TECHNICAL UNIVERSITY OF CRETE

SCHOOL OF PRODUCTION ENGINEERING AND MANAGEMENT

FREEWAY TRAFFIC FLOW MODELING AND CONTROL WITH EMPHASIS ON CONGESTED OFF-RAMP AREAS

Thesis submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
by

ANASTASIA SPILIOPOULOU

CHANIA, GREECE

2015

This Ph.D. thesis is approved by:

Advisory committee:

Papageorgiou Markos (Supervisor)
Professor, School of Production Engineering and Management
Technical University of Crete

Papamichail Ioannis (Co-supervisor)
Assistant Professor, School of Production Engineering and Management
Technical University of Crete

Skabardonis Alexander (Member of the advisory committee)
Professor, Department of Civil and Environmental Engineering
University of California, Berkeley

Thesis committee:

Ampountolas Konstantinos
Lecturer, School of Engineering
University of Glasgow

Delis Anargyros
Associate Professor, School of Production Engineering and Management
Technical University of Crete

Geroliminis Nikolaos
Associate Professor, Department of Civil Engineering
École Polytechnique Fédérale de Lausanne (EPFL)

Nikolos Ioannis
Associate Professor, School of Production Engineering and Management
Technical University of Crete

Acknowledgments

Upon the completion of this Ph.D. thesis I would like to thank all the persons who contributed with their knowledge, experience and work to this effort.

I would like to express my deep gratitude to my supervisor Professor Markos Papageorgiou for his invaluable academic and scientific support as well as for his constructive guidance throughout the course of my doctoral work. I am also thankful to the Assistant Professor Ioannis Papamichail for his guidance, advice and continuous support all these years. Moreover, I would like to thank the members of the examination committee for their valuable comments on the thesis topics. Furthermore, many thanks go to all the friends-members of the Dynamic Systems and Simulation Laboratory for their support and friendship all these years.

Last, but not least, I would like to thank my family and Stelios for being a constant source of support and strength without which I would not have been able to accomplish this goal.

This Ph.D. research has been co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Program: Heraclitus II. Investing in knowledge society through the European Social Fund.

Short Biography

Anastasia Spiliopoulou was born in Patras, Greece, in 1983. She received the Diploma and M.Sc. degrees in Production Engineering and Management from the Technical University of Crete, Greece, in 2007 and 2009, respectively.

Since 2006, she has been a research assistant at the Dynamic Systems and Simulation Laboratory and since 2009 a teaching assistant at the School of Production Engineering and Management. Her main research interests include traffic flow modeling, automatic control and road traffic management. Ms. Spiliopoulou has received several awards for excellence during her undergraduate studies and has also received the post-graduate research scholarship "HERACLITUS II". She has participated in the YEAR (2010) and TRAVISIONS (2014) competitions for early stage researchers in the area of surface transport in which her work was shortlisted within the top 8 ideas. Finally, she is, up to now, author or co-author of six technical papers in scientific journals and thirteen technical papers in scientific conference proceedings and has also participated in various research projects.

Abstract

Most metropolitan cities of developed countries have now an extensive network of urban and peri-urban freeways, which aims to provide virtually unlimited and fast mobility to road users around the metropolitan area. However, the increase of traffic demand, especially during the peak hours, and the occurrence of traffic incidents, leads to daily appearance of recurrent and non-recurrent freeway congestion which results in significant increase of travel times, increased fuel consumption, environmental pollution as well as reduced safety.

The problem of freeway congestion in urban and peri-urban freeways cannot always be faced by expanding the existing infrastructure, for economic and environmental reasons; instead, efficient traffic control measures may be employed to mitigate the problem. However, the development of effective real-time traffic control measures implies the availability of suitable mathematical traffic flow models which may be used for the development and testing of the proposed control strategies.

This thesis investigates the particular, but quite frequent, case of (recurrent) freeway congestion due to saturated off-ramps. This kind of congestion is difficult to deal with, and for this reason this frequent traffic flow degradation is rarely addressed in the traffic control literature. Moreover, within the traffic flow modeling literature there are, so far, no studies undertaking validation and comparison of different traffic flow models regarding the reproduction of traffic conditions in such areas. The aim of this research is to investigate traffic flow modeling and traffic control issues for congested freeway off-ramp areas.

In particular, within this thesis the most popular discrete time-space macroscopic traffic flow models, namely the CTM and the METANET models, were validated and compared regarding the representation of traffic conditions at congested

freeway off-ramp areas. The models were calibrated and validated using real traffic data from Attiki Odos freeway in Athens, and by employing various optimization methods.

Apart from the modeling approach, various innovative real-time traffic control measures were developed for congested freeway off-ramp areas. In particular, two different cases were examined and suitable traffic control strategies were proposed for every case. In the first case, a hypothetical network was simulated, and various route diversion strategies were developed that aim to reroute the drivers through alternative routes, towards the same destination, preventing the off-ramp queue spillover and the creation of mainstream congestion. In the second case, a real traffic network was examined where recurrent freeway congestion is created due to congestion on the surface street network which propagates to the freeway mainstream through a saturated off-ramp. The network was simulated by use of microscopic simulation and a real-time merging traffic control algorithm was proposed that aims to maximize the surface street network throughput and at the same time to prevent the off-ramp queue from spilling back into the freeway mainstream. The simulation results, for both investigated cases, showed that the proposed traffic control measures can improve the prevailing traffic conditions, preventing the formation of mainstream congestion. Thus, they are both very promising for a field implementation.

Περίληψη

Οι περισσότερες μεγαλουπόλεις των ανεπτυγμένων χωρών διαθέτουν ένα εκτεταμένο δίκτυο αστικών και περι-αστικών αυτοκινητοδρόμων, το οποίο στοχεύει στην ουσιαστικά ανεμπόδιστη και γρήγορη κίνηση των οχημάτων γύρω από τα αστικά κέντρα. Εντούτοις, η αύξηση της κυκλοφοριακής ζήτησης, ιδιαίτερα κατά τις ώρες αιχμής, σε συνδυασμό με την εμφάνιση διαφόρων περιστατικών, οδηγεί στην καθημερινή δημιουργία κυκλοφοριακής συμφόρησης η οποία έχει ως αποτέλεσμα τη σημαντική αύξηση των χρόνων διαδρομής των οχημάτων, την αύξηση της κατανάλωσης καυσίμων, την περιβαλλοντική ρύπανση και τη μειωμένη οδική ασφάλεια.

Το πρόβλημα της συμφόρησης στους αστικούς και περι-αστικούς αυτοκινητοδρόμους δεν μπορεί πάντοτε να αντιμετωπιστεί με την επέκταση των ήδη υπάρχοντων υποδομών, για οικονομικούς αλλά και περιβαλλοντικούς λόγους. Αντιθέτως, κατάλληλα μέτρα ελέγχου κυκλοφορίας μπορούν να εφαρμοστούν για να αντιμετωπίσουν το πρόβλημα. Η ανάπτυξη αποτελεσματικών μέτρων ελέγχου κυκλοφορίας προϋποθέτει την ύπαρξη κατάλληλων μαθηματικών μοντέλων κυκλοφοριακής ροής τα οποία μπορούν να χρησιμοποιηθούν για την ανάπτυξη και δοκιμή των προτεινόμενων στρατηγικών ελέγχου.

Η εργασία αυτή διερευνά την ιδιαίτερη, αλλά συχνά εμφανιζόμενη, περίπτωση της κυκλοφοριακής συμφόρησης σε αυτοκινητοδρόμους η οποία οφείλεται σε κορεσμένες ράμπες εξόδου. Αυτό το είδος συμφόρησης είναι δύσκολο να αντιμετωπιστεί και για το λόγο αυτό δεν υπάρχουν πολλές αναφορές στη διεθνή βιβλιογραφία σχετικά με κατάλληλα μέτρα ελέγχου κυκλοφορίας για αυτές τις περιοχές. Επιπλέον, στην αντίστοιχη βιβλιογραφία που αφορά τη μοντελοποίηση της κυκλοφοριακής ροής σε οδικά δίκτυα, δεν υπάρχουν μελέτες που να πραγματοποιούν σύγκριση και αξιολόγηση διαφορετικών μοντέλων

κυκλοφοριακής ροής σχετικά με την ικανότητά τους να αναπαριστούν τις κυκλοφοριακές συνθήκες σε αυτές τις περιοχές. Ο στόχος της παρούσας έρευνας είναι να διερευνήσει θέματα μοντελοποίησης και ελέγχου της κυκλοφοριακής ροής σε περιοχές ραμπών εξόδου αυτοκινητοδρόμων υπό συμφόρηση.

Ιδιαίτερα, στην παρούσα διατριβή τα δύο πιο δημοφιλή, μακροσκοπικά μοντέλα κυκλοφοριακής ροής, συγκεκριμένα το CTM και το METANET, αξιολογούνται και συγκρίνονται σχετικά με την ικανότητά τους να αναπαριστούν τις κυκλοφοριακές συνθήκες σε περιοχές ραμπών εξόδου αυτοκινητοδρόμων υπό συμφόρηση. Τα μοντέλα βαθμονομούνται και αξιολογούνται κάνοντας χρήση πραγματικών δεδομένων κυκλοφορίας από ένα τμήμα του αυτοκινητοδρόμου της Αττικής Οδού, στην Αθήνα.

Εκτός από τα θέματα μοντελοποίησης, διαφορετικά μέτρα ελέγχου κυκλοφορίας προτείνονται για την περίπτωση περιοχών ραμπών εξόδου αυτοκινητοδρόμων υπό συμφόρηση. Συγκεκριμένα, δύο διαφορετικές περιπτώσεις εξετάζονται και κατάλληλες στρατηγικές ελέγχου κυκλοφορίας προτείνονται για κάθε περίπτωση. Στην πρώτη περίπτωση που εξετάζεται, ένα υποθετικό δίκτυο χρησιμοποιείται και διαφορετικές στρατηγικές καθοδήγησης πορείας αναπτύσσονται, οι οποίες στοχεύουν να καθοδηγήσουν τους οδηγούς μέσω εναλλακτικών διαδρομών. Σκοπός των στρατηγικών ελέγχου είναι να προστατεύσουν την κορεσμένη ράμπα εξόδου και να αποτρέψουν την υπερχείλιση της ουράς οχημάτων της ράμπας στον αυτοκινητόδρομο και τη δημιουργία συμφόρησης στο κυρίως ρεύμα του αυτοκινητοδρόμου. Στη δεύτερη περίπτωση, ένα πραγματικό δίκτυο εξετάζεται όπου η συμφόρηση στον αυτοκινητόδρομο οφείλεται στην συμφόρηση που έχει δημιουργηθεί στο παράπλευρο αστικό δίκτυο, η οποία εισέρχεται στον αυτοκινητόδρομο μέσω μίας κορεσμένης ράμπας εξόδου. Το δίκτυο προσομοιώνεται μέσω του

μικροσκοπικού προσομοιωτή οδικών δικτύων AIMSUN και μία στρατηγική ελέγχου πραγματικού χρόνου προτείνεται η οποία έχει στόχο να μεγιστοποιήσει τη ροή οχημάτων στο παράπλευρο δίκτυο και παράλληλα να εμποδίσει την είσοδο της ουράς οχημάτων της ράμπας εξόδου στο κυρίως ρεύμα του αυτοκινητοδρόμου. Τα αποτελέσματα των διερευνήσεων, και για τις δύο περιπτώσεις που εξετάστηκαν, έδειξε ότι τα προτεινόμενα μέτρα ελέγχου κυκλοφορίας μπορούν να βελτιώσουν τις κυκλοφοριακές συνθήκες που επικρατούν στο δίκτυο, εμποδίζοντας τη δημιουργία της συμφόρησης στο κυρίως ρεύμα του αυτοκινητοδρόμου. Το γεγονός αυτό καθιστά τις προτεινόμενες στρατηγικές ελέγχου πολλά υποσχόμενες σε περίπτωση ενδεχόμενης εφαρμογής τους στο πεδίο.

Contents

List of Figures

List of Tables

1 Introduction

This first chapter introduces the reader to the problem under study. Section 1.1 states the motivation of this work. Section 1.2 presents the objectives of the study and the adopted approach. Finally, Section 1.3 provides the outline of this thesis.

1.1 Motivation

During the last decades, freeway congestion has been a major problem especially at urban freeways and peri-urban ring-roads. Recurrent traffic congestion is usually encountered at freeway on-ramp areas or freeway-to-freeway merging areas, but, quite frequently, also close to freeway off-ramp areas, leading to infrastructure underutilization, long delays, increased fuel consumption and reduced safety. Since expanding the existing infrastructure is not always a feasible option, for economic and environmental reasons, traffic control has been proposed and employed as an efficient way to mitigate the problem of freeway congestion.

Although various traffic control measures have been proposed for cases of congested freeway merging areas or other types of recurrent active bottlenecks, there is very limited technical literature (and, to the best of the author's knowledge, no practical systems) addressing appropriate control measures for cases where recurrent freeway congestion originates from off-ramp areas; the main reason probably being that there is no direct way, from the freeway side, to control the freeway exit flow, so as to avoid the resulting strong reduction of the freeway capacity and related deterioration of the mainstream traffic conditions. The development of innovative traffic control measures that are able to face the problem of congestion due to saturated freeway off-ramps requires the existence of accurate traffic flow models that are able to reproduce the traffic conditions at such areas with satisfactory accuracy. Within literature a high number of traffic

flow models have been proposed over the last decades, but, to the best of the author's knowledge, none of them has ever been validated and tested for congested freeway off-ramp areas.

1.2 Objectives and approach

The objective of this research is twofold. First, it aims to identify suitable macroscopic traffic flow models that can represent the traffic conditions at congested freeway off-ramp areas with sufficient accuracy. To this end, different macroscopic traffic flow models are validated and compared using real traffic data from a freeway stretch in Athens, Greece, where recurrent traffic congestion is created due to a saturated off-ramp. In particular, the models are first calibrated and the optimal parameter values are estimated by use of suitable optimization algorithms. Then, the models are validated and compared regarding their accuracy in representing the prevailing traffic conditions. The model that achieves the highest accuracy is utilized in the second part of the thesis.

The second objective of this research includes the development of innovative real-time traffic control measures for congested freeway off-ramp areas. In particular, two different cases are examined, that are often encountered in reality, and suitable traffic control strategies are proposed for every case. In the first case, it is considered that there is a freeway off-ramp with limited capacity, e.g., due to its layout or due to a traffic light placed at the end of the off-ramp, at its intersection with a surface street, and there is no possibility to increase the off-ramp capacity. During the peak hours the off-ramp may not serve the arriving demand thus queue is formed on the off-ramp which, eventually, spills-back into the freeway mainstream creating congestion. Moreover, it is considered that there are nearby off-ramps that could lead towards the same destination and could be utilized to divert a portion of the drivers in order to protect the saturated off-ramp and

prevent the off-ramp queue spill-over into the freeway mainstream. The above situation is simulated for a hypothetical network, by use of the macroscopic traffic flow model selected in the first part of the thesis, and various route diversion strategies are developed that aim to reroute the drivers through alternative routes, preventing the off-ramp queue spillover and the creation of mainstream congestion.

In the second case, it is considered that there is a freeway off-ramp exiting to a surface street and recurrent freeway congestion is created due to congestion on the surface street network which propagates to the freeway mainstream through a saturated off-ramp. The reason for congestion on the surface street network is the high arriving flow, from the surface street and the off-ramp, combined with strong weaving phenomena on the surface street merge area. The outlined situation is actually appearing in a real network in Santiago, Chile. This network is emulated by use of microscopic simulation and a real-time merging traffic control algorithm is proposed that aims to maximize the surface street network throughput and at the same time to prevent the off-ramp queue spill over into the freeway mainstream.

The main contributions of this thesis include:

- The validation and comparison of macroscopic traffic flow models in the reproduction of traffic conditions at congested freeway off-ramp areas.
- The development and testing of various real-time route diversion policies that aim to reroute the drivers through nearby off-ramps towards the same destination, preventing the off-ramp queue spill-over and the creation of mainstream congestion.
- The development and testing of a real-time merging traffic control strategy which aims to maximize the throughput at a surface street network and at

the same time prevent the off-ramp queue spill over into the freeway mainstream and the resulting freeway congestion.

1.3 Thesis outline

The thesis is organized as follows: Chapter 2 presents the state-of-the art in traffic flow modeling and traffic control with emphasis on congested freeway off-ramp areas. Chapter 3 describes the model calibration procedure, presents the macroscopic traffic flow models that are employed and compared regarding the representation of traffic conditions at congested freeway off-ramps areas and presents suitable optimization methodologies that can be applied to solve the parameter estimation problem. Chapter 4 includes the calibration results of the selected traffic flow models for a particular freeway site, as well as, the validation and comparison of the models using real traffic data. Chapter 5 develops, investigates and demonstrates real-time traffic control strategies to mitigate the problem of freeway congestion due to saturated off-ramps and examines two particular cases. Finally, Chapter 6 summarizes the findings and results of this thesis and discusses future extensions with respect to traffic flow modeling and traffic control for congested freeway off-ramp areas.

2 Background

This chapter provides an overview of the two areas within the field of traffic engineering which are related to the contents of this thesis, i.e. traffic flow modeling and traffic control. In particular, Section 2.1 presents the evolution of traffic flow models over the last decades, with emphasis on macroscopic traffic flow modeling. The models are classified and various issues such as accuracy, applicability and limitations are discussed. Section 2.2 summarizes the traffic control measures and methodologies that have been proposed so far, to face the problem of recurrent freeway congestion created due to saturated off-ramps.

2.1 Traffic flow modeling

One important aspect of the traffic flow theory concerns the development of mathematical models for the representation of the road traffic flow. Traffic flow models may be used for the planning of new, upgraded or modified road infrastructures; for the development and testing of traffic flow estimation and prediction algorithms; as well as for the design and testing of traffic control strategies and other traffic engineering tasks [1]. The need for accurate and robust traffic flow models combined with the complexity and non-linearity of the traffic phenomena resulted in a broad number of proposed traffic flow models during the last decades. In the following sections, first, a classification of different modeling approaches is presented, followed by the description of the most popular macroscopic traffic flow models. Furthermore, at the end of this section, some model discretization and calibration issues are discussed.

2.1.1 Classification of traffic flow models

The traffic flow models are classified as microscopic, macroscopic or mesoscopic depending on the level of detail they use. The microscopic models describe the time-space behavior of individual drivers (vehicles) by use of dynamic variables such as position and velocity. In contrast to the microscopic approach, the macroscopic models consider the traffic flow as a compressible fluid with specific characteristics. The traffic variables used to describe the dynamics of this fluid are the mean speed, the density and the traffic flow (or volume). Finally, the mesoscopic models use a medium level of detail, combining microscopic and macroscopic approaches to a hybrid model.

Microscopic models are complex, non-analytic (they are simulation tools rather than closed-form mathematical models) and computationally expensive; hence their suitability for on-line traffic operations, such as estimation prediction and control, is limited. Moreover, the calibration of the microscopic models is not an easy task due to the high number of parameters they include and the lack of real 'microscopic' traffic data. Finally, microscopic models are known to produce unrealistic vehicles' behavior under particular circumstances (e.g. close to merging areas) within the simulation environment. On the other hand, macroscopic traffic flow models include lower number of parameters compared to microscopic models, thus require less calibration effort; also, they have an analytical form, which allows their usage for various significant traffic engineering tasks (estimation, prediction, control strategy design) beyond simulation. Finally, they are computationally less demanding, thus they are suitable for on-line traffic operations and for the simulation of large traffic networks.

Within this thesis two macroscopic traffic flow models are employed and compared regarding the representation of traffic conditions at a congested freeway

stretch due to a saturated off-ramp. In the following, an elaborate description of the most popular macroscopic traffic models is presented.

2.1.2 Conservation equation

All macroscopic traffic flow models are based on the conservation of vehicles equation, also known as the continuity equation:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \tag{2.1}$$

where x denotes the spatial coordinate in the direction of flow, t is the time, ρ is the traffic density and q denotes the traffic flow. This equation implies that in any traffic system the traffic flow is conserved, thus no vehicles are created or destroyed in the system. The conservation equation is complemented by the following fundamental relation

$$q(x,t) = \rho(x,t) \cdot v(x,t) \tag{2.2}$$

where v denotes the mean speed. The above two equations include three unknown variables thus solution is impossible. Therefore, an additional equation or an assumption has to be supplied and this has lead to a variety of continuum traffic flow models. For example, the assumption of functional flow-density (or equivalently speed-density) relationship leads to the first-order models, while the addition of momentum equations results in high-order continuum models. These models are reviewed below.

2.1.3 First-order models

Lighthill and Whitham [2] and, independently, Richards [3], proposed the first, and so far the most popular first-order model, the so-called LWR model. The LWR

model considers a static relation between mean speed and density (known as the fundamental diagram):

$$v(x,t) = V^e[\rho(x,t)\,] \tag{2.3}$$

The non-linear first-order partial differential equation that results by replacing (2.3) and (2.2) into (2.1) is:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial}{\partial x}(\rho(x,t)V^e[\rho(x,t)\,]) = 0 \tag{2.4}$$

This simple traffic flow model is able to reproduce not only free-flow conditions but also wave formation and propagation under congested conditions. On the other hand, the model fails to capture other important traffic phenomena.

The major drawbacks of the original LWR model lie on the fact that it considers a static relation between mean speed and density. As a consequence, the model doesn't allow for mean speed variations, other than those implied by the fundamental diagram, thus it is not suitable for the description of non-equilibrium situations occurring at on-ramp areas, lane-drop areas, or stop-and-go traffic. Moreover, it does not take into account factors such as the drivers' reaction time and the traffic hysteresis phenomena, which have been observed in real traffic flow [4] and indicate that vehicle acceleration and deceleration are not symmetric processes. Furthermore, due to the requirement of continuity made on the fundamental diagram, the model cannot reproduce the capacity drop phenomenon which is observed at congested freeway areas, thus its usage for the design of traffic control strategies is limited (see also [5]).

Several researchers tried to deal with the above model limitations proposing various extensions of the LWR model. As an example, Newell [6] addressed the traffic hysteresis phenomena by suggesting a fundamental diagram with multiple

branches. However, most researchers have followed the direction of incorporating a momentum conservation equation, to describe the dynamics of mean speed.

2.1.4 High-order models

High-order models include the conservation equation (2.1) and one (or more) PDE(s) to describe the dynamics of mean speed. Here below, several high-order traffic flow models are presented.

Payne (1971)

The first, and still the most widely used, second order model was proposed by Payne [7]. This model was derived from car-following theory by means of Taylor's expansion. Payne' momentum equation has the following form:

$$\underbrace{\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}}_{acceleration} = \underbrace{\frac{1}{\tau}[V^e(\rho) - v]}_{relaxation} + \underbrace{\frac{1}{2\tau\rho}\frac{dV^e(\rho)}{d\rho}\frac{\partial \rho}{\partial x}}_{anticipation} \tag{2.5}$$

where τ is the relaxation time and $V^e(\rho)$ is the equilibrium speed-density relation. This equation consists of an *acceleration* term, at the left-hand side of (2.5), a *relaxation* term, which represents the tendency of drivers to adjust their speed to the equilibrium speed-density relation $V^e(\rho)$, and an *anticipation* term which reflects the effect of the downstream traffic conditions to the drivers' reaction.

Payne's model achieves to overcome some deficiencies of the LWR-type models; for example it takes into account the vehicles' acceleration capabilities and the drivers' reaction time. Moreover it is able to predict traffic instabilities such as stop-and-go waves. The most severe criticism of Payne's model was formulated by Daganzo [8]. The criticism includes, first, the fact that the model allows vehicles to be influenced by the upstream traffic conditions while in reality vehicles primarily react to the downstream traffic conditions. Second, it allows slower vehicles to be influenced by faster vehicles, which is unrealistic; and finally, in some cases, it

may estimate negative speeds and flows, also known as *wrong-way travel* phenomenon.

During the last decades, several researchers have proposed various modifications-extensions of Payne's momentum equitation, in order to fix-improve some of the models' deficiencies. In the following, the most popular second-order traffic models are presented.

Phillips (1979)

Based on the kinetic theory Philips [9] proposed the following momentum equation:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = \frac{1}{\tau(\rho)}[V^e(\rho) - v] - \frac{1}{\rho}\frac{dP}{d\rho}\frac{\partial \rho}{\partial x} \tag{2.6}$$

where the relaxation time τ is density-dependent and the variable P is the traffic pressure approximated as the product of density ρ and the speed variance $\Theta(\rho)$, i.e. $P(\rho) = \rho\Theta(\rho)$. The speed variance $\Theta(\rho)$ can be estimated as $\Theta(\rho) = \Theta_0(1 - \rho/\rho_{max})$ with ρ_{max} corresponding to the maximum traffic density. This model accounts for the fact that at low densities small deviations from equilibrium are expected while at high densities large deviations from equilibrium are possible. However, according to this formula the density-gradient $dP/d\rho$ of the traffic pressure will be negative in a certain density range.

Kühne (1984) and Kerner and Konhäuser (1993)

Kühne [10] as well as Kerner and Konhäuser [11],[12] proposed the following momentum equation:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = \frac{1}{\tau}[V^e(\rho) - v] - \frac{\Theta_0}{\rho}\frac{\partial \rho}{\partial x} + \eta\frac{\partial v^2}{\partial x^2} \tag{2.7}$$

where here $\Theta(\rho) = \Theta_0$, thus Θ is considered as a constant and cannot be interpreted as speed variance anymore. In this way, the problem of negative values of density-gradient $dP/d\rho$, mentioned above, is avoided. Moreover, this model includes an additional high order anticipation viscosity term $\eta\, \partial v^2/\partial x^2$, where η is a constant value at Kuhne's model, while at Kerner and Konhauser model $\eta(\rho) = \eta_0/\rho$, with η_0 a positive constant. This viscosity term, essentially, smoothes out the shock wave fronts, which is desirable from empirical and numerical points of view.

Papageorgiou (1990)

In order to account for merging and lane-changing phenomena close to on-ramp and lane-drop areas, Papageorgiou [13] proposed two extra terms at Payne's momentum equation (2.5). In particular, the additional terms are $-\delta v r/\rho$ and $-\varphi \Delta\lambda \rho v^2/\rho_{cr}$, where δ and φ, are model parameters, r is the incoming on-ramp flow, $\Delta\lambda$ are the number of lanes being dropped and ρ_{cr} is the critical density.

Zhang (1998)

Zang [14] proposed a momentum equation similar to Payne's equation (2.5):

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = \frac{1}{\tau}[V^e(\rho) - v] - \rho\left(\frac{dV^e(\rho)}{d\rho}\right)^2 \frac{\partial \rho}{\partial x} \qquad (2.8)$$

The main difference between these two models lies in the anticipation term – i.e. the anticipation in (2.5) is proportional to density but in (2.8) is proportional to the inverse of density. By this formulation, Zang addresses the undesirable property of high-order models of "wrong-way travel" as, according to (2.8), traffic disturbances are always propagated against the traffic stream.

Treiber (1999)

Treiber et al. in [15] proposed a macroscopic gas-kinetic-based traffic model that was derived from a microscopic model of vehicle dynamics of the following form:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = \frac{1}{\tau}[V^0 - v] - \frac{1}{\rho}\frac{\partial(\rho A(\rho)v^2)}{\partial x} - \frac{V^0 A(\rho)}{\tau A(\rho_{max})}\left(\frac{\rho_\alpha T v}{1-\rho_\alpha/\rho_{max}}\right)^2 B(\delta_v) \tag{2.9}$$

where V^0 is the desired speed, $A(\rho)$ is a density-dependent function, T is the time headway, ρ_α is the density at an advanced "interaction point" and $B(\delta_v)$ is a non-local interaction term. In contrast to other macroscopic models, equation (2.9) depends on the density and speed at two different locations $(\rho, v, \rho_\alpha, v_\alpha)$, thus introducing non-locality. The non-locality has smoothing properties like the viscosity term used in Kühne and Kerner and Konhäuser models, but its effect is anisotropic. There is no smoothing in the forward direction, which would imply that cars would react on density or velocity gradients of the vehicles behind them. Moreover, according to the authors, the non-locality of the model leads to a more favorable numerical stability behavior.

General form of Payne-type models

As presented above, most macroscopic high-order traffic models include the continuity equation (2.1) and a momentum equation of the following form:

$$\underbrace{\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}}_{acceleration} = \underbrace{\frac{1}{\tau}[V^e(\rho) - v]}_{relaxation} - \underbrace{\frac{1}{\rho}\frac{\partial P}{\partial x}}_{anticipation} \tag{2.10}$$

where P denotes the traffic pressure. Different settings of the traffic pressure P, the relaxation time τ, and the equilibrium speed-density relation $V^e(\rho)$ results in different macroscopic traffic models. In particular:

- Payne's model and Papageorgiou's model are derived for $P(\rho) = -\frac{V^e(\rho)}{2\tau}$ and constant relaxation time τ.
- Philip's model is obtained for $P(\rho) = \rho\Theta(\rho)$ with $\Theta(\rho) = \Theta_0(1 - \rho/\rho_{max})$ and a density-dependent relaxation time $\tau(\rho)$.
- Kühne's and Kerner's and Konhäuser's models is resulted for $P = \rho\Theta_0 - \eta\frac{\partial v}{\partial x}$, where Θ_0 is a positive constant and η is a viscosity coefficient.
- Zhang's model is derived for $P = \frac{1}{3}\rho^3 V^{e\prime 2}(\rho)$ with $V^{e\prime}(\rho) = \frac{dV^e(\rho)}{d\rho}$.
- Treiber's model is resulted for $P = \rho A(\rho)v^2$ where A is a density dependent function, and the equilibrium speed-density relation is given by

$$V^e(\rho, v, \rho_a, v_a) = V^0\left\{1 - \frac{A}{2A(\rho_{jam})}\left[\frac{\rho_a T v}{1 - \rho_a/\rho_{jam}}\right]^2 B(\delta_v)\right\}$$

- Finally, it should be noted that for $\tau \to 0$ and $P = 0$, the LWR model is obtained.

2.1.5 Discretization of continuum macroscopic models

The original model PDEs cannot be directly computed in digital computers, which calls for the employment of appropriate numerical schemes. From an engineering application point of view, the final space-time discretized models should be as simple as possible and have nice analytical properties (e.g. have an explicit state-space form, contain continuous and differentiable functions), which would allow for simple and transparent computation codes, convenient discretization intervals, short computation times; as well as for direct application of powerful mathematical methods (e.g. Kalman filtering, optimization, optimal control) ([1], [5]). Since the original PDEs are largely empirical, it may not be necessary to apply special effort and employ complex numerical schemes for their accurate discretization. Instead, an approximate, but explicit and analytical, space-time discretized model may first be derived from the PDEs; to be used eventually as a

self-contained modeling tool for practical applications (rather than the original PDEs).

For the LWR model the most popular numerical solutions include the CTM (Cell Transmission Model), which was proposed by Daganzo (see[16], [17]) and it is a discretized (and simplified) version of the LWR model; and the model proposed by Lebacque [18] who applied the Godunov-scheme to the LWR model. Regarding the Payne-type models, FREFLO [19] which is a discretized version of Payne model and METANET ([20], [21]), which is a discretized and enhanced version of Payne model are the most popular simulation tools.

Within this thesis two macroscopic models (a first-order and a second-order model) are validated and compared regarding the representation of traffic conditions at congested freeway off-ramp areas. The selected first-order model is CTM, while the selected second-order model is METANET. These models fulfill the simplicity and convenience requirements mentioned earlier: they have a space-time discrete, explicit, analytical state-space form and allow for convenient discretization intervals. Hence, it is not surprising that they are by far the most frequently utilized macroscopic traffic flow models and have been used by multiple research groups for a variety of traffic engineering tasks, such as simulation, dynamic traffic assignment, estimation, optimization, optimal control of freeway (and, for CTM, also of urban road) traffic.

2.1.6 Calibration of macroscopic traffic flow models

As presented in the previous sections, the traffic flow models include a set of parameters, whose values may differ for different freeway sites and depend on factors such as the network geometry, the drivers' behavior at the specific network, the percentage of trucks, the weather conditions, etc. Thus, before employing a traffic flow model (either a first- or a higher-order model) in practice,

it is important to first calibrate it against real traffic data. The calibration procedure aims to appropriately specify the model parameter values, so that the representation of the network and traffic flow characteristics is as accurate as the model structure allows.

Within the vast literature on macroscopic traffic flow modeling, there are surprisingly few studies addressing or actually conducting model calibration and validation against real traffic data. In particular, Grewal and Payne [22] identify the Payne model parameters using traffic data from a microscopic simulator; Cremer and May [23] propose an extension of Payne model and validate it using real traffic data from a freeway stretch in California; Helbing [24] calibrates a gas-kinetic model using traffic data from a Dutch freeway; Sanwal et al. [25] propose an extension of Papageorgiou model and validate it using data from a freeway stretch in California; Kotsialos et al. [26] calibrate METANET model for a large-scale motorway network around Amsterdam; Muñoz et. al. [27] validate the CTM model for a congested freeway stretch in California; Ngoduy et al. [28] approximate the Payne model using three different numerical schemes and validate it with traffic data from a Dutch freeway; Monamy et al. [29] propose a node model based on the LWR model and calibrate it for a ringroad in Paris; Ngoduy and Maher [30] calibrate the Treiber model for a motorway stretch in UK; Poole and Kotsialos [31] validate METANET model for a UK motorway stretch. Furthermore, there is a very limited number of studies undertaking, in addition to validation, a comparison of different models. In particular, Cremer and Papageorgiou [32] validate an improved version of Payne model and compare it with simplified versions of the new model; Papageorgiou et al. [33] compare three macroscopic traffic models (one first-order and two second-order models); Michalopoulos et al. [34] validate and compare five different traffic models – one first-order model and four second-order models.

In this thesis, the first-order model CTM and the second order model METANET are calibrated and compared using real traffic data from a particular freeway stretch, in Athens, Greece. Chapter 3 contains a detailed description of the utilized models and the calibration procedure, while the calibration and validation results are presented in Chapter 4.

2.2 Traffic control measures at congested freeway off-ramp areas

Freeway traffic congestion at saturated off-ramp areas is usually created either due to the high freeway exit flow, higher than the off-ramp flow capacity; or due to the spillback of an off-ramp queue into the freeway mainstream. In the latter case, an off-ramp queue may have been created due to a capacity-reducing circumstance, e.g. due to a downstream urban traffic light. Traffic congestion originating from off-ramp areas is a particular, but quite frequent case of (recurrent) congestion, appearing usually at urban or peri-urban freeways during peak periods. This kind of congestion is difficult to deal with, since there is no direct way to control the freeway exit flow; and this is probably the reason why this frequent traffic flow degradation is rarely addressed in the traffic control literature.

The methodologies proposed so far either focus on affecting the behavior of the freeway drivers, e.g. by eliminating the lane changing maneuvers near the off-ramps; or suggest the increase of the off-ramps' exit flow; or propose a combination of the above. In [35], some thoughts are expressed on how to increase the capacity at congested freeway off-ramp areas via fixed-time or dynamic lane assignment on the freeway mainstream, according to the drivers' destinations, by use of Variable Message Signs (VMS). The potential closure of a congested off-ramp and drivers' rerouting through nearby off-ramps is also proposed. In [36] and [37], it is suggested to ban the lane-changing maneuvers near off-ramp areas, e.g.

via pavement markings, combined with increasing the off-ramp exit flow when needed, without though considering the effects to the surface street network. As noted in [36], sometimes banning lane-changing maneuvers merely results in shifting the bottleneck to another upstream location where lane-changing maneuvers are allowed. In [38], it is also suggested to abruptly increase the off-ramp exit flow when the queue on the off-ramp is about to spill back to the freeway, admitting that this action may cause problems to the surface street traffic. In [39], the detouring of a part of vehicles moving on the surface street network is proposed, in order to enable the increase of the off-ramp's exit flow, thus benefiting the freeway users at the expense of surface street users. Finally, in [40] a two-stage control model is employed that optimize the signal plans on the surface street network in order to prevent the off-ramp queue spillback into the freeway mainstream. Note that all the above proposed control measures were demonstrated using specific case studies and real or hypothetical traffic networks. This fact indicates that this is a particular type of congestion and different freeway sites may call for different traffic control measures, depending on the network layout, the prevailing traffic conditions, the expected drivers compliance, the available traffic control equipment, etc. As a result, in the field, each situation should be viewed as a particular case and the available traffic control measures should be adjusted to the corresponding network characteristics.

Within this thesis two innovative traffic control measures are proposed for cases where freeway congestion is created due to an over-spilling off-ramp queue. In particular, Chapter 5 examines two different cases, and suitable traffic control measures are proposed which aim to avoid the off-ramp queue spill-over and the resulting mainstream congestion. In the first case the proposed traffic control measures include various route diversion policies which aim to reroute the drivers through nearby off-ramps towards the same destination. The second case demonstrates the application of a real-time merging traffic control strategy which

aims to maximize the throughput at the surface street network and at the same time to prevent the off-ramp queue spill-back into the freeway mainstream and the resulting mainstream congestion.

3 Calibration of Macroscopic Traffic Flow Models

The model calibration procedure aims to appropriately specify the model parameter values, so that the representation of the network and traffic flow characteristics is as accurate as the model structure allows. The most common approach is to minimize the discrepancy between the model's estimations and the real traffic data, by use of appropriate optimization tools. In the following, Section 3.1 presents the selected traffic flow models, Section 3.2 describes the model calibration procedure and Section 3.3 presents suitable optimization methods that can be employed to solve the parameter estimation problem.

3.1 Selected macroscopic traffic flow models

As indicated in Chapter 2, a first-order model and a second-order model, namely the CTM and METANET models, are validated and compared regarding the representation of traffic conditions in a real freeway network. Next sections provide a detailed description of the models' equations.

3.1.1 The Cell Transmission Model (CTM)

As mentioned before, the CTM model ([16], [17]) is a discretized and simplified version of the LWR model. The simplification lies on the fact that a triangular fundamental diagram is considered, as shown in Figure 3.1. Under CTM, the freeway is divided into consecutively numbered sections (cells) of length L_i, where i is the section index. Each section may have an on-ramp and off-ramp near its upstream or downstream boundary, respectively, as shown in Figure 3.2. Time is also discretized into uniform intervals of duration T, with a discrete time index $k = 0,1,2,\dots,K$ where K is the time horizon. The state variable for section i is the

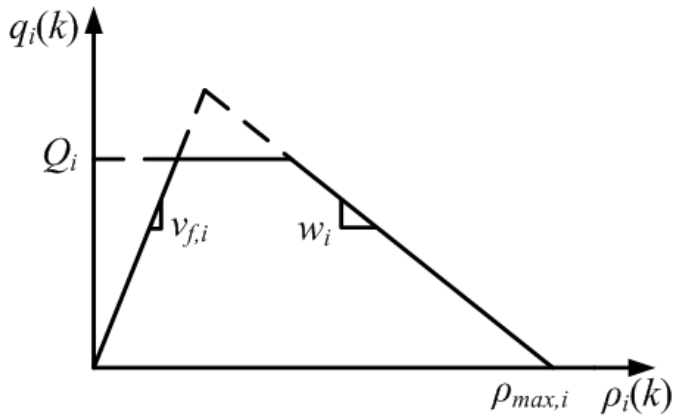

Figure 3.1 Fundamental diagram within CTM.

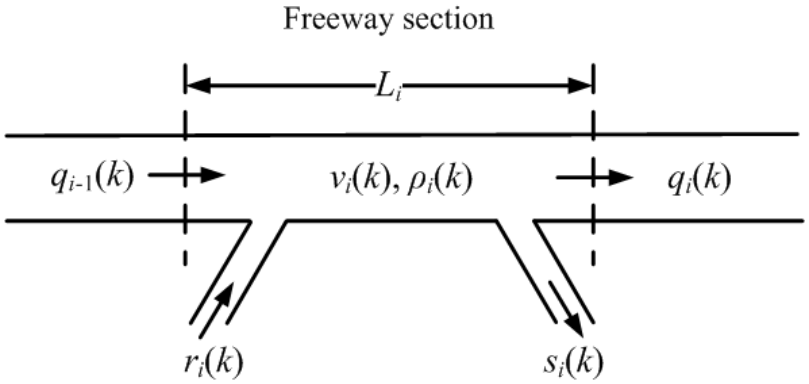

Figure 3.2 Freeway discretization.

density $\rho_i(k)$, which corresponds to the number of vehicles included in section i at the time instant kT, which is calculated as follows:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i \lambda_i}[q_{i-1}(k) - q_i(k) + r_i(k) - s_i(k)] \qquad (3.1)$$

This is a conservation-of-vehicles equation, where $q_i(k)$ is the traffic flow exiting section i and entering section $i+1$, $r_i(k)$ is the traffic flow entering the freeway section i from an on-ramp and $s_i(k)$ is the traffic flow exiting the freeway section from an off-ramp, and equals to $s_i(k) = \beta_i(k) q_i(k)/[1 - \beta_i(k)]$ where $\beta_i(k)$ is the splitting ratio, and λ_i is the number of lanes of section i.

There are three different types of freeway sections, i.e. ordinary sections, merge sections and diverge sections. In case of ordinary sections, i.e. no on-ramps or off-ramps are present between two sections, i and $i+1$, the flow $q_i(k)$ exiting each section i is estimated as the minimum of two quantities:

$$q_i(k) = \min\{S_i(k),\ R_{i+1}(k)\} \tag{3.2}$$

where $S_i(k) = \min\{v_{f,i}\rho_i(k)\lambda_i,\ Q_i\}$ is the maximum flow that can be supplied by section i, during the time interval k and $R_{i+1}(k) = \min\{Q_{i+1},\ w_{i+1}[\rho_{max,i+1} - \rho_{i+1}(k)]\lambda_{i+1}\}$ is the maximum flow that can be received by section $i+1$ over the same time interval. Moreover, Q_i and Q_{i+1} are the flow capacities of sections i and $i+1$, respectively; $\rho_{max,i+1}$ is the maximum density of section $i+1$; and $v_{f,i}$ and w_{i+1} are the slopes of the free flow and congested portions of the triangular fundamental diagram at sections i and $i+1$, respectively (see Figure 3.1). It should be noted that, for the purposes of this study, the flow $q_i(k)$ is calculated in veh/h, not in veh/T as in the original equations of Daganzo [16]. Furthermore the freeway sections may have similar, but not necessarily equal lengths, in contrast to the original CTM which considers strictly equal length sections, which should satisfy, for stability reasons, the following relation:

$$v_{f,i}T \leq L_i \tag{3.3}$$

i.e. the sections length must be longer than the free flow distance.

In case of merge sections, i.e. an on-ramp intervenes between two sections, i and $i+1$, there are two cases, one where the downstream section can receive both the supply flow $S_i(k)$ from the upstream section i and the on-ramp demand $r_{i+1}(k)$, and one where the combined supply flow and on-ramp demand exceed the maximum receiving flow $R_{i+1}(k)$. For each case, the flow $q_i(k)$ exiting section i is estimated as follows:

$$q_i(k) = \begin{cases} S_i(k), & if\ \ S_i(k) + r_{i+1}(k) \le R_{i+1}(k) \\ \max\{0, R_{i+1}(k) - r_{i+1}(k)\}, & \text{otherwise} \end{cases} \quad (3.4)$$

In the second case, the total flow entering section $i+1$ is equal to $R_{i+1}(k)$.

Finally, in case of diverge sections, i.e. an off-ramp intervenes between two sections i and $i+1$, then the flow $q_i(k)$ exiting section i is given (see also [17]) by:

$$q_i(k) = q_{i,total}(k)[1 - \beta_i(k)] \quad (3.5)$$

$$q_{i,total}(k) = min\left\{S_i(k), \frac{R_{i+1}(k)}{1-\beta_i(k)}, \frac{R_J(k)}{\beta_i(k)}\right\} \quad (3.6)$$

where $q_{i,total}(k)$ is the total flow exiting section i and $R_J(k)$ is the maximum flow that can be received by the off-ramp. In this way, the available space on the off-ramp is taken into account, and, in case the mainstream flow wishing to exit via the off-ramp is higher than the available space on the off-ramp, the off-ramp queue may spill over onto the freeway mainstream and create congestion at section i. Finally the mean speed $v_i(k)$ at every section i, is computed, using the fundamental relation (2.2), as:

$$v_i(k) = q_{i,total}(k)/\rho_i(k)\lambda_i\ . \quad (3.7)$$

3.1.2 The METANET model

The METANET model [20] is a discretized and enhanced variation of the Payne model. As with the previous model, it considers that the freeway is divided into consecutively numbered sections i, with respective lengths L_i (which should also satisfy (3.3)) and number of lanes λ_i, as shown in Figure 3.2. Time is also discretized into uniform intervals of duration T, with a discrete time index $k = 0,1,2,\dots,K$ where K is the time horizon. The state variables for section i are

the density $\rho_i(k)$ and the mean speed $v_i(k)$ at the time instant kT, which are calculated according to the following equations:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i \lambda_i}[q_{i-1}(k) - q_i(k) + r_i(k) - s_i(k)] \tag{3.8}$$

$$v_i(k+1) = v_i(k) + \frac{T}{L_i} v_i(k)[v_{i-1}(k) - v_i(k)] + \frac{T}{\tau}[V^e[\rho_i(k)] - v_i(k)] - \frac{\nu T[\rho_{i+1}(k) - \rho_i(k)]}{\tau L_i[\rho_i(k) + \kappa]} \tag{3.9}$$

where, similar to the previous model, $q_i(k)$ is the traffic flow exiting section i and entering section $i+1$, $r_i(k)$ is the traffic flow entering the freeway section i from an on-ramp and $s_i(k)$ is the traffic flow exiting the freeway section from an off-ramp, and equals to $s_i(k) = \beta_i(k) q_i(k)/[1 - \beta_i(k)]$ where $\beta_i(k)$ is the splitting ratio. Moreover, τ (a time constant), ν (an anticipation constant) and κ are model parameters while function $V^e[\rho_i(k)]$ corresponds to the fundamental diagram, calculated using the following equation:

$$V^e[\rho_i(k)] = v_{f,i} \exp\left[-\frac{1}{a_i}\left(\frac{\rho_i(k)}{\rho_{cr,i}}\right)^{a_i}\right] \tag{3.10}$$

where $v_{f,i}$ is the free flow speed, $\rho_{cr,i}$ is the critical density (for which the flow at section i is maximized) and a_i is a further model parameter for section i. Moreover, the mean speed calculated by the model is truncated if it is below a minimum value v_{min}. Papageorgiou [13] proposed two additional terms for more accurate modeling of merging and lane-drop phenomena. In particular, the impact on mainstream speed due to an on-ramp merging flow is considered by adding the term $-\delta T r_i(k) v_i(k)/L_i \lambda_i [\rho_i(k) + \kappa]$ at the right hand side of (3.9) for the merging section, where δ is a model parameter. This term is not used if there is a lane gain downstream of the on-ramp, i.e., if there is a dedicated lane for entering vehicles. In order to take into account the impact on speed due to intensive lane-

changing at lane-drop areas, the term $-\varphi T \Delta\lambda \rho_i(k) v_i(k)^2 / L_i \lambda_i \rho_{cr,i}$, is added to (3.9) for the section immediately upstream of the lane drop, where φ is a model parameter and $\Delta\lambda$ is the number of dropped lanes.

At bifurcation locations (e.g. off-ramps), a downstream density $\rho_{i+1}(k)$ is needed in (3.9) for the section i entering the bifurcation; this density reflects the upstream influence of the downstream traffic conditions. However, as we have at least two downstream sections at bifurcations, the following formula was proposed in [20]:

$$\rho_{i+1}(k) = \sum_{\mu \epsilon O_i} {\rho_\mu}^2(k) \Big/ \sum_{\mu \epsilon O_i} \rho_\mu(k) \tag{3.11}$$

where $\rho_{i+1}(k)$ is the virtual density downstream of section i, which is used in (3.9); and $\rho_\mu(k)$ is the density of each section downstream of section i, O_i being the set of exiting sections. The quadratic average used in (3.11) accounts for the fact that congestion may spill back to a section i from any of its downstream sections (e.g., in case of spillback from a saturated off-ramp), even if the rest downstream sections are not congested. Notice that (3.11) does not include any parameter to be calibrated. Finally, the flow $q_i(k)$ exiting section i and entering section $i + 1$ is computed, using the fundamental relation (2.2), as:

$$q_i(k) = v_i(k) \rho_i(k) \lambda_i [1 - \beta_i(k)]. \tag{3.12}$$

As presented above, the traffic flow models include a set of parameters whose values are unknown (and differ for different freeway sites). Thus, in order to achieve a fair comparison of the selected models, their optimal parameter values should be appropriately specified through a calibration exercise.

3.2 Model calibration procedure

The model parameter calibration (or parameter estimation) procedure aims at enabling a macroscopic traffic flow model to represent the traffic conditions of a freeway network with the highest achievable accuracy. The estimation of the unknown model parameters is not a trivial task, since the system equations are highly nonlinear in both the parameters and the state variables. Consider that a macroscopic discrete-time state-space model is described by the following state equation

$$\begin{aligned} &\mathbf{x}(k+1) = \boldsymbol{f}[\mathbf{x}(k), \mathbf{d}(k), \mathbf{p}] \qquad\qquad k = 0,1,\dots,K-1 \\ &\mathbf{x}(0) = \mathbf{x_0} \end{aligned} \tag{3.13}$$

where k is the discrete time index; $\mathbf{x}$ stands for the state vector, $\mathbf{d}$ corresponds to the external variable (disturbance) vector and $\mathbf{p}$ is the model parameter vector. In particular, the state vector $\mathbf{x}$ includes the section densities (and for higher order models also the mean speeds etc), the external variable (disturbance) vector $\mathbf{d}$ consists of all known boundary conditions such as the network inflows, the turning rates at bifurcations, and the network downstream densities; and $\mathbf{p}$ includes the unknown model parameters that need to be specified.

If the initial state $\mathbf{x_0}$ is given and the external vector $\mathbf{d}(k)$ is known over a time horizon $k = 0,1,\dots,K-1$, then the parameter estimation problem can be formulated as a nonlinear least-squares output error problem which aims at the minimization of the discrepancy between the model calculations and the real traffic data by use of the following cost function,

$$J(\mathbf{p}) = \sqrt{\frac{1}{\mathrm{K}}\sum_{k=1}^{K}[\mathbf{y}(k) - \mathbf{y^m}(k)]^2} \tag{3.14}$$

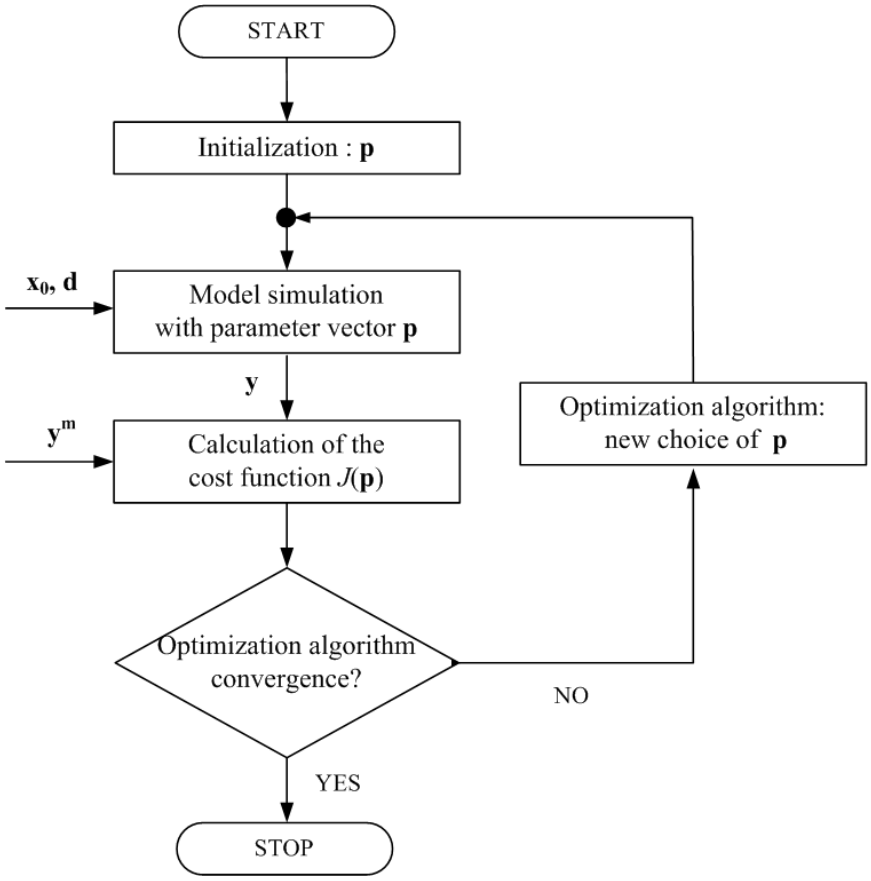

Figure 3.3 Model calibration procedure.

subject to (3.13); where $\mathbf{y}(k) = \mathbf{g}[\mathbf{x}(k)]$ is the measurable model output vector (typically consisting of flows and mean speeds at various network locations) and $\mathbf{y^m}(k)$ includes the real measured traffic data (consisting of flows and speeds at the corresponding network locations). The model parameter values are selected from a closed admissible region of the parameter space, which may be defined on the basis of physical considerations and previous experience. The determination of the optimal parameter set must be performed by means of a suitable nonlinear programming routine, whereby for each choice of a new parameter vector **p,** the value of the performance index (PI) (3.14) may be computed by a simulation run of the model equations as shown in Figure 3.3.

After the calibration procedure, the resulting traffic flow models must be validated before their potential use in a real implementation. Model validation aims to

ensure that the resulting model reflects reliably the traffic characteristic of the specific network, thus it may reproduce its typical traffic conditions. To this end, the model is applied to the same freeway site, albeit by using different data for the disturbance vector $\mathbf{d}$ and initial state $\mathbf{x_0}$, than those used for its calibration, and the model output $\mathbf{y}$ is compared to the corresponding real traffic data $\mathbf{y^m}$. In other words, the calibration procedure is carried out using real traffic data from a specific date, while for the validation procedure traffic data from different dates are used.

The nonlinear, non-convex least-squares optimization problem of parameter calibration is known to have multiple local minima (see [30] for an illustration), and hence gradient-based solution algorithms are not an option. In the investigations presented in Chapter 4, various global optimization algorithms are employed, both deterministic and stochastic. Next section, briefly describes all utilized algorithms.

3.3 Global optimization algorithms

Three derivative-free optimization algorithms are selected to solve the parameter estimation problem examined within this thesis. In particular, the deterministic Nelder-Mead algorithm, a stochastic genetic algorithm and the stochastic cross-entropy method. In the following, the selected algorithms are shortly described along with their potential advantages and weak points.

3.3.1 Nelder-Mead algorithm

The Nelder-Mead method ([41], [42]) is one of the best known algorithms for multidimensional unconstrained optimization. The method does not require any derivative information, which makes it suitable for problems with non-linear, discontinuous or stochastic cost function.

The method uses a simplex, i.e. a n-dimensional geometrical shape with $n+1$ vertices. Every vertex $\mathbf{p}_i$, where $i = 1, \ldots, n+1$, corresponds to a potential solution which in turn corresponds to a cost function value, $J(\mathbf{p}_i)$. The algorithm starts with an initial working simplex and then performs a sequence of transformations aiming at reducing the cost function value at its vertices. In particular, at each iteration the algorithm orders the simplex' vertices with respect to the corresponding cost function values e.g. $J(\mathbf{p}_1) \leq J(\mathbf{p}_2) \leq \cdots \leq J(\mathbf{p}_{n+1})$ and calculates the centroid $\mathbf{p}_c$ of all vertices excluding the worst vertex $\mathbf{p}_{n+1}$. Then, it computes the new working simplex from the current one as follows. First, an attempt is made to replace only the worst vertex $\mathbf{p}_{n+1}$ with a better point by using *reflection*, *expansion* or *contraction*. If this succeeds, the accepted point becomes the new vertex of the working simplex. Otherwise, the algorithm *shrinks* the simplex towards the best vertex $\mathbf{p}_1$. In this case, n new vertices are computed. Simplex transformations are controlled by four parameters: ζ for *reflection*, χ for *contraction*, γ for *expansion* and σ for *shrinkage*. Note that there is a low need for fine-tuning the algorithm parameters since the parameter values proposed in the original papers seem to work good in a broad number of applications. The above procedure continues until the working simplex becomes sufficiently small or when the function values $J(\mathbf{p}_i)$ are close enough to each other.

In contrast to other direct search methods which call, at each iteration, for multiple cost function evaluations, Nelder-Mead typically requires only one or two function evaluations, except when performing the shrinkage transformation which is, actually, quite rare in practice. As a result, the method typically gives significant ameliorations of the cost function value quite fast. On the other hand, in some cases the method may perform a large number of iterations without significant improvement of the cost function value. To cope with this problem, restarting the algorithm several times, with reasonably small number of allowed iterations per each run, may prove helpful. Generally, the evolution of the

working simplex and the produced best solution are dependent on the initial working simplex, since the algorithm searches for new points using the vertices of the working simplex, and this may lead to different paths for different initial simplexes. Such different evolution paths may or may not lead to the same final best solution. To face this fact, multiple algorithm runs may be carried out using different initial vertices for the working simplex and checking the corresponding obtained solutions.

3.3.2 Genetic algorithms

A genetic algorithm ([43], [44]) is a heuristic search method which belongs to the larger class of evolutionary algorithms. The genetic algorithm (GA) mimics the process of biological evolution and uses techniques inspired by natural selection, mutation and crossover. It is suitable for a variety of optimization problems, in which the objective function is discontinuous, non-differentiable, stochastic, or highly nonlinear.

The method uses a population of candidate solutions to an optimization problem and evolves it towards better solutions. The evolution starts from an initial *population* of randomly generated *individuals* (solutions) which are evaluated through their respective cost function values (fitness). At each iteration, called *generation*, the algorithm selects individuals (*parents*) from the current generation and uses them to produce the individuals (*offspring*) for the next population. To do so, the GA uses three main types of rules:

- *Selection rules* select individuals (parents), with probabilities proportional to their fitness; the selected parents contribute to the population of the next generation. Some of the individuals in the current population, which have best fitness, are chosen as *elite*. These elite individuals are passed directly and unchanged to the next population.

- *Crossover rules* combine (random) couples of parents to form offspring for the next generation, thus exchanging information between two candidate solutions.
- *Mutation rules* apply random changes to individual parents, which may introduce new features (i.e. new parameter space regions) to the population.

Through the stochastic operations of *selection*, *crossover* and *mutation,* the population "evolves", over successive generations, towards potentially better solutions, and the algorithm stops when one of the stopping criteria is met, e.g. when no significant improvement in the cost function value is achieved over successive iterations (generations), or when the maximum allowed number of iterations is reached.

The main advantage of GA is its flexibility to search complex solution spaces; thanks to its stochastic operations, it is less likely to restrict the search to a bad local minimum area, in contrast to point-to-point movement optimization techniques. On the other hand, each iteration requires many cost function evaluations, which increases substantially the computational cost, especially for problems with a computationally expensive cost function or problems which require large population size. It is worth noting that, since the evaluation of the cost function for each individual is independent of all others, the parallelization of GA is an option. Finally, it is important to tune the algorithm's parameters, i.e. the population size, the elite rate, the crossover probability and the mutation rate in order to find appropriate and efficient settings for the specific problem being examined.

3.3.3 Cross-entropy method

The cross-entropy method ([45], [46]) is a general Monte-Carlo approach to combinatorial and continuous multi-extremal optimization and importance

sampling problems. The method originates from the field of rare event simulation, where very small probabilities need to be accurately estimated.

The algorithm starts from an initial population of potential solutions generated using a continuous, usually uniform, distribution h_0. At each iteration t, the solutions are evaluated through the cost function and sorted into ascending order; and the best b% solutions comprise the elite sample. The probability density function, $\hat{g}_t$, of this elite sample is estimated, e.g. using a Kernel density estimator as proposed by [30], and the probability distribution of the population is updated using the equation:

$$\hat{h}_{t+1} = (1 - \varepsilon)\hat{h}_t + \varepsilon\hat{g}_t \tag{3.15}$$

where ε is a smoothing parameter, typically in the range [0.7, 0.9]. The updated density equation $\hat{h}_{t+1}$ is used in the next iteration to generate the new random sample of solutions. The algorithm continues leading, over iterations, to increasingly more spiked shapes of the population probability distribution; and it stops when one of the stopping criteria is met, e.g. when the shape of the probability density function becomes very spiked (i.e. concentrated around the optimal value) or when the maximum allowed number of iterations is reached.

As with the previous algorithms, the CE method does not require any derivative information, thus it may be applied to problems where the objective function is discontinuous, non-differentiable or highly nonlinear. In contrast to other stochastic methods, the selection of the potential solutions is not a completely random process, since the utilized distribution is affected by the best solutions of each iteration. The main disadvantage of the method is that it requires as many cost function evaluation as the size of the population, resulting in large computational cost and slow convergence. Again, it is important to tune the

algorithm's parameters, i.e. the population size, the elite rate b and the smoothing parameter ε in order to find appropriate and efficient settings for the specific problem being examined.

The following chapter presents the calibration and validation results of the selected macroscopic traffic flow models by employing the above optimization methods.

4 Calibration and Validation Results

Two well-known macroscopic traffic flow models, namely the CTM (Cell Transmission Model) and the METANET model, are validated and compared in the special, but quite frequently occurring, case where congestion is created due to saturated freeway off-ramps. In the following sections, Section 4.1 presents the considered freeway stretch and the utilized traffic data, Section 4.2 describes the calibration settings for both models and all optimization methods, Section 4.3 displays the calibration results and Section 4.4 includes the sensitivity investigations and the validation of the resulting models. Finally, Section 4.5 summarizes the results and conclusions of the calibration exercise.

4.1 Freeway test site and real traffic data

The freeway stretch considered in this study is a part of Attiki Odos freeway (34th to 28th km, direction from the Airport to Elefsina) in Athens, Greece. This freeway stretch includes three on-ramps and three off-ramps, as shown in Figure 4.1. Figure 4.2 represents the examined freeway stretch in terms of nodes and links. Each node (N0–N8) illustrates a bifurcation point or a junction or any location marking a change of the network geometry; whereas the homogeneous road stretches between these locations are represented by links (L1–L8). Each network link is subdivided in model sections of equal length; see for example link L1 which is divided in 3 sections, with the vertical short lines denoting the section borders. Using this representation, the network sections are well-defined, and the model equations presented in Section 3.1 are directly applied to these sections. Moreover, Figure 4.2 displays the length, number of sections and number of lanes for each link; the exact location of the on-ramps and off-ramps; as well as the location of the available detector stations, which are depicted by bullets.

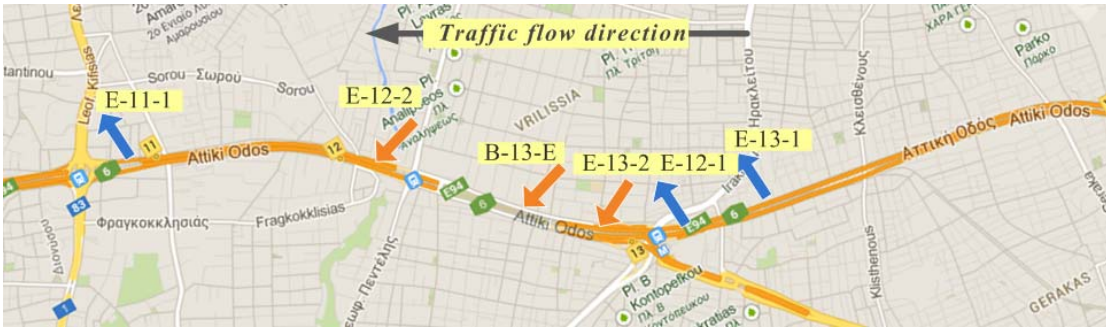

Figure 4.1 Attiki Odos freeway stretch.

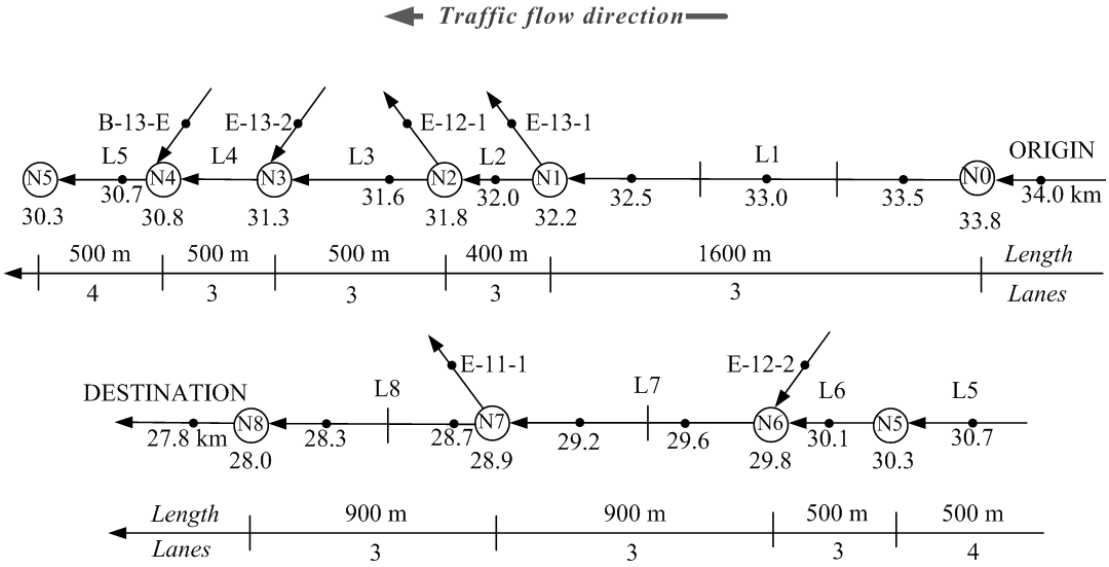

Figure 4.2 Representation of the considered freeway stretch.

The real traffic data used in this study were provided by ATTIKES DIADROMES S.A., which is the freeway operating company. In particular, the provided traffic data includes flow and speed measurements at the corresponding detector station locations, with a time resolution of 20 seconds, for the time period May-June 2009. The traffic data analysis showed that, within this particular freeway stretch, recurrent traffic congestion is formed during the morning peak hours.

Figure 4.3 illustrates the space-time diagram of speed measurements for 4 different days: 26/05/2009, 16/06/2009, 23/06/2009 and 25/06/2009. It is observed that congestion is created during 8-10 a.m.; the congestion originates at the 29th km of

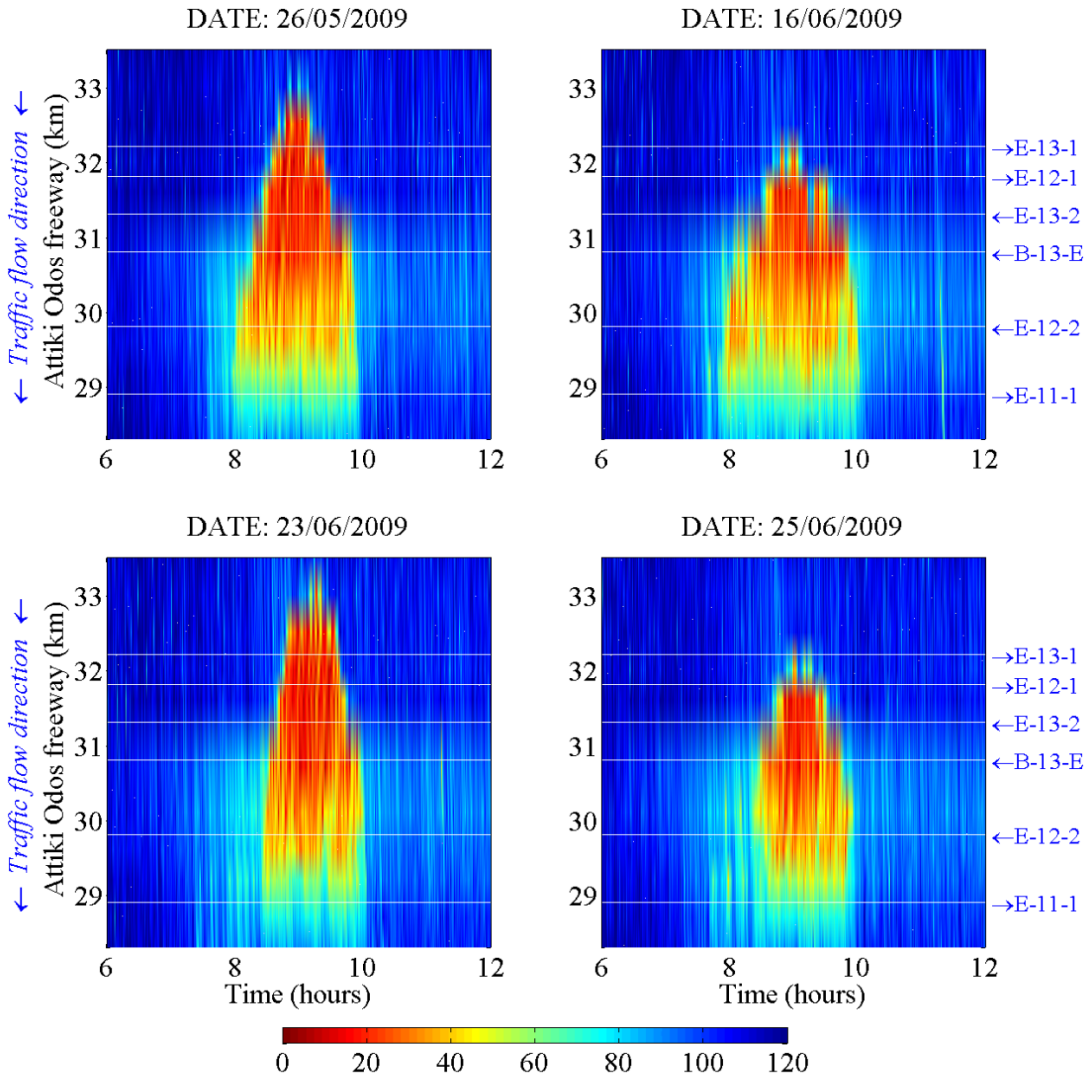

Figure 4.3 Time-space diagram of real speed measurements for 26/05/2009, 16/06/2009, 23/06/2009 and 25/06/2009, for the time period 6-12am.

the freeway stretch and spills back several kilometers upstream, up to the 32nd km, and on some days up to the 33rd km.

From Figure 4.2 it may be seen that the congestion creation area is actually a diverge area, with the off-ramp E-11-1 receiving high exit flow during the morning peak hours, according to real traffic data. The high exit flow rate, in combination with the limited capacity of the off-ramp, leads to the creation of congestion, which propagates upstream for several kilometers on the freeway

mainstream. The test network and traffic data presented above, are used to calibrate and validate the selected traffic flow models. It should be noted that the main criterion for selecting these 4 days was that, during the morning hours 6-12 a.m., no incident and no detector failure occurred at the examined freeway stretch, which can, of course, not be reproduced by any traffic flow model.

4.2 Calibration settings

The calibration procedure, as presented in Section 3.2, is applied to the CTM and the METANET models. Both models can readily assume the state-space form of (3.13) for any freeway network. The state vector $\mathbf{x}$ includes: the section densities in the case of CTM; the section densities and mean speeds in the case of METANET.

The parameter vector $\mathbf{p}$ for the CTM model consists of the free flow speed v_f, the maximum density ρ_{max}, the congestion wave speed w and the capacity flow Q. It is important to stress that these parameters are common for all network links, i.e. one single fundamental diagram was considered for all mainstream sections. Note that the fundamental diagram is deemed to reflect infrastructure and general traffic conditions; hence it should be the same for sections with similar characteristics. Specifying a separate fundamental diagram for each section may lead to fake results, e.g. when important model parameters, such as capacity or critical density, are not visible in the local data which may cover only a limited part of the real fundamental diagram. In any case, calibration tests, where different fundamental diagrams were considered for each section or group of sections, did not lead to a considerable improvement of the performance index compared to the case of one single fundamental diagram for all sections.

In addition to the above, the off-ramps are considered as sections with specific maximum available space. In particular, the off-ramp E-11-1 (see Figure 4.2) has

congestion wave speed and maximum density, denoted as $w_{off-ramp}$ and $\rho_{\max\_off-ramp}$, respectively, while for the rest off-ramps a high value for the maximum density was considered, since no congestion originates from these off-ramps. Based on the above considerations, the parameter vector $\mathbf{p_1}$ for the CTM model is $\mathbf{p_1} = [v_f \;\; \rho_{max} \;\; w \;\; Q \;\; \rho_{\max\_off-ramp} \;\; w_{off-ramp}]$.

Regarding the METANET model, the parameter vector $\mathbf{p}$ consists of the free flow speed v_f, the critical density ρ_{cr} and the parameters α, τ, ν, δ and φ which are common for all the freeway sections. Thus, one single fundamental diagram is considered also in this model. Moreover, the model includes two extra parameters which are known from previous validation exercises to be of minor importance and are, therefore, given constant values, in order to reduce the dimension of the parameter vector. In particular, κ is set equal to 10 veh/km/lane and v_{min} is set to 7 km/h. Considering the above, the parameter vector $\mathbf{p_2}$ for the METANET model is $\mathbf{p_2} = [v_f \;\; \rho_{cr} \;\; a \;\; \tau \;\; \nu \;\; \delta \;\; \varphi]$.

The external variable vector $\mathbf{d}$ for the CTM model consists of the origin inflows (mainstream inflow at the network origin and on-ramp inflows (Figure 4.2)), turning rates at bifurcations (i.e. at the off-ramps), and densities (at the downstream end of the network and at the considered off-ramps (see again Figure 4.2)); while, in addition, the origin speeds are also needed for the METANET model. The initial state vector $\mathbf{x_0}$ includes the initial densities at all network sections for the CTM model; while in addition to this, the initial speed measurements for all network sections are also required for the METANET model. These initial values are estimated as the mean value of the measurements, at the corresponding detector locations, for the first three minutes of the simulation period. Finally, for both models, the vectors $\mathbf{y}$ and $\mathbf{y^m}$ (i.e. the model output and the real traffic data, respectively), which are utilized for the calculation of the Performance Index (PI) (see 3.14), include only the speed values at all detector

locations. This is because traffic densities are difficult to measure directly; on the other hand, experience from previous validation procedures has shown that the calculation of reasonably accurate flows is not a major problem for a macroscopic traffic flow model, since the conservation equation guarantees that, whatever flows in will eventually flow out; in contrast, it is much more challenging to correctly model the time evolution of the mean speeds in each section.

Both models were calibrated using real traffic data from 16/06/2009 and a simulation step T = 5 sec. Three different optimization methods were employed to estimate the models' parameters for the examined freeway stretch; namely, the deterministic Nelder-Mead algorithm, a stochastic genetic algorithm and the stochastic cross-entropy method. All simulations were performed using a desktop computer with 2.4 GHz CPU and 2.0 GB of RAM. The calibration procedure, including the traffic flow models and the optimization algorithms, has been programmed in MATLAB (R2010a). It should be noted that, for each utilized algorithm, various initial calibration tests were carried out using different values for the algorithms' parameters, which helped to fine-tune the algorithms parameters for this particular problem, and these values were used in the investigations presented in the following sections. In particular, the Nelder-Mead algorithm was implemented using the following parameters: ζ = 1, χ =2, γ = –0.5 and σ = 0.5. Moreover the utilized termination criteria were the cost function convergence or the working simplex convergence, with tolerance equal to 0.1 and the maximum allowed number of iterations which was set equal to 1000. The genetic algorithm was employed with *population size* equal to 500, *elite rate* equal to 0.01, *crossover rate* equal to 0.8 and *mutation rate* equal to 0.1. The utilized termination criteria were again the cost function convergence and the maximum allowed number of iterations (generations) which was set equal to 1000. Finally, the cross-entropy method was applied using *population size* equal to 500, *elite rate* 0.05 and smoothing parameter ε equal to 0.8. The utilized termination criteria

were the bandwidth of the kernel estimation function, which was set equal to 0.1 and the maximum allowed number of iterations which was set to 1000. See Section 3.3 for a description of the algorithms parameters.

4.3 Calibration results

The aim of the current study is to test and compare two macroscopic traffic flow models in the particular case where recurrent freeway congestion is triggered by a saturated off-ramp. In order to achieve a fair comparison, the models are first calibrated, i.e. the optimal parameter values are specified for the examined freeway test network, using a day's measured data. The calibration results for each examined model and the estimated parameter values are presented in the following sections.

4.3.1 CTM model calibration

Three different optimization methods were employed to calibrate the CTM model. Figure 4.4 presents the convergence of each utilized algorithm over iterations. It is observed that all three methods finally converge to a low PI value, i.e. 14.4 for the Nelder-Mead algorithm, 14.7 for the genetic algorithm and 14.4 for the CE method.

Table 4-1 presents the performance of all algorithms in terms of various criteria, i.e. the total number of iterations, total number of cost function evaluations and computation time. It is shown here, that although the Nelder-Mead algorithm took a large number of iterations to converge, it actually requires much less cost function evaluations, 609, compared to the genetic algorithm and the CE method which needed 36000 and 18500, respectively. As a result, the computation time of the Nelder-Mead algorithm is considerably lower compared to the other two algorithms, since it converged in just 0.8 min, in contrast to the other two

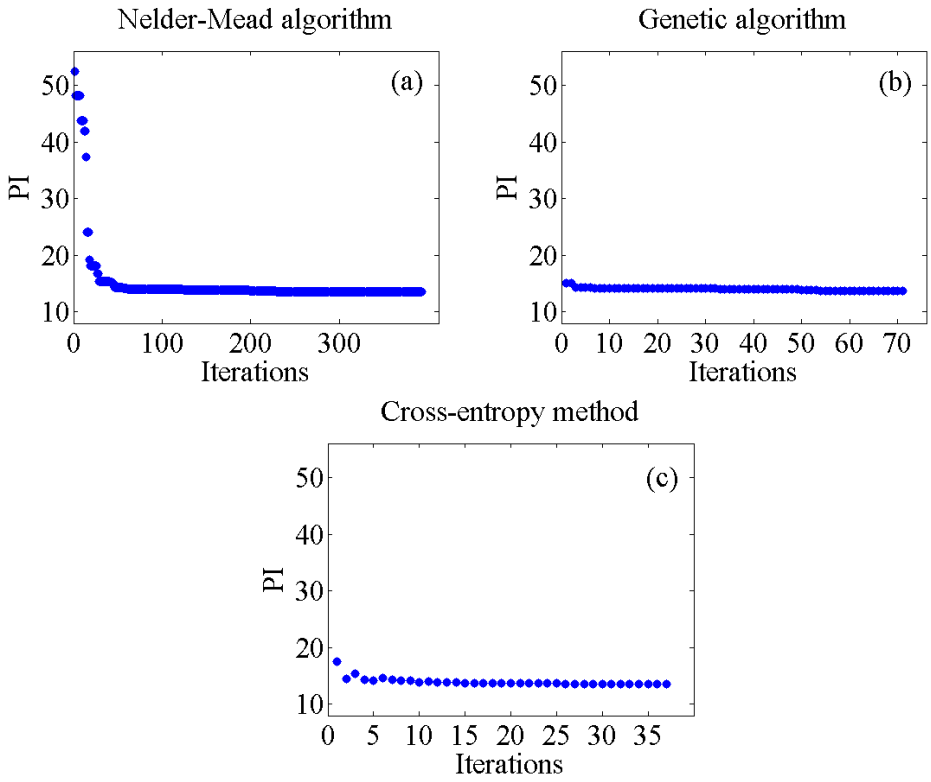

Figure 4.4 CTM model calibration: performance index value over iterations using (a) the Nelder-Mead algorithm, (b) the genetic algorithm and (c) the cross-entropy method.

methods which required 34.6 min and 19.7 min, respectively. Finally, comparing the genetic algorithm with the CE method it may be seen that, both algorithms achieve similar PI values with the CE method being some 15 minutes, faster than the genetic algorithm.

Each utilized algorithm converged to a different optimal parameter set resulting to three different models. Table 4-2 presents the optimal parameter values estimated by each employed method, where Model 1.1 denotes the model produced using the Nelder-Mead algorithm, Model 1.2 is the model obtained by employing the genetic algorithm and Model 1.3 is the model resulted by use of the CE method. It is observed that all three methods estimated the very same value for the free flow speed, v_f, while similar values were estimated for the rest model parameters. Figure 4.5 illustrates the estimated fundamental diagram (FD) (3.2) for the freeway for all

three obtained models. It is shown here, that the estimated FD is very similar for all three models. Moreover, it should be noted that all specified optimal parameter values are reasonably reflecting their respective physical significance, which indicates that the basic model structure is accordingly suitable to describe real traffic phenomena.

Figure 4.6 presents the space-time diagram of the real speed measurements for the calibration date (16/06/2009) and the corresponding speed estimations of all three produced models. It is observed that the estimations of all three models are very similar to the real traffic data, thus they are able to reproduce the traffic conditions of this freeway network for this particular date with high accuracy, creating the occurring congestion at the right time and place and for the right duration and extent. In fact, the estimations of all three models are very close to

Table 4-1 CTM model calibration: optimization algorithms' performance.

Optimization method	**Iterations**	**Cost function evaluations**	**Computation time (min)**
Nelder-Mead algorithm	393	609	0.8
Genetic algorithm	71	36000	34.6
Cross-entropy method	37	18500	19.7

Table 4-2 CTM optimal parameter values estimated by use of different optimization algorithms.

Model	**CTM parameters**					
	v_f **(km/h)**	ρ_{max} **(veh/km/lane)**	w **(km/h)**	Q **(veh/h)**	$\rho_{max\_off\text{-}ramp}$ **(veh/km/lane)**	w_{off_ramp} **(km/h)**
Model 1.1 (Nelder-Mead algorithm)	100.4	142.6	22.6	2273	126.3	19.7
Model 1.2 (genetic algorithm)	100.3	148.9	21.5	2247	124.7	21.0
Model 1.3 (Cross-entropy method)	100.4	153.8	19.8	2268	123.5	20.5

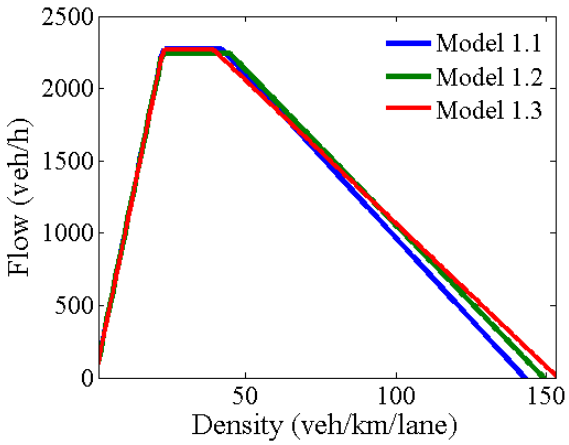

Figure 4.5 CTM model calibration: estimated fundamental diagram (FD) for all three obtained models.

each other, which is actually expected due to the fact that they include similar optimal parameter values.

In more detail, Figure 4.7 and Figure 4.8 present the time-series of the real speed and flow measurements and the corresponding speed and flow estimations of Model 1.1 at various detector locations. It is shown that congestion originates, indeed, from the 29th km and propagates upstream up to the 32nd km persisting between 8 and 10 a.m. It should be noted that downstream of the congestion creation area (29th km), the model estimates free flow speed (see Figure 4.7), since, by its very structure, it does not take into account the acceleration of vehicles exiting the congested freeway area. The corresponding time-series of the speed and flow estimations of Model 1.2 and Model 1.3 may be found in the Appendix (Figure A.1 – Figure A.4)

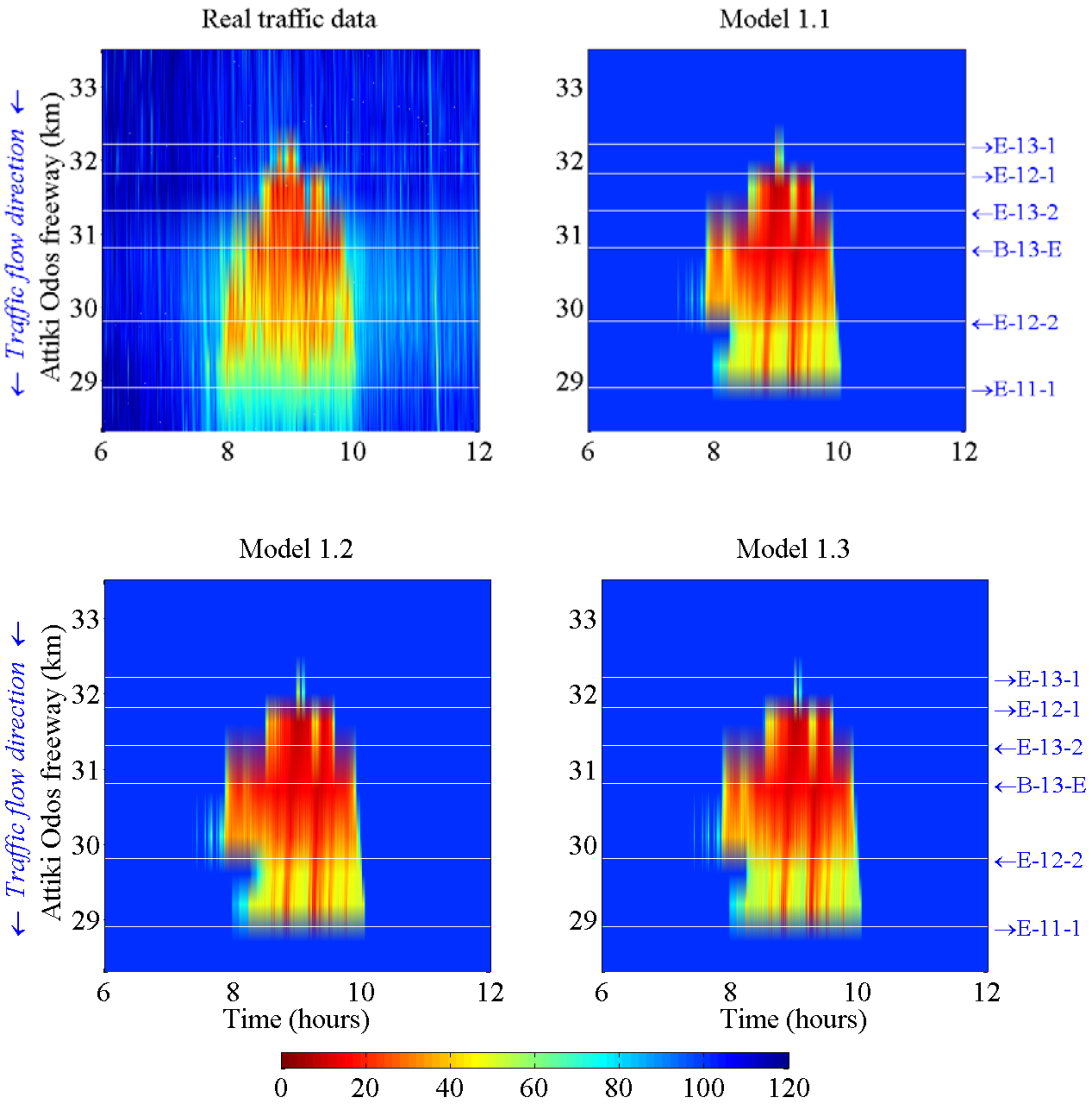

Figure 4.6 CTM model calibration: space-time diagrams of measured speeds and the models' estimation of speed for 16/06/2009.

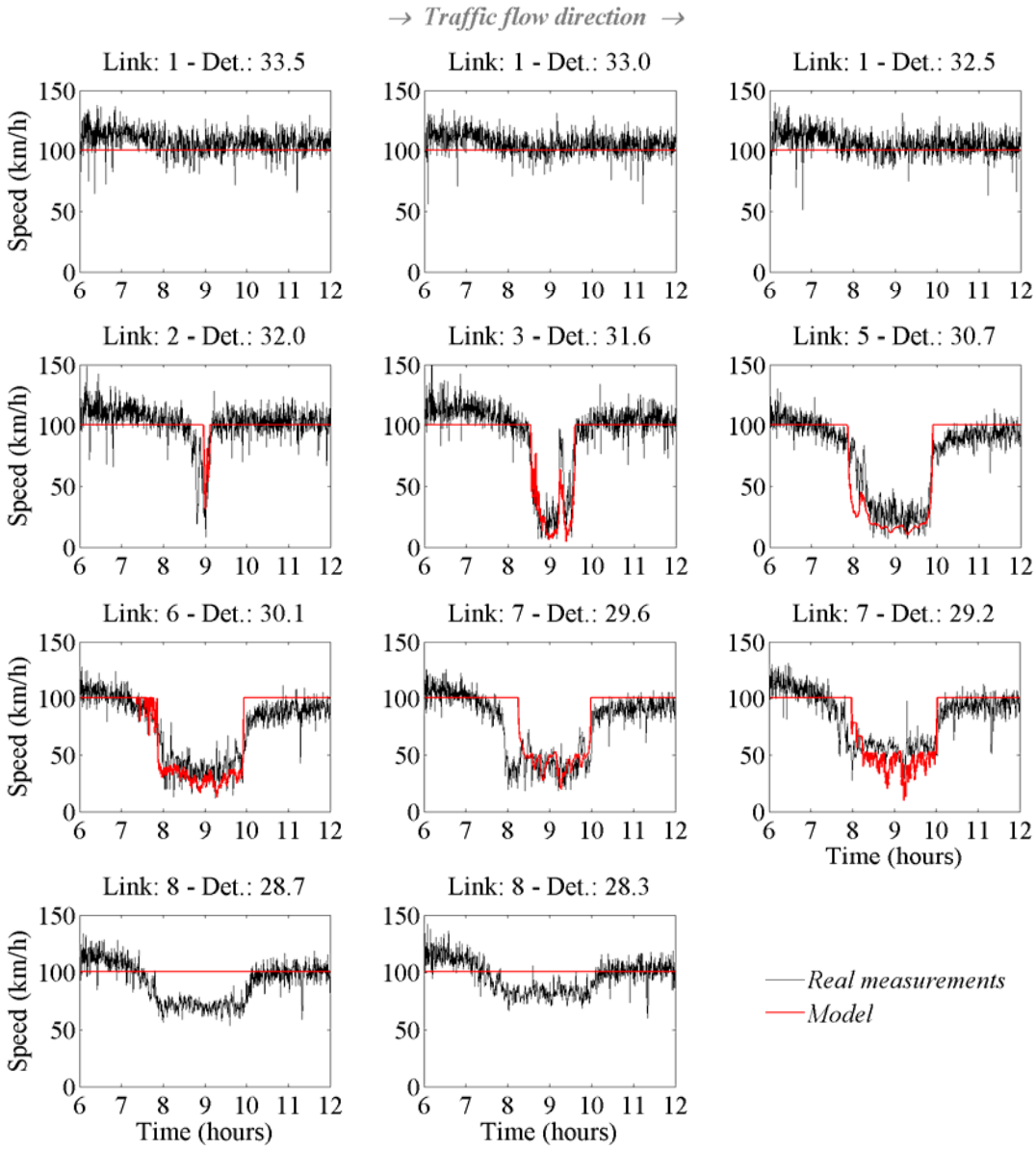

Figure 4.7 CTM model calibration: time-series of the real speed measurements and Model 1.1 estimation of speed at various detector locations for 16/06/2009.

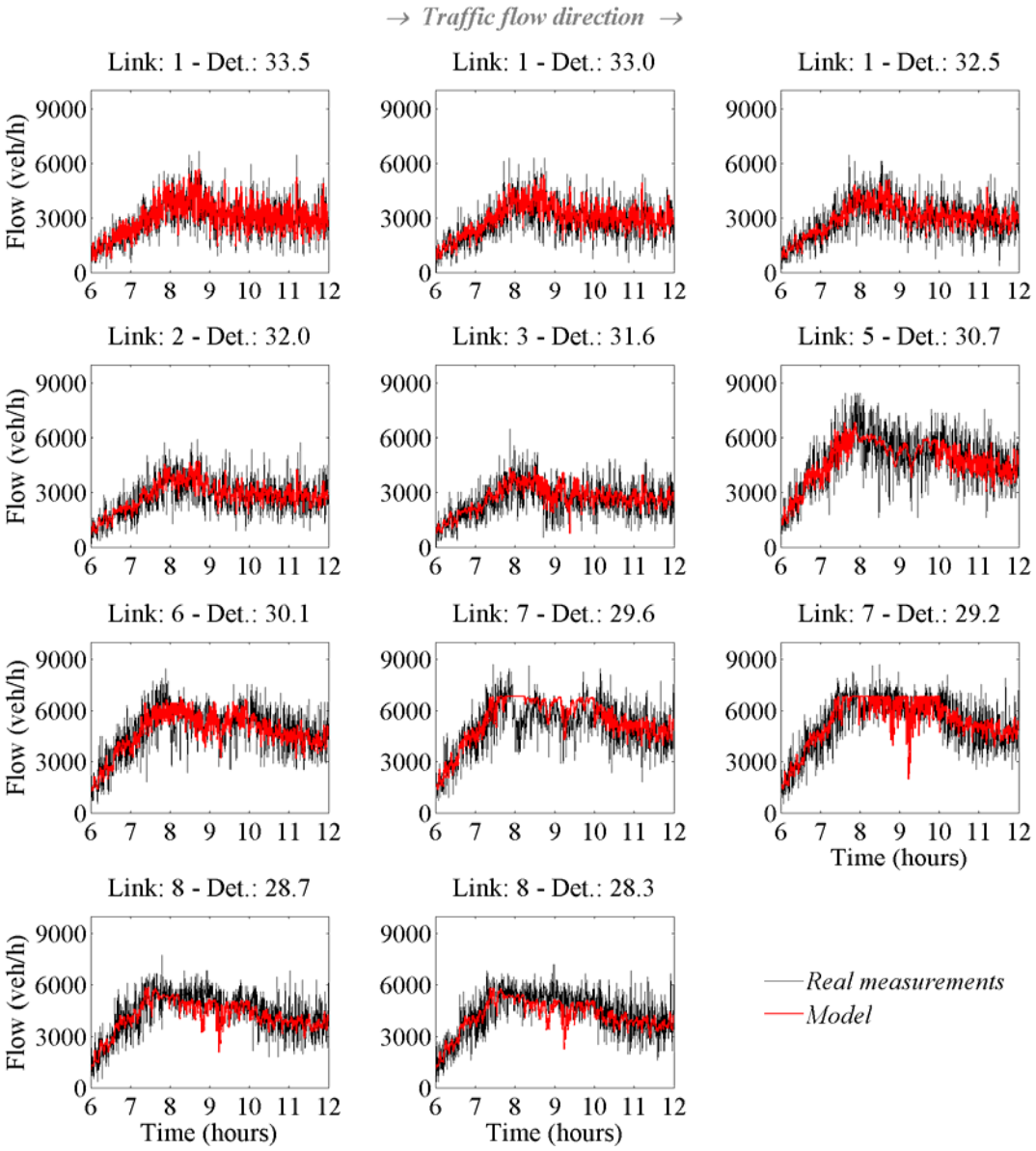

Figure 4.8 CTM model calibration: time-series of the real flow measurements and Model 1.1 estimation of flow at various detector locations for 16/06/2009.

4.3.2 METANET model calibration

As with the CTM model, METANET model was calibrated using the same three optimization methods. Figure 4.9 presents the convergence of each utilized algorithm over iterations. Again, all three methods finally converge to a low PI value, i.e. 10.1 for the Nelder-Mead algorithm, 9.8 for the genetic algorithm and 9.9 for the CE method. Table 4-3 presents the performance of all three algorithms in terms of different criteria. As with the previous model, the Nelder-Mead algorithm required a large number of iterations to converge, compared to the other two methods, albeit by requiring much less cost function evaluations, 317, compared to the genetic algorithm and the CE method that needed 26000 and 42500, respectively, which corresponds to considerably lower computation time;

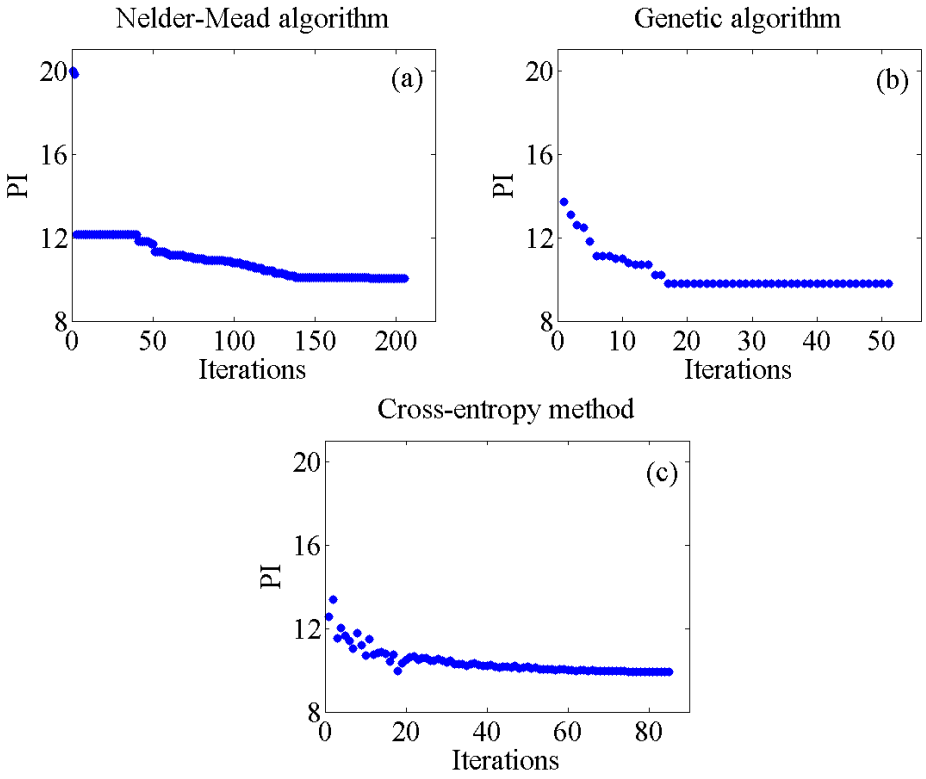

Figure 4.9 METANET model calibration: performance index value over iterations using (a) the Nelder-Mead algorithm, (b) the genetic algorithm and (c) the cross-entropy method.

Table 4-3 METANET model calibration: optimization algorithms' performance.

Optimization method	Iterations	Cost function evaluations	Computation time (min)
Nelder-Mead algorithm	204	317	0.5
Genetic algorithm	51	26000	122.9
Cross-entropy method	85	42500	197.8

Table 4-4 METANET optimal parameter values estimated by use of different optimization algorithms.

Model	METANET parameters						
	v_f (km/h)	ρ_{cr} (veh/km/lane)	a	τ (s)	ν (km^2/h)	δ (h/km)	φ (h/km)
Model 2.1 (Nelder-Mead algorithm)	117.8	35.5	1.5	18.6	24.5	1.2	1.1
Model 2.2 (genetic algorithm)	118.1	36.2	1.4	18.1	21.1	0.2	1.5
Model 2.3 (cross-entropy method)	118.8	34.4	1.5	27.2	33.1	0.5	1.0

just 0.5 min for the Nelder-Mead algorithm while the genetic algorithm and the CE method needed 122.9 and 197.8 min, respectively. Finally, comparing the genetic algorithm with the CE method it may be seen that, both algorithms achieve similar PI values with the genetic algorithm being some 75 minutes, faster than the CE method.

Table 4-4 presents the three optimal parameter sets estimated by use of the selected optimization methods. It is observed that all three algorithms estimated very similar, but not exactly the same, parameter values. Again, the specified optimal parameter values are reasonably reflecting their respective physical significance (or have similar values as in previous calibration exercises), which indicates that the basic model structure is accordingly suitable to describe real traffic phenomena. Table 4-4 also indicates the very close proximity of,

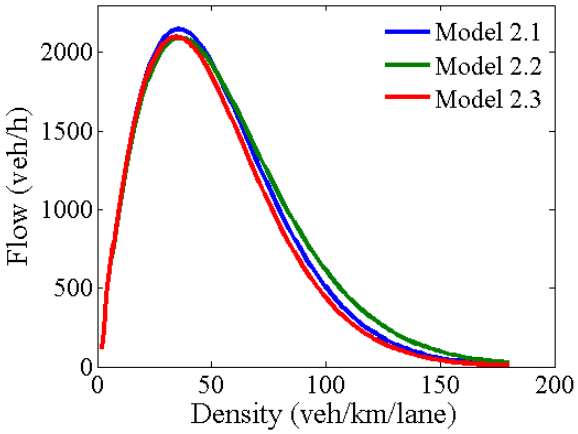

Figure 4.10 METANET model calibration: estimated fundamental diagram (FD) for all three obtained models.

particularly, the optimal parameter values which are involved in the fundamental diagram (FD) (3.10); Figure 4.10 traces the 3 respective FDs, indicating that they are virtually identical. It should also be noted that the capacity that corresponds to the estimated fundamental diagram parameters for all three models is around to 2100 veh/h which is close to the capacity estimated by all three CTM models, which is about 2250veh/h. Regarding the parameters ν and τ, it is known from previous model validation work (e.g. [13]) that the calibration PI features low sensitivity around the optimum if the parameters ν and τ are changing values simultaneously. This is confirmed with the results of Table 4-4, where the ratio ν/τ may be calculated to be 1.32, 1.17, 1.22 for the three respective optimization methods, despite the stronger deviation of the underlying absolute parameter values. More observations on the sensitivity of the model parameters maybe found in Section 4.4.

Figure 4.11 displays the space-time diagram of the real speed measurements for the calibration date (16/06/2009) and the corresponding speed estimations of all three produced models. It may be seen that all three models are able to reproduce the traffic phenomena on this freeway stretch, for this particular day, sufficiently.

In more detail now, Figure 4.12 and Figure 4.13 present the time-series of the real speed and flow measurements and the corresponding model estimations, using Model 2.1. It is observed, in both figures, that the model calculations are very close to the real traffic data. In addition, METANET is able to reflect more realistically the vehicle acceleration downstream of the congestion creation area (29th km), since this model acknowledges the limited acceleration ability of vehicles (see Figure 4.12). The corresponding time-series of the speed and flow estimations of Model 2.2 and Model 2.3 may be found in the Appendix (Figure A.5 – Figure A.8).

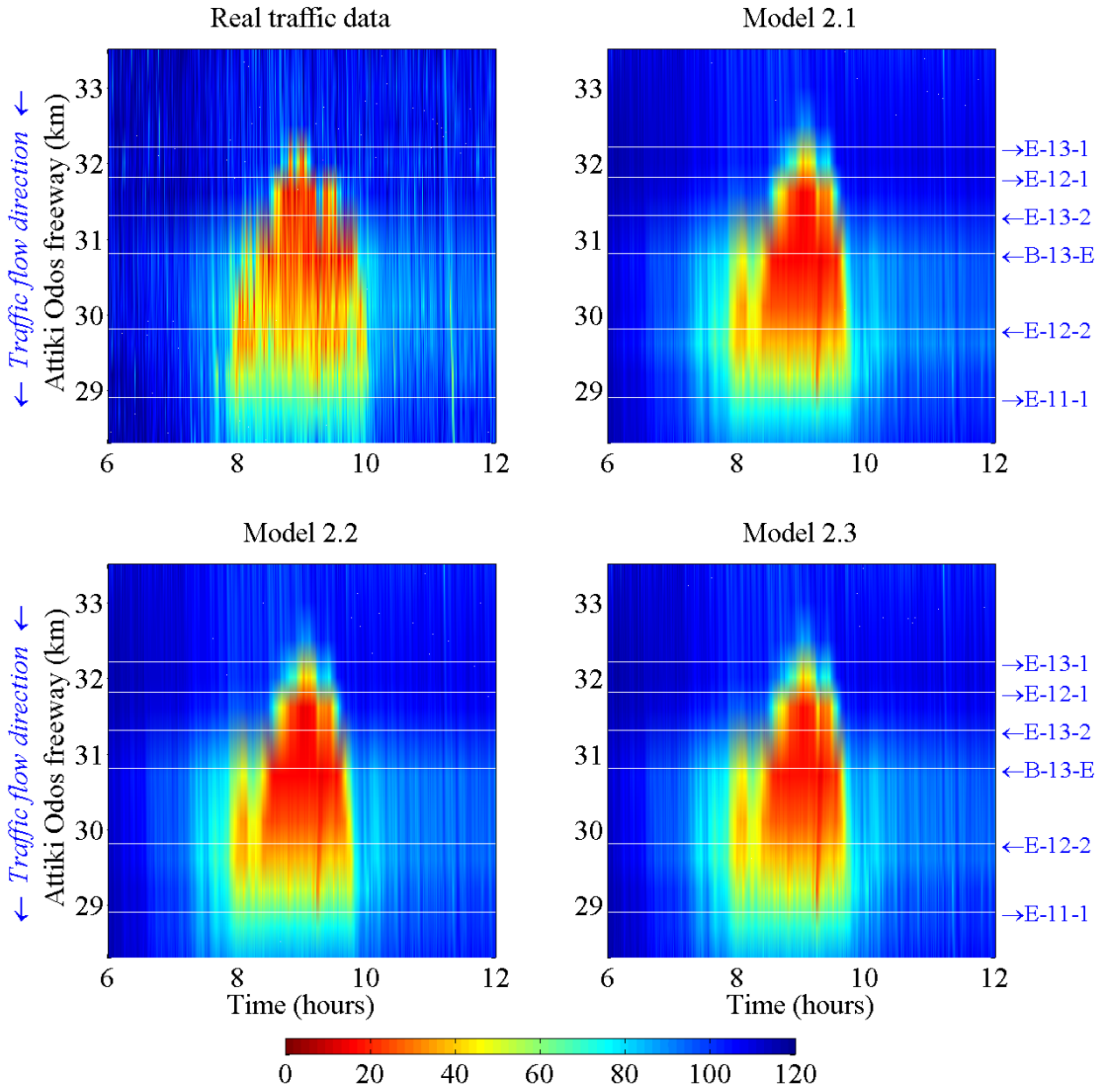

Figure 4.11 METANET model calibration: space-time diagrams of measured speeds and the models' estimation of speed for 16/06/2009.

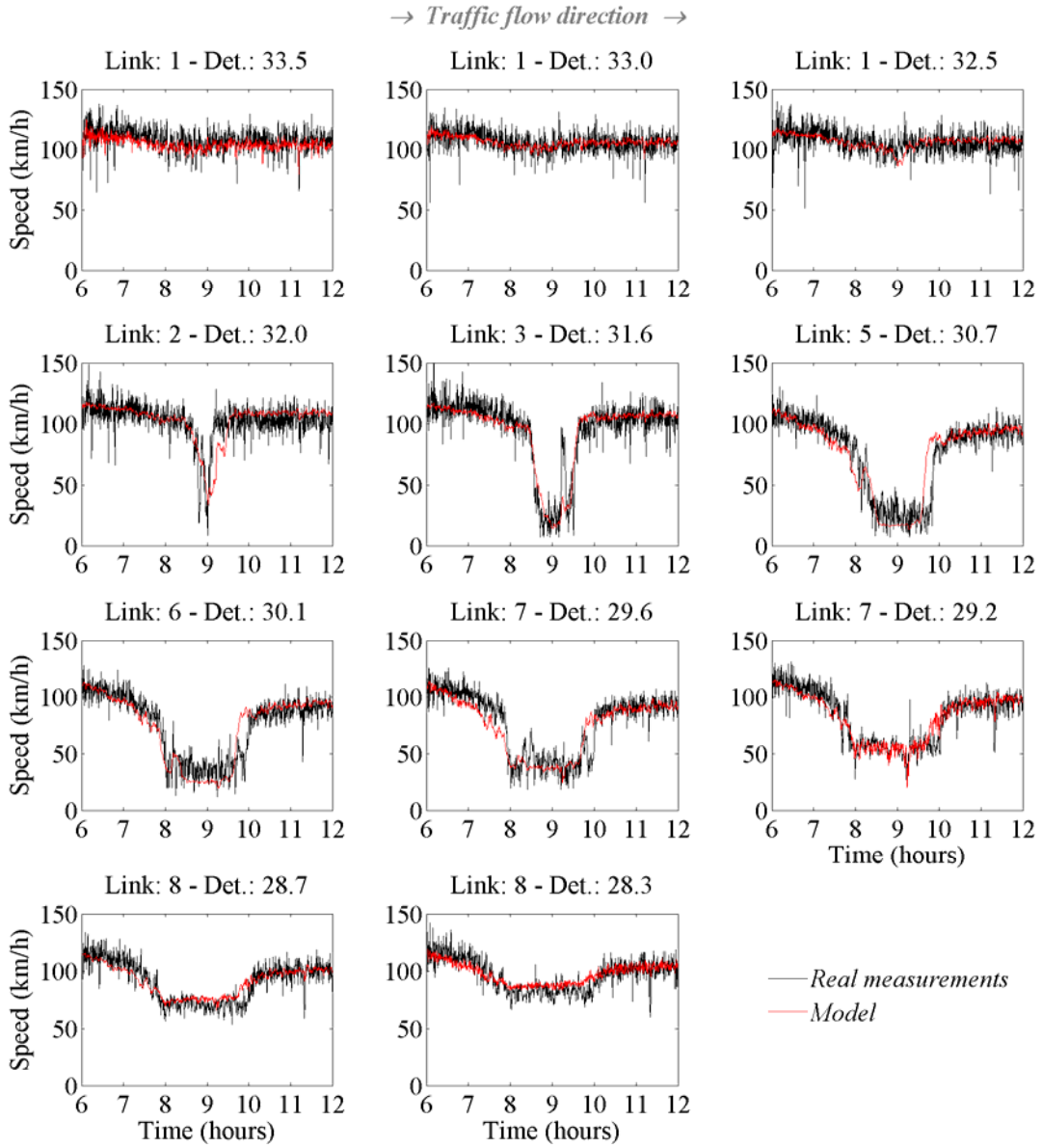

Figure 4.12 METANET model calibration: time-series of the real speed measurements and Model 2.1 estimation of speed at various detector locations for 16/06/2009.

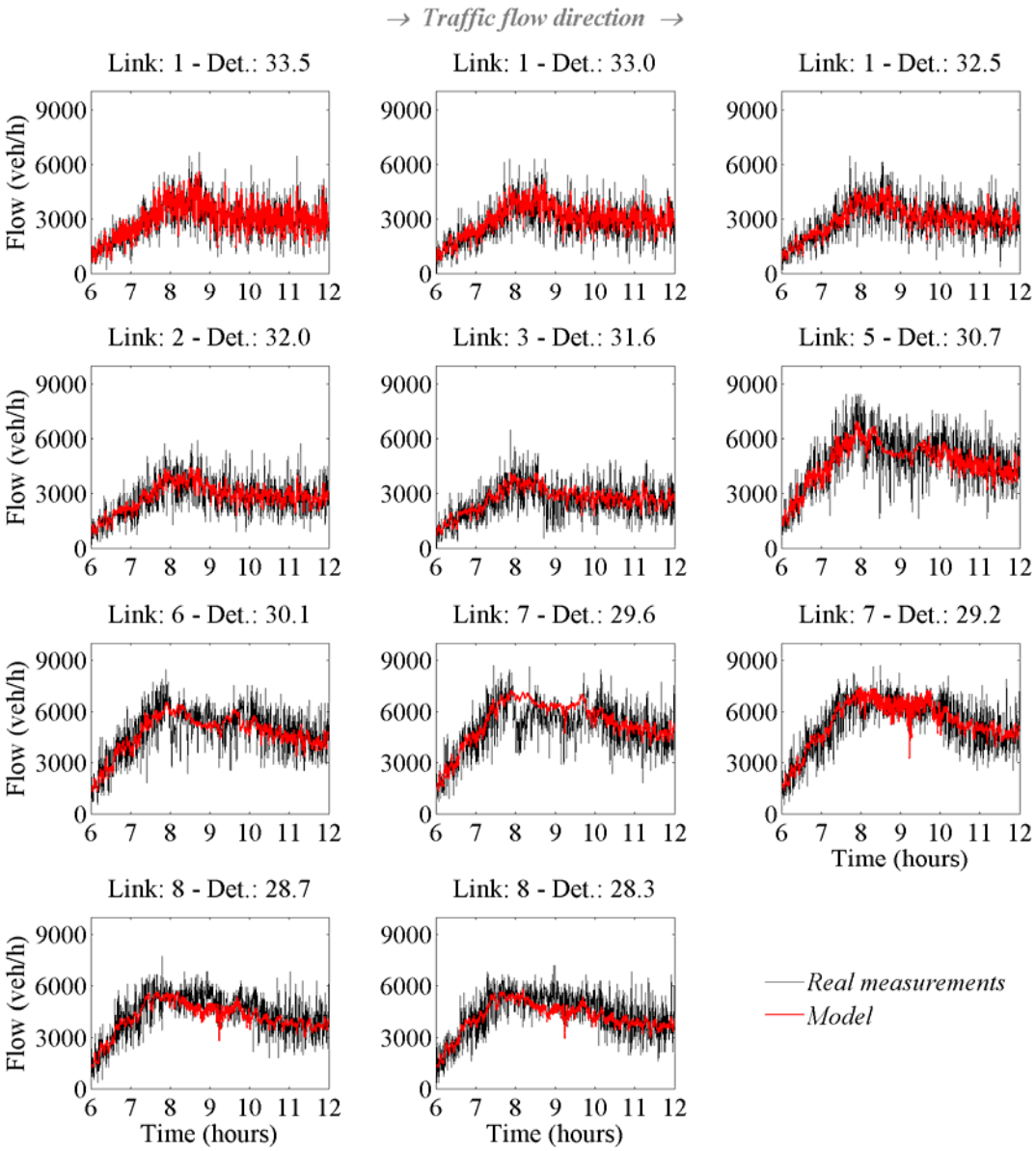

Figure 4.13 METANET model calibration: time-series of the real flow measurements and Model 2.1 estimation of flow at various detector locations for 16/06/2009.

4.4 Sensitivity investigations and validation

After performing a model calibration exercise, the produced models should be tested in terms of effectiveness and robustness. To this end, Section 4.4.1, presents some sensitivity investigations that were carried out, which aim to identify the sensitivity of the models to changes at their parameter values. Then, in Section 4.4.2 the models are validated, i.e. are applied to the same network using traffic data from different days, to ensure that they can reliably reproduce the typical traffic conditions of the examined freeway site.

4.4.1 Sensitivity investigations

Sensitivity analysis is useful to test how the variation of the model parameter values may affect the effectiveness of the models. In the following, the sensitivity investigations for Model 1.1 (CTM) are presented first, followed by the corresponding investigations for Model 2.1 (METANET).

Figure 4.14 presents the sensitivity diagrams of Model 1.1, in terms of PI value,

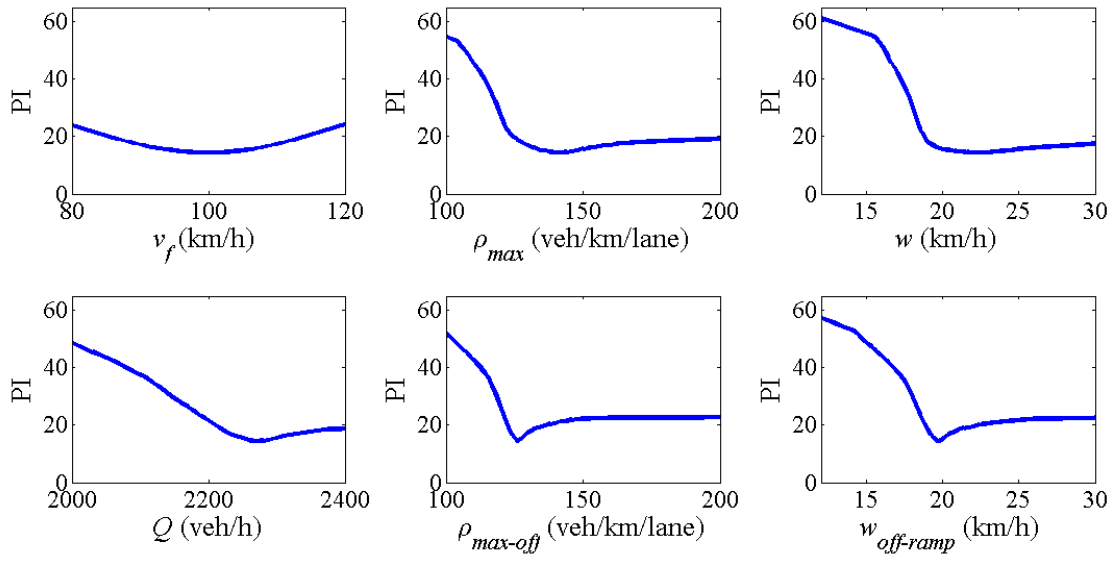

Figure 4.14 CTM model (Model 1.1) sensitivity to changes of the model parameters, in terms of PI value.

obtained by changing one parameter at a time, within a feasible range, while keeping the rest parameter values fixed and equal to the optimal values. As an example, in the first diagram various simulation runs of Model 1.1 were carried out using different values of the parameter v_f, within the range [80 120], and the corresponding PI values were obtained while the rest model parameters were kept equal to their nominal values included in Table 4-2. Figure 4.14 illustrates that the most sensitive parameters of Model 1.1 are the maximum density at the off-ramp E-11-1, $\rho_{max\ off-ramp}$, and the corresponding congestion wave speed, $w_{off-ramp}$. This fact was rather expected as, within this model, a saturated off-ramp may spill-back and create mainstream congestion thanks to (3.6) which includes these two parameters.

Extra sensitivity investigations were also carried out by changing two model parameters at a time while considering the rest parameter values fixed and equal to the optimal values. In this way, the dependence and correlation between the model parameters can be observed. Figure 4.15 includes the corresponding diagrams for all couples of model parameters. As an example, the first diagram of Figure 4.15 presents the correlation between the congestion wave speed w and the maximum density ρ_{max}. It is observed that these two parameters are correlated as different values of parameter w can lead to equally low PI values as long as the parameter ρ_{max} also changes value. Figure 4.15 shows that the most correlated parameters are the $w_{off-ramp}$ and $\rho_{max\ off-ramp}$ as a small change to the value the first parameter requires an immediate change to the value of the second parameter, so as to maintain the low PI value. Moreover, the parameter v_f does not seem very sensitive as it may give equally good results for a big range of values around 100 km/h. Finally, parameter Q is mostly sensitive to changes of the parameters $w_{off-ramp}$ and $\rho_{max\ off-ramp}$ as well as the parameters w and ρ_{max}.

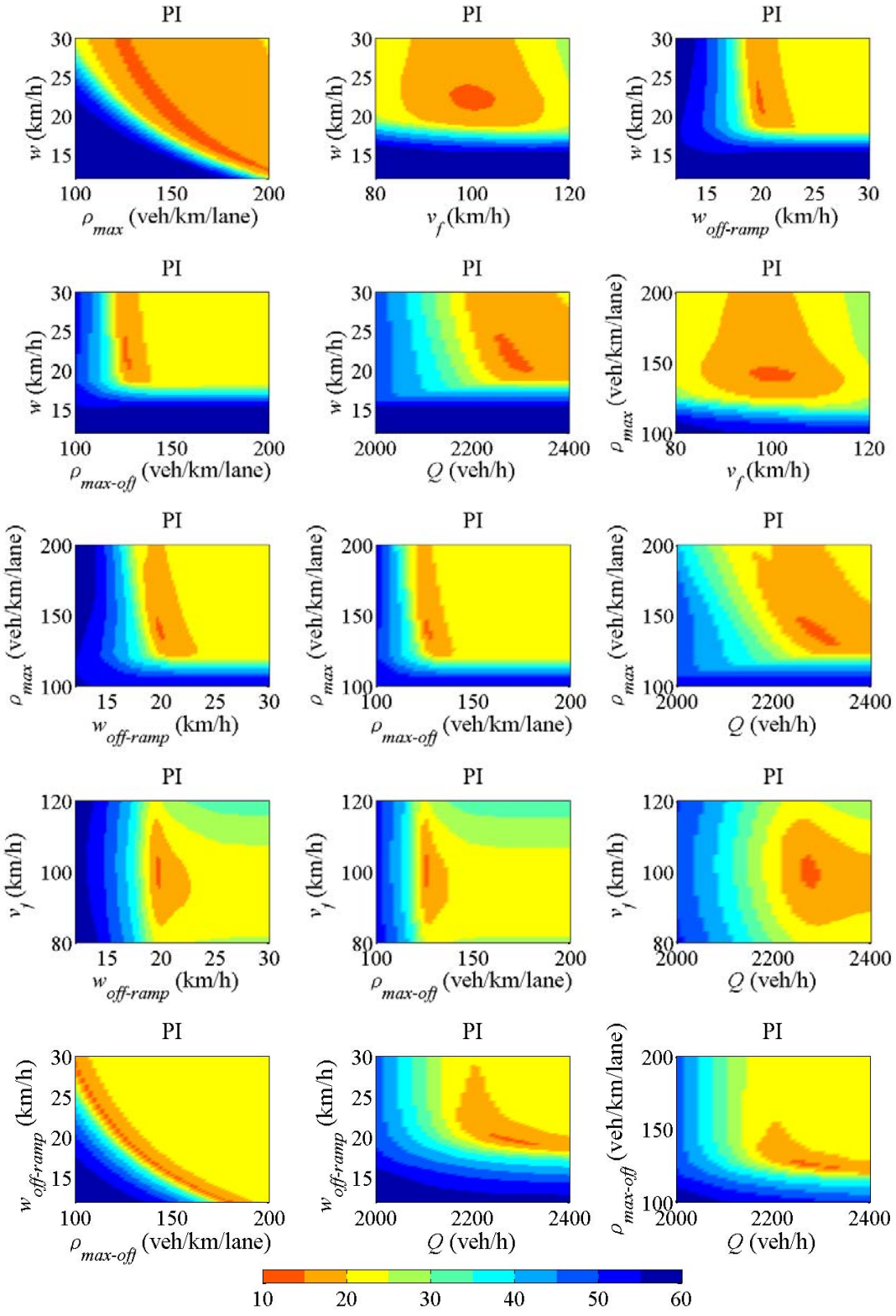

Figure 4.15 CTM model (Model 1.1): sensitivity to changes of the model parameters, in terms of PI value.

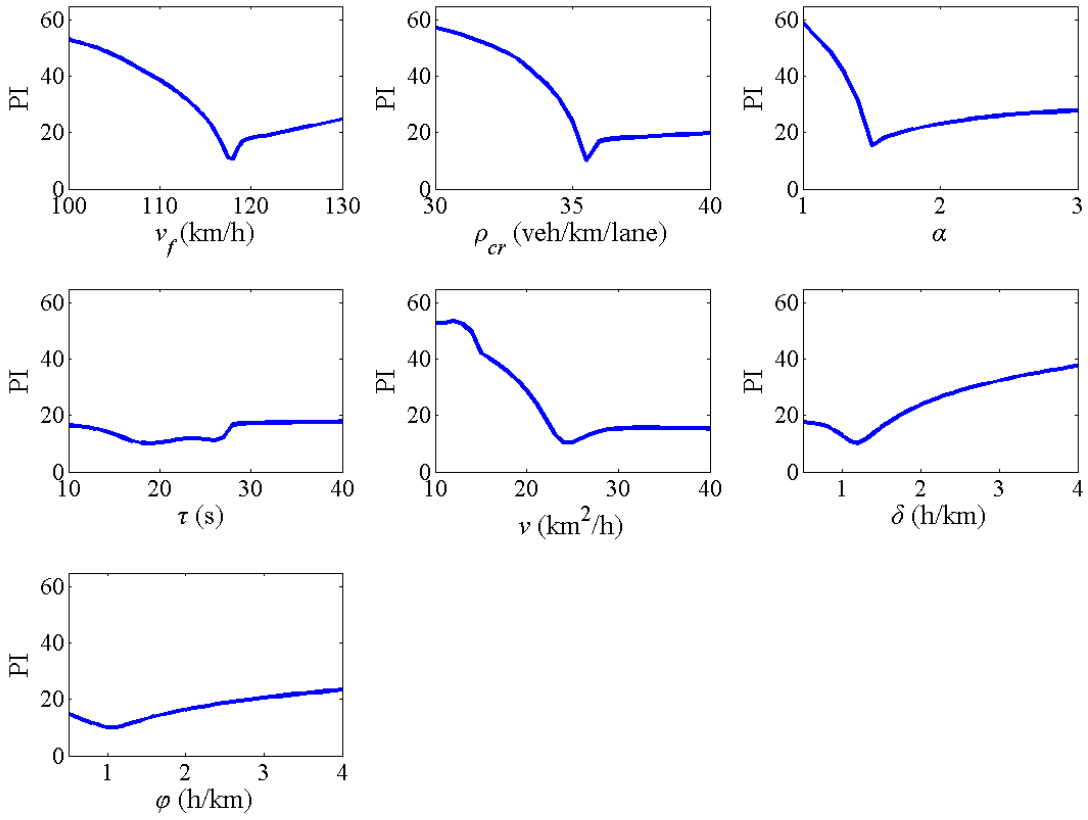

Figure 4.16 METANET model (Model 2.1): sensitivity to changes of the model parameter, in terms of PI value.

Similar investigations were carried out using the METANET model (Model 2.1). Figure 4.16 presents the sensitivity diagrams of Model 2.1, in terms of PI value, for all model parameters. It is observed that in this model the most sensitive parameters are the parameters of the fundamental diagram (FD), i.e. the parameters v_f, ρ_{cr} and α. Extra sensitivity investigations between the model parameters were also carried out for Model 2.1 and Figure 4.17 displays the obtained results. It is seen here that the parameters of the FD are strongly correlated thus they are sensitive to small changes in their values. Moreover, it is observed that the parameters τ and ν are also correlated as well as the parameter δ with all three FD parameters plus parameter φ.

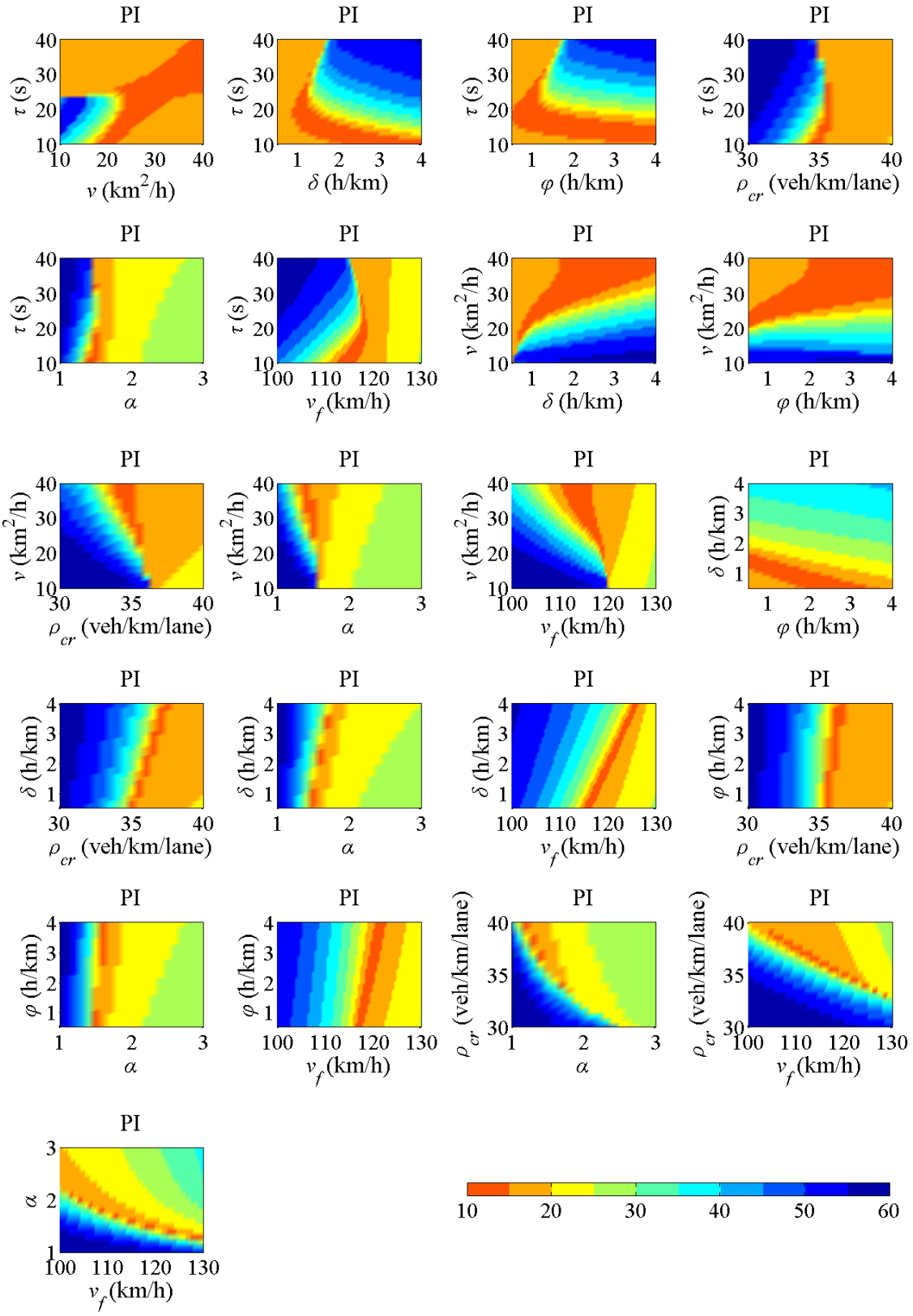

Figure 4.17 METANET model (Model 2.1) sensitivity to changes of the model parameters, in terms of PI value.

4.4.2 Validation and comparison

The resulting traffic flow models should reflect reliably the traffic characteristics of the considered network, thus they should be able to reproduce its typical traffic conditions. In order to test the accuracy and robustness of the produced models, the models are validated, i.e. are applied using different traffic data sets (from the same freeway stretch) than the one used for their calibration. To this end, the models were applied using traffic data from 26/05/2009, 23/06/2009 and 25/06/2009.

Table 4-5 presents the validation results in terms of PI values for all six models and all utilized traffic data sets. In particular Table 4-5 shows, that all three CTM models, Model 1.1, Model 1.2 and Model 1.3, achieve an average PI value, over all dates, equal to 16.1, 16.8 and 16.2, respectively, while all three METANET models achieve average PI value equal to 10.8. Moreover, Figure 4.18 and Figure 4.19 present the space-time diagrams of the real measured speeds and the corresponding models' estimation of speed for all considered dates. It is observed that all models are able to reproduce the traffic conditions of other days with sufficient accuracy, creating the congestion at the right time period and for the right space-time extent; although the CTM models are seen to produce slightly longer congestion duration for some dates. However all six models achieve low PI values for both the calibration and the validation dates.

Table 4-5 Validation results: performance index value for all models and all dates.

Model		Validation results (PI)				
		16/06/2009	26/05/2009	23/06/2009	25/06/2009	Average
CTM	Model 1.1	14.4	18.9	16.4	14.8	**16.1**
	Model 1.2	14.7	19.7	16.8	16.3	**16.8**
	Model 1.3	14.4	19.0	16.3	14.9	**16.2**
METANET	Model 2.1	10.1	12.1	12.4	8.4	**10.8**
	Model 2.2	9.8	12.3	11.8	9.1	**10.8**
	Model 2.3	9.9	12.6	12.4	8.3	**10.8**

Table 4-6 Validation results: performance index value for the whole network and per network link for the CTM model (Model 1.1) and the METANET model (Model 2.1) for all investigated dates.

Model	Date	Validation results (PI)								Flow RMSE
		L1-L8	L1	L2	L3	L5	L6	L7	L8	L1-L8
CTM (Model 1.1)	16/06/09	14.4	10.7	13.5	13.3	14.5	13.6	16.9	17.4	713.2
	26/05/09	18.9	17.7	19.8	22.5	23.7	18.7	18.8	15.7	759.8
	23/06/09	16.4	17.7	17.7	18.1	17.6	16.2	13.5	14.6	758.3
	25/06/09	14.9	10.3	13.2	18.8	19.9	17.4	15.0	14.7	707.8
Average		**16.1**	**14.1**	**16.0**	**18.2**	**18.9**	**16.5**	**16.1**	**15.6**	**734.8**
METANET (Model 2.1)	16/06/09	10.1	7.3	12.0	11.4	13.6	12.4	10.3	8.1	695.2
	26/05/09	12.1	15.3	14.7	13.3	11.8	11.0	9.1	6.9	709.6
	23/06/09	12.4	13.9	15.5	14.2	14.3	12.6	10.1	7.2	729.9
	25/06/09	8.4	6.9	9.6	10.0	9.1	9.6	9.5	6.7	684.7
Average		**10.8**	**10.8**	**12.9**	**12.2**	**12.2**	**11.4**	**9.8**	**7.2**	**704.9**
Comparison (%)		**-33**	**-23**	**-19**	**-33**	**-35**	**-31**	**-39**	**-54**	**-4**

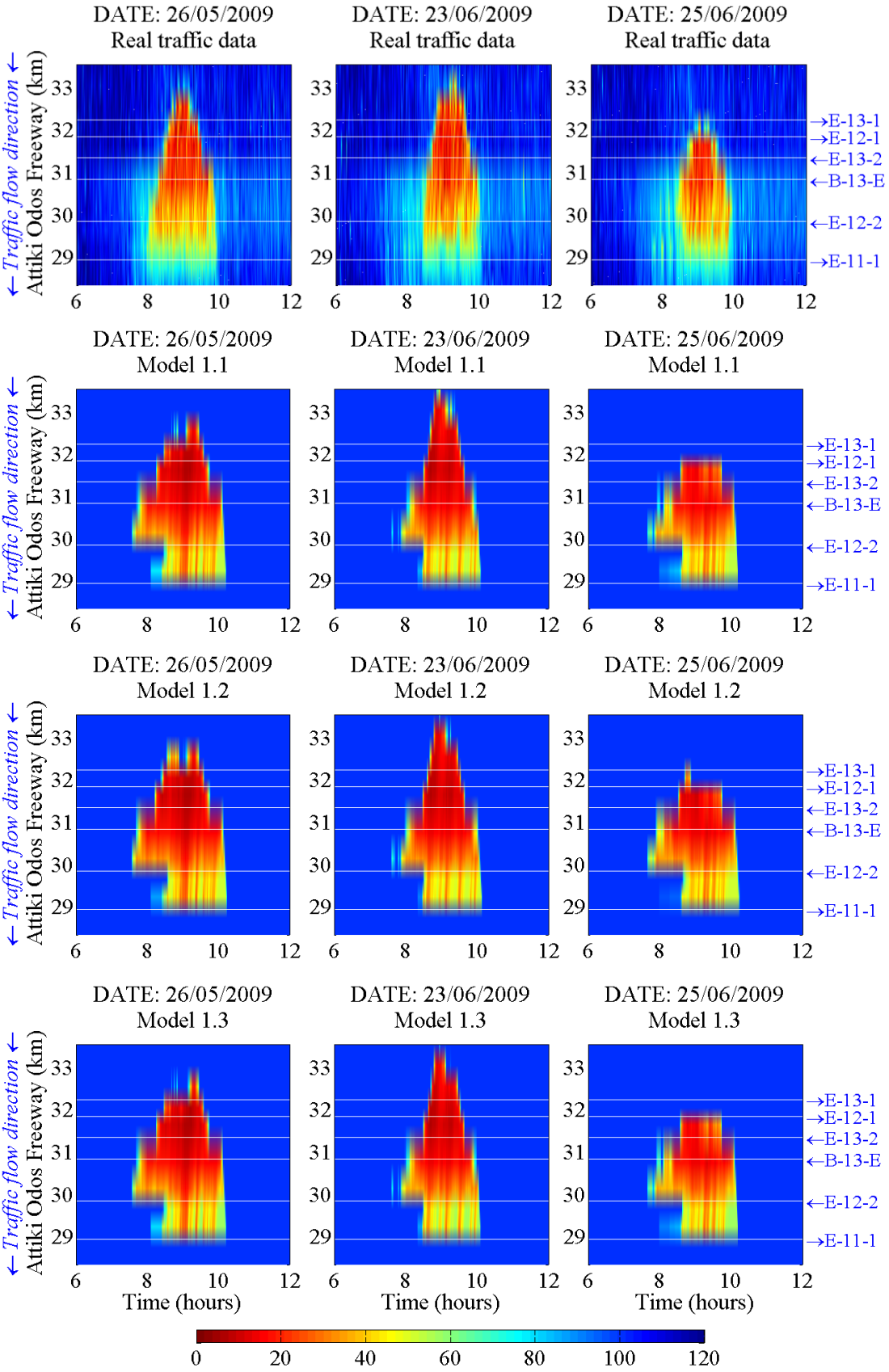

Figure 4.18 CTM model validation: space-time diagrams of measured speeds and the models' estimation for 26/05/2009, 23/06/2009 and 25/06/2009.

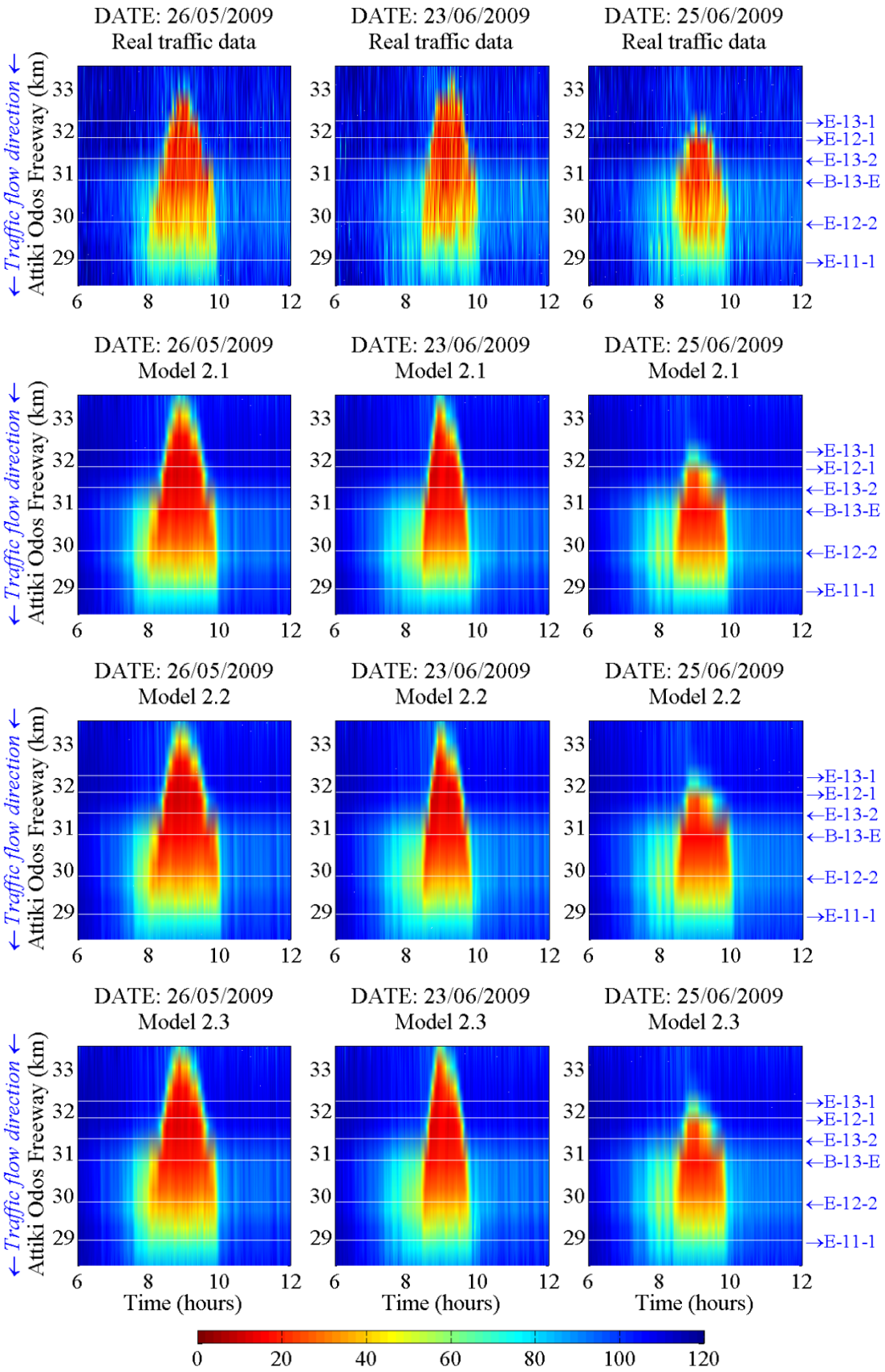

Figure 4.19 METANET model validation: space-time diagrams of measured speeds and the models' estimation for 26/05/2009, 23/06/2009 and 25/06/2009.

In more detail now, Figure 4.20 and Figure 4.21 illustrate the time-series of the real and estimated speeds of Model 1.1 (CTM) and Model 2.1 (METANET) for one particular validation date (23/06/2009). It is shown here that both models may replicate the real traffic conditions of other days with sufficient accuracy. Table 4-6 includes the validation results for Model 1.1 and Model 2.1 with regard to the achieved PI values, calculated for the whole network and for each freeway link separately, for the calibration and the validation dates. Note that there are no available data for link 4; while for links with multiple detector stations, a corresponding average PI value is displayed in Table 4-6. It is observed that METANET (Model 2.1) achieves lower PI values comparing to the CTM model (Model 1.1) for all utilized dates. In particular, it is seen that METANET acquires lower PI for all freeway links, and, especially, for link L8, where it achieves 54% lower PI compared to CTM. The main reason for this is that METANET takes into account the limited acceleration of vehicles downstream of the congestion head; while CTM predicts free speeds in the areas downstream of the congestion. Moreover, CTM creates time-longer congestion at links L3, L5 and L6 resulting to higher average PI values at the corresponding links, 33%, 35% and 31% higher than METANET. Finally, with respect to link L7, which is the congestion creation link, METANET is again 39% more accurate than CTM regarding the estimation of speed. Finally, the last column of Table 4-6 presents the root-mean-square error (RMSE) for the corresponding flow measurements calculated for the whole network and for each investigated date. It is shown that both models achieve similar accuracy in the estimation of flows with METANET being 4% more accurate than CTM model. Similar results are obtained for the rest traffic models.

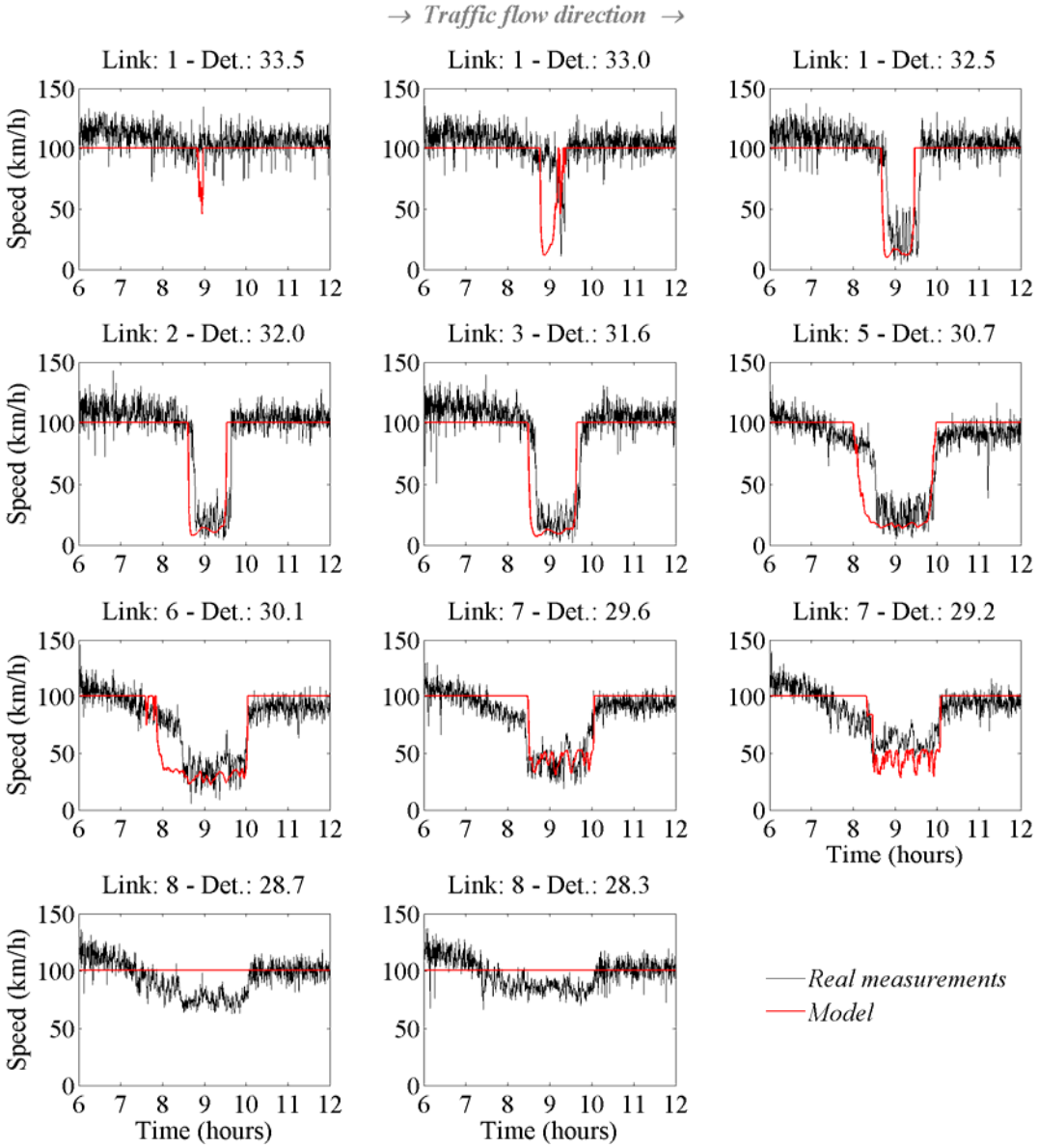

Figure 4.20 CTM model validation: time-series of the real speed measurements and Model 1.1 estimation of speed at various detector locations for 23/06/2009.

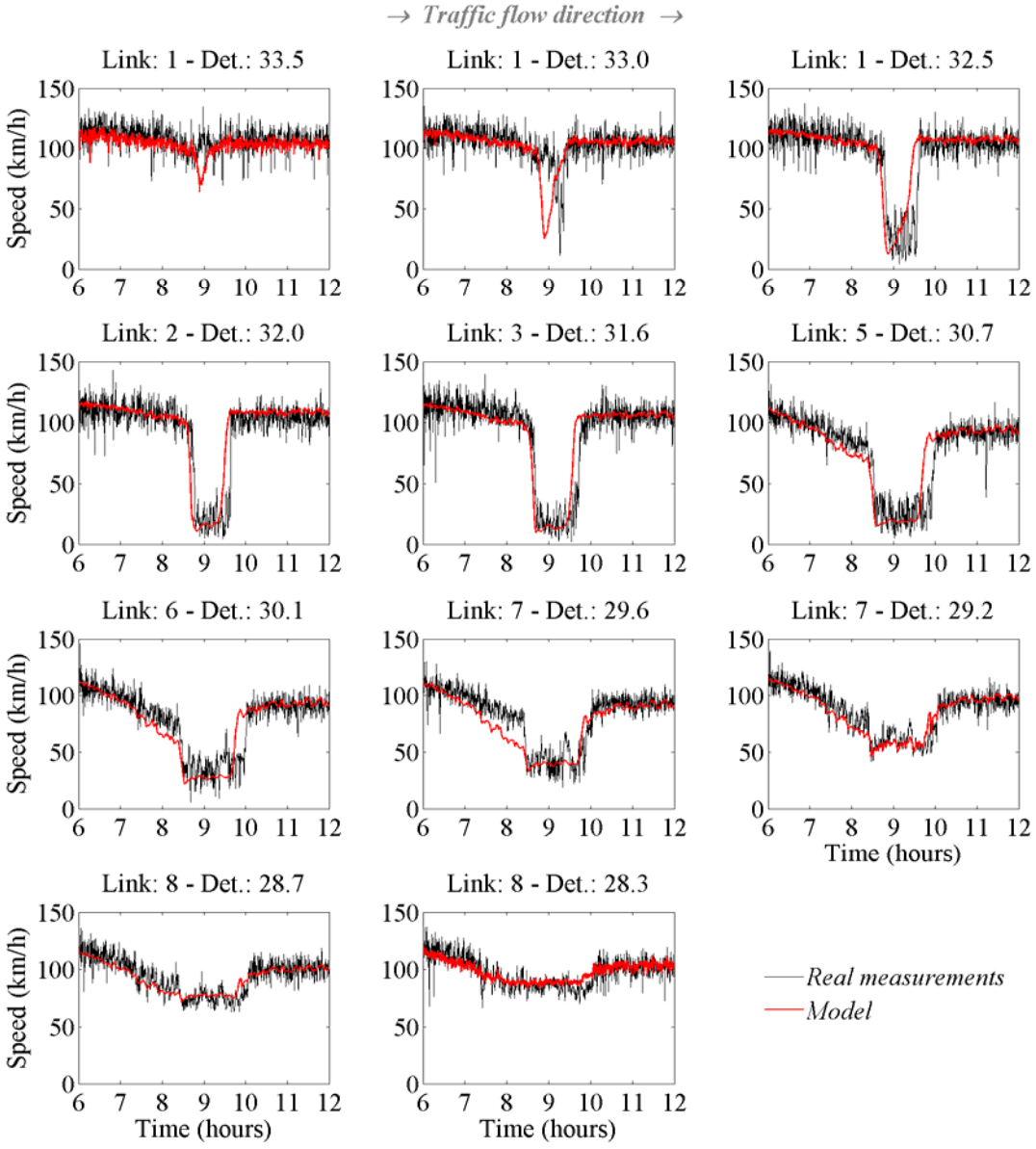

Figure 4.21 METANET model validation: time-series of the real speed measurements and Model 2.1 estimation of speed at various detector locations for 23/06/2009.

4.5 Conclusions and remarks

Two space-time discrete macroscopic traffic flow models, the first-order model CTM and the second-order model METANET, were compared regarding the representation of traffic congestion created due to a saturated freeway off-ramp. The models were first calibrated using real traffic data from Attiki Odos freeway in Athens, and by employing three different optimization methods; i.e. the deterministic Nelder-Mead algorithm, the stochastic genetic algorithm and the stochastic cross-entropy method. Then, the resulted models were tested in terms of sensitivity to their parameter values and were also validated and compared using different traffic data sets from the same freeway site.

The calibration results showed that all three optimization methods estimated similar parameter values for the CTM and also the METANET model, which achieve a satisfactory reproduction of the network traffic conditions for the calibration date. Moreover the sensitivity investigations showed that the CTM model is sensitive to the parameters related to the characteristics of the saturated off-ramp while METANET model seems to be more sensitive to the parameters of the fundamental diagram (FD). Finally, the validation of the CTM and METANET model indicated that they are both able to reproduce the traffic conditions of this particular network also for other dates, with METANET model offering a more accurate representation of the prevailing traffic conditions.

Next chapter presents innovative real-time traffic control measures for cases were freeway congestion is created due to saturated off-ramps.

5 Real-Time Traffic Control Measures for Congested Freeway Off-Ramp Areas

This chapter develops, investigates and demonstrates real-time traffic control strategies to mitigate the problem of freeway congestion due to saturated off-ramps. As already indicated in Chapter 2, this is a particular type of congestion and different freeway sites may call for different traffic control measures, depending on the network layout, the prevailing traffic conditions, the expected drivers compliance, the available traffic control equipment etc. As a result each situation should be viewed as a particular case and the corresponding network characteristics should be taken into account during the development of traffic control strategies.

Within this thesis two different cases are examined and suitable real-time traffic control measures are proposed. In the first case, presented in Section 5.1, it is considered that freeway congestion is created due to the limited capacity of an off-ramp which results in the off-ramp queue spill back into the freeway mainstream, during the peak hours. In this examined case, the application of various route diversion policies is proposed and demonstrated by use of macroscopic simulation. In the second case, examined in Section 5.2, recurrent freeway congestion is created due to congestion on a surface street network which propagates to the freeway mainstream through a saturated off-ramp. In this examined case, a real network is simulated by use of microscopic simulation and a real-time merging traffic control algorithm is proposed that aims to maximize the surface street network throughput and at the same time to prevent the off-ramp queue spill over into the freeway mainstream.

5.1 Case 1: Real-time route diversion control

This section proposes real-time route diversion policies in cases where recurrent freeway traffic congestion is created due to a saturated off-ramp. In particular, the proposed route diversion policies attempt to avoid the off-ramp queue spill-over onto the freeway mainstream and the resulting freeway congestion by appropriately re-routing the freeway vehicles through nearby off-ramps towards the same destination. The proposed route diversion concepts are based on feedback control laws and are tested for a hypothetical, but quite typical, network infrastructure and several traffic scenarios by use of macroscopic simulation. Based on the results of Chapter 4, the METANET model is selected for the simulation investigations. The obtained simulation results are compared to the case where no route guidance is applied to the network and reveal interesting features and the potential for significant improvements.

In particular, Section 5.1.1 presents the proposed route diversion concepts and the control strategies employed to enable the rerouting decisions. Section 5.1.2 describes the hypothetical network and the traffic conditions considered for the investigations. Section 5.1.3 presents the simulation results for all investigated traffic scenarios and, finally, Section 5.1.4 concludes with the main remarks and finding of the study.

5.1.1 Dynamic route diversion concept

Route guidance systems aim to provide the drivers with information or guidance related to their route choice decisions in case of non-recurrent events, e.g. incidents, but also in cases of low-reliability recurrent congestion conditions. For example, the system may guide, in real-time, the drivers through alternative routes during maintenance works, as proposed in [47], or in case of incidents as suggested in [48]. This study investigates the application of route diversion

measures aiming at avoiding the formation of recurrent freeway congestion due to an over-spilling off-ramp.

5.1.1.1 Problem description

Consider a freeway stretch and a bifurcation point (such as N1 in Figure 5.1 on page 96) for two alternative routes, a primary, which is the distance-shorter route, and a secondary, towards the same destination (e.g. D3 in Figure 5.1). Each alternative route includes a different freeway off-ramp that a vehicle may use in order to reach the desired destination. Assuming that the primary-route off-ramp (e.g. L5 in Figure 5.1) has limited capacity, which is not possible to be increased, recurrent traffic congestion may appear on the mainstream, during the peak hours, due to the saturated off-ramp. The concept of this study is to divert a portion of the vehicles through the secondary route, when and to the extent needed, in order to protect the saturated off-ramp and avoid mainstream congestion due to queue spill-back.

Since drivers are free to ignore messages that they perceive incompatible with their own criteria, the objective of the route guidance system cannot be simply based on the system-optimal conditions, but must mainly target user-optimal conditions, i.e. suggest the alternative route only if this route is equivalent or time-shorter than the primary route. Thus, three different cases arise, depending on the network topology and traffic conditions. In the first case, the user-optimal conditions may be achieved before the off-ramp queue spill-over and creation of mainstream congestion; thus the route guidance system may propose an alternative route without any disbenefit for the compliant drivers. In the second case, the user-optimal conditions are achieved only after the off-ramp queue spills back to the freeway mainstream; thus the route diversion system will have to consider the expected compliance to the proposed route choice or be based on mandatory actions, such as temporary off-ramp closures. Finally, in the third case,

the user-optimal conditions may not be achievable, due to the prevailing traffic conditions (e.g. congestion) on the surface network. In this case, the expected compliance of the drivers to the proposed route indications would be low, thus the route diversion system should decide for the temporary off-ramp closure when and to the extent needed. In the following sections, all cases are examined and various real-time route diversion strategies are proposed for each investigated case.

5.1.1.2 Case 1: User-optimal conditions may be achieved before the off-ramp queue spill-over

As mentioned above, in the first examined case the user-optimal conditions may be achieved without the creation of mainstream congestion, which means that, during peak hours, the secondary route becomes time-shorter, before the queue on the primary-route off-ramp spills back to the mainstream. However, due to the inherent randomness in the traffic demand patterns and traffic flow behavior, particularly under saturated traffic conditions, the spill-over phenomenon at a specific ramp may have low reliability, i.e. occur at different times on different days, and perhaps even not occur at all on some days. Under these conditions, the drivers may not be able to make the best route decisions based only on their own past experiences, but they may need real-time information, e.g. from the route guidance system. Based on these assumptions, two alternative policies are proposed, which a route guidance system may utilize in order to divert vehicles towards competitive routes. The first policy bases its decisions on real-time estimations of the (reactive) travel time for the two alternative routes; while the second policy uses the estimated queue length on the primary-route off-ramp as the pertinent real-time information for deciding on driver diversion. The corresponding control strategies actually decide on the percentage of vehicles that should be re-routed, and may convey the information to the drivers either

through VMS or through vehicle-to-infrastructure communication. The two policies are described in more detail in the following.

Dynamic route diversion based on reactive travel time estimation

This policy may be utilized, in a potential field implementation, when reasonably accurate real-time estimations of instantaneous (or reactive) travel times along the alternative routes are available; this implies the availability of a sufficient number of detectors or of a sufficient sample of probe vehicles along both alternative routes. Instantaneous (or reactive) travel time $\tau(k)$ is the travel time needed by a vehicle to drive along the route, if the mean speeds along the route are frozen to the values they have at time k, which is the time at which the vehicle starts its trip on the route. At this point, the reader is referred to [49], [50] for more details on feedback-based route guidance notions that are employed in the following sections. Feedback route guidance is a simple but efficient method for route diversion; the dynamics of the corresponding closed-loop system are fast, but may gradually deteriorate if the involved travel times become excessively long. Related field applications of feedback-based route guidance may be found in [51], [52], [53], [54].

For the present application, the splitting rate $\beta \in [0,\ 1]$, i.e. the percentage of the drivers that should follow the primary route, may be calculated by a variety of feedback control strategies. This real-time decision is taken at every period (or control interval) T (typically equal to 1-6 min), and $\beta(k)$ denotes the splitting rate to be applied during $[kT, (k+1)T]$, where $k = 1,2, \dots, K$, is the discrete time index, and K is the time horizon. The feedback regulators attempt to keep the travel time difference $\Delta\tau(k)$ on the two alternative routes close to zero; where $\Delta\tau(k) = \tau_s(k) - \tau_p(k)$, with $\tau_s(k)$, $\tau_p(k)$ being the estimated (reactive) travel times through the secondary and primary route, respectively. Since in this study

only two alternative routes are considered, the percentage of vehicles that should follow the secondary route is simply $1-\beta$.

There are several ways for implementing the desired splitting rates $\beta(k)$ in practice. To start with, a first possibility, which does not require any of the feedback regulators discussed below, is to simply display both estimated travel times on a VMS, which is positioned just upstream of the bifurcation point; trusting that the drivers' routing behavior in view of this information will be sufficient to equalize the travel times on both alternative routes without further ado. A second possibility arises if the splitting rate $\beta(k)$ only takes the values 0 and 1 (as in the bang-bang controller presented below); then the VMS may explicitly propose the corresponding directions to the drivers. On the other hand, if the splitting rate $\beta(k)$ can take any value within the range [0,1], then a pulse-modulation technique may be used; whereby the VMS proposes the main route for a duration $\beta(k)T$ of the corresponding period; and the alternative route for the remaining duration $[1-\beta(k)]T$. Finally, if a communication system with the vehicles is in place, then it is possible to split the vehicles at will, via corresponding individual messages.

If a bang-bang (or on-off or all-or-nothing) controller is employed, then, at each control interval k, the splitting rate is calculated as follows

$$\beta(k)=\begin{cases}1 & if\ \Delta\tau(k)\geq 0\\ 0 & otherwise\end{cases} \tag{5.1}$$

The main advantage of the bang-bang strategy is its simplicity, since no regulation parameters need to be specified. However, due to its switching nature, it creates oscillations of the travel time difference $\Delta\tau(k)$. Since the amplitude of the oscillations depends on the control time step, an appropriate trade-off should be established in particular applications.

Another feedback strategy, that may be employed, is a PI-type (proportional-integral) controller, which calculates, at each control interval k, the splitting rate as follows

$$\beta(k) = \beta(k-1) + K_p[\Delta\tau(k) - \Delta\tau(k-1)] + K_i\Delta\tau(k) \qquad (5.2)$$

The calculated splitting rate β is truncated if it exceeds the range $[0,\ 1]$. The PI-strategy includes two parameter values that should be specified, K_p and K_i; and it leads to smoother trajectories for the splitting rates and travel time differences compared to the bang-bang strategy.

Dynamic route diversion based on off-ramp queue length estimation

This policy makes use of real-time estimation of the queue length on the off-ramp of interest; thus, in contrast to the previous policy, it calls for fewer detector measurements. Also in this case, the splitting rate β towards the primary route may be calculated by a variety of feedback control strategies, which attempt to maintain the estimated queue length on the primary off-ramp, w, close to a pre-specified level. As for the previous policy, the splitting rate towards the secondary route is simply calculated as $1 - \beta$.

If a bang-bang strategy is employed then, at each control interval k, the splitting rate β towards the primary route is calculated as follows

$$\beta(k) = \begin{cases} 1 & if\ w(k) < \widehat{w} \\ 0 & otherwise \end{cases} \qquad (5.3)$$

where $\widehat{w}$ is the desired set-point value for the queue length on the primary off-ramp. In addition to its simplicity (no parameters to be specified), the potential field implementation of the bang-bang strategy only requires one single detector, appropriately placed on the off-ramp of interest to reflect the value of the set-point $\widehat{w}$. However, as already mentioned, this simple regulator may create

oscillations in the trajectories of the splitting rates and the off-ramp's queue length due to its switching nature.

If the PI-strategy is applied, the splitting rate β towards the primary route is calculated, at each control interval k, according to the following equation,

$$\beta(k) = \beta(k-1) + K_p[w(k-1) - w(k)] + K_i[\hat{w} - w(k)] \qquad (5.4)$$

Also here, the calculated splitting rate β is truncated if it exceeds the range $[0,\ 1]$. As mentioned above, the PI-strategy includes two parameter values to be specified, K_p and K_i, and calls for a more accurate real-time estimation of the off-ramp queue length. The queue length estimation may be obtained via three appropriately placed detectors according to [55]. The PI-regulator leads to smoother trajectories for the calculated splitting rates and the off-ramp's queue length compared to the bang-bang strategy.

5.1.1.3 Case 2: User-optimal conditions may be achieved only after the off-ramp queue spill-over

In this case, the implementation of a route guidance policy that aims to reach user-optimal conditions, i.e. suggests the alternative route only if this route is equivalent or time-shorter than the primary route, may not succeed in avoiding mainstream congestion; thus, two alternative policies are proposed in order to divert vehicles towards alternative (but time-longer) routes, and avoid the creation of mainstream congestion. The first policy uses the estimated queue length on the primary-route off-ramp as the real-time information for deciding on driver diversion and is, actually, the very same policy described in the previous section. Although, in this case, the alternative route is time-longer for all selected set-point values, it is assumed that the expected compliance of the drivers to the rerouting decisions is not too low, e.g. because the travel time reliability from day to day is low or the travel time difference between the alternative routes is minor. The

second proposed policy is based on the assumption that the drivers will not follow the system's rerouting suggestions sufficiently; in such cases, a temporary off-ramp closure (in real time), whenever the off-ramp queue length is about to exceed the off-ramp bounds, is seen as the only possibility to divert traffic from the critical off-ramp and avoid queue spill-over onto the freeway mainstream.

Again, all above diversion control policies actually decide on the percentage of vehicles that should be re-routed and may convey the information to the concerned drivers either via VMS or via vehicle-to-infrastructure communication. In the following, only the second route diversion policy is presented, since the first policy is already described in a previous section.

Dynamic route diversion through temporary off-ramp closures

This policy decides in real time on temporary off-ramp closure, when needed, specifically whenever the estimated queue length on the primary off-ramp, w, exceeds a pre-specified value $\widehat{w}$. Practical possibilities for closing the off-ramp include moving physical barriers or VMS indication (sufficiently enforced), placed appropriately at the off-ramp entrance. To facilitate the intended route diversion, a VMS must also be placed upstream of the alternative off-ramp exit (e.g. upstream of N1 in Figure 5.1) to alert the concerned drivers about the saturated off-ramp closure. Note that the actual off-ramp closure should be effectuated with a time delay, so that concerned drivers who passed the VMS before the closure announcement, can still exit via the saturated off-ramp; this time delay should be apparently in the order of the travel time between both off-ramps. A similar time delay may also be applied while (re-) opening the off-ramp.

Whenever a ramp closure is decided and announced, the concerned drivers (or a large portion of them) are forced to divert via the alternative route; and this corresponds to a splitting rate β towards the primary route equal to zero. Thus, the

impact of this action is fully reflected via a bang-bang feedback strategy (without the need to consider explicitly the aforementioned switching delay), which decides, at each control interval k, about the off-ramp opening or closure, thus determining the splitting rate β as follows:

$$\beta(k) = \begin{cases} 1 & if\ w(k) < \widehat{w} \\ & \\ 0 & otherwise \end{cases} \tag{5.5}$$

In addition to its simplicity (no parameters to be specified), the potential field implementation of the bang-bang strategy only requires one detector, appropriately placed on the off-ramp of interest to reflect the value of the set-point $\widehat{w}$.

5.1.1.4 Case 3: User-optimal conditions may not be achieved

In this case, due to the prevailing traffic conditions on the surface network, the alternative route is, always, the time-longer route. Thus, the application of a route guidance policy that aims to travel time equalization of the two alternative routes is not an option. Moreover, a policy that aims to keep the primary off-ramp queue close to a desired queue level by suggesting the drivers to take the alternative (time-longer) route will not succeed since, in this case, the expected compliance of the drivers to the rerouting decisions is very low. Therefore, the most appropriate measure is the application of temporary off-ramp closures, when and to the extent needed, in order to avoid the off-ramp queue spill-over and the formation of mainstream congestion. This policy is, actually, the very same policy presented in the previous section.

5.1.2 Test network and traffic demand scenarios

A hypothetical, but quite typical, road network (Figure 5.1) is considered to test and demonstrate the proposed concepts and regulators' action. The network is

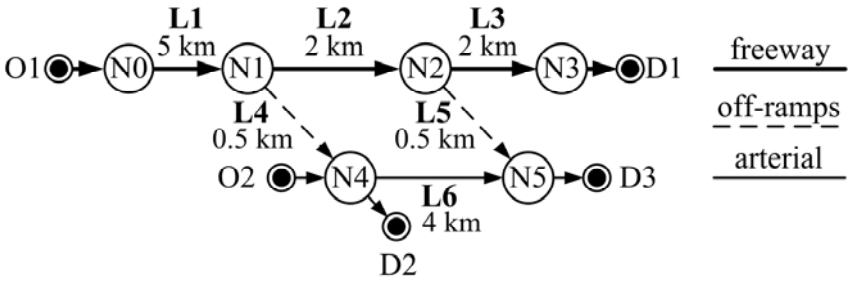

Figure 5.1 Hypothetical traffic network.

modeled by use of the macroscopic traffic flow simulator METANET, which was validated successfully in Chapter 4 for a congested freeway off-ramp area. The hypothetical traffic network is represented through nodes and links, and includes 6 nodes (N0–N5) and 6 links (L1–L6), as shown in Figure 5.1. More specifically, the network includes a 3-lane and 9 km long freeway stretch (L1–L3), two off-ramps (L4, L5) of 1 lane and 0.5 km length, and a 1-lane and 4 km long adjacent surface street (L6). Moreover, the traffic flow enters the simulated network through two origin links, O1 and O2, and exits the network from three destinations, D1–D3. The off-ramp L5 is modeled as a store-and-forward link with constant outflow capacity [20].

All investigated cases, described in Section 5.1.1, consider the same demand profile for the origin link O1, but different demand profiles for the origin O2, so as to create the corresponding differences among the three cases. In particular, Figure 5.2(a) presents the considered demand for the first examined case, while Figure 5.2(b) shows the demand profile for the second and third cases. Moreover, in the first examined case, the adjacent road (L6) is considered as an arterial, with free flow speed at 90 km/h; while in the second and third cases it is considered as an urban street with free flow speed at 40 km/h and 28 km/h, respectively. Figure 5.2(c) shows for each network origin the percentage of flow that exits from each destination and is common for all investigated cases. Moreover, in all cases, it is considered that the maximum exit flow from the off-ramp L5 is fixed to 520 veh/h

for the whole simulation horizon, e.g. due to a traffic light placed at the off-ramp's exit. Since the timing and appearance of excessive ramp queues in the off-ramp L5 may be stochastic from day to day, it is assumed that all freeway drivers, bound for destination D3, are using this off-ramp, unless otherwise informed. Under these conditions, a queue is formed on the off-ramp L5 during the peak hours, which eventually spills back into the freeway mainstream, creating congestion that mounts upstream for several kilometers. These simple traffic scenarios reproduce the three traffic cases described in Section 5.1.1 and enable us to test different route diversion policies, which aim to guide the drivers through alternative routes in real time during the rush hours, in order to prevent the formation of mainstream congestion. In the current investigations, it is assumed that the off-ramp L4 as well as link L6 may accommodate the diverted traffic flow without major problems.

The utilized performance indices to compare the proposed policies are the Total

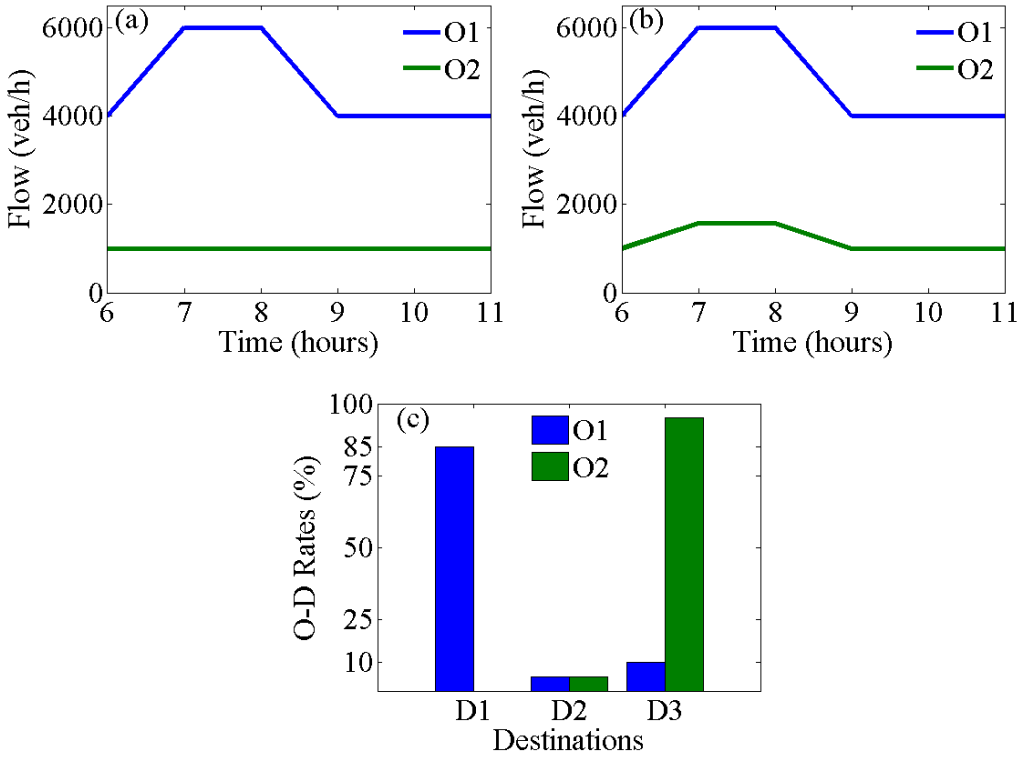

Figure 5.2 (a) Traffic demand at origins for Case 1; (b) traffic demand at origins for Case 2 and Case 3; (c) origin-destination rates.

Time Spent (TTS) in the network in veh·h; and the Total Disbenefit (TD) in veh·h, which reflects the total vehicle-hours wasted on time-longer routes, and is calculated, for the node-destination couple (N1,D3), as follows:

$$TD = \sum_{k=0}^{K} [q(k)\beta^{l}(k)(\tau^{ex,l}(k) - \tau^{ex,sh}(k))]\, T \tag{5.6}$$

where $q(k)$ corresponds to the flow arriving at node N1 that is bound for destination D3, $\beta^{l}(k)$ is the percent of flow that is directed to the time-longer route, $\tau^{ex,l}(k)$ is the experienced travel time over the time-longer route and $\tau^{ex,sh}(k)$ is the experienced travel time over the time-shorter route.

5.1.3 Simulation investigations

As described in Section 5.1.1, this study aims to propose various route diversion policies in order to prevent the formation of freeway congestion, triggered by an overspilling off-ramp. To this end, three different cases are examined. In the first case, the user-optimal conditions may be achieved before the off-ramp queue spill-over; in the second case the user-optimal conditions are achieved only after the off-ramp queue spills back into the freeway mainstream; and in the third case the user-optimal conditions cannot be achieved due to the prevailing conditions on the alternative route (e.g. the alternative route is very long or congested). In the following, various route diversion policies are tested and demonstrated for each investigated case.

5.1.3.1 Case 1: User-optimal conditions may be achieved before the off-ramp queue spill-over

In this case, the alternative route becomes time-shorter before the queue of the primary route off-ramp spills back into the freeway mainstream. Thus, the route guidance system suggests to the drivers the alternative route only if it features equal or shorter travel time, compared to the primary route. Two route guidance

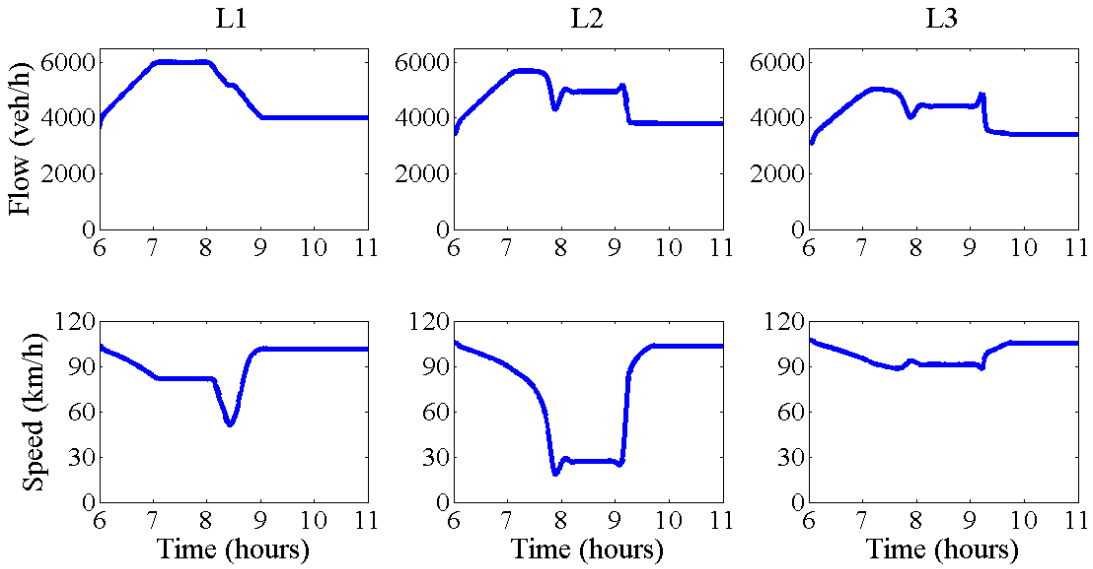

Figure 5.3 No route diversion (Case 1.1): Flow and speed measurements over simulation time.

policies are proposed for this investigated case. The first policy bases its calculations on real-time estimations of the travel time for the two alternative routes; while the second one uses the estimated queue length on the primary off-ramp. To start with, the no-route-guidance case is presented first; followed by the results obtained for both proposed route diversion policies.

A. No route diversion

In absence of a route diversion indication, all the drivers travelling on the freeway bound for destination D3, are assumed to exit the freeway from the off-ramp L5, which belongs to the distance-shorter route, e.g. because a diversion would represent a risk due to low day-to-day reliability of travel times. Figure 5.3 shows, for each mainstream freeway link, the (emulated) flow and speed measurements over the simulation time. It is observed, that, around 7:45 a.m., the flow in the mainstream link L2 drops abruptly; while at the same time there is also a steep speed drop occurring at the same link, which propagates upstream to link L1, without reaching the upstream end of the freeway stretch. Mainstream congestion

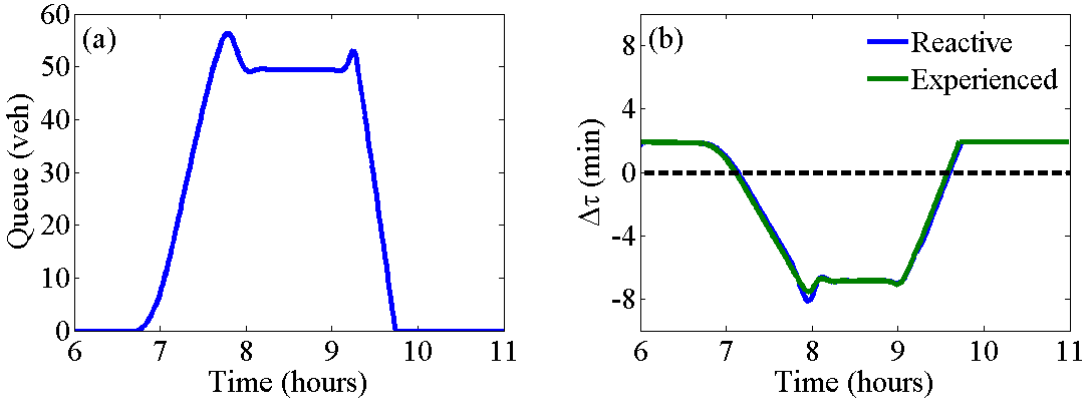

Figure 5.4 No route diversion (Case 1.1): (a) queue length at the primary off-ramp (L5); and (b) travel time difference for the two alternative routes; over simulation time.

lasts up to around 9:15 a.m.; after which it is dissolved due to lower traffic demand entering the network from O1 (see Figure 5.2).

Figure 5.4 (a) presents the queue length on the off-ramp L5 over time, and confirms that a queue is formed around 6:50 a.m., which increases gradually and eventually reaches, around 7:45 a.m., the maximum vehicle storage capacity of the off-ramp, which is around 55 veh. After this time, the off-ramp queue spills back into the mainstream creating congestion. It is also shown here, that around 9:15 a.m. the queue length starts to decrease, which has a prompt positive impact on the mainstream traffic conditions (see Figure 5.3). Figure 5.4 (b) displays the reactive travel time difference $\Delta\tau$ of the two alternative routes, as defined in Section 5.1.1, over simulation time, as well as the experienced travel time difference. The experienced travel time is the real (simulated) travel time of vehicles, which becomes known only after the completion of the corresponding trip. It is observed that the two lines are similar over the whole simulation time. In particular, Figure 5.4(b), shows that, at the beginning of the simulation, the primary route is the time-shorter route, up until 7:10 a.m. After this time, the secondary route becomes the time-shorter, due to the increasing queue on the primary off-ramp and, eventually, due to the queue spillover and mainstream

congestion. Finally, the situation changes again, at around 9:40 a.m., when the queue formed on the off-ramp is gradually decreasing and the primary route becomes again the time-shorter route. Table 5-1 presents the network performance in terms of Total Time Spent (TTS) in the network and Total Disbenefit (TD). These values will be compared to the cases where route diversion policies are applied.

5.1.3.1.1 Dynamic route diversion based on reactive travel time estimation

According to this policy, the route diversion system calculates the splitting rates by use of feedback control strategies, which attempt to keep the travel time difference $\Delta\tau(k)$ on the two alternative routes close to zero. In the following, the application results are presented, in case that a bang-bang or a PI feedback strategy is employed.

The bang-bang strategy was tested with a control interval equal to 2 min. Figure 5.5 displays the queue length at the off-ramp L5, the travel time difference of the two alternative routes and the strategy's calculated splitting rates towards the primary route, over simulation time, under the operation of the bang-bang strategy. It is observed that the strategy manages to maintain the travel time differences close to zero (see Figure 5.5(b)) via occasional recommendation of the secondary route for small periods of time; and by doing this, the queue length on the off-ramp L5 is kept well below the 55 veh, which is its maximum vehicle

Table 5-1 Performance criteria for Case 1.

CASE 1			**PERFORMANCE CRITERIA**	
Route Diversion Policy		***Feedback Strategy***	**TTS (veh·h)**	**TD(veh·h)**
1.	***No route diversion***	–	3228	110
2.	***Route diversion based on travel times***	*Bang-bang*	2462	1.6
		PI	2473	0.1
3.	***Route diversion based on queue length***	*Bang-bang*	2478	6.1
		PI	2500	13.7

storage capacity (Figure 5.5 (a)), thus no congestion is created on the freeway mainstream. An operational inherent disadvantage of the bang-bang strategy is the visible oscillations of all involved quantities.

Table 5-1 displays the performance of the strategy in terms of TTS and TD. It is shown that the control strategy achieves 24% reduction of TTS and 98% less vehicle-hours wasted on time-longer routes, compared to the no route diversion case. It should be noted that different control intervals may be applied according to the infrastructure characteristics and the control equipment capabilities, leading to similar results.

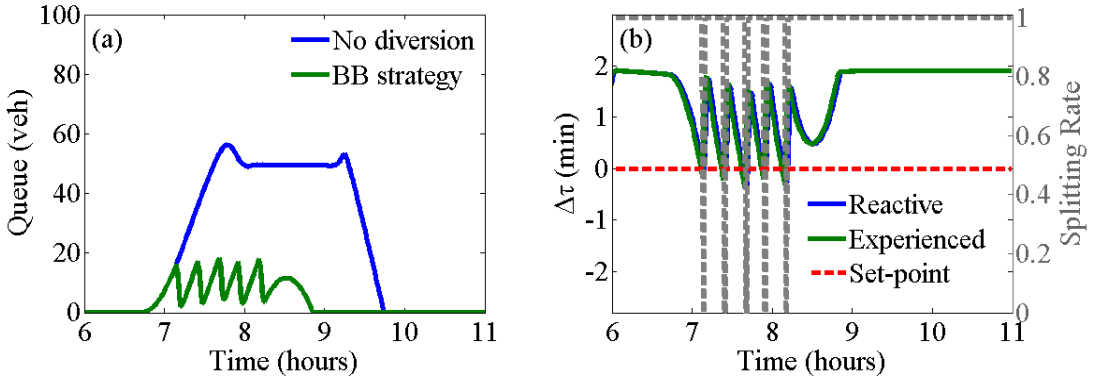

Figure 5.5 Bang-bang strategy (Case 1.2): (a) queue length at the primary off-ramp (L5); and (b) travel time difference for the two alternative routes and calculated spitting rate; over time.

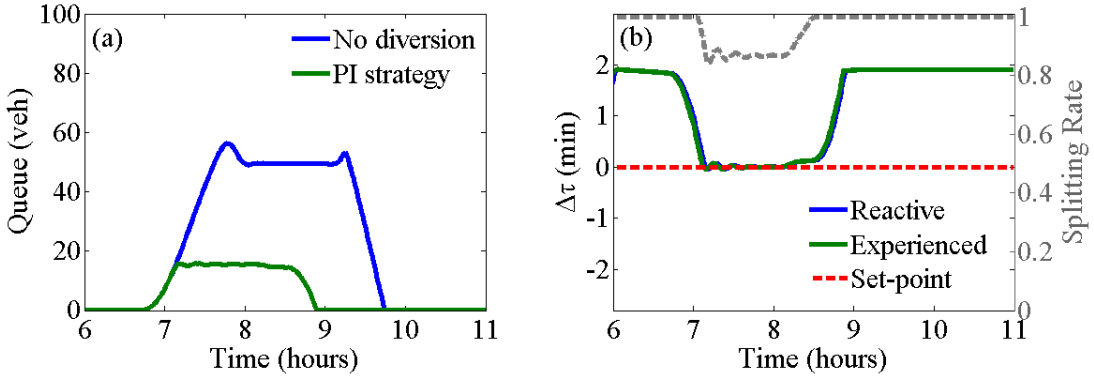

Figure 5.6 PI-strategy (Case 1.2): (a) queue length at the primary off-ramp (L5); and (b) travel time difference for the two alternative routes and calculated splitting rate; over time.

The PI-strategy was employed with a control interval of 2 min and parameter values $K_p = 18$ min^{-1} and $K_i = 6$ min^{-1} (specified via trial-and-error). Figure 5.6(b) indicates that the strategy succeeds in achieving virtually perfect travel time equalization on the two alternative routes by appropriately splitting of the relevant traffic sub-flow; and at the same time maintains the off-ramp queue length at a low level (Figure 5.6(a)). As expected, the PI-strategy results in smoother trajectories for the travel time differences and also for the off-ramp queue length compared to the bang-bang strategy. Table 5-1 shows that the PI-strategy achieves 23% reduction of TTS and almost 100% improvement of the TD, which actually corresponds to the user-optimal conditions. It should be noted that, also in this case, different control intervals may be applied (albeit with accordingly modified regulator parameters), leading to similar results.

A question that may arise is how the compliance rate (CR) of the drivers to the routing instructions may affect the controllers' operation. To address this issue, Figure 5.7 and Figure 5.8 reflect the operation of the control strategies considering various compliance rates, modelled according to [9, 10]. Figure 5.7(a) and Figure 5.8(a) present, for each control strategy, the experienced travel time differences for the two alternative routes; Figure 5.7(b) and Figure 5.8(b) display the splitting rate β calculated by each control strategy; and Figure 5.7(c) and Figure 5.8(c) show the splitting rate that is actually implemented according to the level of the drivers' compliance; over the simulation time. Note that compliance rate equal to 1 corresponds to the results presented in the previous sections, while compliance rate equal to zero corresponds to the no-diversion case. It is observed that, thanks to their feedback character, both control strategies succeed in maintaining the travel time difference close to zero by automatically adapting their control decisions to the (unknown) level of drivers' compliance. Of course, if CR becomes much smaller, then the reaction of the feedback regulators will be accordingly slower; and for CR=0, the process becomes literally uncontrollable.

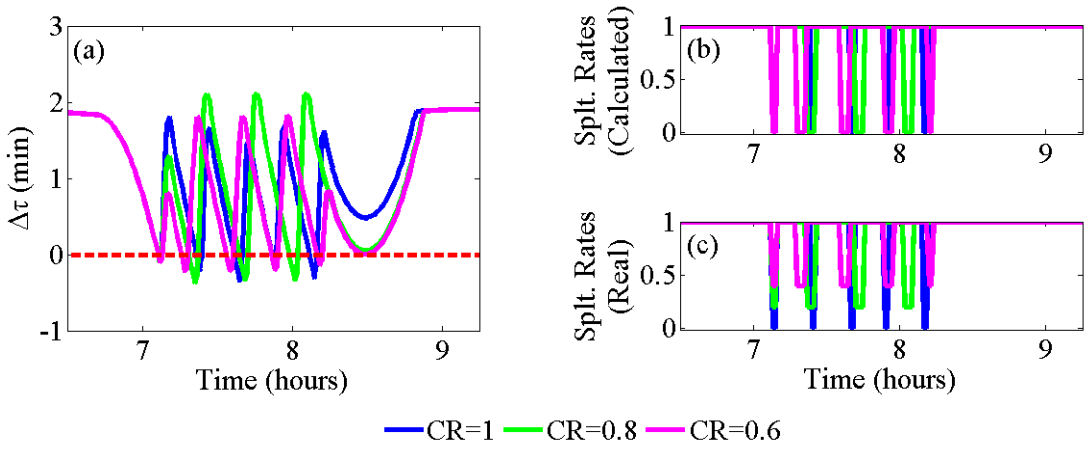

Figure 5.7 Compliance rate investigation for the bang-bang strategy (Case 1.2): (a) travel time difference for the two alternative routes; (b) calculated splitting rates; and (c) real splitting rates, over time.

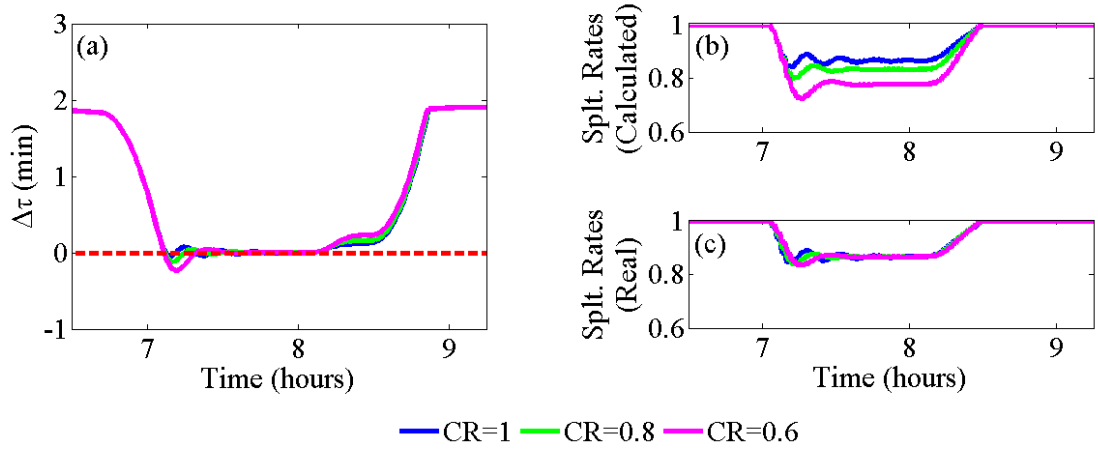

Figure 5.8 Compliance rate investigation for the PI-strategy (Case 1.2): (a) travel time difference for the two alternative routes; (b) calculated splitting rates; and (c) real splitting rates, over time.

5.1.3.1.2 Dynamic route diversion based on off-ramp queue length estimation

According to this policy, the route guidance system calculates the splitting rates by use of feedback control strategies, which aim to keep the primary off-ramp's queue length at a pre-specified level, so that congestion does not spill back to the

mainstream. In the following, the application results are presented, in case a bang-bang or a PI feedback strategy is employed.

Figure 5.9 displays the operation of the bang-bang strategy, for a control interval equal to 2 min and queue length set-point equal to 25 veh; based on previous observations, this set-point value is suitable for achieving approximately user-optimal traffic conditions and is also low enough to prevent spill-over of the off-ramp queue. Indeed, it is observed that the strategy manages to maintain the off-ramp queue length close to the set-point value (Figure 5.9(a)), by calculating appropriate splitting rates (Figure 5.9(a)), thus no congestion is created on the freeway. Moreover, whenever the strategy decides on rerouting, the secondary route is indeed time-shorter (Figure 5.9(b)), which means that the compliance rate of the drivers to the route diversion system decision is expected to be high. The situation where diverted vehicles have a strictly lower travel time than non-diverted vehicles, is sometimes called a weak user optimum.

Table 5-1 includes the performance indices of the strategy in terms of TTS and TD. It is shown that the control strategy achieves 23% reduction of TTS and 94% less vehicle-hours wasted on time-longer routes compared to the no route diversion case. Note that in this case the wasted time occurs on the primary route, since diverted vehicles experience a shorter travel time. It should also be noted that similar results can be obtained for set-point values from a reasonably wide range, provided that the queue on the off-ramp does not affect the freeway mainstream, possibly at the expense of a slight increase of the incurred disbenefit in Table 5-1. Moreover, different control intervals may be applied with similarly successful results.

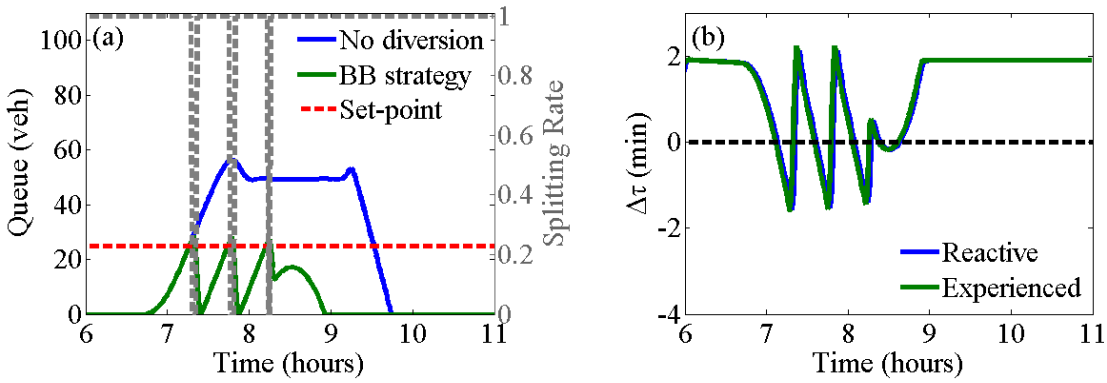

Figure 5.9 Bang-bang strategy (Case 1.3): (a) queue length at the primary off-ramp (L5) and calculated splitting rate; and (b) travel time difference for the two alternative routes; over time.

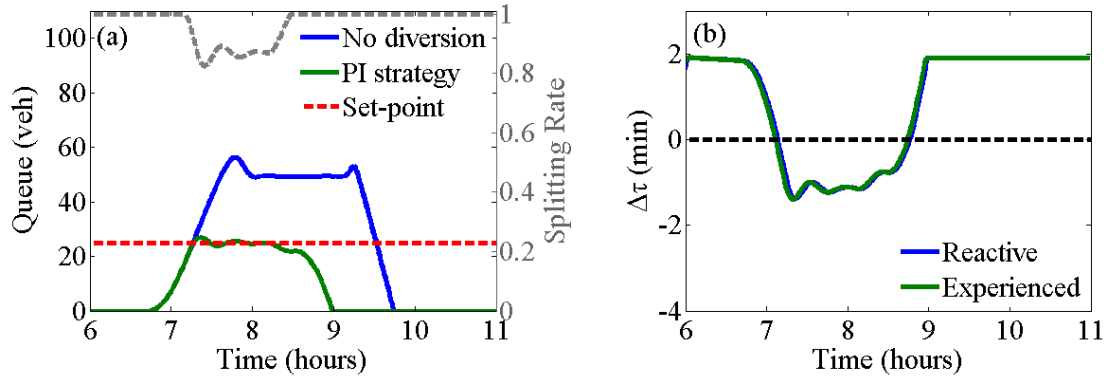

Figure 5.10 PI-strategy (Case 1.3): (a) queue length at the off-ramp (L5) and calculated splitting rate; and (b) travel time difference for the two alternative routes; over time.

The PI-strategy is tested for a control interval equal to 2 min, and parameter values $K_p = 0.02$ veh^{-1} and $K_i = 0.005$ veh^{-1} (specified via trial-and-error). Figure 5.10(a) presents the operation of the PI-strategy which aims to maintain the queue length at the off-ramp L5 close to the set-point value $\widehat{w} = 25$ veh. It is observed that the controller manages to keep the queue length close to the desired value (Figure 5.10(a)) by calculating proper splitting rates for the two alternative routes (Figure 5.10(a)). It should be mentioned that the drivers that are rerouted through the secondary route, experience shorter travel times (Figure 5.10(b)) (weak user optimum), thus the compliance level is expected to be high. Similar results can be

obtained for different control intervals and various set-point values, provided the queue on the off-ramp does not affect the freeway mainstream. As expected, the operation of the PI-strategy leads to smoother traffic conditions on the off-ramp L5 compared to the implementation of the bang-bang strategy. Table 5-1 includes the performance indices for the PI-strategy, which achieves 22% reduction of the TTS and 87% improvement of the TD, compared to the no route diversion case. Again, the wasted time occurs on the primary route.

Figure 5.11 and Figure 5.12 present the controllers' operation considering various drivers' compliance levels to the route diversion system recommendations. It is observed that both strategies succeed in maintaining the off-ramp queue length close to the set-point value (Figure 5.11(a), Figure 5.12(a)), thanks to their feedback nature, by automatically adapting their decisions to the (unknown) compliance of the drivers (Figure 5.11(b), Figure 5.12(b)).

5.1.3.2 Case 2: User-optimal conditions may be achieved only after off-ramp queue spill-over

In this case, the network geometry and traffic conditions lead to user-optimal conditions, i.e. equal travel time for the two alternative routes, only after the off-ramp queue spills back into the mainstream. Thus a route guidance policy that proposes alternative routes based on shorter travel times will result in mainstream congestion. In the following, first the no route diversion case is addressed, followed by the application results of the two route diversion policies presented in Section 5.1.1.

A. No route diversion

As in the first examined case, if no route diversion is applied, all freeway drivers bound for destination D3, exit the freeway from the off-ramp L5, which belongs to the distance-shorter route. Since both investigated cases consider the same

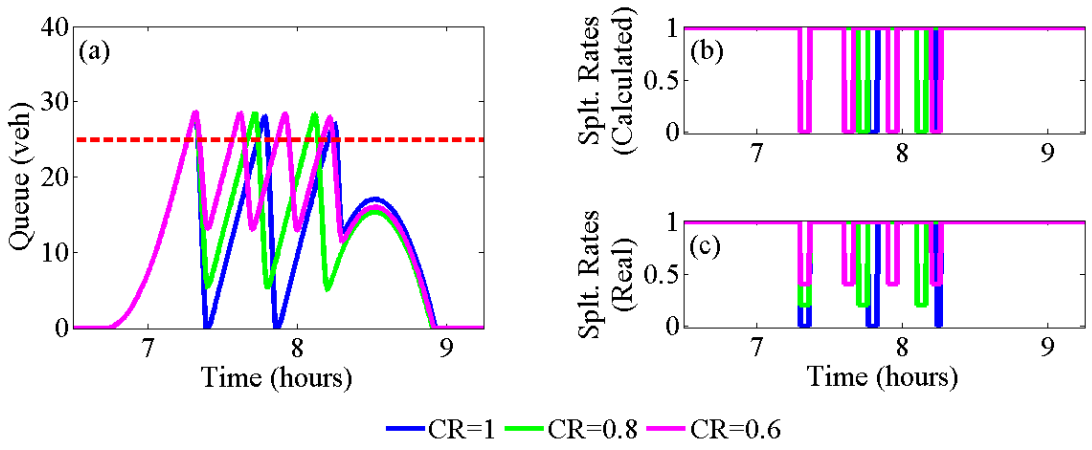

Figure 5.11 Compliance rate investigation for the bang-bang strategy (Case 1.3): (a) travel time difference for the two alternative routes; (b) calculated splitting rates; and (c) real splitting rates; over time.

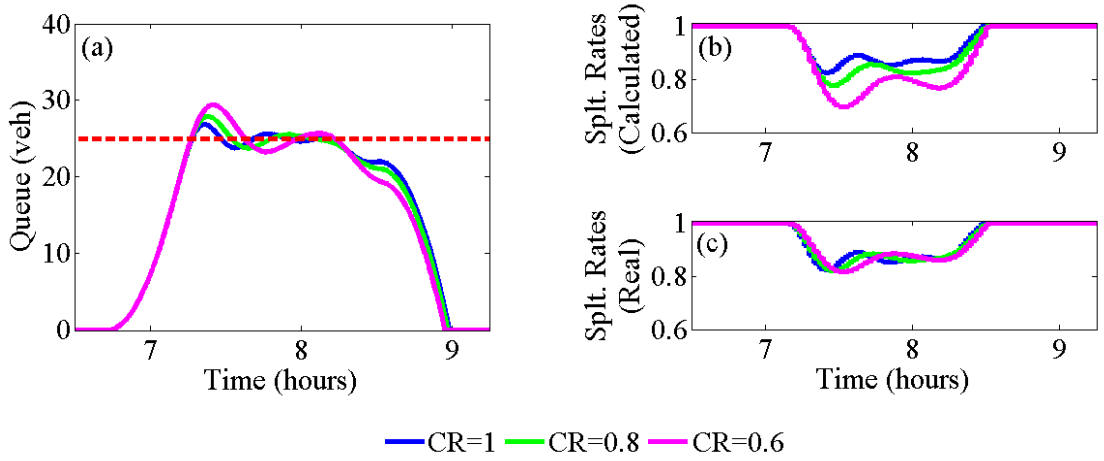

Figure 5.12 Compliance rate investigation for the PI-strategy (Case 1.3): (a) queue length at the primary off-ramp (L5); (b) calculated splitting rates; and (c) real splitting rates; over time.

freeway and primary off-ramp conditions, the (emulated) flow and speed measurements over the simulation time for the mainstream links are exactly the same (as presented in Figure 5.3).

Figure 5.13(a) displays the queue length on the off-ramp L5 over time, and Figure 5.13(b) depicts the reactive and experienced travel time differences for the two alternative routes. It is seen that the off-ramp queue exceeds the off-ramp bounds

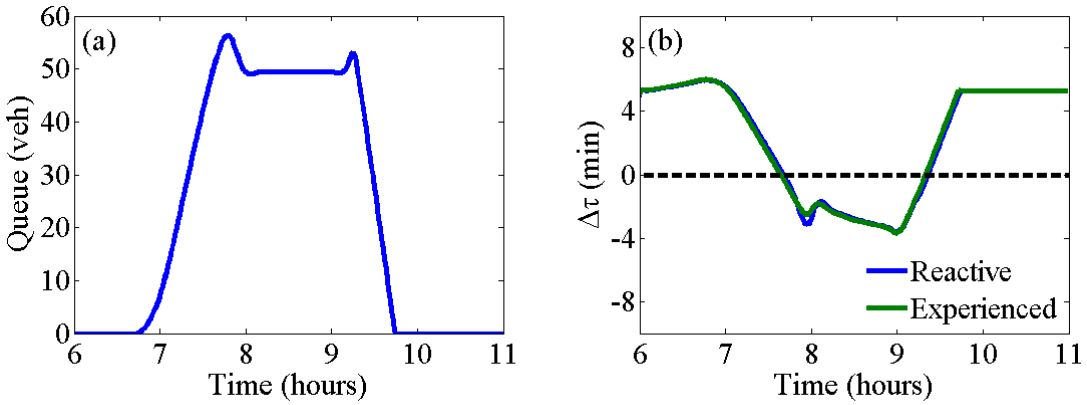

Figure 5.13 No route diversion (Case 2.1): (a) queue length at the primary off-ramp (L5); and (b) travel time difference for the two alternative routes; over time.

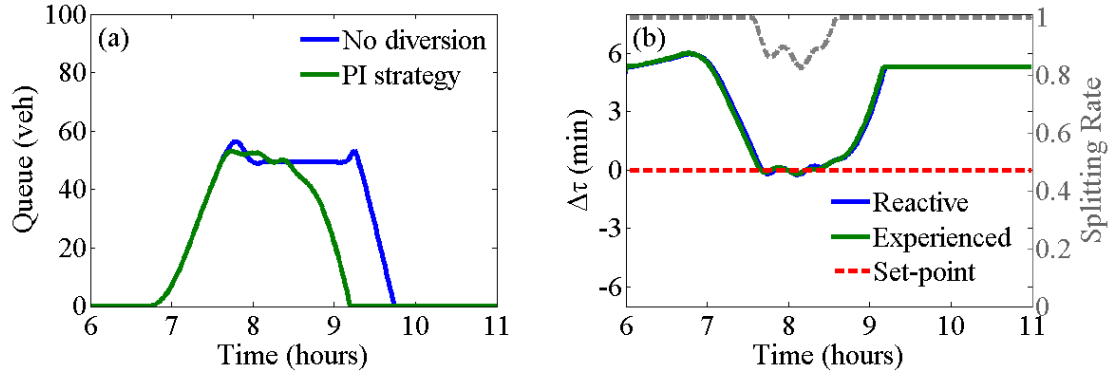

Figure 5.14 PI-strategy (Case 2.2): (a) queue length at the off-ramp (L5); and (b) travel time differences for the two alternative routes and calculated splitting rate; over time.

between 7:45 a.m. and 9:15 a.m., triggering mainstream congestion. Moreover, the equalization of travel time ($\Delta\tau = 0$) for the two alternative routes is achieved around 7:45 a.m., by which time the off-ramp queue has entered into the freeway mainstream. Table 5-2 presents the TTS in the network and the TD. These values will be compared to the cases where route diversion policies are applied.

5.1.3.2.1 Dynamic route diversion based on reactive travel time estimation

This policy calculates the splitting rates by use of a feedback control strategy, which aims at the equalization of the (reactive) travel time ($\Delta\tau = 0$) on the two alternative routes. This means that the utilized strategy decides on rerouting only if the secondary route is competitive. The reported application results are obtained

with the PI-strategy, using a control interval of 2 min, and parameter values $K_p = 8$ min^{-1} and $K_i = 2$ min^{-1}. While the strategy manages to maintain $\Delta\tau(k)$ close to zero, as indicated in Figure 5.14(b), by calculating appropriate splitting rates towards the primary route (Figure 5.14(b)), the off-ramp queue exceeds the maximum storage capacity (see Figure 5.14(a)), leading to mainstream congestion, which is, though, restricted to link L2 (compare Figure 5.3 and Figure 5.15). Table 5-2 shows that this policy achieves 16% reduction of TTS and 99% improvement of TD compared to the no route diversion case. It should be noted that different control intervals may be applied (albeit with accordingly modified regulator parameters), leading to similar results. Moreover, a bang-bang feedback regulator, instead of the PI-regulator, may also be utilized.

5.1.3.2.2 Dynamic route diversion based on off-ramp queue length estimation

According to this policy, the route diversion system calculates the splitting rates by use of a feedback control strategy, which aims to keep the primary off-ramp queue length at a pre-specified level, such that congestion does not spill back to the mainstream. Due to the network and traffic characteristics considered, this policy may actually divert the vehicles to time-longer routes, in order to prevent

Table 5-2 Performance criteria for Case 2.

CASE 2			**PERFORMANCE CRITERIA**	
Route Diversion Policy		***Feedback Strategy***	**TTS (veh·h)**	**TD (veh·h)**
1.	***No route diversion***	–	3637	34.3
2.	***Route diversion based on travel times***	*PI*	3044	0.4
3.	***Route diversion based on queue length***	*PI*	2927	5.9
4.	***Route diversion through temporary off-ramp closures***	*Bang-bang*	2916	8.3

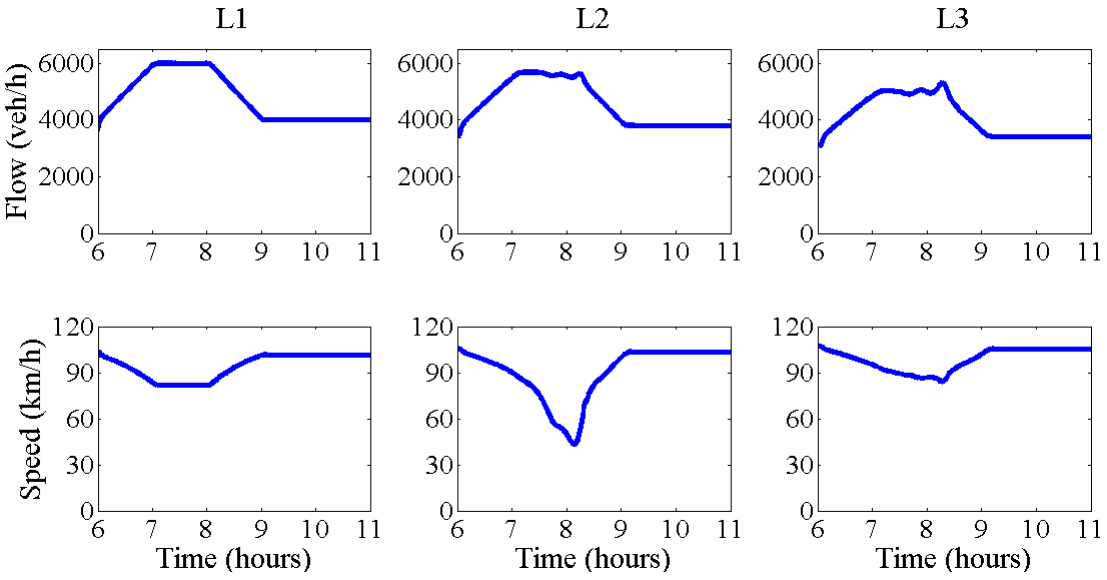

Figure 5.15 Route diversion (Case 2.2): Flow and speed measurements over simulation time.

the formation of mainstream congestion. Since the drivers are free to ignore the route indications, this policy may be applied in the field if the drivers compliance is expected to be at reasonable levels, e.g. because the travel time difference between the alternative routes is minor; or because the exact traffic conditions are not known to the drivers due to low reliability from day to day.

In the following, the application results are presented in case a PI feedback strategy is employed with a control interval equal to 2 min and parameter values $K_p = 0.02$ veh^{-1} and $K_i = 0.05$ veh^{-1}. Figure 5.16(a) reflects the operation of the PI-strategy which manages to maintain the queue length at the off-ramp L5 close to the set-point value $\hat{w} = 25$ veh, by calculating appropriate splitting rates towards the primary route (Figure 5.16(a)). Figure 5.16(b) shows that the drivers that are rerouted through the secondary route, experience longer travel time, actually some 4 min longer than on the primary route. Similar results can be obtained for different control intervals and various set-point values, provided the queue on the off-ramp does not affect the freeway mainstream. Note that an

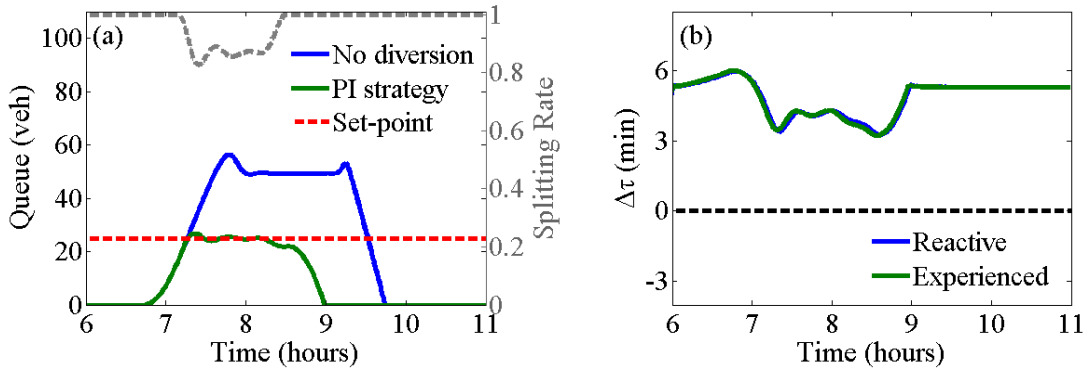

Figure 5.16 PI-strategy (Case 2.3): (a) queue length at the primary off-ramp (L5) and calculated splitting rate; and (b) travel time difference for the two alternative routes; over time.

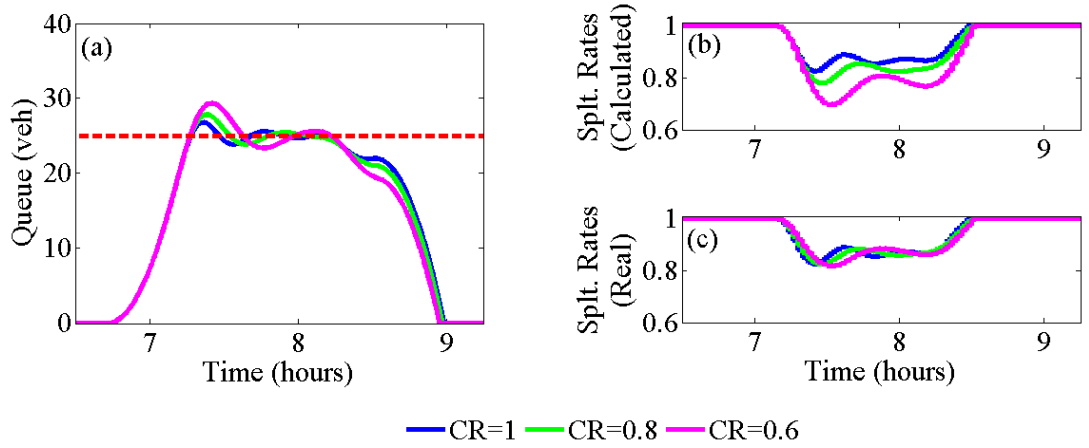

Figure 5.17 Compliance rate investigation for the PI-strategy (Case 2.3): (a) queue length at the primary off-ramp (L5); (b) calculated splitting rates; and (c) real splitting rates; over time.

increase of the queue length set-point would decrease the travel time difference among the two alternative routes. Finally, a bang-bang feedback regulator, instead of the PI-regulator, may also be utilized. Table 5-2 includes the performance indices for the queue control strategy, which achieves 20% reduction of the TTS and 83% improvement of the TD compared to the no route diversion case.

Regarding the influence of the compliance rate (CR) of the drivers to the controller's recommendations, Figure 5.17 reflects the operation of the queue control strategy considering various compliance rates. Figure 5.17(a) presents the

off-ramp queue length on the off-ramp L5, Figure 5.17(b) displays the splitting rates β calculated by the control strategy, and Figure 5.17(c) shows the splitting rate that is actually implemented according to the level of the drivers' compliance, over the simulation time. Note that compliance rate equal to 1 corresponds to the results presented above. It is observed that, thanks to its feedback character, the control strategy succeeds in maintaining the off-ramp queue length close to the set-point value (Figure 5.17(a)), by automatically adapting its decisions to the (unknown) compliance of the drivers (Figure 5.17(b)).

5.1.3.2.3 Dynamic route diversion through temporary off-ramp closures

This policy considers that the drivers will not be eager to follow a time-longer route in order to benefit the rest of freeway users, thus a mandatory action, i.e. the off-ramp closure, must be applied. For this purpose, the feedback bang-bang strategy is employed with a control interval equal to 2 min. Figure 5.18(a) demonstrates the operation of the bang-bang strategy which decides for the primary-route off-ramp closure whenever the off-ramp queue exceeds the desired set-point value $\widehat{w} = 25$ veh. Figure 5.18(b) indicates that the secondary route is indeed the time-longer route for the whole simulation horizon; and that the travel time difference among the two alternative routes is comparable with Figure

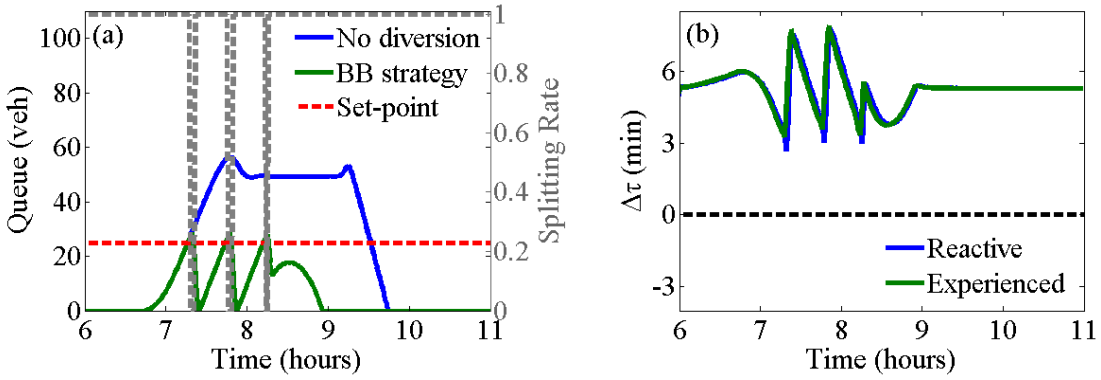

Figure 5.18 Bang-bang strategy (Case 2.4): (a) queue length at the primary off-ramp (L5) and calculated splitting rate; and (b) travel time difference for the two alternative routes; over time.

5.18(b). It should be noted that similar results can be obtained for different control intervals and various set-point values, provided the queue on the off-ramp does not affect the freeway mainstream.

Table 5-2 includes the performance indices for the bang-bang strategy, which achieves 20% reduction of the TTS and 76% improvement of the TD compared to the no route diversion case. Note that, although, in both strategies, a portion of drivers is diverted through a time-longer route, in practice it would not be every day that the same drivers will be proposed a diversion. Moreover, the driver delay would be much higher in absence of the route diversion system due to the formation of heavy mainstream congestion.

5.1.3.3 Case 3: User-optimal conditions cannot be achieved

In this case, due to the prevailing traffic conditions, the alternative route is always the time-longer route, even during the peak hours, hence the expected drivers' compliance to the re-routing indications would be very low. Therefore, the route diversion system should be based on mandatory actions, i.e. temporary off-ramp closures, in order to prevent the off-ramp queue spill-over and the formation of mainstream congestion. In the following, first the no route diversion case is presented, followed by the results obtained for the proposed route diversion policy described in Section 5.1.1.

A. No route diversion

As in the previous cases, if no route diversion is applied, congestion is created in the freeway mainstream, during the peak hours, as shown in Figure 5.3, due to the primary off-ramp queue spill-over (see Figure 5.19(a)). Figure 5.19(b) shows that, in this investigated case, the alternative route is the time-longer route throughout the simulation ($\Delta\tau > 0$). Table 5-3 displays the TTS in the network and the TD

criterion, which is equal to zero since all drivers bound for destination D3, exit the freeway from the primary off-ramp which is always the time-shorter route.

5.1.3.3.1 Dynamic route diversion through temporary off-ramp closures

Considering that in this case, the expected compliance of the drivers to the rerouting indications would be very low, the route diversion system should be based on mandatory actions, i.e. temporary off-ramp closures when and to the extent needed. For this purpose, the bang-bang feedback strategy is employed with a control interval equal to 2 min. Figure 5.20(a) shows the operation of the bang-bang strategy which decides for the primary-route off-ramp closure whenever the off-ramp queue exceeds the desired set-point value $\widehat{w} = 25$ veh.

The diverted vehicles will experience about 9 min longer travel time compared to

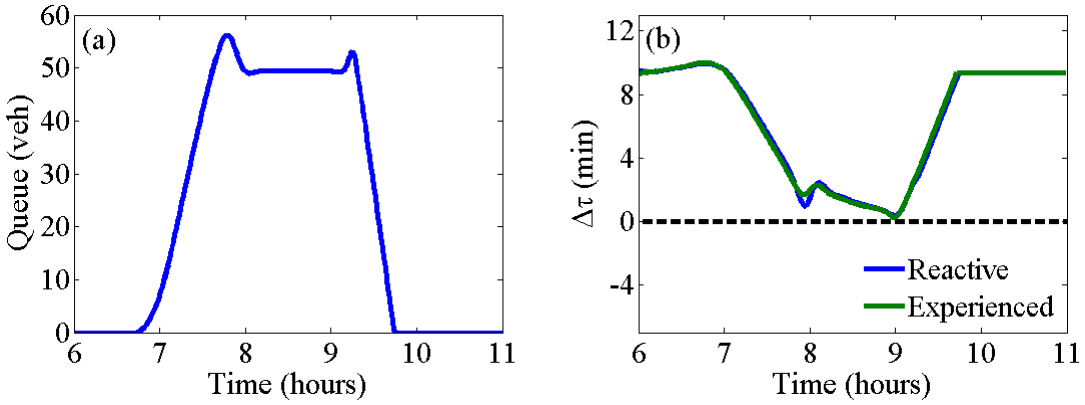

Figure 5.19 No route diversion (Case 3.1): (a) queue length at the primary off-ramp (L5); and (b) travel time difference for the two alternative routes; over simulation time.

Table 5-3 Performance criteria for Case 3.

CASE 3			**PERFORMANCE CRITERIA**	
Route Diversion Policy		***Feedback Strategy***	**TTS (veh·h)**	**TD (veh·h)**
1.	***No route diversion***	–	3979	0
2.	***Route diversion through temporary off-ramp closures***	*Bang-bang*	3263	14.5

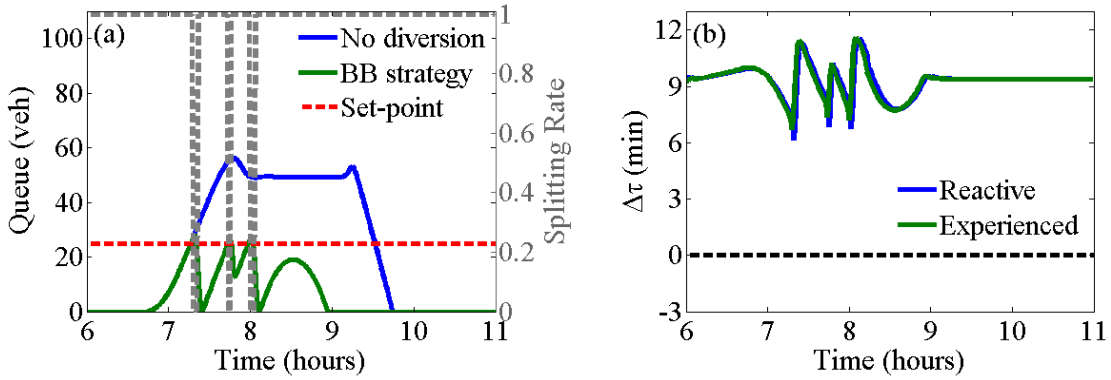

Figure 5.20 Bang-bang strategy (Case 3.2): (a) queue length at the primary off-ramp (L5) and calculated splitting rate; and (b) travel time difference for the two alternative routes; over time.

the non-diverted vehicles (see Figure 5.20(b)). Again, similar results can be obtained for different control intervals and various set-point values, provided the off-ramp does not affect the freeway mainstream. Table 5-3 includes the performance indices for the bang-bang strategy, which achieves 18% reduction of the TTS and 14.5 veh·h wasted on time-longer routes compared to the no control case.

5.1.4 Conclusions

Section 5.1 proposes various route diversion policies that aim to prevent recurrent freeway congestion which is triggered by a saturated off-ramp. The proposed policies employ simple but efficient feedback laws and attempt to reroute the drivers, who would typically exit from the saturated off-ramp, through alternative routes, in order to avoid the off-ramp spill-over and the resulting mainstream congestion. In particular, three different traffic cases are examined for a given typical network topology. In the first case, the user-optimal conditions may be achieved without off-ramp queue spill-over and creation of mainstream congestion; thus the route guidance system may propose an alternative route without any disbenefit for the compliant drivers. In the second case, the user-

optimal conditions may be achieved only after the off-ramp queue spills back to the freeway mainstream; thus the route diversion system will have to assume sufficient compliance to the proposed route choice; or be based on mandatory actions, such as temporary off-ramp closures. Finally, in the third examined case, the user-optimal conditions cannot be achieved, due to the traffic conditions on the alternative route, thus the route diversion system should decide for the temporary off-ramp closure, when and to the extent needed, in order to prevent the formation of mainstream congestion. The simulation results showed that, in all investigated cases, the proposed policies succeed in maintaining the off-ramp queue length within the off-ramp bounds, thus improving the traffic conditions on the freeway mainstream substantially, compared to the case that no route diversion is applied to the network.

5.2 Case 2: Real-time merging traffic control

This section presents a real-time merging traffic control algorithm to mitigate the problem of freeway congestion due to an over-spilling off-ramp. The proposed control algorithm aims at maximizing the surface street merge area outflow and at the same time preventing the off-ramp queue spill-over into the freeway mainstream and the resulting freeway congestion. The potential benefits obtained by the application of the proposed control concept are demonstrated by use of microscopic simulation applied to a real freeway network where recurrent traffic congestion is created due to an over-spilling off-ramp.

5.2.1 Real-time merging traffic control concept

Real-time merging traffic control aims at improving the traffic conditions at a merge bottleneck by appropriately regulating the inflows to a merge area with limited capacity. In the past, merging traffic control has been successfully used in the form for ramp-metering ([56], [57], [58]), but it has also been proposed for

mainline metering, e.g. at freeway toll-plaza and work-zone areas ([59], [60]), as well as at other types of freeway bottlenecks [61]. In this study, merging traffic control is proposed to face the problem of freeway congestion triggered by a saturated off-ramp which merges with a parallel arterial. In the following sections, the problem examined is described first, followed by the description of the proposed control concept.

5.2.1.1 Problem description

Consider a freeway stretch, as presented in Figure 5.21, with an off-ramp flow exiting to a surface street network. This exiting flow q_{off} merges with the flow moving on the surface street network, q_{srf}, in the merge area. The merge area is a potential bottleneck location, which may be activated during the peak period; the congestion created on the surface street, may spill back into the freeway mainstream through the saturated off-ramp (see Figure 5.22). The outlined situation is actually appearing in real networks, e.g. as reported in [39].

The merge area may be a bottleneck location due to a number of reasons:

- High arriving demand (including the surface street and the off-ramp demand).
- Infrastructure layout, e.g. lane drop.
- Strong weaving of traffic streams.
- Downstream urban traffic lights.
- Other capacity reducing events, such as incidents.

Figure 5.23 displays a typical flow-density diagram for a merge area, also known as the Fundamental Diagram (FD), where q_{out} is the exit flow from the merge area and N is the number of vehicles included in the merge area. As long as the total arriving merging flow, q_{off} plus q_{srf}, is lower than the flow capacity of the merge area, Q_{cap}, the merging efficiency is satisfactory and there is no need for external

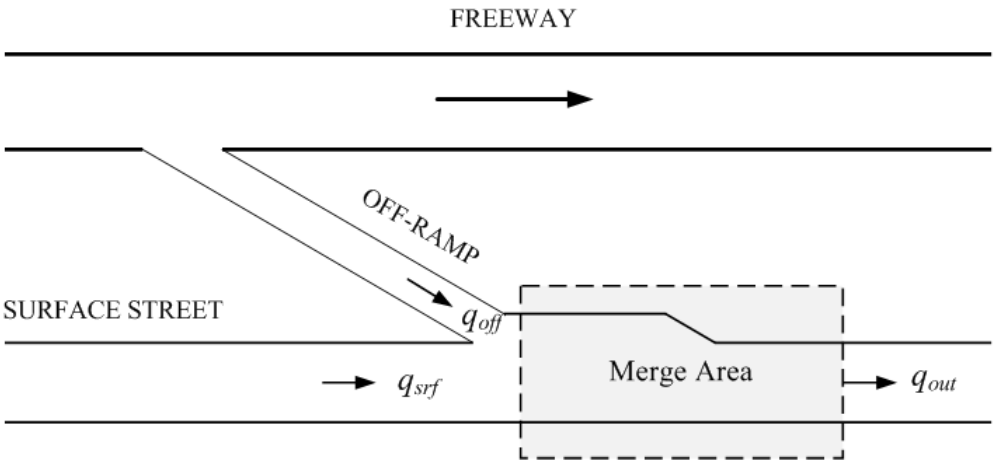

Figure 5.21 Typical freeway off-ramp area.

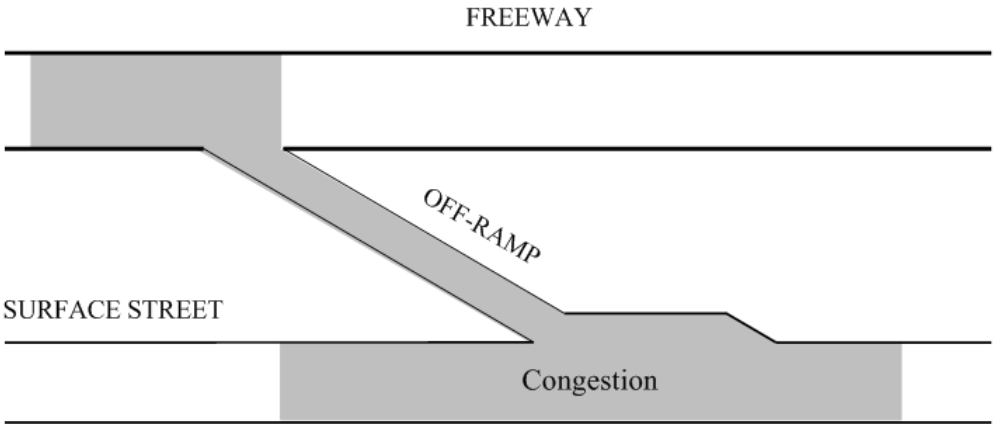

Figure 5.22 Congestion at the surface street network spilling back to the freeway mainstream through the saturated off-ramp.

intervention. If the total arriving merging flow reaches or exceeds the capacity of the merging area, i.e. N increases beyond N_{cr}, congestion is created in the merge area and the exiting flow is reduced to lower values, Q_c, where $Q_{cap} - Q_c$ is the capacity drop due to congestion. Several empirical investigations with real traffic data ([62], [63]) indicate that the reduced outflow Q_c may be 5-20% lower than the nominal capacity, Q_{cap}, leading to a corresponding serious degradation of the road infrastructure and increased delays.

This study proposes a framework for real-time merging traffic control, which aims to maximize the surface street merge area outflow, prevent the off-ramp queue

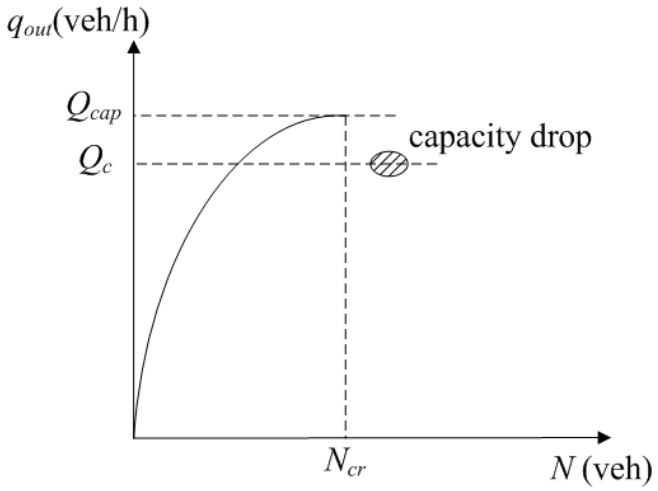

Figure 5.23 Fundamental diagram of a merge area.

spillback into the freeway mainstream and, consequently, also the formation of congestion on the freeway mainstream. The proposed control concept is based on the general real-time merging traffic control framework presented in a previous work [59] and tested for toll-plaza and work-zone areas. In this study, the general merging traffic control concept is combined with queue management techniques and is applied, through microscopic simulation, to a real traffic network where recurrent traffic congestion is created on the freeway mainstream due to an over-spilling off-ramp queue.

5.2.1.2 Traffic control algorithm

As indicated above, the aim of the proposed control concept is to maximize the surface street merge area outflow and prevent the off-ramp queue spill-over into the freeway mainstream. To this end, two control strategies are employed, the ALINEA feedback strategy and the Queue Override strategy which are described in the following sections.

ALINEA control strategy

ALINEA ([62], [63]) is a well-known feedback control strategy that has been successfully applied to many ramp-metering installations and has, also, been

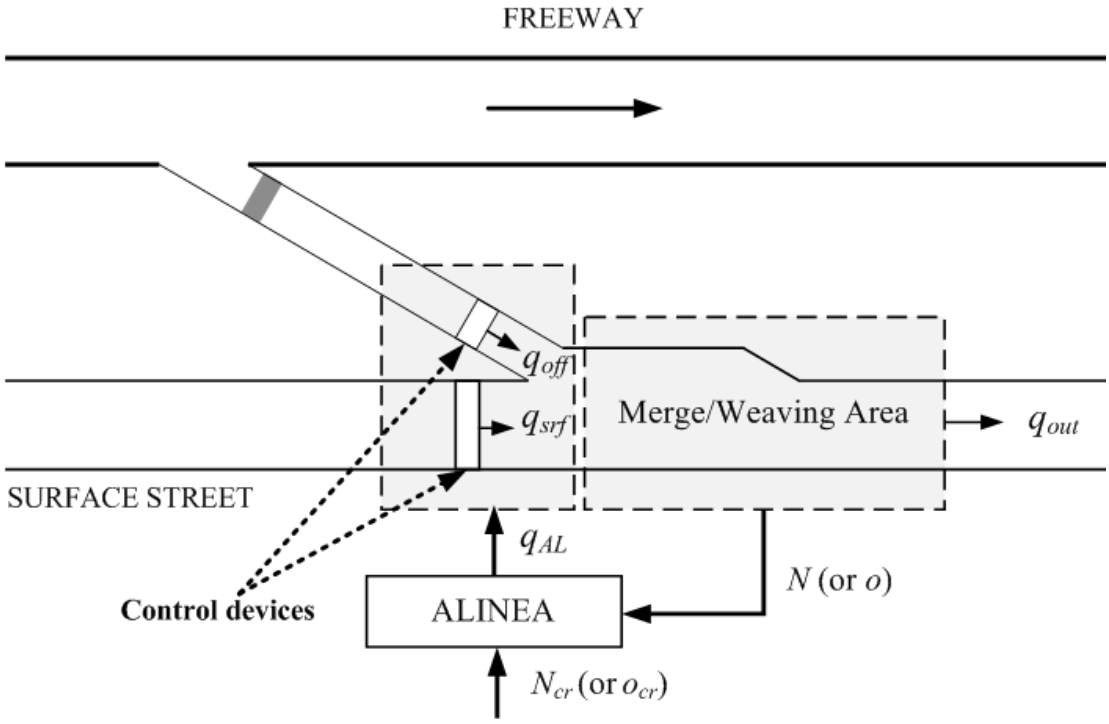

Figure 5.24 Illustration of ALINEA strategy operation.

proposed for mainstream merging traffic control, e.g. at toll-plaza and work-zone areas ([59], [60]). Figure 5.24 illustrates the operation of the ALINEA strategy. In particular, at every period (or control interval) T_c, ALINEA receives real-time measurements of the number of vehicles N (or occupancy measurements o) collected from the merge area and calculates the total flow that should enter in the merge area so that $N \approx N_{cr}$ (or $o \approx o_{cr}$) is maintained, thus maximizing the merge area outflow (see also Figure 5.23).

In this study, an extension of ALINEA is utilized; in particular, a proportional-integral (PI-type) regulator is used [66] which reads:

$$q_{AL}(k) = q_{AL}(k-1) - K_P[N(k) - N(k-1)] + K_I\left[\widehat{N} - N(k)\right] \qquad (5.7)$$

where $k = 1,2, \ldots$, is the discrete time index; $q_{AL}(k)$ is the controlled entering flow (in veh/h) in the merge area to be implemented during the new period k; $K_P > 0$ and $K_I > 0$ denote the regulator parameters for the proportional and integral

terms, respectively; $N(k)$ and $N(k-1)$ are the number of vehicles in the merge area at time kT_c and $(k-1)T_c$, respectively; and $\widehat{N}$ is a set (desired) value for the number of vehicles in the merge area, which may be set equal to the critical value N_{cr}, mentioned earlier, for maximum throughput. Note that the same equation can be used if the occupancy percentage o is measured instead of the number of vehicles N. The calculated $q_{AL}(k)$ is eventually truncated if it exceeds the range $[q_{min}, q_{max}]$, where q_{max} is the capacity of the controlled lanes and q_{min} is a minimum admissible flow value. In a potential field implementation, the estimation of the number of vehicles N in the merge area may be carried out by use of ordinary loop detectors placed at appropriate positions [55]. Alternatively, the occupancy measurements o, may be obtained from detectors placed at or upstream from the location where serious vehicle decelerations (congestion) appear first.

As presented above, at every control interval, ALINEA delivers the total flow q_{AL} to be implemented at the exit of the controlled lanes using appropriate control devices, e.g. traffic lights (see Figure 5.24). The question that arises is how this total flow should be distributed among the individual controlled lanes. There is no unique answer to this question, and the decision on the flow distribution policy may depend on both the infrastructure and demand characteristics and specific pursued control goals. In the investigated infrastructure, the ALINEA flow order is equally distributed among the surface street and off-ramp controlled lanes.

Queue Override strategy

Since the arriving demand exceeds the bottleneck capacity during the peak periods, the application of ALINEA may lead to the formation of queues at the controlled lanes. In order to avoid over-long queues, a Queue Override strategy may be employed that overrides the flow control decisions when and to the extent needed. Figure 5.25 presents the operation of the Queue Override strategy for the

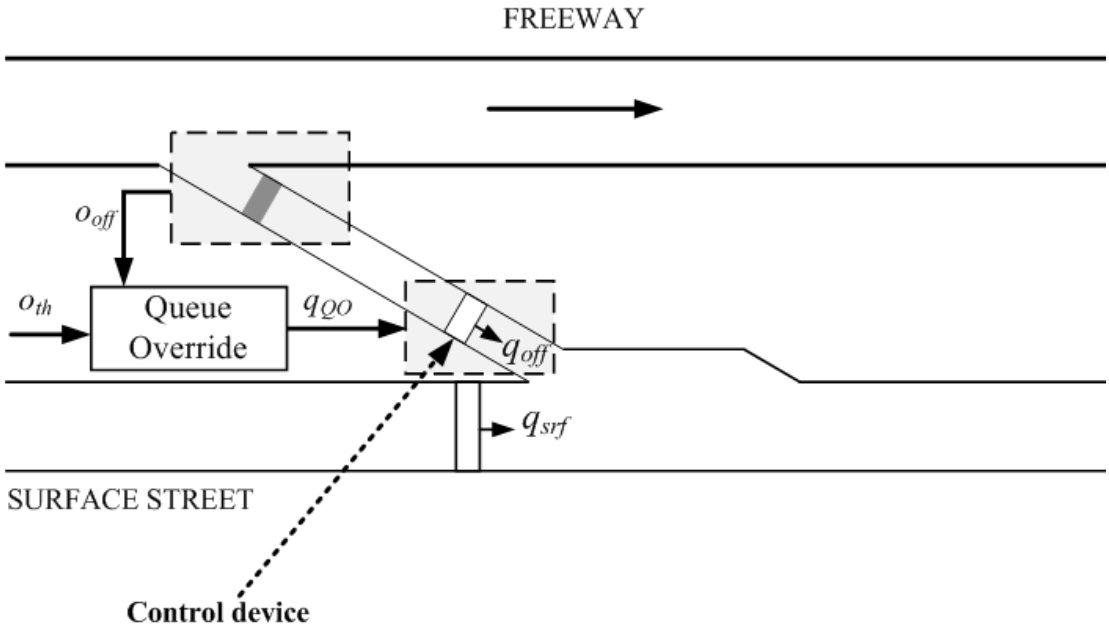

Figure 5.25 Illustration of the Queue Override strategy operation.

investigated implementation. Specifically, at every control interval T_c, the activation of the Queue Override strategy is decided based on occupancy measurements from a detector located close to the upstream end of the off-ramp. If the measured occupancy $o_{off}(k)$ exceeds a pre-specified threshold, o_{th}, the queue override exit flow, $q_{QO}(k)$, is set to a pre-specified high flow value, q_{over}; otherwise, the queue override exit flow is set to zero. This leads to

$$q_{QO}(k) = \begin{cases} q_{over}, & if\ \ o_{ff}(k) > o_{th} \\ 0, & otherwise. \end{cases} \tag{5.8}$$

In the current investigation, the Queue Override policy is only applied at the off-ramp controlled lanes and aims to prevent the off-ramp queue spill-back into the freeway mainstream and the creation of mainstream congestion. However, under a different policy, the application of queue management could also be decided for the surface street controlled lanes, in order to avoid over-long queues that could affect the upstream urban network. Moreover, instead of the Queue Override strategy, other queue management techniques could also be applied [56].

Final flow decision

Based on the decisions of the employed control strategies, the final flows to be implemented during the next control interval at the controlled lanes of the surface street, $q_{srf}(k)$, and the off-ramp, $q_{off}(k)$, are calculated as follows:

$$q_{srf}(k) = q_{AL}(k)\lambda_{srf}/\lambda_{total} \quad (5.9)$$

$$q_{off}(k) = \max\left\{\frac{q_{AL}(k)\lambda_{off}}{\lambda_{total}}, q_{QO}(k)\right\} \quad (5.10)$$

where λ_{srf} and λ_{off} are the number of controlled lanes at the surface street and the off-ramp, respectively, while λ_{total} is the total number of controlled lanes. Note that, as long as the occupancy, o_{off}, is below the threshold o_{th}, the total flow, q_{srf} plus q_{off}, ordered for implementation, is equal to the ALINEA flow order q_{AL}; while, when the occupancy o_{off} exceeds the threshold o_{th}, i.e. the Queue Override strategy is activated, the total flow ordered for implementation is higher than the ALINEA flow order, leading to a temporary increase of the number of vehicles in the merge area, given of course sufficient arriving demand. However, thanks to the feedback operation of ALINEA, the increase of the number of vehicles in the merge area is immediately detected, and ALINEA adapts its control decisions so that $N$is maintained close to the set-point $\widehat{N}$.

Translation of control decisions

The above calculated flows are implemented via appropriate operation of the control devices, i.e. traffic lights, located at the surface street and the off-ramp, upstream of the merge area, as shown in Figure 5.24. There are different possible metering policies to translate the flow decisions of the control algorithm into corresponding traffic light settings; e.g. one-car-per-green, n-cars-per-green, full traffic cycle, discrete release rates etc., see [67] for an overview. A full traffic cycle policy is employed here so as to maximize the resulting flow capacity of the traffic light, and also because it is more appropriate for urban street traffic lights. In

particular, for a fixed traffic cycle T_c, the final flow to be implemented at each traffic light, is translated into a corresponding green phase duration G_{srf} and G_{off} (in s), respectively, via the following equations

$$G_{srf}(k) = q_{srf}(k)T_c/S_{srf}\lambda_{srf} \quad (5.11)$$

$$G_{off}(k) = q_{off}(k)T_c/S_{0ff}\lambda_{off} \quad (5.12)$$

where S_{srf} and S_{off} (in veh/h) are the saturation flows (by lane) of the surface street and the off-ramp, respectively. Finally, the calculated green phase durations are truncated if they exceed pre-specified bounds $[G_{min}\,,\ G_{max}]$.

5.2.2 Network description and traffic demand pattern

The utilized network is a part of the Autopista Central and the adjacent surface street network, in Santiago, Chile, as presented in Figure 5.26. Based on the analysis of Gunther et al. [39], during the morning peak hours, congestion is created on the surface street network which propagates to the freeway mainstream through the saturated off-ramp. More specifically, the reason for congestion is the limited capacity of the surface street merge section, which may not accommodate both the freeway off-ramp exit flow and the surface street demand during the peak periods. Moreover, strong lane-changing maneuvers (weaving) are observed in the merge area as the majority of vehicles exiting the freeway wish to turn right at the junction located downstream of the merge area; while most of the vehicles moving on the surface street wish to turn left or go straight to access the freeway further downstream.

This real freeway stretch is utilized to test and demonstrate the application of the proposed traffic control algorithm, by use of microscopic simulation. Figure 5.27 presents the simulated network within the AIMSUN simulator [68]. The total length of the simulated freeway stretch is about 7 km, while the simulated surface street network is about 1.52 km and they are both sufficiently long to

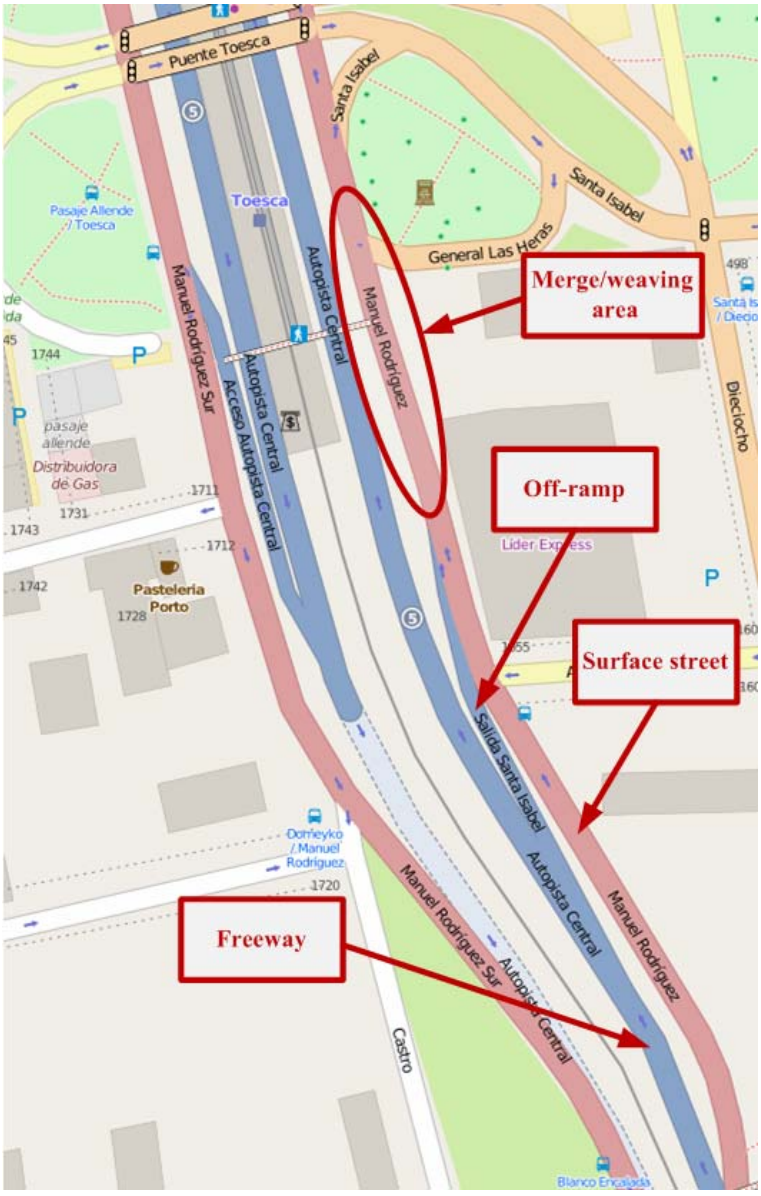

Figure 5.26 Examined freeway stretch, in Santiago, Chile.

accommodate any forming queue length. In order to collect real-time measurements, for control and monitoring purposes, detectors have been placed at several locations of the simulated network.

The utilized traffic demand scenario is stochastic. Figure 5.28 shows the considered average demand, which lasts for about 2.5 hours. In particular, the freeway traffic demand is trapezoidal, with maximum average flow at 4000 veh/h. The off-ramp demand, i.e. the portion of traffic flow that exits the freeway

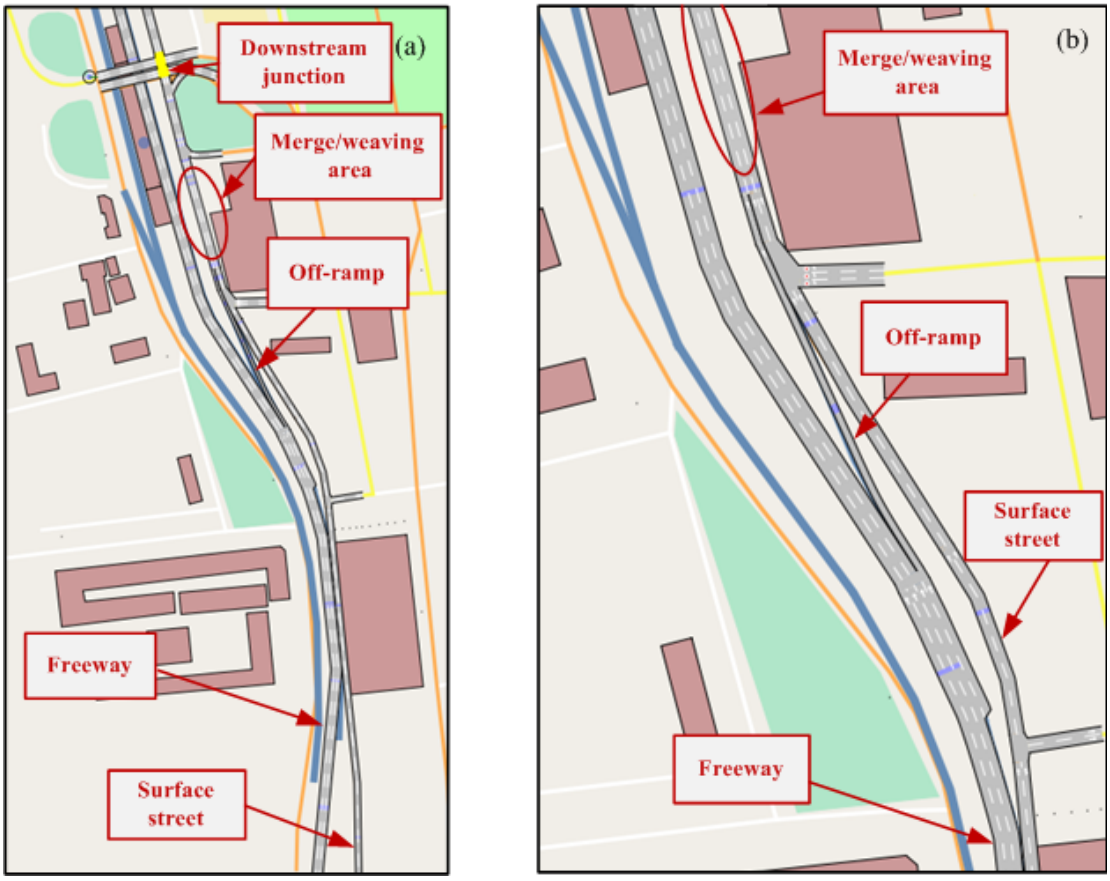

Figure 5.27 Network representation within AIMSUN; (a) network top view (b) off-ramp area top view.

through the off-ramp, is also trapezoidal with maximum average flow at 1500 veh/h, while the surface street traffic demand is constant for the whole simulation time, and equal to 1600 veh/h on average. Note that the above utilized maximum average values are similar to the corresponding real traffic demand, appearing during the morning peak hours, as presented in [39]. Moreover, for simplification, the surface street demand is entering the network only from the upstream end of the primary road, while no flow is introduced from the two secondary roads (see Figure 5.26).

The simulated traffic demand includes two vehicle types, i.e. cars and trucks. The trucks represent an average of 2% of the total freeway traffic demand, while 5% of trucks are included in the surface street traffic demand. The selected traffic

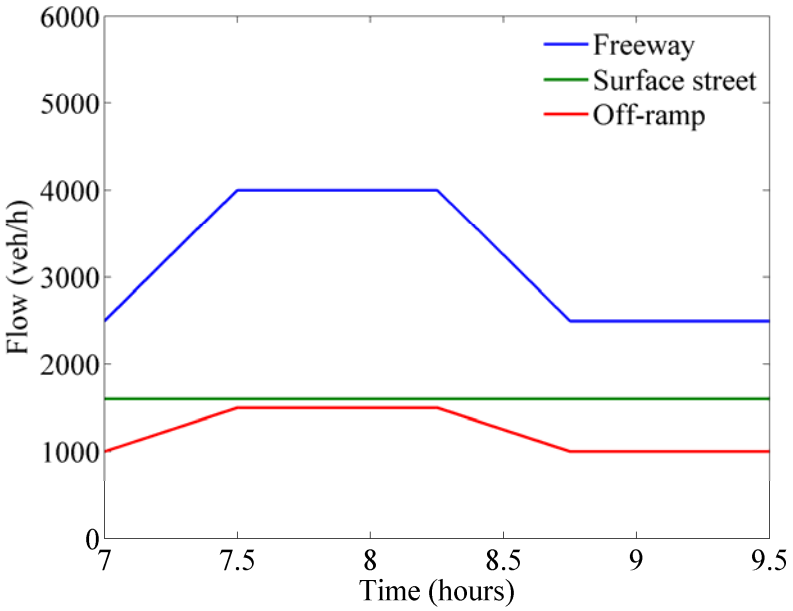

Figure 5.28 Average traffic demand at the examined network.

demand mimics the real traffic conditions, creating congestion on the surface street network, during the peak hours, which propagates to the freeway mainstream, and enables us to test the proposed control concept and demonstrate the potential benefits of its application.

5.2.3 Simulation investigations

The described infrastructure and traffic demand scenario were simulated by use of the microscopic simulator AIMSUN v.8.0, using a simulation step $T = 1\,s$. AIMSUN is a stochastic simulator, thus different simulation runs (replications), with different random seeds, may produce quite different results. To address this issue, 10 replications were carried out for each examined scenario, and the obtained results are presented here below. In the following, the no-control case is presented first, followed by the results of the proposed control algorithm.

Figure 5.29 No control case: (a) network top view (b) off-ramp area top view.

5.2.3.1 No control case

In the no-control case, as long as the total arriving demand at the merge area, i.e. the off-ramp and surface street demand, is low, the vehicles travel through the merge area without serious problems. When the demand increases (peak period) beyond the merge area capacity, vehicle merging conflicts are observed, that lead to vehicle decelerations and formation of congestion; first on the surface street network; and, soon after, also on the freeway mainstream due to the off-ramp queue spill-back, see Figure 5.29.

Figure 5.30(a) presents the average vehicle delay (AVD) (in s/veh/km) and Figure 5.30(b) the average harmonic speed (AHS) (in km/h) in the network for all 10 replications. It is observed that the resulting mean AVD of the ten replications is equal to 24.3 s/veh/km while the mean AHS is equal to 51.9 km/h. Within the

investigated network, there exist three different traffic groups. The first group includes the traffic flow entering the network from the upstream origin of the freeway and exiting the network from the downstream end of the freeway; the second group includes the traffic flow entering the network from the upstream origin of the freeway and exiting the network from the downstream end of the surface street network, using the off-ramp; and the third group includes the traffic flow entering the network from the upstream origin of the surface street network and exiting the network from the downstream end of the surface street network.

Figure 5.30(c) and Figure 5.30(d) present the mean values (over 10 replications) of the AVD and AHS for the whole network and, also, for all three traffic groups. It is observed that the second traffic group, including vehicles traveling from the freeway to the surface street network, face the biggest mean delay (35.1 s/veh/km); the first traffic group, including vehicles traveling on the freeway, also face significant mean delay (24.0 s/veh/km) due to the formation of mainstream congestion; while the third traffic group, including vehicles traveling on the surface street network, face relatively lower mean delay (17.1 s/veh/km).

Figure 5.31(a) and Figure 5.31(b) present the number of vehicles on the surface street merge area and the outflow from the merge area, for a particular replication (replication #2) with AVD = 25.6 s/veh/km, which is very close to the mean AVD value of the 10 replications. It is observed that between 7:00 a.m. and 7:35 a.m. the number of vehicles N in the merge area is slowly increasing (as a consequence of the increasing demand), while the merge area outflow follows the increase of arriving demand. At around 7:40 a.m., the number of vehicles in the merge area increases steeply, resulting in serious merging conflicts and congestion, and this situation becomes stationary until about 9:10 a.m. The outflow and the number of vehicles in the merge area, during this time period, are about 2850 veh/h and 29 veh, on average, respectively. After 9:15 a.m., when the queue dissolves, the

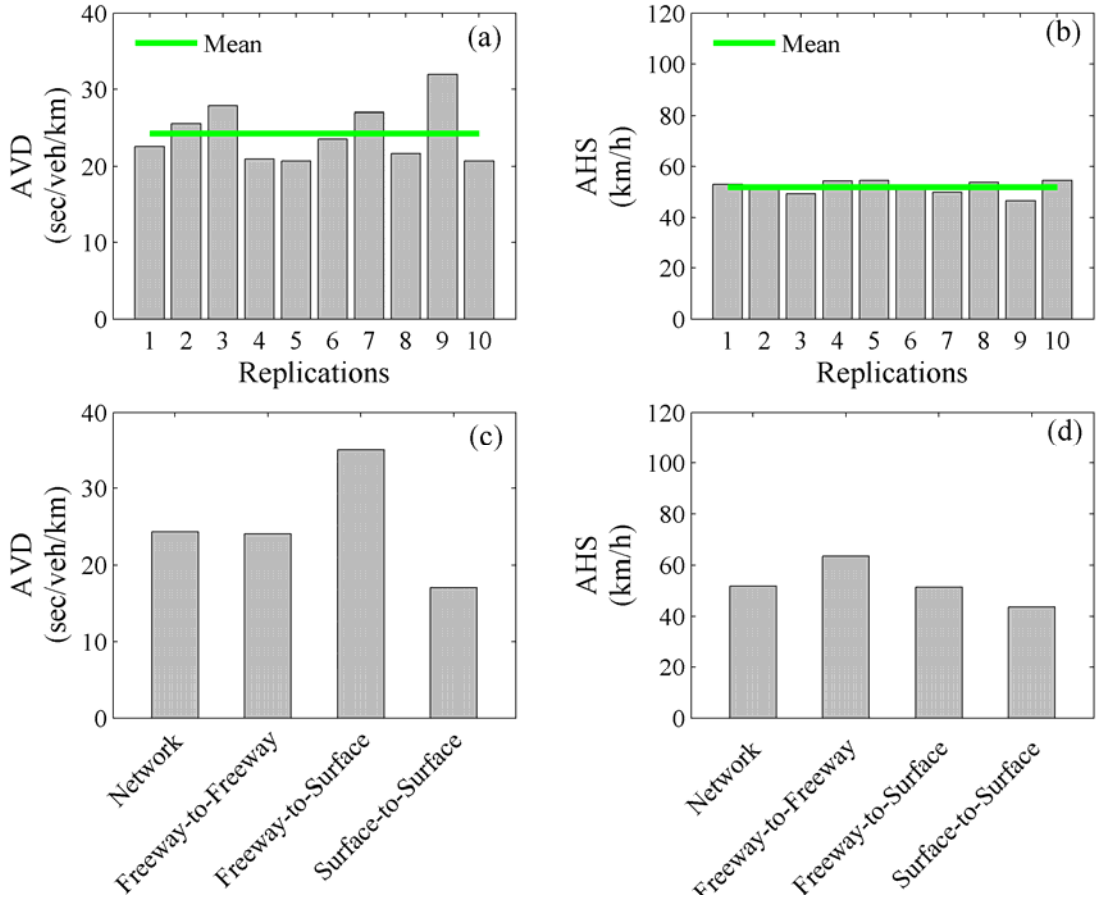

Figure 5.30 No control case: (a) Average vehicle delay for 10 replications; (b) Average harmonic speed for 10 replications; (c) mean average vehicle delay for the network and all three traffic groups; (d) mean average harmonic speed for the network and all three traffic groups.

number of vehicles in the merge area drops, and the outflow reduces to lower values due to the decreased demand.

Figure 5.32 illustrates, for the same replication, the space-time diagram of speed on the freeway. Considering that the value zero on the y-axis corresponds to the location of the off-ramp, it is observed that a speed breakdown appears upstream of the off-ramp, at around 7:55 a.m., due to the congestion on the surface street network and the off-ramp queue spill-over. The created congestion propagates upstream for about 1.6 km, lasts up to 9:05 a.m., after which the congestion

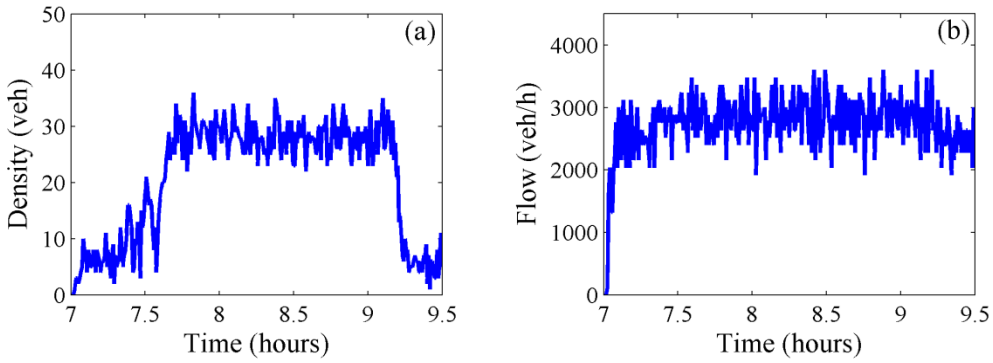

Figure 5.31 No control case: (a) Density on surface street merge area; (b) outflow from the surface street merge area.

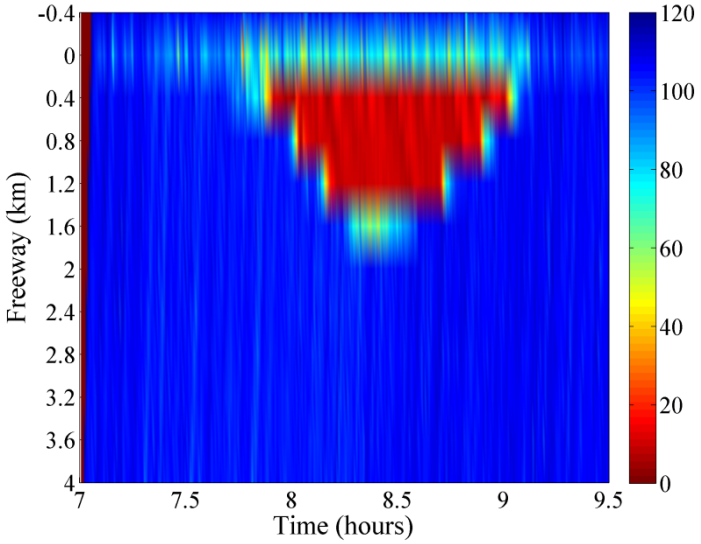

Figure 5.32 No control case. Space-time diagram of measured speed on the freeway.

dissolves due to the decreased freeway mainstream demand. Similar results are obtained for the other replications.

5.2.3.2 Merging traffic control case

The proposed control concept, as described in Section 5.2.1.2, is applied to the investigated network. In particular, the ALINEA strategy is activated every $T_c = 30$ s and receives real-time measurements of the number of vehicles N included in the surface street merge area. The regulator parameters, K_P and K_I, were manually fine-tuned using some practical trial-and-error rules from Control

Engineering and were set equal to 110 h^{-1} and 80 h^{-1}, respectively; once appropriate values have been found, the regulator is known to be little sensitive to related moderate variations [59]. The set-value $\widehat{N}$ for the number of vehicles in the merge area, should be selected such that the merge area outflow is maximized according to Figure 5.23. In the following sections the impact of the $\widehat{N}$ value is thoroughly investigated. Finally, the calculated flow q_{AL} is not allowed to exceed the range [600, 4800] veh/h, i.e. a minimum and maximum flow, respectively.

The Queue Override policy is also implemented every $T_c = 30$ s and takes its decisions based on occupancy measurements located close to the upstream end of the off-ramp. The pre-specified occupancy threshold, o_{th}, was set equal to 25%, and the high off-ramp exit flow value, q_{over}, was set to 1600 veh/h, i.e. equal to the off-ramp capacity.

The estimated final flows q_{srf} and q_{off} (see equations (5.9)-(5.10)), are implemented through the control devices, i.e., two traffic lights, placed upstream of the merge area, at the surface street and the off-ramp, respectively, as shown in Figure 5.33. As described in Section 5.2.1.2, a full traffic cycle policy is employed, with a fixed traffic cycle which is equal to the control interval T_c. The green phase duration for each traffic direction is calculated using equations (5.11)-(5.12), where $S_{srf} = 1600$ veh/h, $S_{off} = 1600$ veh/h, $\lambda_{srf} = 2$ and $\lambda_{off} = 1$. Moreover, the ALINEA minimum flow order $q_{min} = 600$ veh/h, results in a minimum green phase duration $G_{min} = 4$ s, while the maximum green phase G_{max} is equal to the traffic cycle, i.e. $G_{max} = 30$ s (all green). Furthermore, in case that the calculated green phase is in the range [28, 30) s, then, for safety reasons, the green phase duration is set to 28 s, so that the drivers will not face a very short red-phase duration, less than 2 s. Finally, the traffic lights, of the surface street network and the off-ramp operate with an offset at their traffic cycle start, in order to enable (to the extent possible) a continuous flow and reduce simultaneous vehicle departures

Figure 5.33 Control case: (a) network top view; (b) off-ramp area top view.

(or no departures during red). The implementation of the above control strategy was done via the AIMSUN API (Application Programming Interface), which allows the user to emulate a real-time control environment.

As already mentioned, the choice of the set-point $\widehat{N}$ value for ALINEA controller is crucial for maximizing the merge area outflow, which also corresponds to minimization of delays. In a field investigation, this may be achieved by gradually incrementing $\widehat{N}$ and monitoring the measured outflow, until a maximum throughput is obtained. In the current investigations, a series of simulation experiments were carried out using different (integer) $\widehat{N}$ values within the range $\widehat{N} \in [14, 30]$ veh and the corresponding mean AVD values were obtained. Figure 5.34 displays, for every investigated $\widehat{N}$ value, the corresponding AVD values for

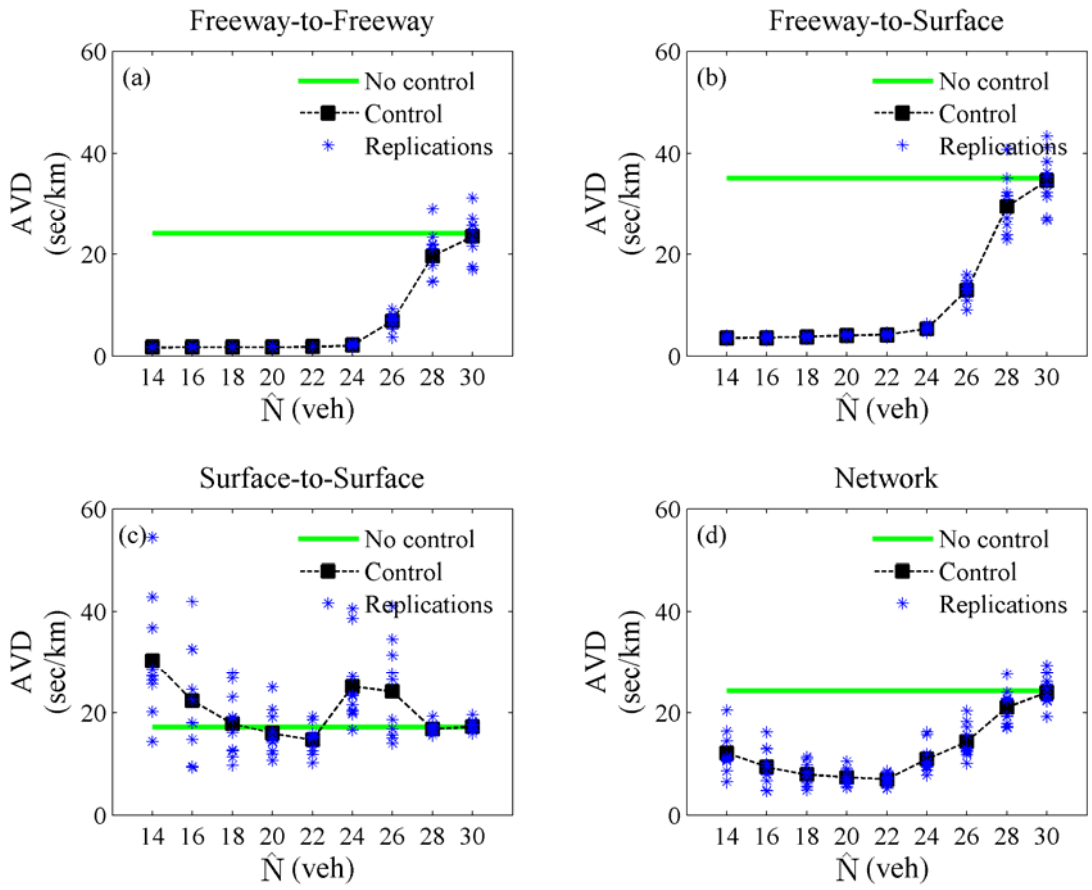

Figure 5.34 Control case. Average vehicles delay versus $\hat{N}$ for: (a) the whole network; (b) the first traffic group; (c) the second traffic group; and (d) the third traffic group.

the 10 replications as well as the mean AVD of all replications. Moreover, the mean AVD value for the no-control case is also displayed on the same figure for comparison.

Figure 5.34(a)–Figure 5.34(c) present the mean AVD values for all three traffic groups. It is observed that, for the first and the second traffic groups (Figure 5.34(a) and Figure 5.34(b)), the mean AVD is minimum for $\hat{N}$ values in the range $[14, 24]$ veh. This is because the utilized control algorithm prevents the formation of freeway congestion for any set-point value within this range, thanks to the operation of the Queue Override strategy. For high set-point values, the number of vehicles in the merge area is too high, creating congestion and reduced throughput on the surface street merge area. In this case, the operation of the

Queue Override policy is not sufficient to prevent the off-ramp queue spill-over and the formation of congestion on the freeway. Finally, for $\widehat{N}$ equal to 30 veh, the mean AVD value is equal to the mean AVD value of the no-control case, as the system behaves similarly to the case where no control is applied.

Figure 5.34(c) shows the mean AVD values for all investigated $\widehat{N}$ values, for the third traffic group. It is observed that for low $\widehat{N}$ values, e.g. in the range $[14, 18]$ veh, the mean AVD value is high, and actually higher or equal to the mean AVD of the no-control case. This is because the system operates at under-critical conditions (see Figure 5.23) and the surface street merge area "starves for flow"; for $\widehat{N}$ values in the range $[20, 22]$ veh, the mean AVD is minimized, and, particularly for $\widehat{N}$ equal to 22 veh, it takes the lowest value, which corresponds to the critical value mentioned earlier. Note that for this range of $\widehat{N}$ values the mean AVD is actually lower than the mean AVD of the no control case. For higher set-points, e.g. in the range $[24, 26]$, the merging conflicts are increasing, leading to congestion and reduced throughput. Furthermore, for even higher $\widehat{N}$ values the traffic conditions are similar to the no control case (which corresponds to 29 veh in the merge area on average, as shown in Figure 5.31), leading to similar mean AVD values. It is observed, that the operation of the control strategy for set-point values in the range [24, 26] leads to higher mean AVD values compared to the no control case, although in both cases the merge area outflow is similarly low. This is because the operation of the strategy results in queue formation upstream of the traffic lights which causes extra delays to the surface-street drivers compared to the no-control case, where no queues are formed.

Finally, Figure 5.34(d) shows that the mean AVD value for the network is minimized for $\widehat{N}$ values in the range $\widehat{N} \in [20, 22]$ veh, and, particularly for $\widehat{N}$ equal to 22 veh, it takes the lowest value. It is also noteworthy that the mean AVD for almost all investigated $\widehat{N}$ values is significantly lower than the corresponding

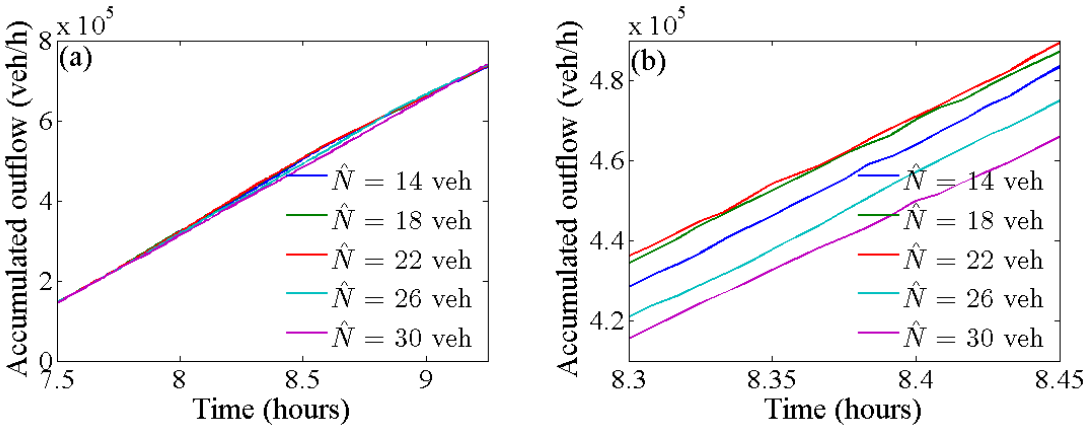

Figure 5.35 Control case. Accumulated merge area outflow for various set-points $\hat{N}$: (a) during peak period; (b) for a shorter period within the peak period.

mean AVD value of the no-control scenario; while for very high $\hat{N}$ values, e.g. over 28 veh, the traffic conditions are similar to the no control case resulting to similar AVD value.

In addition to the minimization of mean delays for $\hat{N}$ values in the range $\hat{N} \in [20, 22]$ veh, Figure 5.35 verifies that for $\hat{N} = 22\ veh$ the maximization of the merge area outflow is achieved. In particular, Figure 5.35(a) presents the accumulated outflow from the merge area during the peak period for various set-point values and Figure 5.35(b) shows the same accumulated merge area outflows, zooming on a shorter period within the peak period. It is observed that the outflow from the merge area is maximized for $\hat{N} = 22\ veh$.

Figure 5.36 (a) presents the AVD and Figure 5.36(b) the AHS in the network for all 10 replications, for $\hat{N} = 22$ veh. It is observed that the resulting mean AVD of the ten replications is equal to 6.9 s/veh/km; while the mean AHS is equal to 69.0 km/h. Figure 5.36(c) and Figure 5.36(d) present the mean values (over 10 replications) of the AVD and AHS for the whole network and, also, for all three traffic groups, for $\hat{N} = 22$ veh. It is observed that the third traffic group, including vehicles traveling on the surface street network, face the biggest mean delay (14.7

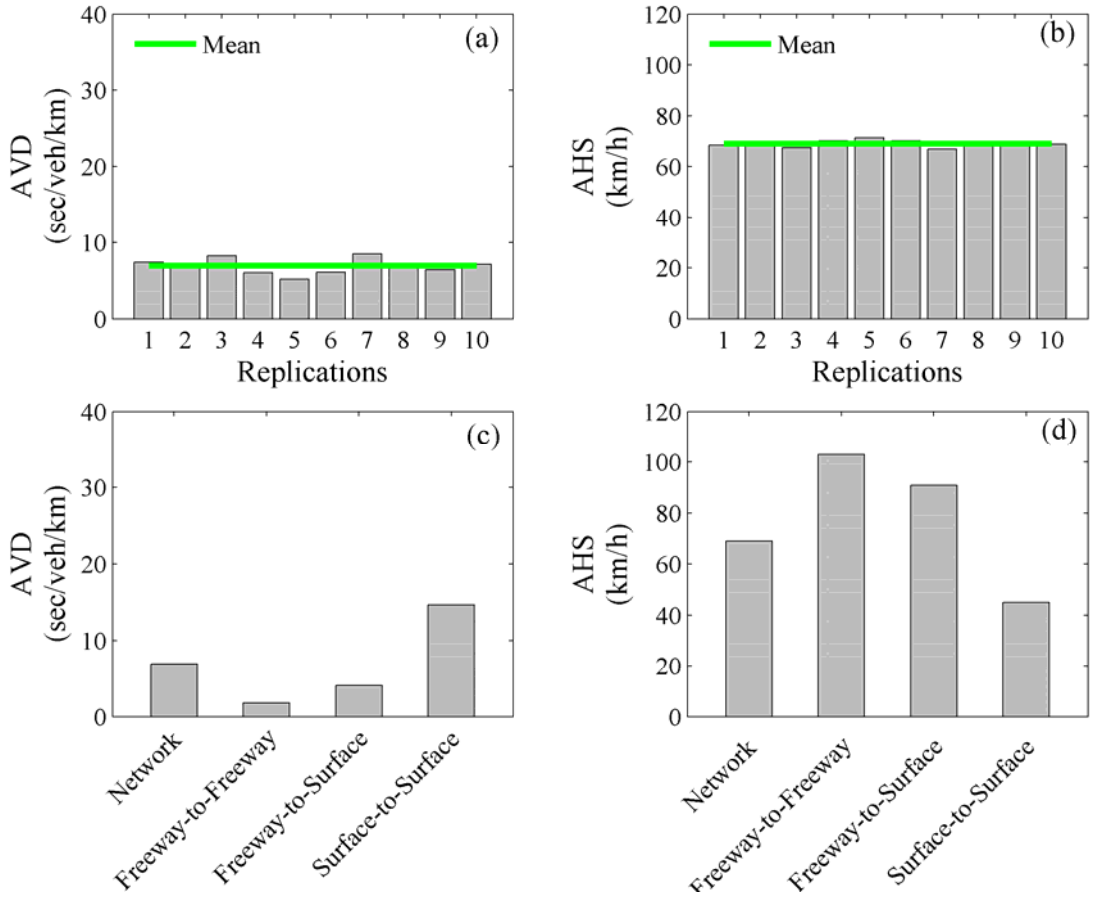

Figure 5.36 Control case ($\widehat{N}$=22 veh): (a) Average vehicle delay for 10 replications; (b) Average harmonic speed for 10 replications; (c) mean average vehicle delay for the network and all three traffic groups; (d) mean average harmonic speed for the network and all three traffic groups.

s/veh/km), while the first and the second traffic groups experience significantly lower mean delays (1.8 and 4.1 s/veh/km, respectively).

Table 5-4 summarizes the obtained results and compares the mean AVD of the control (for $\widehat{N} = 22$ veh) and the no-control cases. It is shown here, that the application of the proposed control algorithm improves the average vehicle delay for the whole network, achieving 71.6% reduction of the mean AVD, but also for each of the three traffic groups, and in particular for the first and second traffic group for which it achieves 92.5% and 88.3% reduction of the mean AVD, respectively, compared to the no control case. Finally, regarding the third traffic

Table 5-4 Mean AVD (in s/veh/km) for the no-control and control case.

	Network	Freeway-to-Freeway	Freeway-to-Surface	Surface-to-Surface
No Control	24.3	24.0	35.1	17.1
Control (set-point = 22 veh)	6.9	1.8	4.1	14.7
% Difference	-71.6	-92.5	-88.3	-14.0

groups, it is shown here that, although the reduction of the mean AVD is lower, compared to the other two traffic groups, the operation of the proposed control strategy benefits also the surface street users, reducing their mean AVD by 14%.

Figure 5.37 (a) and Figure 5.37 (b) present the number of vehicles on the surface street merge area and the outflow from the merge area, using $\widehat{N} = 22$ veh, for a particular replication (replication #2) with AVD = 6.9 s/veh/km, which is equal to the mean AVD value of the 10 replications. It is observed that, during the peak period, the controller manages to maintain, on average, the number of vehicles in the merge area around the set-point value, and, by doing so, the merge area outflow is maximized, reaching 3200 veh/h, on average, for some time periods. Moreover Figure 5.37 (c) and Figure 5.37 (d) display the queue (in veh) formed upstream of the surface street and the off-ramp traffic lights over the simulation time. It is shown that the operation of ALINEA control strategy leads to the temporary formation of a queue on the surface street, during the peak period, which, however never exceeds the 45 vehicles. Regarding the off-ramp, it is seen that during the peak period there is an occasional queue formation, which never exceeds the available off-ramp storage space, thanks to the activation of the Queue Override strategy.

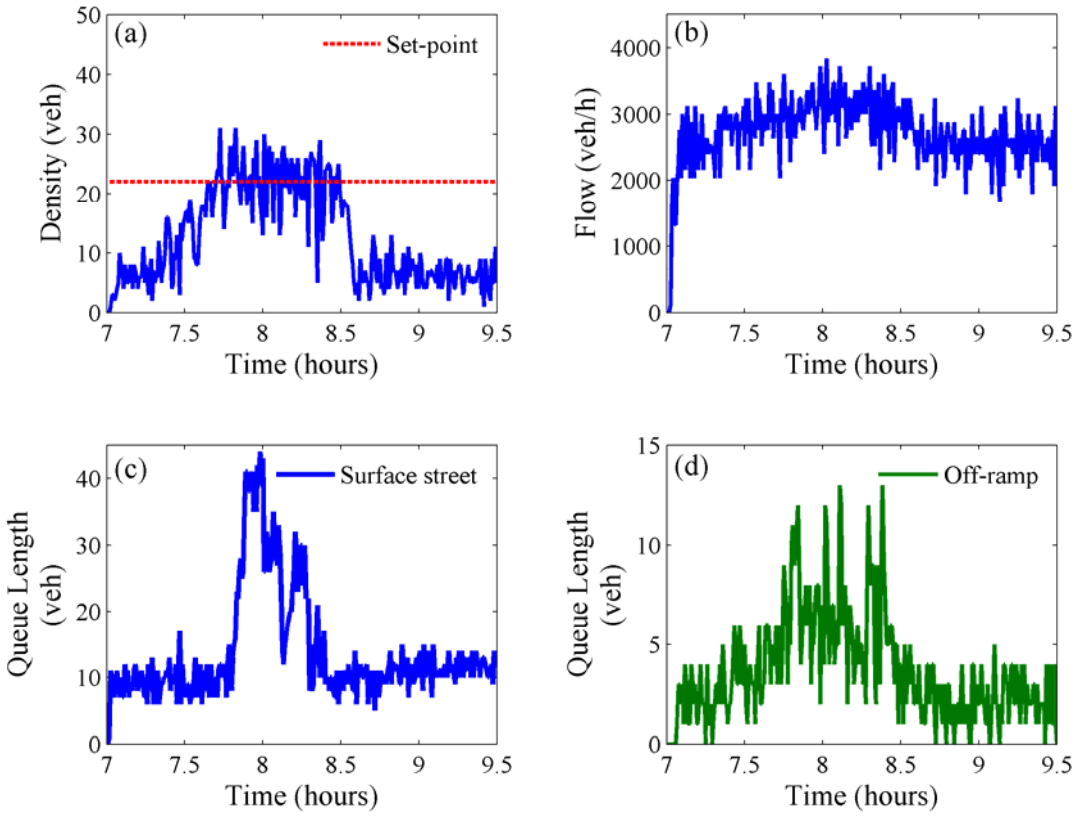

Figure 5.37 Control case ($\widehat{N}$=22 veh): (a) Density on surface street merge area; (b) outflow from the surface street merge area; queue length on (c) the surface street and (d) the off-ramp.

Figure 5.38 displays the flow calculated by the ALINEA control strategy (using equation (5.7)), the time-periods of the Queue-Override strategy activation (based on equation (5.8)), and the total flow order (q_{srf} plus q_{off}) that is calculated by the control algorithm for implementation (using equation (5.9) plus (5.10)), over the whole simulation time. Moreover, the total flow that is actually exiting the surface street and off-ramp traffic lights is also displayed for comparison. It is observed that as long as the Queue Override policy is not activated, the total flow that is calculated by the control algorithm for implementation is equal to the ALINEA flow order; while at the time periods that the Queue Override is activated, the total flow to be implemented is higher than the ALINEA flow order. Moreover, it is observed that, during the peak period, the total flow that is actually

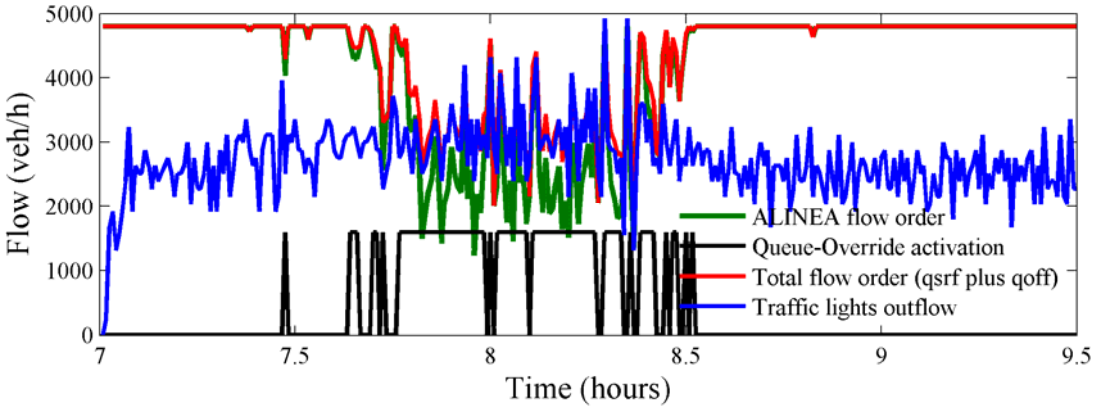

Figure 5.38 Control case ($\widehat{N}$=22 veh). ALINEA flow order, Queue Override activation, final flow order and traffic lights outflow over time.

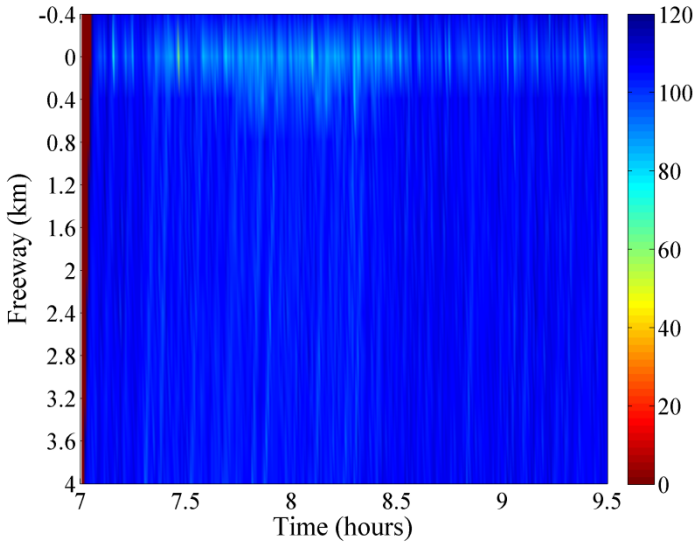

Figure 5.39 Control case ($\widehat{N}$=22 veh). Space-time diagram of measured speed on the freeway.

exiting the traffic lights is equal to the total flow calculated by the control algorithm, while for the rest time it is equal to the arriving traffic demand.

Figure 5.39, illustrates, for the same replication, the space-time diagram of speed on the freeway. It is shown here that the operation of the employed control strategy prevents the off-ramp queue spill-over and the formation of mainstream freeway congestion. Note that all ten replications provide similar results.

5.2.4 Conclusions

Section 5.2 addresses the problem of freeway congestion due to an over-spilling off-ramp. In particular, a control framework was proposed, which aims to maximize the surface street merge area outflow and at the same time to prevent the off-ramp queue spill-over into the freeway mainstream and the resulting freeway congestion. The proposed control concept was demonstrated via microscopic simulation, using a real traffic network. The simulation results showed that the proposed control algorithm may improve the prevailing traffic conditions, preventing the formation of congestion and benefiting both the freeway drivers and the surface street users.

6 Conclusions and Future Work

This final chapter summarizes the findings and results of this thesis. In particular, Section 6.1 gives a summary of the study and highlights the main results and contributions of the thesis, while Section 6.2 indicates future research aspects that could be considered to extend the investigation results.

6.1 Concluding remarks

Traffic congestion originating from off-ramp areas is a particular, but quite frequent case of (recurrent) congestion, appearing usually at urban or peri-urban freeways during the peak periods. This kind of congestion is difficult to deal with, and for this reason this frequent traffic flow degradation is rarely addressed in the traffic control literature. Moreover, within the traffic flow modeling literature there are, so far, no studies undertaking validation and comparison of different traffic flow models regarding the reproduction of traffic conditions at congested freeway off-ramp areas. The emergence of traffic flow models that are able to reproduce such cases with satisfactory accuracy is deemed important as it may trigger the development of innovative traffic control strategies that face this particular type of freeway congestion. This gap in the literature was addressed within this thesis, which focuses on traffic flow modeling and traffic control issues for congested freeway off-ramp areas.

In particular, in the first part of the thesis (Chapter 3–Chapter 4), the two most popular space-time discrete macroscopic traffic flow models, namely the CTM and the METANET models, were compared regarding the representation of traffic conditions at congested freeway off-ramp areas. The models were first calibrated using real traffic data from Attiki Odos freeway in Athens, and by employing three different optimization methods; i.e. the deterministic Nelder-Mead

algorithm, the stochastic genetic algorithm and the stochastic cross-entropy method. Then, the resulted models were tested in terms of sensitivity to their parameter values and were also validated and compared using different traffic data sets from the same freeway site.

The calibration results showed that all three optimization methods estimated similar parameter values for the CTM and also the METANET model, which achieve satisfactory reproduction of the network traffic conditions for the calibration date. Moreover the sensitivity investigations showed that the CTM model is sensitive to the parameters related to the characteristics of the saturated off-ramp, while the METANET model seems to be more sensitive to the parameters of the fundamental diagram (FD). Finally, the validation of the CTM and METANET models indicated that they are both able to reproduce the traffic conditions of the network also for other dates, with the METANET model offering a more accurate representation of the prevailing traffic conditions.

The second part of the thesis (Chapter 5) includes the development and testing of innovative real-time traffic control measures for congested freeway off-ramp areas. In particular, two different cases were examined and suitable traffic control strategies were proposed for every case. In the first case, a hypothetical network was simulated, by use of the macroscopic traffic flow model selected in the first part of the thesis (METANET), and various route diversion strategies were developed that aim to reroute the drivers through alternative routes, towards the same destination, preventing the off-ramp queue spillover and the creation of mainstream congestion. The simulation results showed that, in all investigated cases, the proposed policies succeed in maintaining the off-ramp queue length within the off-ramp bounds, thus improving the traffic conditions on the freeway substantially, compared to the case that no route diversion is applied to the network.

In the second case, a real traffic network was examined where recurrent freeway congestion is created due to congestion on the surface street network which propagates to the freeway mainstream through a saturated off-ramp. The network was simulated by use of microscopic simulation and a real-time merging traffic control algorithm was proposed that aims to maximize the surface street network throughput and at the same time to prevent the off-ramp queue spill over into the freeway mainstream. The simulation results showed that the proposed control algorithm may improve the prevailing traffic conditions, preventing the formation of congestion and benefiting both the freeway drivers and the surface street users.

Considering the above, the main contributions of the thesis can be summarized as follows:

- The two most popular macroscopic traffic flow models (the CTM and METANET model) were validated and compared regarding the reproduction of traffic conditions at congested freeway off-ramp areas. The validation results showed that both models are able to reproduce the traffic conditions in such networks, with the METANET model offering a more accurate representation of the prevailing traffic conditions. Moreover the sensitivity analysis offered a better understanding of the models when applied to this type of networks.
- Two different cases of congested freeway off-ramp areas were examined and innovative traffic control measures were proposed for each investigated case. The simulation results showed that in both cases the proposed traffic control strategies manage to prevent the off-ramp queue spill-over and the creation of mainstream congestion thus they are both very promising in case of potential field implementation.

6.2 Further research

There are various ways to extend the investigations presented in this thesis. Few of them are listed in the following:

- The utilized traffic flow models, and in particular the CTM model can be extended and improved to increase the achieved accuracy, e.g. using different fundamental diagram (FD).
- More macroscopic traffic flow models can be employed and compared against the utilized models.
- The first examined traffic control case, for congested freeway off-ramp areas, can be extended in order to account for multiple routes considering more off-ramps located further downstream or further upstream.
- In the second traffic control case, a bigger surface street network can be considered taking also into account possible restrictions that may apply due to signalized junctions. Moreover real traffic data can be utilized instead of a hypothetical demand scenario.
- Finally, field trial of the proposed traffic control strategies would provide more evidence about the achievable level of benefits.

7 Bibliography

[1] A. Kotsialos and M. Papageorgiou, “The importance of traffic flow modeling for motorway trafficc control,” *Networks and Spatial Economics*, vol. 1, no. 1–2, pp. 179–203, 2001.

[2] M. J. Lighthill and G. B. Whitham, “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*, 1955, vol. 229, no. 1178, pp. 317–345.

[3] P. I. Richards, “Shock waves on the highway,” *Operations Research*, vol. 4, no. 1, pp. 42–51, 1956.

[4] J. Treiterer and J. A. Myers, “The hysteresis phenomena in traffic flow,” in *Proceedings of the 6th International Symposium on Transportation and Traffic Theory, (D.J. Buckley Ed.)*, 1974, pp. 13–38.

[5] M. Papageorgiou, “Some remarks on macroscopic traffic flow modelling,” *Transportation Research Part A: Policy and Practice*, vol. 32, no. 5, pp. 323–329, Sep. 1998.

[6] G. F. Newell, “Instability in dense highway traffic, a review,” in *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow, (J. Almond Ed.)*, 1965, pp. 73–83.

[7] H. J. Payne, “Models of freeway traffic and control,” *Mathematical models of public systems. Simulation Council Proceedings*, vol. 1, pp. 51–61, 1971.

[8] C. Daganzo, “Requiem for second-order fluid approximations of traffic flow,” *Transportation Research Part B: Methodological*, vol. 29, no. 4, pp. 277–286, 1995.

[9] W. F. Phillips, “A kinetic model for traffic flow with continuum implications,” *Transportation Planning and Technology*, vol. 5, pp. 131–138, Mar. 1979.

[10] R. Kühne, “Macroscopic freeway model for dense traffic - stop-start waves and incident detection,” in *Proceedings of the 9th International Symposium on Transportation and Traffic Theory*, 1984, pp. 21–42.

[11] B. Kerner and P. Konhäuser, “Cluster effect in initially homogeneous traffic flow,” *Physical Review E*, vol. 48, no. 4, pp. 2335–2338, 1993.

[12] B. Kerner and P. Konhäuser, “Structure and parameters of clusters in traffic flow,” *Physical Review E*, vol. 50, no. 1, pp. 54–83, 1994.

[13] M. Papageorgiou, J.-M. Blosseville, and H. Hadj-Salem, “Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris: Part I: Modelling,” *Transportation Research Part A: General*, vol. 24, no. 5, pp. 345–359, 1990.

[14] H. M. Zhang, “A theory of nonequilibrium traffic flow,” *Transportation Research Part B: Methodological*, vol. 32, no. 7, pp. 485–498, 1998.

[15] M. Treiber, A. Hennecke, and D. Helbing, “Derivation, properties, and simulation of a gas-kinetic-based, nonlocal traffic model,” *Physical Review E*, vol. 59, no. 1, pp. 239–253, Jan. 1999.

[16] C. F. Daganzo, “The cell transmission model. Part I: a simple dynamic representation of highway traffic,” *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994.

[17] C. Daganzo, “The cell transmission model, part II: network traffic,” *Transportation Research Part B: Methodological*, vol. 29B, no. 2, pp. 79–93, 1995.

[18] J. P. Lebacque, “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 (Lesort, J. B., Editor)*, 1996, pp. 647–678.

[19] H. J. Payne, “FREFLO: A macroscopic simulation model of freeway traffic,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 722, pp. 68–77, 1979.

[20] A. Messmer and M. Papageorgiou, “METANET: A macroscopic simulation program for motorway networks,” *Traffic Engineering and Control*, vol. 31, no. 8–9, pp. 466–470, 1990.

[21] M. Papageorgiou, I. Papamichail, A. Messmer, and Y. Wang, “Traffic simulation with METANET,” in *Fundamentals of Traffic Simulation*, J. Barceló, Ed. New York, NY: Springer, 2010, pp. 399–430.

[22] M. S. Grewal and H. J. Payne, "Identification of parameters in a freeway traffic model," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 6, no. 3, pp. 176–185, 1976.

[23] M. Cremer and A. D. May, "An extended traffic flow model for inner urban freeways," in *Proceedings of the 5th IFAC/IFIP/IFORS International Conference on Control in Transportation Systems*, 1986, pp. 383–388.

[24] D. Helbing, "Derivation and empirical validation of a refined traffic flow model," *Physica A: Statistical Mechanics and its Applications*, vol. 233, no. 1, pp. 253–282, 1996.

[25] K. K. Sanwal, K. Petty, J. Walrand, and Y. Fawaz, "An extended macroscopic model for traffic flow," *Transportation Research Part B: Methodological*, vol. 30, no. 1, pp. 1–9, Feb. 1996.

[26] A. Kotsialos, M. Papageorgiou, C. Diakaki, Y. Pavlis, and F. Middelham, "Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 4, pp. 282–292, 2002.

[27] L. Muñoz, X. Sun, D. Sun, G. Gomes, and R. Horowitz, "Methodological calibration of the cell transmission model," *In Proc. of the American Control Conference*, vol. 1, pp. pp. 798–803, 2004.

[28] D. Ngoduy, S. Hoogendoorn, and H. Van Zuylen, "Comparison of numerical schemes for macroscopic traffic flow models," *Transportation Research Record*, vol. 1876, no. 1, pp. 52–61, Jan. 2004.

[29] T. Monamy, H. Haj-Salem, and J.-P. Lebacque, "A macroscopic node model related to capacity drop," in *Proceedings of the 15th Meeting of the EURO Working Group on Transportation*, 2012, vol. 54, pp. 1388–1396.

[30] D. Ngoduy and M. J. Maher, "Calibration of second order traffic models using continuous cross entropy method," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 102–121, Oct. 2012.

[31] A. Poole and A. Kotsialos, "METANET model validation using a genetic algorithm," in *Proceedings of the 13th IFAC Symposium on Control in Transportation Systems*, 2012, vol. 13, no. 1, pp. 7–12.

[32] M. Cremer and M. Papageorgiou, "Parameter identification for a traffic flow model," *Automatica*, vol. 17, no. 6, pp. 837–843, 1981.

[33] M. Papageorgiou, J. M. Blosseville, and H. Hadj-Salem, “Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris,” *Transportation Research Part B: Methodological*, vol. 23, no. 1, pp. 29–47, 1989.

[34] P. G. Michalopoulos, P. Yi, and A. S. Lyrintzis, “Continuum modelling of traffic dynamics for congested freeways,” *Transportation Research Part B: Methodological*, vol. 27, no. 4, pp. 315–332, Aug. 1993.

[35] C. Daganzo, J. Laval, and J. Muñoz, “Ten strategies for freeway congestion mitigation with advanced technologies,” UCB-ITS-PRR-2002-3, 2002.

[36] J. Rudjanakanoknad, “Capacity change mechanism of a diverge bottleneck,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2278, no. -1, pp. 21–30, Dec. 2012.

[37] X. Di, X. Zhang, H. Zhang, and H. Liu, “Application of pavement marker to avoid queue-jumping and traffic spillback at off-ramp of expressways,” *Transportation Research Board 92nd Annual Meeting*, no. 13–3298, 2013.

[38] Z. Tian, K. Balke, R. Engelbrecht, and L. Rilett, “Integrated control strategies for surface street and freeway systems,” *Transportation Research Record: Journal of the Transportation Research Board*, no. 1811, pp. 92–99, 2002.

[39] G. Günther, J. E. Coeymans, J. C. Munoz, and J. C. Herrera, “Mitigating freeway off-ramp congestion: A surface streets coordinated approach,” *Transportation Research Part C: Emerging Technologies*, vol. 20, pp. 112–125, Jan. 2012.

[40] X. Yang, Y. Cheng, and G. L. Chang, “Integrating off-ramp spillback control with the decomposed arterial signal optimization model,” *Transportation Research Record: Journal of the Transportation Research Board*, 2015.

[41] J. A. Nelder and R. Mead, “A simplex method for function minimization,” *Computer Journal*, vol. 7, pp. 308–313, 1965.

[42] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, “Convergence properties of the Nelder-Mead simplex method in low dimensions,” *SIAM Journal on Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[43] D. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley Longman Publishing Co., USA, 1989.

[44] J. Holland, *Adaptation in natural and artificial systems*. MIT Press,USA., 1992.

[45] R. Y. Rubinstein and D. P. Kroese, *The cross-entropy method: a unified approach to combinatorial optimization, monte-carlo simulation, and machine learning*. Springer-Verlag, USA, 2004.

[46] P.-T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A Tutorial on the cross-entropy method," *Annals of Operations Research*, vol. 134, pp. 19–67, 2005.

[47] B. H. Heutinck, M. Van den Berg, J. Hellendoorn, and L. H. Immers, "Dynamic route guidance during maintenance works: A case study," in *11th IFAC Symposium on Control in Transportation Systems*, 2006, pp. 380–385.

[48] Y. Wang, M. Papageorgiou, G. Sarros, and W. J. Knibbe, "Feedback route guidance applied to a large-scale express ring road," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1965, pp. 79–88, 2006.

[49] M. Papgeorgiou and A. Messmer, "Dynamic network traffic assignment and route guidance via feedback regulation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1306, pp. 49–58, 1991.

[50] Y. Pavlis and M. Papageorgiou, "Simple decentralized feedback strategies for route guidance in traffic networks," *Transportation Science*, vol. 33, pp. 264–278, 1999.

[51] H. Haj Salem, S. Cohen, E. Sididki, and M. Papageorgiou, "Field trial results of VMS travel time display on the corridor peripherique of Paris," in *Proc. 4th Intern. ASCE Conference on Applications of Advanced Technologies in Transportation Engineering*, 1995, pp. 368–372.

[52] S. Mammar, A. Messmer, P. Jensen, M. Papageorgiou, H. Haj-Salem, and L. Jensen, "Automatic control of variable message signs in Aalborg," *Transportation Research Part C: Emerging Technologies*, vol. 4, pp. 131–150, 1996.

[53] A. Messmer, M. Papageorgiou, and N. Mackenzie, "Automatic control of variable message signs in the interurban Scottish highway network," *Transportation Research Part C: Emerging Technologies*, vol. 6, pp. 173–187, 1998.

[54] C. Diakaki, M. Papageorgiou, and T. McLean, "Integrated traffic-responsive urban corridor control strategy in Glasgow, Scotland - Application and evaluation," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1727, pp. 101–111, 2000.

[55] G. Vigos, M. Papageorgiou, and Y. Wang, "Real-time estimation of vehicle-count within singalized links," *Transportation Research Part C: Emerging Technologies*, vol. 16, pp. 18–35, 2008.

[56] A. Spiliopoulou, D. Manolis, I. Papamichail, and M. Papageorgiou, "Queue Management Techniques for Metered Freeway On-Ramps," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2178, pp. 40–48, Dec. 2010.

[57] I. Papamichail, M. Papageorgiou, V. Vong, and J. Gaffney, "Heuristic Ramp-Metering Coordination Strategy Implemented at Monash Freeway, Australia," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2178, pp. 10–20, Dec. 2010.

[58] L. Faulkner, F. Dekker, D. Gyles, I. Papamichail, and M. Papageorgiou, "Evaluation of HERO Coordinated Ramp Metering Installation at the M1 / M3 Freeway in Queensland , Australia," *93rd Transportation Research Board Annual Meeting*, vol. 14–0992, 2014.

[59] M. Papageorgiou, I. Papamichail, A. Spiliopoulou, and A. F. Lentzakis, "Real-time merging traffic control with applications to toll plaza and work zone management," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 535–553, Oct. 2008.

[60] A. Tympakianaki, A. Spiliopoulou, A. Kouvelas, I. Papamichail, M. Papageorgiou, and Y. Wang, "Real-time merging traffic control for throughput maximization at motorway work zones," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 242–252, 2014.

[61] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal mainstream traffic flow control involving variable speed limits and ramp metering," *Transportation Science*, vol. 18, pp. 193–210, 2010.

[62] K. Chung, J. Rudjanakanoknad, and M. J. Cassidy, "Relation between traffic density and capacity drop at three freeway bottlenecks," *Transportation Research Part B: Methodological*, vol. 41, no. 1, pp. 82–95, 2007.

[63] M. J. Cassidy and J. Rudjanakanoknad, "Increasing the capacity of an isolated merge by metering its on-ramp," *Transportation Research Part B: Methodological*, vol. 39, no. 10, pp. 896–913, 2005.

[64] M. Papageorgiou, H. Hadj-Salem, and J. M. Blosseville, "ALINEA: a local feedback control law for on-ramp metering," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1320, pp. 58–64, 1991.

[65] Papageorgiou, H. M., Haj-Salem, and F. Middelham, "ALINEA local ramp metering: summary of field results," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1603, pp. 90–98, 1997.

[66] Y. Wang, E. Kosmatopoulos, M. Papageorgiou, and I. Papamichail, "Local ramp metering in the presence of a distant downstream bottleneck: Theoretical analysis and simulation study," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–16, 2014.

[67] I. Papageorgiou, M. Papamichail, "Overview of traffic signal operation policies for ramp metering," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2047, pp. 28–36, 2008.

[68] "TSS – Transport Simulation Systems, 2013. AIMSUN Users Manual Version 8, Barcelona, Spain."

8 Appendix

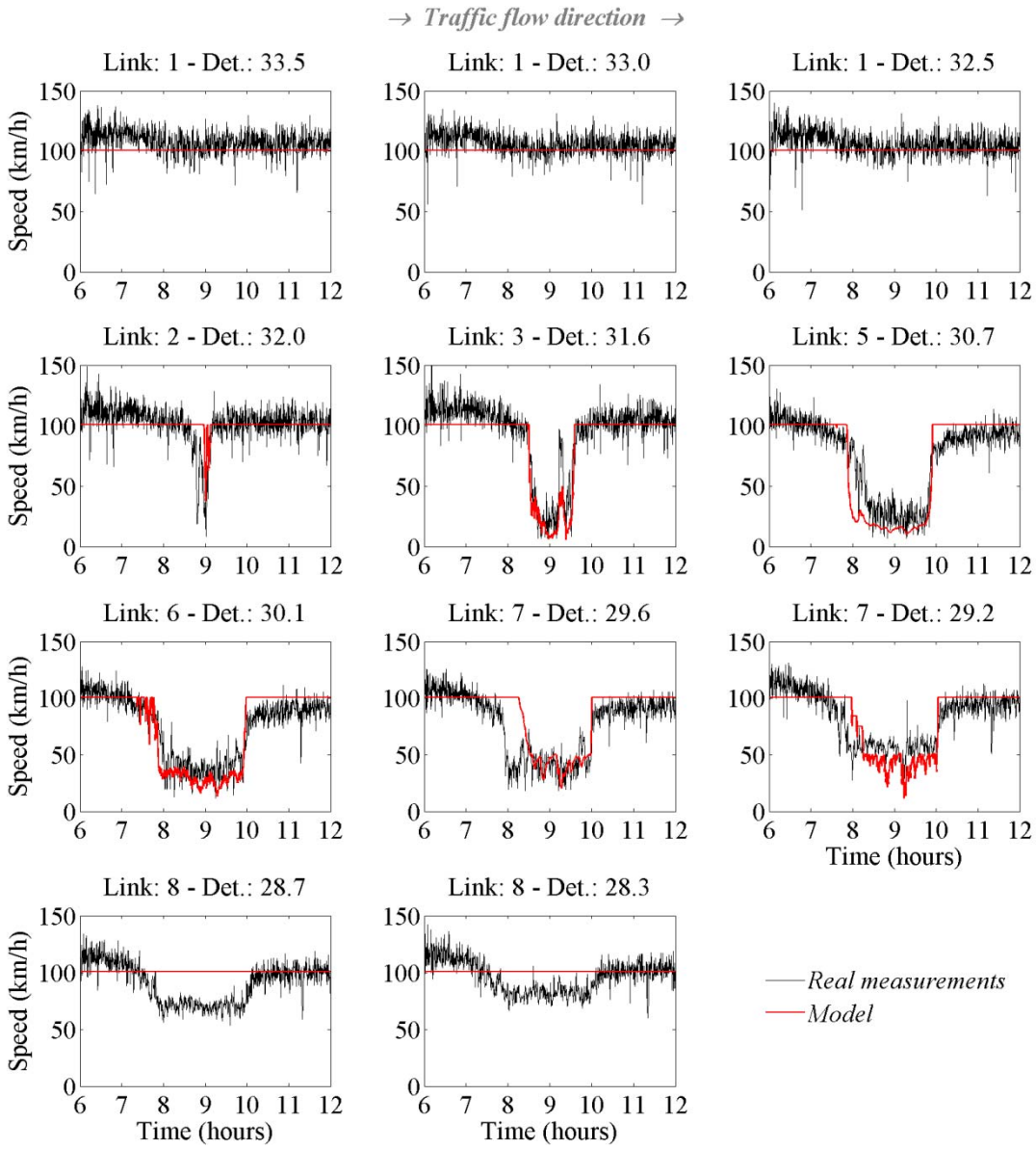

Figure A.1 CTM model calibration: time-series of the real speed measurements and Model 1.2 estimation of speed at various detector locations for 16/06/2009.

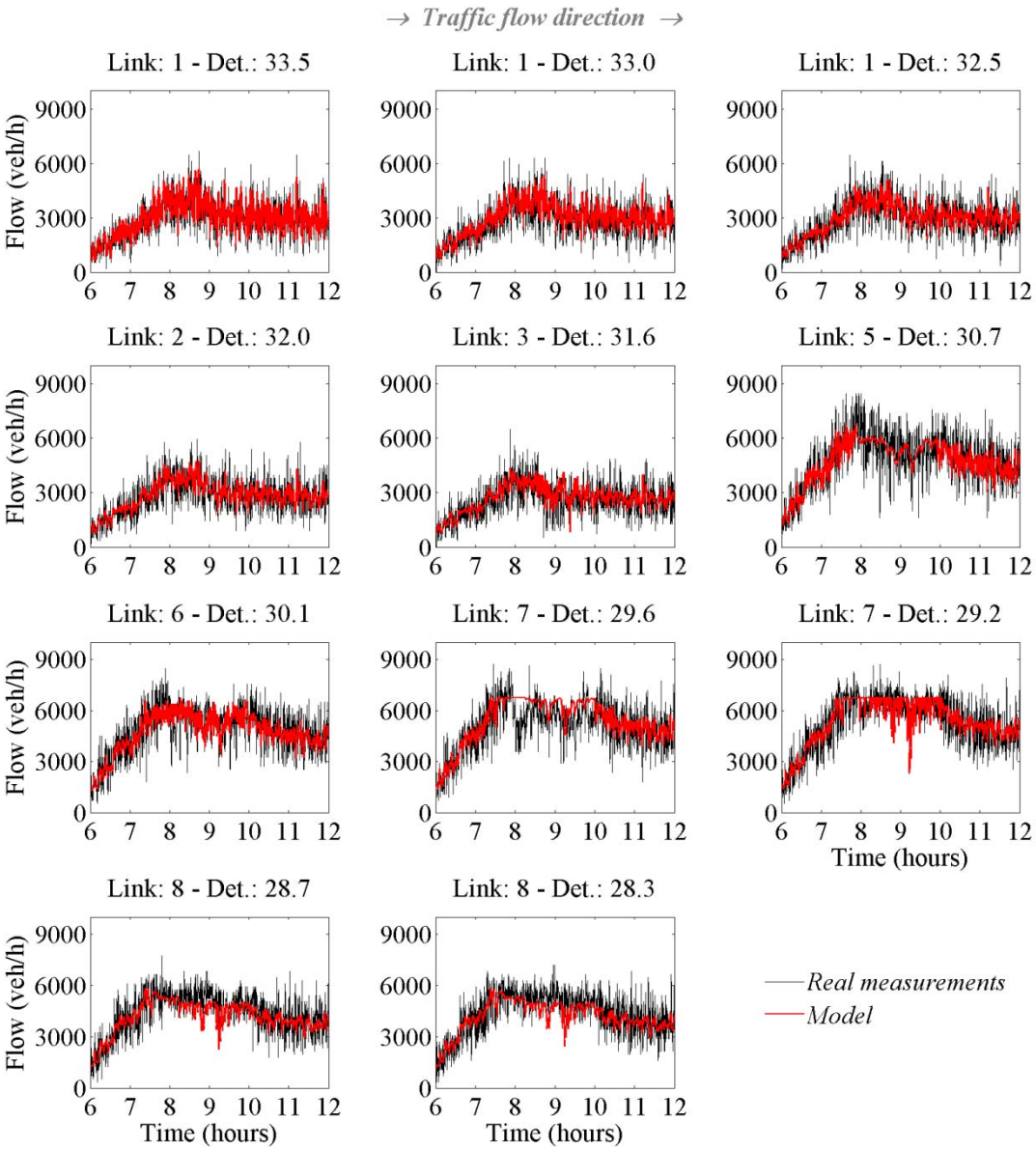

Figure A.2 CTM model calibration: time-series of the real flow measurements and Model 1.2 estimation of flow at various detector locations for 16/06/2009.

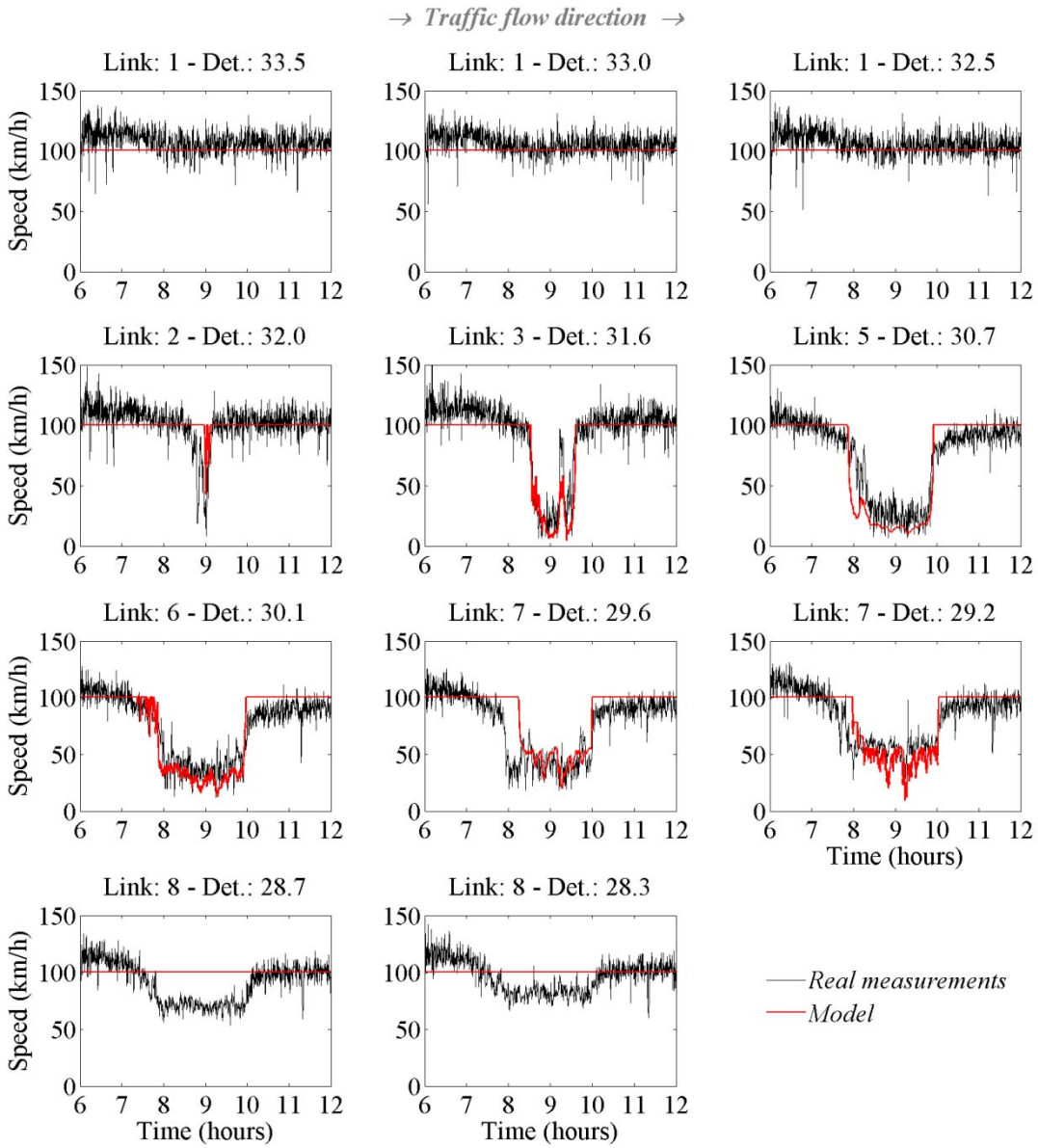

Figure A.3 CTM model calibration: time-series of the real speed measurements and Model 1.3 estimation of speed at various detector locations for 16/06/2009.

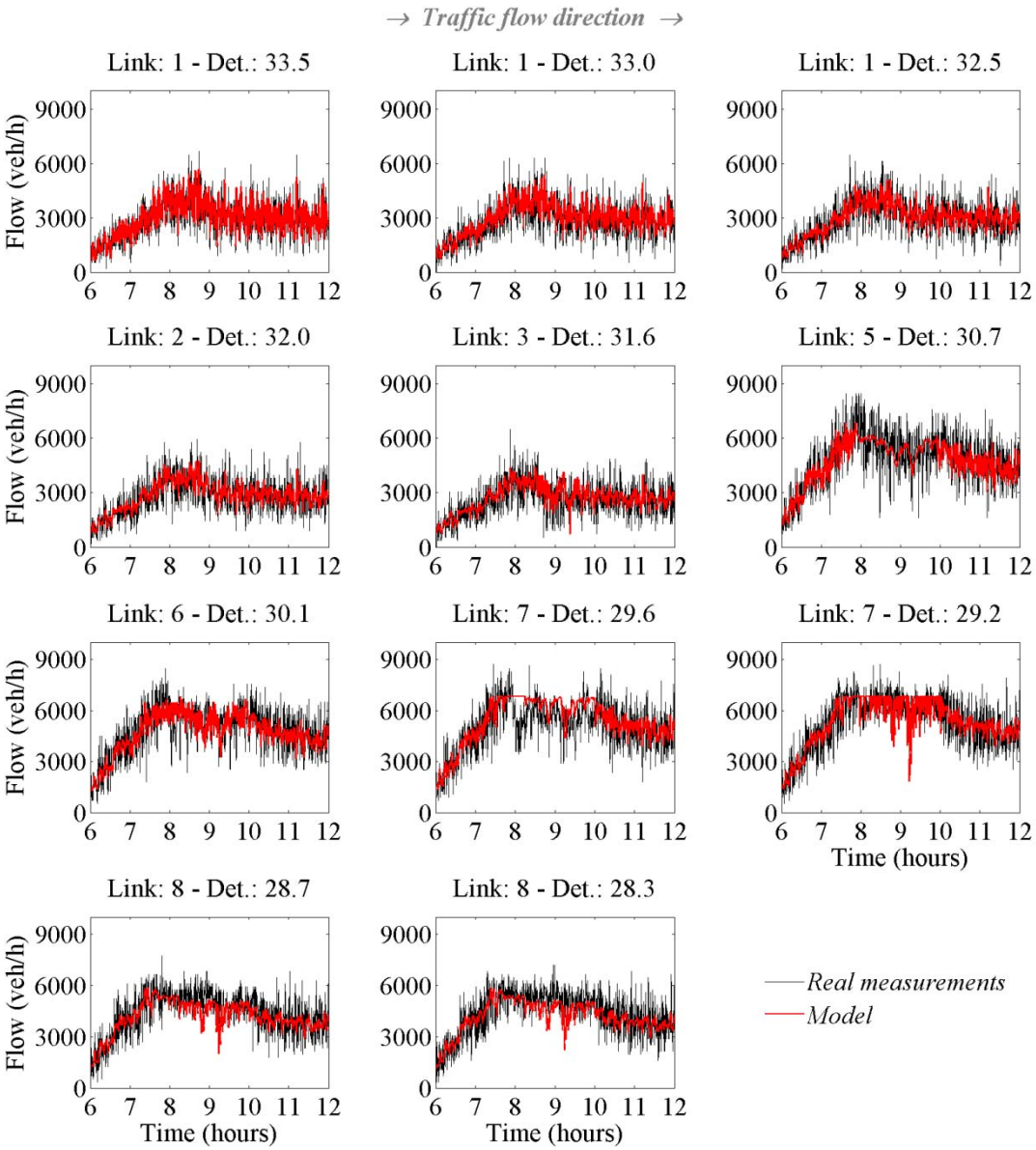

Figure A.4 CTM model calibration: time-series of the real flow measurements and Model 1.3 estimation of flow at various detector locations for 16/06/2009.

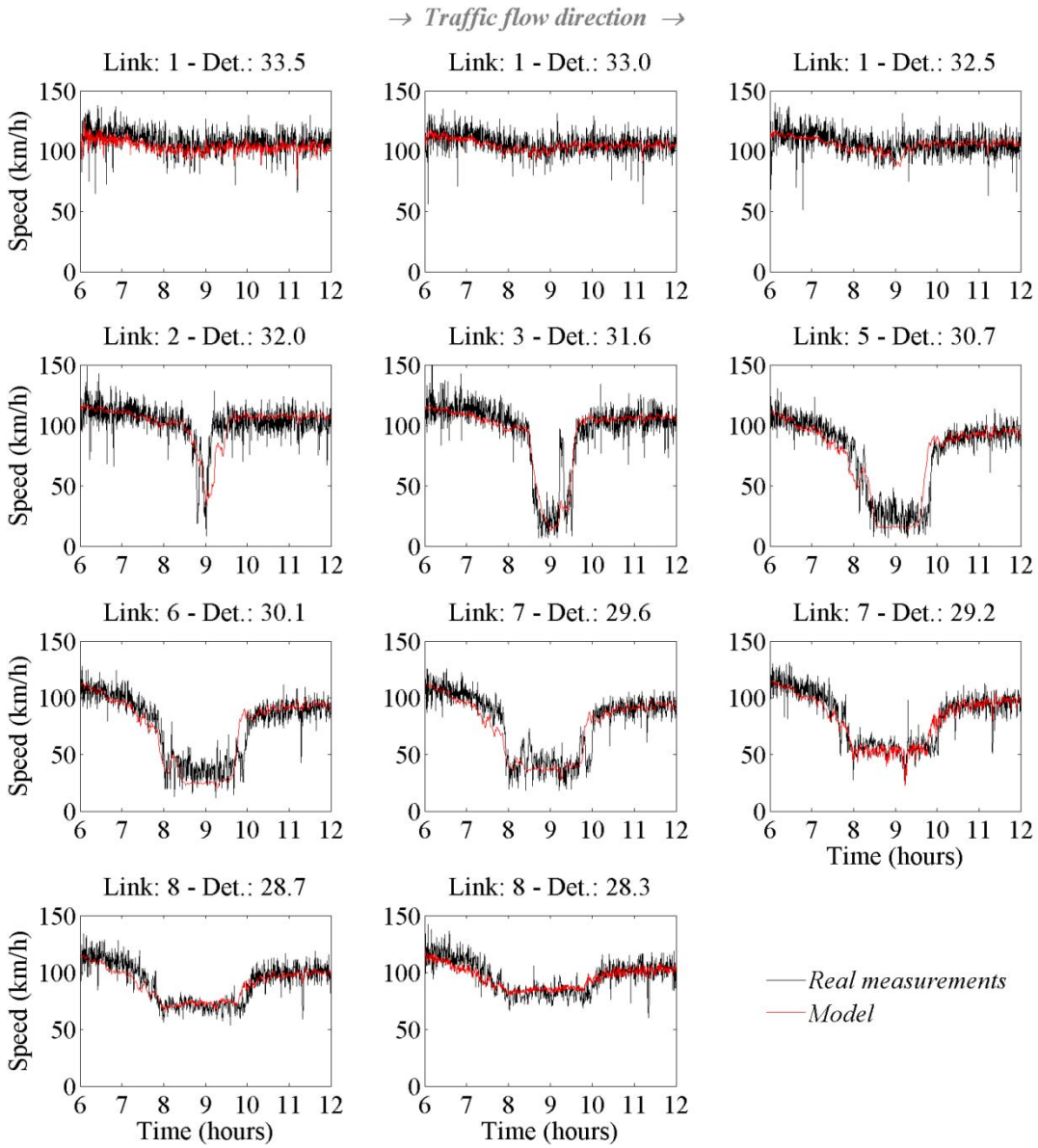

Figure A.5 METANET model calibration: time-series of the real speed measurements and Model 2.2 estimation of speed at various detector locations for 16/06/2009.

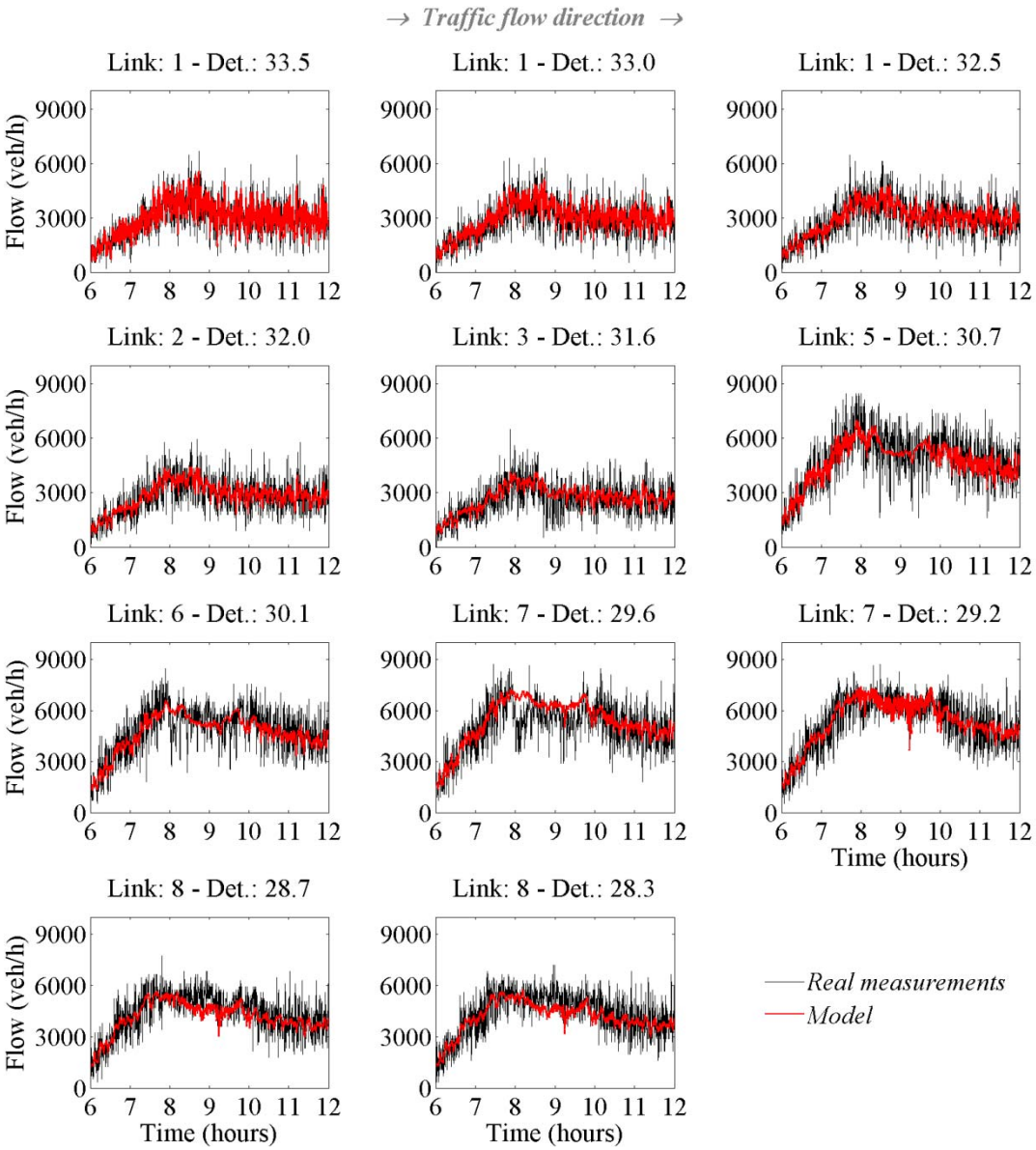

Figure A.6 METANET model calibration: time-series of the real flow measurements and Model 2.2 estimation of flow at various detector locations for 16/06/2009.

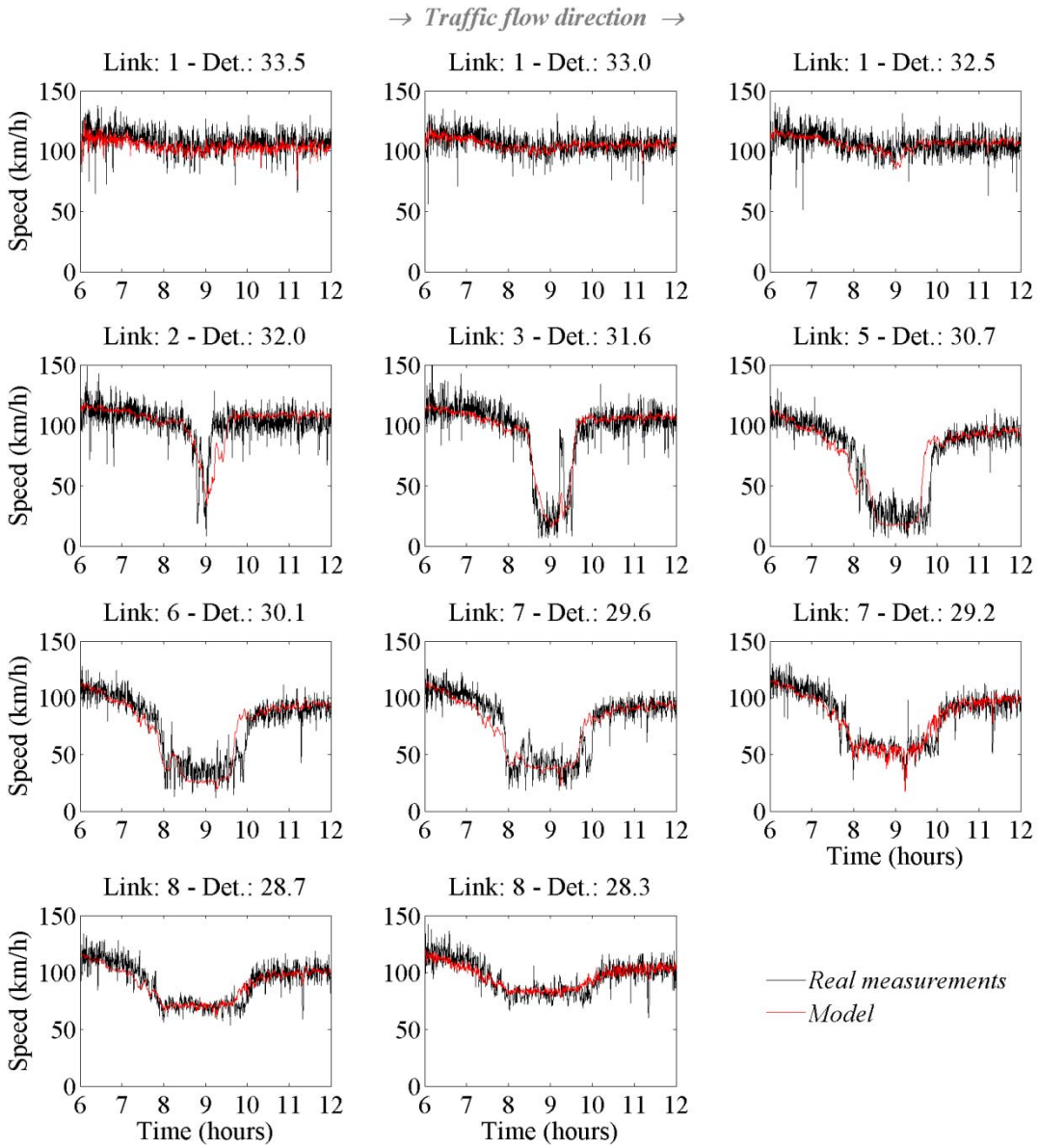

Figure A.7 METANET model calibration: time-series of the real speed measurements and Model 2.3 estimation of speed at various detector locations for 16/06/2009.

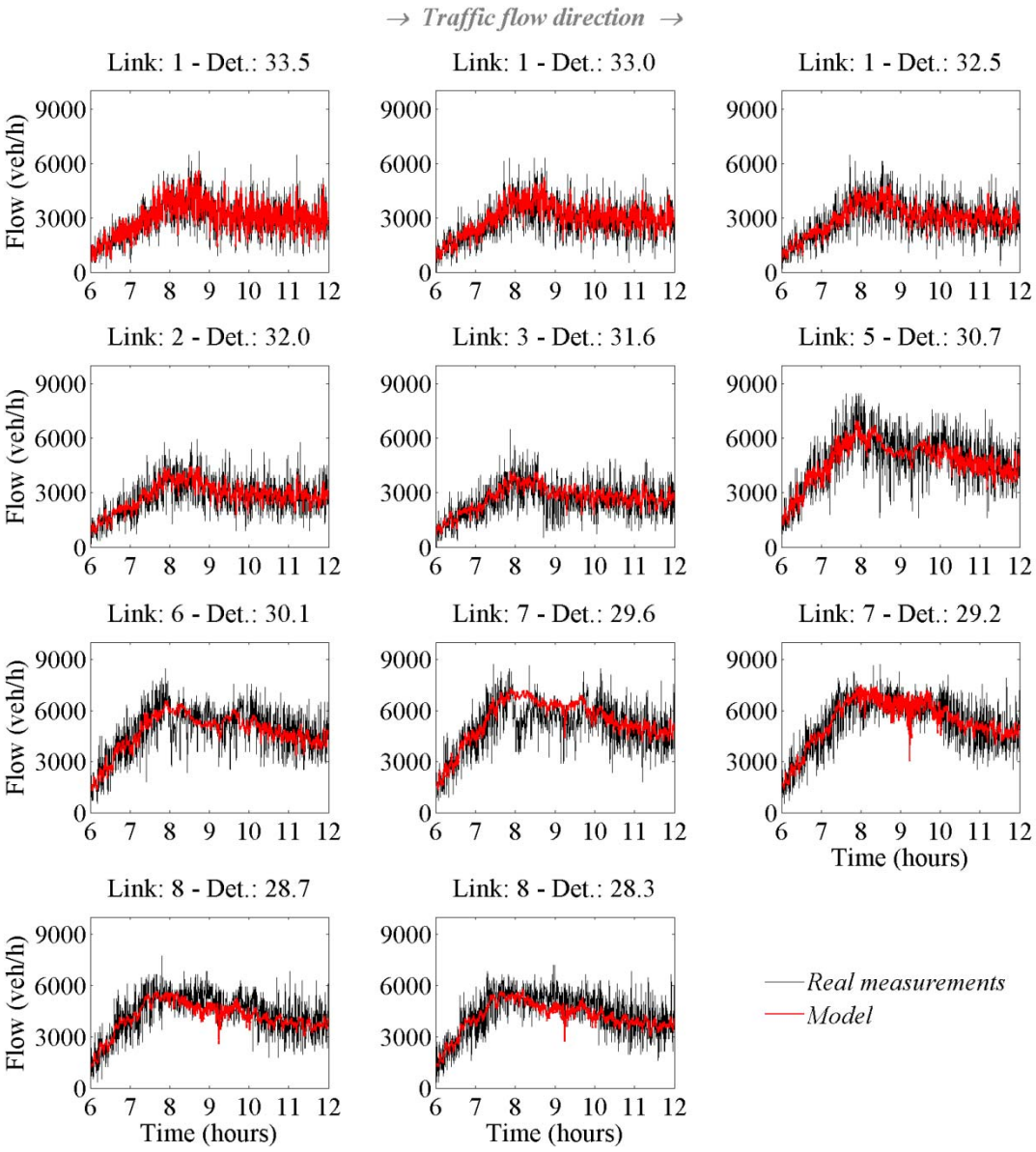

Figure A.8 METANET model calibration: time-series of the real flow measurements and Model 2.3 estimation of flow at various detector locations for 16/06/2009.