. "Motorway Incident Detection and Automatic Signalling". URL http://www.midas-data.org.uk/.

[2] Highways Agency. 2004. M25 Controlled Motorways: Summary Report. URL http://www.standardsforhighways.co.uk/pilots_trials/files/ha2004.pdf.

[3] Department for Transport. 2009. Road Transport Forecasts 2009 Results from the Department for Transport's National Transport Model. URL http://www.dft.gov.uk/pgr/economics/ntm/forecasts2009/.

[4] Highways Agency. 2009. Managed Motorways Implementation Guidance– Hard shoulder running, interim advice note. URL http://www.standardsforhighways.co.uk/ha/standards/ians/pdfs/ian111_09.pdf.

[5] Highways Agency. 2010. M4 Bus Lane: Before and after monitoring outline specification. URL http://assets.highways.gov.uk/roads/road-projects/m4-junction-3-to-junction-2-bus-lane-suspension-scheme/Highway_Agency_M4_Bus_Lane_Before_and_After_Monitoring_Outline_Specification.pdf.

[6] Bellemans, T., De Schutter, B., and De Moor, B. 2006. Model predictive control for ramp metering of motorway traffic: a case study. *Control Engineering Practice*, 14(7), pp. 757–767.

[7] Ben-Tal, A. and Nemirovski, A. 1998. Robust convex optimization. *Mathematics of Operations Research*, 23(4), pp. 769–805.

[8] Ben-Tal, A. and Nemirovski, A. 1999. Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1), pp. 1–13.

[9] Bertsimas, D., Brown, D., and Caramanis, C. 2011. Theory and applicaitons of robust optimization. *SIAM Review*, 53(3), pp. 464–501.

[10] Bickel, P. J., Chen, C., Kwon, J., Rice, J., van Zwet, E., and Varaiya, P. 2007. Measuring traffic. *Statistical Science*, 22(4), pp. 581–597.

[11] Biham, O., Middleton, A. A., and Levine, D. 1992. Self organization and a dynamical transition in traffic flow models. *Physical Review A*, 46(10), pp. 6124–6127.

[12] Brackstone, M. and McDonald, M. 1999. Car-following: a historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2(4), pp. 181–196.

[13] Brilon, W., Geistefeldt, J., and Regler, M. 2005. Reliability of freeway traffic flow: A stochastic concept of capacity. In *Proceedings of the 16th International Symposium on Transportation and Traffic Theory: Flow, Dynamics and Human Interaction (ISTTT)*, July 19–21, College Park, Maryland, USA, pp. 125–144.

[14] Carey, M. and Bowers, M. 2012. A review of properties of flow–density functions. *Transport Reviews: A Transnational Transdisciplinary Journal*, 32(1), pp. 49–73.

[15] Carlson, R. C., Papamichail, I., Papageorgiou, M., and Messmer, A. 2010. Optimal motorway traffic flow control involving variable speed limits and ramp metering. *Transportation Science*, 44(2), pp. 238–253.

[16] Casas, J., Ferrer, J. L., Garcia, D., Perarnau, J., and Torday, A. 2010. Traffic simulation with Aimsun. *Fundamentals of Traffic Simulation*. pp. 173–232. Springer, New York.

[17] Cassidy, M. J. and Bertini, R. L. 1999. Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, 33(1), pp. 25–42.

[18] Chandler, R. E., Herman, R., and Montroll, E. W. 1958. Traffic dynamics: Studies in car following. *Operations Research*, 6(2), pp. 165–184.

[19] Chow, A. H. F., Dadok, V., Dervisoglu, G., Gomes, G., Horowitz, R., Kurzhanskiy, A., Kwon, J., Lu, X.-Y., Muralidharan, A., Norman, S., and others, . 2008. TOPL: Tools for operational planning of transportation networks. In *ASME 2008 Dynamic Systems and Control Conference*, pp. 1035–1042.

[20] Chow, A. H. F., Lu, X.-Y., Qiu, Z.-J. T., Shladover, S. E., and Yeo, H. 2010. An empirical study of traffic speed drop for freeway management. In *Proceedings of the 89th Annual Meeting of the Transportation Research Board* ($Paper: 10-1967$), January 10–14, Washington, DC, USA.

[21] Cohen, S. 2004. Using the hard shoulder and narrowing lanes to reduce traffic congestion some lessons from an experience on the Paris motorway network. In *Proceedings of the 12th IEE International Conference Road Transport Information & Control*, April 20–22, London, UK, pp. 149–153.

[22] Courant, R., Friedrichs, K., and Lewy, H. 1928. Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische Annalen*, 100(1), pp. 32–74.

[23] 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.

[24] Daganzo, C. F. 1995. The cell transmission model, part II: network traffic. *Transportation Research Part B: Methodological*, 29(2), pp. 79–93.

[25] Daganzo, C. F. 1995. Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B: Methodological*, 29(4), pp. 277–286.

[26] Dervisoglu, G., Gomes, G., Kwon, J., Horowitz, R., and Varaiya, P. 2009. Automatic calibration of the fundamental diagram and empirical observations on capacity. In *Proceedings of the 88th Annual Meeting of the Transportation Research Board* ($Paper: 09-3159$), January 11–15, Washington, DC, USA.

[27] Doan, K. and Ukkusuri, S. V. 2012. On the holding-back problem in the cell transmission based dynamic traffic assignment models. *Transportation Research Part B: Methodological*, 46(9), pp. 1218–1238.

[28] Eddington, R. 2006. The Eddington Transport Study – The case for action-Sir Rod Eddington's advice to Government–December 2006. URL `http://trid.trb.org/view.aspx?id=820379`.

[29] Edie, L. C. 1961. Car-following and steady-state theory for noncongested traffic. *Operations Research*, 9(1), pp. 66–76.

[30] Gazis, D. C., Herman, R., and Potts, R. B. 1959. Car-following theory of steady-state traffic flow. *Operations Research*, 7(4), pp. 499–505.

[31] Geistefeldt, J. 2012. Operational experience with temporary hard shoulder running in Germany. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 2278, pp. 67–73.

[32] Gipps, P. G. 1981. A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, 15(2), pp. 105–111.

[33] Godunov, S. K. 1959. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, Volume. 47(89), pp. 271–306.

[34] Gomes, G. and Horowitz, R. 2006. Optimal freeway ramp metering using the asymmetric cell transmission model. *Transportation Research Part C: Emerging Technologies*, 14(4), pp. 244–262.

[35] Greenberg, H. 1959. An analysis of traffic flow. *Operations Research*, 7(1), pp. 79–85.

[36] Greenshields, B. D. 1935. A study of traffic capacity. In *Highway Research Board 14th Annual Meeting (Part I)*, December 6–7, Washington, DC, USA.

[37] Greenshields, B. D., Thompson, J. T., Dickinson, H. C., and Swinton, R. S. 1934. The photographic method of studying traffic behavior. In *Highway Research Board 13th Annual Meeting (Part I–Report of Research Committees and Papers)*, December 7–8, Washington, DC, USA, pp. 382–399.

[38] Hadiuzzaman, M. and Qiu, T. Z. 2013. Cell transmission model based variable speed limit control for freeways. *Canadian Journal of Civil Engineering*, 40(1), pp. 46–56.

[39] Haj-Salem, H., Poirier, P., Heylliard, J.-F., and Peynaud, J.-P. 2001. ALINEA: a local traffic responsive strategy for ramp metering. Field results on A6 motorway in Paris. In *Proceedings of the 2001 International IEEE Conference on Intelligent Transportation Systems (ITSC)*, August 25-29, Oakland, CA, USA, pp. 106–111.

[40] Han, K., Liu, H.-C., Gayah, V. V., Friesz, T. L., and Yao, T. 2015. A robust optimization approach for dynamic traffic signal control with emission considerations. *Transportation Research Part C: Emerging Technologies*, pp. In press.

[41] Harbord, B. 1998. M25 controlled motorway results of the first two years. In *Proceedings of the 9th International Conference on Road Transport Information and Control* (*Paper* : 454), April 21-23, London, UK, pp. 149–154.

[42] Hegyi, A. 2004. Model predictive control for integrating traffic control measures. Doctoral Dissertation, Delft University of Technology, Delft, Netherlands.

[43] Hegyi, A., De Schutter, B., and Hellendoorn, H. 2005. Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies*, 13(3), pp. 185–209.

[44] Hegyi, A., De Schutter, B., and Hellendoorn, J. 2005. Optimal coordination of variable speed limits to suppress shock waves. *IEEE Transactions on Intelligent Transportation Systems*, 6(1), pp. 102–112.

[45] Heydecker, B. 1987. Uncertainty and variability in traffic signal calculations. *Transportation Research Part B: Methodological*, 21(1), pp. 79–85.

[46] Heydecker, B. G. and Addison, J. D. 2011. Analysis and modelling of traffic flow under variable speed limits. *Transportation Research Part C: Emerging Technologies*, 19(2), pp. 206–217.

[47] Hoogendoorn, S. P. and Bovy, P. H. L. 2001. State-of-the-art of vehicular traffic flow modelling. In *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, pp. 283–303.

[48] Jagannathan, R., Petrovic, S., Powell, G., and Roberts, M. 2014. Predicting road accidents based on current and historical spatio-temporal traffic flow data. In *Proceedings of the 5th International Conference on Computational Logistics (ICCL)*, September 24–26, Valparaíso, Chile, pp. 83–97.

[49] Jin, S., Wang, D.-H., Huang, Z.-Y., and Tao, P.-F. 2011. Visual angle model for car-following theory. *Physica A: Statistical Mechanics and its Applications*, 390 (11), pp. 1931–1940.

[50] Kerner, B. S. 2008. On-ramp metering based on three-phase traffic theory downstream off-ramp and upstream on-ramp bottlenecks. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 2088, pp. 80–89.

[51] Kometani, E. and Sasaki, T. 1959. Dynamic behaviour of traffic with a non-linear spacing-speed relationship. In *Proceedings of the Symposium on Theory of Traffic Flow*, pp. 105–119.

[52] Kotsialos, A. and Papageorgiou, M. 2001. The importance of traffic flow modeling for motorway traffic control. *Networks and Spatial Economics*, 1(1-2), pp. 179–203.

[53] Kotsialos, A. and Papageorgiou, M. 2004. Nonlinear optimal control applied to coordinated ramp metering. *IEEE Transactions on Control Systems Technology*, 12(6), pp. 920–933.

[54] Kurzhanskiy, A. A. and Varaiya, P. 2010. Active traffic management on road networks: a macroscopic approach. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368(1928), pp. 4607–4626.

[55] Kurzhanskiy, A. A. and Varaiya, P. 2012. Guaranteed prediction and estimation of the state of a road network. *Transportation Research Part C: Emerging Technologies*, 21(1), pp. 163–180.

[56] Lax, P. and Wendroff, B. 1960. Systems of conservation laws. *Communications on Pure and Applied Mathematics*, 13(2), pp. 217–237.

[57] Lebacque, J.-P. 1996. The Godunov scheme and what it means for first order traffic flow models. In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, July 24–26, Lyon, France, pp. 647–677.

[58] Lee, C., Hellinga, B., and Saccomanno, F. 2006. Evaluation of variable speed limits to improve traffic safety. *Transportation Research Part C: Emerging Technologies*, 14(3), pp. 213–228.

[59] Levinson, D. and Zhang, L. 2004. Chapter 9 Evaluating the effectiveness of ramp meters. *Assessing the Benefits and Costs of ITS*, Dordrecht, The Netherlands: Kluwer Academic, pp. 145–166.

[60] Li, J.-Q. 2011. Discretization modeling, integer programming formulations and dynamic programming algorithms for robust traffic signal timing. *Transportation Research Part C: Emerging Technologies*, 19(4), pp. 708–719.

[61] Lighthill, M. J. and Whitham, G. B. 1955. On kinematic waves II. A theory of traffic flow on long crowded roads. In *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, May 10, 229(1178), pp. 317–345.

[62] Liu, H.-C., Han, K., Gayah, V. V., Friesz, T. L., and Yao, T. 2015. Data-driven linear decision rule approach for distributionally robust optimization of on-line signal control. *Transportation Research Part C: Emerging Technologies*, Volume. 59, pp. 260–277.

[63] Liu, R.-H. 2010. Traffic simulation with DRACULA. *Fundamentals of Traffic Simulation.* pp. 295–322. Springer, New York.

[64] Lo, H. K. 1999. A novel traffic signal control formulation. *Transportation Research Part A: General*, 33(6), pp. 433–448.

[65] Lo, H. K. 2001. A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. *Transportation Science*, 35(2), pp. 148–164.

[66] Lo, H. K. and Chow, A. H. F. 2002. Adaptive traffic control system: Control strategy, prediction, resolution, and accuracy. *Journal of Advanced Transportation*, 36(3), pp. 323–347.

[67] Luenberger, D. G. and Ye, Y. 2008. *Linear and Nonlinear Programming.* Springer US.

[68] Masher, D. P., Ross, D. W., Wong, P. J., Tuan, P. L., Zeidler, H. M., and Petracek, S. Stanford Research Institute. 1975. Guidelines for design and operation of ramp control systems. URL `http://trid.trb.org/view.aspx?id=33060`.

[69] May, A. D. and Harmut, E. M. K. 1967. Non-integer car-following models. *Highway Research Record*, Volume. 199, pp. 19–32.

[70] Nagel, K. and Nelson, P. 2005. A critical comparison of the kinematic-wave model with observational data. In *Proceedings of the 16th International Symposium on Transportation and Traffic Theory: Flow, Dynamics and Human Interaction (ISTTT)*, July 19–21, College Park, Maryland, USA, pp. 145–163.

[71] Nagel, K. and Schreckenberg, M. 1992. A cellular automaton model for freeway traffic. *Journal de Physique I*, 2(12), pp. 2221–2229.

[72] Highway Capacity Manual (Special Report). 2000. National Research Council, Washington, DC, USA.

[73] Newell, G. F. 1993. A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B: Methodological*, 27(4), pp. 281–287.

[74] Ngoduy, D. and Liu, R. 2007. Multiclass first-order simulation model to explain non-linear traffic phenomena. *Physica A: Statistical Mechanics and its Applications*, 385(2), pp. 667–682.

[75] Ngoduy, D. 2011. Multiclass first-order traffic model using stochastic fundamental diagrams. *Transportmetrica*, 7(2), pp. 111–125.

[76] Ozaki, H. 1993. Reaction and anticipation in the car-following behavior. In *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, July 21-23, Berkeley, CA, USA, pp. 349–366.

[77] Papageorgiou, M. 1998. Some remarks on macroscopic traffic flow modelling. *Transportation Research Part A: General*, 32(5), pp. 323–329.

[78] Papageorgiou, M. and Kotsialos, A. 2000. Freeway ramp metering: an overview. In *Proceedings of the 2000 International IEEE Conference on Intelligent Transportation Systems (ITSC)*, October 1–3, Dearborn, Michigan, USA, pp. 228–239.

[79] Papageorgiou, M. and Vigos, G. 2008. Relating time-occupancy measurements to space-occupancy and link vehicle-count. *Transportation Research Part C: Emerging Technologies*, 16(1), pp. 1–17.

[80] Papageorgiou, M., Blosseville, J. M., and Hadj-Salem, H. 1989. Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris. *Transportation Research Part B: Methodological*, 23(1), pp. 29–47.

[81] Papageorgiou, M., Blosseville, J. M., and Hadj-Salem, H. 1990. Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part I: modelling. *Transportation Research Part A: General*, 24(5), pp. 345–359.

[82] Papageorgiou, M., Hadj-Salem, H., and Blosseville, J. 1991. ALINEA: a local feedback control law for on-ramp metering. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 1320, pp. 58–64.

[83] Papageorgiou, M., Hadj-Salem, H., and Middelham, F. 1997. Alinea local ramp metering: summary of field results. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 1603, pp. 90–98.

[84] Papageorgiou, M., Kosmatopoulos, E., and Papamichail, I. 2008. Effects of variable speed limits on motorway traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 2047, pp. 37–48.

[85] Papamichail, I. and Papageorgiou, M. 2008. Traffic-responsive linked ramp-metering control. *IEEE Transactions on Intelligent Transportation Systems*, 9 (1), pp. 111–121.

[86] Papamichail, I., Kotsialos, A., Margonis, I., and Papageorgiou, M. 2010. Coordinated ramp metering for freeway networks – A model-predictive hierarchical control approach. *Transportation Research Part C: Emerging Technologies*, 18(3), pp. 311–331.

[87] Park, B. and Schneeberger, J. 2003. Microscopic simulation model calibration and validation: case study of VISSIM simulation model for a coordinated actuated signal system. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 1856, pp. 185–192.

[88] Payne, H. J. 1971. Models of freeway traffic and control. *Mathematical Models of Public Systems*, pp. 51–61.

[89] Powell, W. B. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality.* John Wiley & Sons, Inc., Hoboken, New Jersey.

[90] Richards, P. I. 1956. Shock waves on the highway. *Operations Research*, 4(1), pp. 42–51.

[91] Samoili, S., Efthymiou, D., Antoniou, C., and Dumont, A.-G. 2013. Lane flow distribution investigation of hard shoulder running freeways. *Transportation Research Record: Journal of the Transportation Research Board*, pp. 58–64.

[92] Seaman, E. 2010. Modelling the UK's busiest road-using detailed data to inform design at junction 15 on the M25. In *European Transport Conference 2010*, October 11–13, Glasgow, Scotland, United Kingdom.

[93] Skabardonis, A. and Geroliminis, N. 2008. Real-time monitoring and control on signalized arterials. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 12(2), pp. 64–74.

[94] Smulders, S. 1990. Control of freeway traffic flow by variable speed signs. *Transportation Research Part B: Methodological*, 24(2), pp. 111–132.

[95] Sumalee, A., Zhong, R.-X., Pan, T.-L., and Szeto, W. Y. 2011. Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment. *Transportation Research Part B: Methodological*, 45 (3), pp. 507–533.

[96] Sun, X.-T., Muñoz, L., and Horowitz, R. 2003. Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control. In *Proceedings of the 42nd IEEE Conference on Decision and Control (Volume: 6)*, December 9–12, pp. 6333–6338.

[97] Sykes, P. 2010. Traffic simulation with Paramics. *Fundamentals of Traffic Simulation*. pp. 131–171. Springer, New York.

[98] Underwood, R. T. 1961. Speed, volume, and density relationships. *Quality and theory of traffic flow*, pp. 141–188.

[99] van den Hoogen, E. and Smulders, S. 1994. Control by variable speed signs: Results of the Dutch experiment. In *Proceedings of the 7th International Conference on Road Traffic Monitoring and Control* (*Paper* : 391), April 26–28, London, UnK, pp. 145–149.

[100] van Vuren, T., Baker, J., Ogawa, J., Cooke, D., and Unwin, P. 2012. Managed motorways: modelling and monitoring their effectiveness. *Transportation Research Record: Journal of the Transportation Research Board*, Volume. 2278, pp. 85–94.

[101] Vanderbei, R. J. 2007. *Linear Programming: Foundations and Extensions*. International Series in Operations Research & Management Science. 3rd Edition, Springer, New York.

[102] Wattleworth, J. A. Texas Transportation Institute. 1965. Peak-period analysis and control of a freeway system (Number: 24-15). URL `http://d2dtl5nnlpfr0r.cloudfront.net/tti.tamu.edu/documents/24-15.pdf`.

[103] Wolfram, S. 1983. Statistical mechanics of cellular automata. *Reviews of Modern Physics*, 55(3), pp. 601–644.

[104] Wolfram, S. 1984. Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, 10(1-2), pp. 1–35.

[105] Yeo, H. and Skabardonis, A. 2009. Understanding stop-and-go traffic in view of asymmetric traffic theory. In Lam, W. H. K., Wong, S. C., and Lo, H. K., editors, *Proceedings of the 18th International Symposium on Transportation and Traffic Theory: Golden Jubilee (ISTTT)*, July 16–18, Hong Kong, China, pp. 99–115.

[106] Yin, Y.-F. 2008. Robust optimal traffic signal timing. *Transportation Research Part B: Methodological*, 42(10), pp. 911–924.

[107] Yin, Y.-F. and Lawphongpanich, S. 2007. A robust approach to continuous network designs with demand uncertainty. In Allsop, R. E., Bell, M. G. H., and Heydecker, B. G., editors, *Proceedings of the 17th International Symposium on*

Transportation and Traffic Theory (ISTTT), July 23–25, London, UK, pp. 111–126.

[108] Zhang, L. and Levinson, D. 2004. Optimal freeway ramp control without origin–destination information. *Transportation Research Part B: Methodological*, 38(10), pp. 869–887.

[109] Zhang, L. and Levinson, D. 2005. Balacing efficiency and equity of ramp meters. *ASCE Journal of Transportation Engineering*, 131(6), pp. 477–481.

[110] Zhang, L. and Levinson, D. 2010. Ramp metering and freeway bottleneck capacity. *Transportation Research Part A: General*, 44(4), pp. 218–235.

[111] Zhao, D.-B., Bai, X.-R., Wang, F.-Y., Xu, J., and Yu, W.-S. 2011. DHP method for ramp metering of freeway traffic. *IEEE Transactions on Intelligent Transportation Systems*, 12(4), pp. 990–999.

[112] Zhong, R.-X., Sumalee, A., Pan, T.-L., and Lam, W. H. K. 2013. Stochastic cell transmission model for traffic network with demand and supply uncertainties. *Transportmetrica A: Transport Science*, 9(7), pp. 567–602.

[113] Zhong, R.-X., Sumalee, A., Pan, T.-L., and Lam, W. H. K. 2014. Optimal and robust strategies for freeway traffic management under demand and supplyuncertainties: an overview and general theory. *Transportmetrica A: Transport Science*, 10(10), pp. 849–877.

[114] Ziliaskopoulos, A. K. 2000. A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science*, 34 (1), pp. 37–49.