GEORGIA-ROUMPINI IORDANIDOU

TRAFFIC FLOW CONTROL ON MOTORWAYS FOR THROUGHPUT MAXIMIZATION

CHANIA, GREECE

2017

TECHNICAL UNIVERSITY OF CRETE

SCHOOL OF PRODUCTION ENGINEERING AND MANAGEMENT

TRAFFIC FLOW CONTROL ON MOTORWAYS FOR THROUGHPUT MAXIMIZATION

Thesis submitted for the partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
By

GEORGIA-ROUMPINI IORDANIDOU

Chania, Greece, March 2017

Typeset in Microsoft Office Word.

Figures created with Visio.

Plots generated in Matlab

TRAFFIC FLOW CONTROL ON MOTORWAYS FOR THROUGHPUT MAXIMIZATION

Georgia-Roumpini Iordanidou

'This Ph.D. Thesis is approved by:'

Advisory committee:

Ioannis Papamichail (Supervisor)
Associate Professor, School of Production Engineering and Management
Technical University of Crete

Markos Papageorgiou (Co-supervisor)
Professor, School of Production Engineering and Management
Technical University of Crete

Ioannis Nikolos (Member of the advisory committee)
Associate Professor, School of Production Engineering and Management
Technical University of Crete

Thesis committee:

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

Yannis Marinakis
Assistant Professor, School of Production Engineering and Management
Technical University of Crete

Rodrigo Castelan Carlson
Assistant Professor, Department of Automation and Systems, Technological Center, Federal University of Santa Catarina, Florianópolis, Brasil

Claudio Roncoli
Assistant Professor, Department of Built Environment, School of Engineering, Aalto University, Espoo, Finland

To Pavlos.

ACKNOWLEDGEMENTS

I would like to thank my husband, Pavlos Koulouridakis, for his priceless help during this Ph.D. endeavor. I would also like to thank my family for being always next to me. I am grateful to my father Dimitrios, my mother Georgia and my sister Vasiliki.

I feel grateful to my supervisor Prof. Ioannis Papamichail, for his constant support and his knowledge dissemination, my co-supervisor Prof. Markos Papageorgiou, for his support and his inexhaustible ideas, and the member of my advisory committee Prof. Ioannis Nikolos for his support. I am also grateful to the professors of my thesis committee.

I would also like to thank Dr. Claudio Roncoli and Dr. Natasa Spiliopoulou for their essential and valuable help in my Ph.D.

Finally, I would like to thank the members of TRAMAN21, a European Research Council project under the European Union's Seventh Framework Programme (FP/20072013) / ERC Grant Agreement n. 321132. Funding from TRAMAN21 is also gratefully acknowledged. My involvement in this project was an amazing experience.

SHORT BIOGRAPHY

Iordanidou Georgia-Roumpini received the Diploma and M.Sc. degrees with honors in Production Engineering and Management from the Technical University of Crete, Chania, Greece, in March 2011 and October 2012, respectively.

From September 2012, she is a Ph.D. candidate with the Dynamic Systems and Simulation Laboratory, School of Production Engineering and Management, Technical University of Crete, Chania, Greece. She is an author of three technical papers in scientific journals, and two technical papers in scientific conference proceedings. Her main research interests include automatic control and optimization theory and applications to traffic and transportation systems.

Ms. Iordanidou received Scholarships from the Technical University of Crete and from the State Scholarships Foundation for Undergraduate Studies from 2006 to 2011.

Abstract of the Thesis presented to the Technical University of Crete as a partial fulfillment of the requirements for the degree of Doctor of Philosophy.

TRAFFIC FLOW CONTROL ON MOTORWAYS FOR THROUGHPUT MAXIMIZATION

Georgia-Roumpini Iordanidou

March 2017

Supervisor: Prof. Ioannis Papamichail
Co-Supervisor: Prof. Markos Papageorgiou
Keywords: traffic management; feedback control; integrated motorway traffic flow control; mainstream traffic flow control; ramp metering; multiple bottlenecks; variable speed limits; delay balancing
Number of pages: 116

Traffic congestion on motorways is one of the most serious problems of modern societies that leads to a significant reduction in motorway infrastructure capacity [1]. This reduction regularly occurs during peak periods, causing degradation in terms of travel times, traffic safety, fuel consumption, and environmental pollution. So far, several traffic control measures have been proposed to alleviate traffic congestion. However, some of them face limitations; e.g., ramp metering (RM) efficiency is limited by the available storage space at on-ramps [2], whereas route guidance is most valuable under non-recurrent traffic congestion.

Mainstream traffic flow control (MTFC) enabled via variable speed limits (VSLs) has been investigated in previous studies, utilizing various control strategies. In this thesis, an extended feedback control strategy is proposed for MTFC enabled via VSLs, considering multiple-bottleneck locations. The evaluation of the proposed

control strategy, using a second order macroscopic traffic flow simulator, and its comparison with an optimal control approach, for a real network, demonstrates its efficiency. The feedback concept is approaching the performance of optimal control, is more robust (no model or demand predictions are needed), and can be immediately implemented in the field as it considers practical and safety constraints.

The development and deployment of simple, yet efficient, coordinated and integrated control tools for motorway traffic control remains a challenge. In this thesis, a generic integrated feedback-based motorway traffic flow control concept is proposed. It is based on the combination and suitable extension of control algorithms and tools proposed or deployed in other studies, such as RM or VSL-enabled cascade-feedback mainstream traffic flow control, and allows for consideration of multiple bottlenecks. The new controller enables coordination of RM actions at a series of on-ramps, as well as integration with VSL control actions, towards a common control goal, which is bottleneck throughput maximization. While doing this, the approach considers a pre-specified (desired) balancing of the incurred delays upstream of the employed actuators, via a suitably designed knapsack problem. Despite the multitude of the offered configurations, options and possibilities, the generic control algorithm remains simple, efficient and suitable for field implementation. The control algorithm is demonstrated and evaluated using a validated macroscopic traffic flow model for a real infrastructure and has been compared to other control structures. The integrated controller is shown to be superior as it takes advantage of all the available storage capacity required for queueing upstream of the bottlenecks. The feedback controller is robust as there is no need, neither for any predictions of the demand nor for any model calibration or parameter identification.

Περίληψη της Διατριβής που υπεβλήθη στο Πολυτεχνείο Κρήτης για τη μερική ικανοποίηση των απαιτήσεων για την απόκτηση Διδακτορικού Διπλώματος

ΕΛΕΓΧΟΣ ΚΥΚΛΟΦΟΡΙΑΣ ΓΙΑ ΜΕΓΙΣΤΟΠΟΙΗΣΗ ΤΗΣ ΡΟΗΣ ΣΕ ΑΥΤΟΚΙΝΗΤΟΔΡΟΜΟΥΣ

Γεωργία-Ρουμπίνη Ιορδανίδου

March 2017

Επιβλέπων: Καθηγητής Ιωάννης Παπαμιχαήλ
Συν-Επιβλέπων: Καθηγητής Μάρκος Παπαγεωργίου
Λέξεις Κλειδιά: διαχείριση κυκλοφορίας, έλεγχος με ανατροφοδότηση, ολοκληρωμένος έλεγχος κυκλοφοριακής ροής σε αυτοκινητόδρομους, έλεγχος κύριας κυκλοφοριακής ροής, έλεγχος ραμπών εισόδου, πολλαπλά σημεία συμφόρησης, μεταβλητά όρια ταχύτητας, εξισορρόπηση της καθυστέρησης
Αριθμός Σελίδων: 116

Η κυκλοφοριακή συμφόρηση σε αυτοκινητόδρομους αποτελεί ένα από τα πιο σημαντικά προβλήματα των μοντέρνων κοινωνιών, το οποίο οδηγεί σε αξιοσημείωτη μείωση της ικανότητας της υποδομής των αυτοκινητοδρόμων [1]. Αυτή η μείωση ως επί των πλείστον παρατηρείται κατά τη διάρκεια των περιόδων αιχμής, προκαλώντας υποβιβασμό, όσον αφορά το χρόνο ταξιδίου, την κυκλοφοριακή ασφάλεια, την κατανάλωση καυσίμου, και τη μόλυνση του περιβάλλοντος. Μέχρι τώρα, πολλά κυκλοφοριακά μέτρα ελέγχου έχουν προταθεί ούτως ώστε να ανακουφίσουν την κυκλοφοριακή συμφόρηση. Ωστόσο, μερικά από αυτά αντιμετωπίζουν περιορισμούς, π.χ., η αποδοτικότητα του ελέγχου των ραμπών εισόδου περιορίζεται από το διαθέσιμο χώρο αποθήκευσης των ραμπών εισόδου [2], ενώ η καθοδήγηση σε κάποιον προορισμό είναι περισσότερο πολύτιμη σε περιπτώσεις μη επαναλαμβανόμενων κυκλοφοριακών συμφορήσεων.

Ο έλεγχος της κύριας κυκλοφοριακής ροής μέσω μεταβλητών ορίων ταχύτητας έχει διερευνηθεί σε προηγούμενες εργασίες, χρησιμοποιώντας διάφορες στρατηγικές ελέγχου. Σε αυτή τη διατριβή, προτείνεται μια εκτεταμένη στρατηγική ελέγχου με ανατροφοδότηση για έλεγχο σε κύρια ροή με μεταβλητά όρια ταχύτητας, λαμβάνοντας υπόψιν πολλαπλά σημεία συμφόρησης. Η αποτίμηση της προτεινόμενης στρατηγικής ελέγχου, χρησιμοποιώντας ένα μακροσκοπικό προσομοιωτή κυκλοφοριακής ροής δεύτερης τάξης, και η σύγκριση της με μια προσέγγιση βέλτιστου ελέγχου, για ένα πραγματικό δίκτυο, αναδεικνύει την αποδοτικότητα της. Ο έλεγχος με ανατροφοδότηση προσεγγίζει την απόδοση του βέλτιστου ελέγχου, είναι πιο εύρωστος (δε χρειάζεται μοντέλο ή προβλέψεις της ζήτησης), και μπορεί άμεσα να εφαρμοστεί στο πεδίο καθώς λαμβάνει υπόψιν της πρακτικούς περιορισμούς και περιορισμούς ασφάλειας.

Η ανάπτυξη και η αξιοποίηση απλών, αλλά αποτελεσματικών, συντονισμένων και ολοκληρωμένων εργαλείων ελέγχου για κυκλοφοριακό έλεγχο σε αυτοκινητοδρόμους παραμένει μια πρόκληση. Σε αυτή τη διατριβή, προτείνεται μια γενικευμένη ολοκληρωμένη ιδέα βασισμένη σε κυκλοφοριακό έλεγχο αυτοκινητοδρόμων με ανατροφοδότηση. Βασίζεται στο συνδυασμό και την κατάλληλη επέκταση αλγορίθμων και εργαλείων ελέγχου που έχουν προταθεί ή αξιοποιηθεί σε άλλες εργασίες, όπως ο έλεγχος ραμπών εισόδου ή ο κυκλοφοριακός έλεγχος της κύριας ροής με ανατροφοδότηση με χρήση μεταβλητών ορίων ταχύτητας, και επιτρέπει την εξέταση πολλαπλών σημείων συμφόρησης. Ο νέος ελεγκτής επιτρέπει το συντονισμό δράσεων ελέγχου ραμπών εισόδου σε μία σειρά ραμπών, καθώς επίσης και τη συνεργασία με δράσεις ελέγχου με μεταβλητά όρια ταχυτήτων, προς ένα κοινό στόχο, ο οποίος είναι η μεγιστοποίηση της ροής σε σημεία συμφόρησης. Ενώ κάνει αυτό, η προσέγγιση λαμβάνει υπόψιν της μια προκαθορισμένη (επιθυμητή) εξισορρόπηση των προκυπτόμενων καθυστερήσεων ανάντι των απασχολούμενων ενεργοποιητών, μέσω ενός κατάλληλα σχεδιασμένου προβλήματος σακιδίου. Παρά την πληθώρα των προσφερόμενων διαμορφώσεων, επιλογών και δυνατοτήτων, ο γενικευμένος αλγόριθμος ελέγχου παραμένει απλός, αποδοτικός και κατάλληλος για εφαρμογές στο πεδίο. Ο αλγόριθμος ελέγχου παρουσιάζεται και αποτιμάται χρησιμοποιώντας ένα επικυρωμένο μακροσκοπικό μοντέλο κυκλοφοριακής ροής για ένα πλήθος σεναρίων. Ο ελεγκτής με ανατροφοδότηση είναι εύρωστος καθώς δεν υπάρχει ανάγκη, ούτε για

πρόβλεψη της ζήτησης ούτε για βαθμονόμηση μοντέλου ή αναγνώριση των παραμέτρων του.

CONTENTS

List of Figures

1. Introduction

A major community problem which has appeared in the last decades is the daily traffic congestion on motorways. The main reasons are the continuous increase of car ownership and demand that contribute to the daily appearance of recurrent and non-recurrent motorway congestion whose expanse is continuously increasing in space and time.

The motorway throughput decreases because of the traffic congestion that leads to the degradation of the available infrastructure [1]. As a result, the motorway infrastructure is underutilized when it is really needed, i.e., during the peak periods. The implications of the infrastructure degradation are huge considering the economic and social life of the influenced regions: traffic safety reduction, increased fuel consumption and environmental pollution and enormous delays. For instance, the future costs imposed on households by congestion are expected to increase for the UK, France, Germany and the US. These advanced economies are, on average, expected to witness increases in these costs of up to 50% between 2013 and 2030 [3].

The simple queuing systems serving capacity is not influenced by the waiting queue appearance, whereas a motorway congestion influences the nominal motorway capacity and throughput because of two reasons [1]: capacity drop at the head of the forming congestion and blocking of off-ramps due to the congestion body expansion. These two effects lead to infrastructure degradation escalation. This results in an accelerated congestion increase that by itself leads to further infrastructure degradation, further increase of congestion and so on, until a big region of the motorway network is covered by the formed congestion, commonly spilling over from one motorway to another. When generalized congestion exists, the overall arriving demand is generally much lower than the nominal infrastructure capacity, but due to the severe degraded congested infrastructure the demand cannot be served. As a result, congestion exists until the demand values become low enough, at the last phase of the peak period.

Considering all the above, it is obvious that the extended daily congestion on motorways is not caused only due to the enormous demand exceeding the nominal

network capacity. Certainly, the demand may exceed temporarily and locally the motorway capacity, and as a consequence triggers local congestion. Nevertheless, the generalized congestion is due to the unstable escalation produced by the infrastructure degradation in lack of appropriate control measures that would reduce this evolution.

An optimal utilization of the available infrastructure is necessary in order to have a safe, less-polluting and efficient, transportation system on motorways. This can be achieved via appropriate application of many traffic control measures, like ramp metering (RM) [4], [5] variable speed limits (VSLs) [6], [7] route guidance and driver information. There are various methodological approaches incorporating feedback control, optimal control, fuzzy systems, neural networks and expert systems, that have been deployed in the previous years for the design of relevant control strategies. Many previous works have shown that appropriate traffic control measures may give important improvements, under certain conditions [8], [9].

Nevertheless, many of these traffic control measures that have been proposed and partly implemented in motorway networks are known to face limitations:

- RM is a very helpful control measure, but its benefits may be constrained due to limited ramp storage space [2].
- VSLs are helpful for traffic safety but their present usage has hardly any positive effect on throughput increase or average travel times decrease, because the utilized control strategies are quite simplistic and not able to increase traffic flow efficiency [10]. In addition, the applied speed limits range is commonly limited.
- Route guidance and driver information systems are more helpful for non-recurrent cases, e.g., congestion created because of an incident [11].
- Emerging vehicle-infrastructure integration (VII) systems offer a promising technological background for efficient traffic control. In the past there were no specified efficiency-improving implementations and relevant control algorithms, but there have been huge improvements in the last few years. It may take some time before real implementations occur.

From all the above mentioned control measures, the display of VSLs on suitable variable message signs (VMSs), taking into consideration the prevailing traffic

conditions, is an issue that is investigated in this thesis, due to the potential benefits. The integration of VSLs with RM is also studied. Traffic safety enhancement, is a basic aim of VSLs, and thus the choice of VSL installations in many countries is depending on registered accidents frequency. The positive effect of VSL on traffic safety is due to the reduction of speed and speed homogenization that are associated with less accident percentages. According to many investigations performed in the last years, the VSL effect on traffic safety depicts a reduction in accident numbers by as much as 20 to 30 percent after VSL installation [12]. Furthermore, some authorities are foreseeing VSL as means to reduce vehicle emissions and road noise [13], [14]. In contrast, concerning traffic efficiency, until recently there was no evaluation of the VSL impact of available installations that would show a consistent and measurable improvement, e.g., reduced travel times.

1.1. Traffic Problem Description

The following statements tabulate the current state:

- The present traffic condition in metropolitan motorways is characterized by heavy congestion during rush hours, and the associated cost is very high for the economic and social life of metropolitan areas.
- The motorway infrastructure is greatly underutilized because of congestion.
- High demand is only one of the reasons for heavy congestion. In case vehicles are permitted to use the infrastructure according to their will, i.e., without applying any control measure, the constricted original congestion escalates and leads to extended congestion areas and significant infrastructure degradation. The motorway control is necessary for maximum efficiency which is associated with traffic safety improvement and environmental impact reduction.
- Various control measures are envisaged or implemented in some motorway network stretches, but the attainable improvements face limitations.

Concerning VSL, it should be remarked that the ideal exploitation of the chances provided by VSL would be to maintain the safety and environmental benefits offered by the current systems together with an increase of traffic flow efficiency. It does not mean that VSL are not a suitable control measure for traffic flow efficiency

improvement, because the conducted field evaluations could not depict an efficiency increase. As a matter of fact:

- The VSL effect on aggregate traffic behavior, e.g., on the fundamental diagram has not been adequately investigated with real data. Thus, the comprehension of even qualitative effects of VSL is constricted to conjectures and assumptions. As a result, the VSL control strategies cannot achieve traffic flow efficiency increase.
- The VSL installations currently existing employ simple rule-based control strategies for VSL switching. These strategies decisions are based on predefined thresholds of traffic flow, or occupancy, or mean speed. The threshold selection is done in an ad hoc way that does not always exploit potential impact of VSL on traffic flow efficiency.
- There are several simulation studies that proposed practical strategies that are yet to be considered for implementation by the authorities.

In order to design control strategies that may increase traffic flow efficiency, an accurate description of VSL impact on the aggregate (macroscopic) traffic conditions is necessary. In the past there were very few investigations dealing with the exact impact of VSL on aggregate traffic flow behavior. Lately, the impact of VSL on the aggregate traffic flow behavior was investigated in [10] and [15], using traffic data from a VSL-equipped European motorway.

1.2. Research Approach

Considering the problems discussed in the previous section, this thesis proposes and investigates an extended feedback control strategy for MTFC enabled via VSL, considering multiple bottlenecks, and a generic integrated feedback-based motorway traffic flow control concept. MTFC regulates the mainstream traffic flow upstream of a bottleneck location so as to avoid capacity drop [6]. Until recently MTFC was dealing only with single bottlenecks and the integration with ramp metering was limited.

Thus, the following topics are addressed:

- An extended feedback control strategy for MTFC enabled via VSL, considering multiple bottlenecks is proposed. Multiple bottleneck inspiration

and causes are described. The implementation aspects of MTFC, the feedback controller and some practical application aspects are analyzed.

- The feedback controller for MTFC enabled via VSL, considering multiple bottlenecks is evaluated using simulation-based tests for a real motorway stretch using a second-order macroscopic traffic flow simulator, so as to demonstrate its features and compare its efficiency with a sophisticated optimal control approach.
- A generic integrated feedback-based motorway traffic flow control concept is proposed. The feedback control structure is described and the implementation aspects of RM and MTFC via VSL are presented. In addition the delay estimation procedure and the flow distribution for delay balancing are discussed.
- The feedback controller for the integrated feedback-based motorway traffic flow control concept is evaluated using simulation-based tests for a real motorway stretch using a second-order macroscopic traffic flow simulator, so as to demonstrate its features, for a number of scenarios.
- In the proposals of this thesis, the MTFC concept is utilized, thus a brief description of the concept is first demonstrated. RM and VSL (as an MTFC actuator) are used as actuators, so an overview of these actuators is presented.

1.3. Outline of the Thesis

Chapter 2 includes a preliminary discussion on motorway traffic congestion and an overview of the basic motorway traffic control measures. An overview of RM strategies and RM effects is following. In the same chapter a VSL overview is included, along with the fundamental diagram description and VSL effects. MTFC is described below, focusing on the basic MTFC concept and bottleneck applications, MTFC implementation and MTFC integration with RM. The traffic flow model formulation is outlined as well as the optimal control problem formulation. Finally, the model parameter calibration procedure is analyzed.

Chapter 3 incorporates an overview of MTFC-like strategies and the multiple bottleneck inspiration and causes. MTFC control for multiple bottlenecks is then described together with its implementation aspects, the utilized feedback controller and some practical application aspects.

In Chapter 4, a new integrated traffic flow control concept for multiple bottlenecks with delay balancing is proposed. This chapter begins with a literature overview of integrated control concepts. A description of the proposed feedback control structure follows. RM actuator characteristics and MTFC actuator characteristics are described next. The delay estimation and the flow distribution for delay balancing procedures are concluding this chapter.

In Chapter 5, the control concept described in Chapter 3 is evaluated and compared in simulation for a network in Perth, Australia, considering several practical application aspects. The performance of the feedback controller is shown to approach the optimal control results.

In Chapter 6, the control concept described in Chapter 4 is evaluated and compared in simulation for a network in United Kingdom, for a number of scenarios.

Finally, Chapter 7 concludes this thesis and annotates about potential future research directions.

1.4. Publications

In this section, the journals and the conference papers that have been produced from this thesis, are listed.

1.4.1. Journals

- Iordanidou G.R., Roncoli C., Papamichail I. and Papageorgiou M. Feedback-based mainstream traffic flow control for multiple bottlenecks on motorways. IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No. 2, pp. 610-621, April 2015.
- Iordanidou G.R., Papamichail I., Roncoli C. and Papageorgiou M. Feedback-based integrated motorway traffic flow control with delay balancing. IEEE Transactions on Intelligent Transportation Systems, accepted (doi: 10.1109/TITS.2016.2636302).

1.4.2. Conferences

- Iordanidou G.R., Papamichail I., Roncoli C. and Papageorgiou M. A Feedback-Based Approach for Mainstream Traffic Flow Control of Multiple Bottlenecks on Motorways. 19th IFAC World Congress (IFAC'14), Cape Town, South Africa, August 24-29, 2014, pp. 11344-11349.

- Iordanidou G.R., Papamichail I., Roncoli C. and Papageorgiou M. Feedback-Based Integrated Motorway Traffic Flow Control with Delay Balancing. 14th IFAC Symposium on Control in Transportation Systems, Istanbul, Turkey, May 18-20, 2016, pp. 315-322.

2. Theoretical Background

Chapter 2 contains the theoretical background that was used for the developments presented in this thesis. The motorway traffic congestion causes and effects, as well as the control measures, are explained in section 2.1. Section 2.2 reviews RM strategies and effects, whereas section 2.3 reviews the VSL strategies and effects. The MTFC concept, applications, implementation and integration with RM, are analyzed in section 2.4. The basics of traffic flow model formulation are described in section 2.5. Section 2.6 presents the optimal control problem formulation. Finally, section 2.7 describes the model parameter calibration procedure.

2.1. Motorway Traffic Congestion and Control Measures

The location where the upstream flow capacity is higher than the downstream flow capacity q_{cap}^{down}, is the definition of a (latent) bottleneck on a motorway (Figure 1).

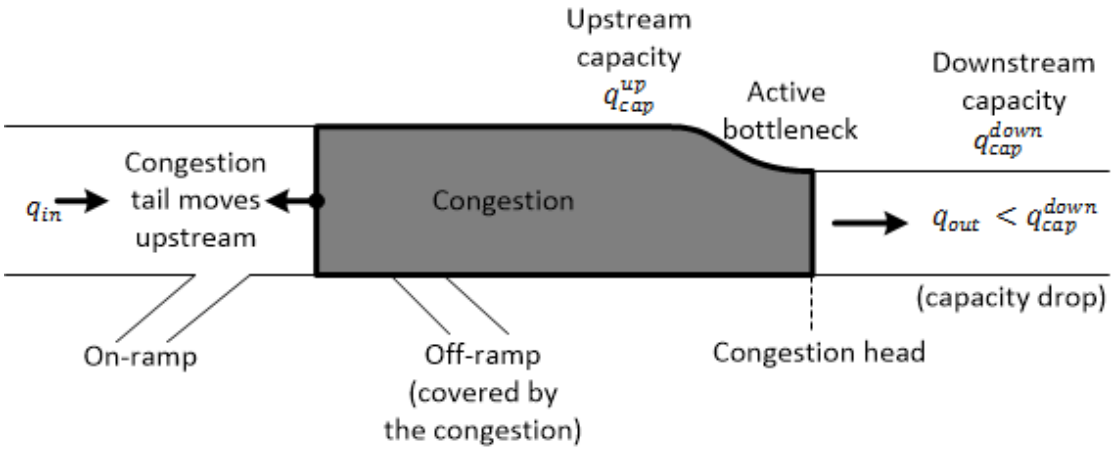

Figure 1: Active bottleneck notions.

A bottleneck can be activated due to various reasons. The most frequent reason is the on-ramp flow that merges with the mainstream flow. Another reason is the motorway infrastructure, namely, lane-drop, curvature, strong grade. Also, due to some special traffic conditions, e.g., strong weaving sections, which are created when a merge junction is followed closely by a diverge junction. The control measures could also be a cause of congestion, e.g., fixed speed limits on specific motorway locations. In

addition, non-recurring events that cause a reduction of roadway capacity may be a reason for congestion, e.g., incidents and reconstruction projects.

The activation of a potential bottleneck, depends on the traffic flow q_{in}, arriving upstream of the bottleneck location. In case q_{in} has the same value with q_{cap}^{down}, the nominal bottleneck capacity q_{cap}^{down} is reached, and the bottleneck is not activated. The value of q_{cap}^{down} is the maximum traffic flow that the bottleneck can accommodate. In the other case, where q_{in} is higher than q_{cap}^{down}, the bottleneck is activated, and as a result congestion is formed. The head of the congestion is located at the bottleneck, whereas the tail of the congestion is moving upstream while the upstream arriving flow is sufficiently high (Figure 1). When a bottleneck is activated, causing congestion, two detrimental effects are appearing, named capacity drop and blocking of off-ramps [1].

1) Capacity drop is a 5%-20% reduction of the active bottleneck outflow q_{out} compared with the nominal capacity q_{cap}^{down} (Figure 1). The capacity drop phenomenon is observed at the congestion head, and the speed reduction upstream of the bottleneck location, is the reason of its appearance. More specifically, the vehicles need to accelerate from lower speeds, in the formed congestion, to higher speeds, downstream of the bottleneck. There are many research publications in which capacity drop is observed in an empirical way, such as [16]. The avoidance of capacity drop at active bottlenecks would ameliorate the traffic flow conditions on motorways, and the throughput would be increased.
2) Blocking of off-ramps (BOR), upstream of active bottlenecks, is created because the tail of a formed congestion usually propagates many kilometers upstream (Figure 1). Several on-ramps and off-ramps upstream of the bottleneck are covered from the congestion. As a result the off-ramp flow is decreased, due to the lower traffic flow around the congested area, compared with the upstream arriving flow. Except of the vehicles within the congestion, also the vehicles that are bound for exits upstream of the active bottleneck are delayed. As the congestion expands, BOR expands also and deteriorate the

traffic conditions. Thus, if congestion mitigates spatially, throughput will be improved.

To deal with traffic congestion, traffic control measures are used, to decrease or ideally avoid traffic congestion effects. Traffic control measures can be categorized as follows: Ramp Metering (RM), Variable Speed Limits (VSL), Route Guidance (RG), and emerging vehicle-infrastructure integration (VII) systems.

Ramp metering (RM), controls the on-ramp inflow to the mainstream, and is implemented traffic lights. The goal of RM is to preserve capacity flow at the mainstream and avoid congestion [1]. Due to the restriction or avoidance of the congestion by RM, the mainstream throughput is increased, and also, the mainstream delays are reduced, leading to smaller total travel times for most drivers, despite the inevitable ramp queue delays. It is a fact that RM can improve mainstream traffic flow efficiency, but on the other hand it has a strict limitation. The limitation is that the RM created queue, should not spill back to the adjacent upstream infrastructure. The ramp storage space is usually restricted, so when the ramp queue has reached its storage capacity, RM is typically released. Despite the offered advantages of RM, which are the delay of the congestion onset, the acceleration of congestion dissolution and the decrease of congestion expansion, it has also some serious disadvantages. One of the disadvantages is that due to full ramps, RM should be de-activated, usually for the most of the peak periods duration or enter a queue management state. The unfair mainstream traffic inclination, compared with the on-ramp inflow, is another disadvantage of RM, which affect the drivers not to select the on-ramps route, considering the traffic lights [17]. For road users the criterion of equity is very important, so as to accept and use a control measure. There are RM strategies, called coordinated RM strategies, which use measurements for a whole region of a network, and control all the metered ramps included in this region [1]. Equity could be achieved, with RM actions coordination, at successive on-ramps, in contrast to local RM actions. Even for coordinated strategies, the storage space is substantial, to avoid mainstream congestion [2]. Sophisticated methods such as multivariable control strategies [18], optimal control strategies ([8], [19], [20], [21], [22]), and further heuristic algorithms [23], have been developed for coordinated RM.

Variable speed limits (VSLs) are speed limits that change based on road, traffic and weather conditions, and are displayed on road-side variable message signs (VMS). The range of VSLs, depends on legislative bodies of national or local governments, but on average the range is a subset of [60,120] km/h. The main impact of VSL is improved traffic safety, due to the homogenization of speeds of individual vehicles and of the mean speeds of different motorway lanes. So far, there was no field implementation evaluation of the VSL impact on traffic flow efficiency improvement, e.g., reduced travel times, except from the SPECIALIST approach ([24], [25]).

Route Guidance (RG) recommends optimal vehicle routes based on current traffic conditions, for travelling from the beacon location toward various destination zones. Usually, RG is important for drivers, when non recurrent events are happening, where the traffic conditions are unpredictable, e.g., accidents [26]. The RG proposed routes, should ensure that the available free capacity is enough, so as to accommodate the drivers.

Vehicle-infrastructure-integration (VII) systems, are at the top of the interest the last few years, and concerted research on these systems is done, in academia and industry [27]. Vehicles in such systems have a lot of capabilities, like computing, communication and sensing capabilities, as well as user interfaces. The position, the speed and the inter-vehicle distance of each vehicle is known, because vehicles are behaving like mobile sensors. Communication between vehicles, and between vehicles and infrastructure, is feasible, using wireless communication technologies. The messages a vehicle can receive are related usually with warnings, traffic information and speed limits that are sometimes automatically enforced. Research and development projects are focused on deployment, testing and demonstration of VII systems. A lot of applications have emerged, exploiting VII systems technologies that enhance traffic safety and efficiency, and also provide new or integrate existing services for drivers and passengers.

2.2. Ramp Metering

The current section contains an overview of RM strategies and RM effects.

2.2.1. Overview of Ramp Metering Strategies

The RM strategies could be categorized into 4 categories: fixed-time strategies, reactive RM strategies, nonlinear optimal RM strategies and integrated freeway network traffic control [1].

Fixed-time RM strategies are applied off-line for specific times of day, taking into account constant historical demands, with no use of real-time measurements. These strategies are based on simple static models. A motorway containing a number of on-ramps and off-ramps is subdivided into sections, where each section includes one on-ramp. So we have

$$q_j = \sum_{i=1}^{j} a_{ij} r_i \tag{1}$$

where q_j is the mainline flow of section j, r_i is the on ramp volume, in veh/h, of section i, and $a_{ij} \in [0,1]$ reflects the (known) portion of vehicles that enter the freeway in section i and do not leave the freeway upstream of section j. In order to avoid congestion the following inequality must hold

$$q_j \le q_{cap,j} \forall j \tag{2}$$

where $q_{cap,j}$ is the capacity of section j. Some more constraints are

$$r_{j,\min} \le r_j \le \min\left\{r_{j,\max}, d_j\right\} \tag{3}$$

where d_j is the demand, whereas $r_{j,\max}$ is the ramp capacity at on-ramp j, and $r_{j,\min}$ is the minimum flow, that always must be permitted to get in the freeway. Wattleworth [28], was the first that proposed this approach. Such formulations can be found in [29], [30], [31].

The objective function that could be used is the maximization of the served vehicles number, that is equivalent to the minimization of the total time spent

$$\sum_j r_j \to \max \tag{4}$$

or the maximization of the total traveled distance

$$\sum_j \Delta_j q_j \to \max \tag{5}$$

where Δ_j is the length of section j, or the balancing of the ramp queues

$$\sum_j \left(d_j - r_j\right)^2 \rightarrow \min \tag{6}$$

The formulations described above result in linear-programming or quadratic-programming problems, which could be solved readily with computer codes which are widely available. In [32], an extension of these methods is proposed that renders the static model **(1)** dynamic, with the introduction of constant travel times for each section.

The basic disadvantage of fixed-time strategies is that their settings rely on historical data and not on real-time data. This might be a bad simplification, because of the following reasons.

- Even within a time of day, demands are not constant.
- Demands are possible to differ at different days.
- In the long term, demands are changing and as a result the optimized settings are becoming old.
- As the demands are changing, also the portions a_{ij} are changing. A reason of the portions' change is the drivers' response to the new optimized signal settings, where the drivers exploit them so as to minimize their individual travel times.
- External disturbances, e.g., incidents, might affect traffic conditions in a non-predictable way.

Thus, fixed-time RM strategies, because of the absence of real-time measurements, lead either to an overload of the mainstream flow (congestion) or to underutilization of the freeway.

Reactive RM strategies are widely used and have as a goal the maintenance of the traffic conditions of the freeway close to preselected set values, relying on real-time measurements. Local RM strategies and multivariable regulator strategies belong to reactive RM strategies.

Local RM strategies compute appropriate RM values, using traffic measurements around a ramp. The demand-capacity strategy [33] applies

$$r(k)=\begin{cases} q_{cap}-q_{in}(k-1), & if \quad o_{out}(k)\le o_{cr} \\ r_{\min}, & else \end{cases} \tag{7}$$

where (see Figure 2) q_{cap} is the freeway capacity downstream of the ramp, q_{in} is the freeway flow measurement upstream of the ramp, o_{out} is the freeway occupancy measurement downstream of the ramp, o_{cr} is the critical occupancy, and $r_{\min}$ is a preselected minimum ramp flow value. The strategy described with (7), tries to add to the measured upstream flow $q_{in}(k-1)$ as much ramp flow $r(k)$ as needed to reach the downstream freeway capacity q_{cap}. In case the downstream measured occupancy $o_{out}(k)$ becomes overcritical (i.e., a possible forming of congestion) the ramp flow $r(k)$ is decreased to the minimum flow $r_{\min}$, so as to avoid or dissolve the congestion. In fact, strategy (7) does not depict a closed-loop strategy but an open-loop disturbance-rejection policy (see Figure 2(a)) that is widely known to be sensitive to various non-measurable disturbances.

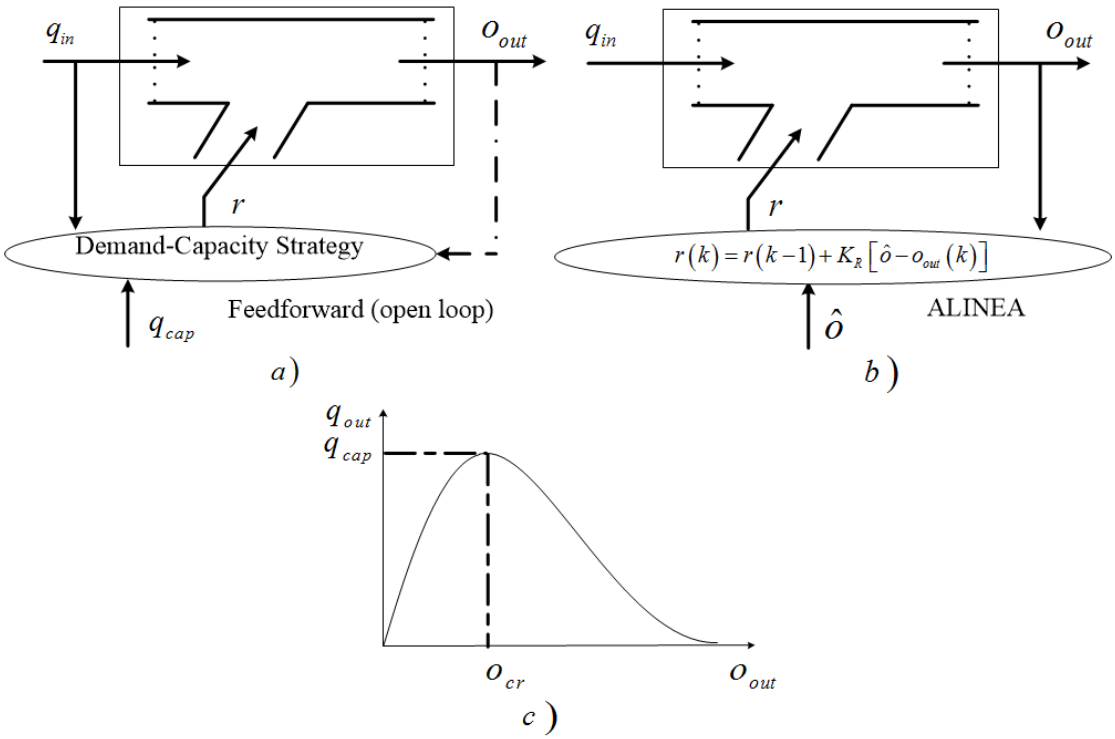

Figure 2: Local RM strategies: a) Demand-capacity, b) ALINEA, c) Fundamental diagram.

The occupancy strategy [33] is based on the same philosophy as the demand-capacity, with the difference that relies on occupancy-based estimation of q_{in}. This might decrease the corresponding implementation cost, under specific conditions.

In [34], an alternative closed loop RM strategy, called ALINEA , is proposed that reads

$$r(k) = r(k-1) + K_R\left[\hat{o} - o_{out}(k)\right] \quad (8)$$

where (see Figure 2(b)) $K_R > 0$ is a regulator parameter and $\hat{o}$ is a set (desired) value for the downstream occupancy. Usually, $\hat{o} = o_{cr}$ may be set, whereby the downstream freeway flow becomes close to q_{cap} (see Figure 2(c)) (the fundamental diagram presented in this Figure is described in more details in section 2.3). The selection of the regulator parameter K_R for ALINEA, is not very sensitive in field implementations. The value $r(k-1)$ of equation (8) must be the value of the previous time step after the bounding so as to prevent the well-known wind-up phenomenon for I-type regulators.

The demand-capacity strategy counteracts in a rather crude way, to enormous occupancies o_{out}, only if a threshold value (o_{cr}) is exceeded, whereas ALINEA counteracts smoothly, also for very small differences $\hat{o} - o_{out}(k)$. So, ALINEA stabilizes the traffic flow at a high throughput level and as a result may avoid congestion. At a stationary state (i.e., if q_{in} is constant), it is obvious that $o_{out}(k) = \hat{o}$ results from (8), even though the measurements of the inflow q_{in} are not explicitly used in the strategy.

The set value $\hat{o}$ can be changed at any time, so ALINEA may be incorporated into a hierarchical control system with the set values of the individual ramps being defined in real time by a superior coordination level or by an operator.

Appropriate ramp volumes r are being calculated by all the control strategies [1]. If traffic-cycle realization of RM is selected, ramp volume r is converted to a green-phase duration, g using

$$g = (r/r_{sat}) \cdot c \quad \textbf{(9)}$$

where c is the fixed cycle time and r_{sat} is the ramp's saturation flow. The constraints of the green-phase duration g are: $g \in [g_{\min}, g_{\max}]$, where $g_{\min} > 0$ in order to prevent ramp closure, and $g_{\max} \le c$. When the one-car-per-green realization holds, a constant-duration green phase allows only one vehicle to pass. So, the ramp volume r is controlled by varying the red-phase between a minimum (zero) and a maximum value. It is noteworthy that ALINEA can be applied directly to the green or red phase duration, by combining **(8)** and **(9)**

$$g(k) = g(k-1) + K_R' \left[\hat{o} - o_{out}(k) \right] \quad \textbf{(10)}$$

where $K_R' = K_R \, c / r_{sat}$. The value $g(k-1)$ of equation **(10)** must be the value of the previous time step after the bounding (i.e., after the constraints application, of the $g_{\min}$ and $g_{\max}$) so as to prevent the wind-up phenomenon in the regulator.

In case a very big ramp queue of vehicles is created, interference with the adjacent street traffic may occur. Appropriately located detectors (upstream of the on-ramp) could detect this case, leading to regulator decisions override, to permit more vehicles to enter the freeway and decrease the ramp queue.

The previously described specifications and constraints are applied in the same way to any RM strategy.

The efficiency of local RM strategies has been estimated and compared in various countries with comparative field tests, see, e.g., [35]. The result from these tests is that ALINEA prevails compared over other local strategies and over the no-control case, taking into account any performance criterion, like total time spent, total travelled distance, mean (daily) congestion duration, mean speed. The local RM strategies improvements of the total time spent (incorporating the ramps' waiting time), can reach 20%.

Multivariable regulator strategies for RM aim at the same objective as local RM strategies, because they are trying to keep the traffic conditions close to some predefined set (desired) values. Local RM is implemented independently for each ramp, relying on local measurements, whereas multivariable regulators exploit all the

available mainstream measurements $o_i(k), i=1,\ldots,n$, on a freeway stretch, to calculate at the same time the ramp volume values $r_i(k), i=1,\ldots,m$ for the ramps that are controlled and included in the same stretch [36]. This difference between local RM and multivariable regulators, leads to multivariable regulators superiority, due to more complete information forecasting and due to coordinated control actions. Some of the reported approaches for RM with multivariable regulators are in [36], [37], [38], [39], [40]. A generalization of ALINEA is the multivariable regulator strategy METALINE, with the difference that the metered on-ramp volumes are calculated from (vectors and matrices are displayed with bold variables)

$$\boldsymbol{r}(k)=\boldsymbol{r}(k-1)-\boldsymbol{K}_1\left[\boldsymbol{o}(k)-\boldsymbol{o}(k-1)\right]+\boldsymbol{K}_2\left[\hat{\boldsymbol{O}}-\boldsymbol{O}(k)\right] \quad \textbf{(11)}$$

where $\boldsymbol{r}=\left[r_i\ldots r_m\right]^T$ is the vector of m controllable on-ramp volumes, $\boldsymbol{o}=\left[o_1\ldots o_n\right]^T$ is the vector of n measured occupancies on the freeway stretch, $\boldsymbol{O}=\left[\hat{O}_1\ldots\hat{O}_m\right]^T$ is a subset of $\boldsymbol{o}$ that comprises m occupancy locations for which predefined set values $\hat{\boldsymbol{O}}=\left[\hat{O}_1\ldots\hat{O}_m\right]^T$ could be given. The number of set-valued occupancies, must not be higher than the number of controlled on-ramps, because of control-theoretic reasons. Usually, one bottleneck location is selected downstream of each controlled on-ramp, for inclusion in the vector $\boldsymbol{O}$. At the end, $\boldsymbol{K}_1$ and $\boldsymbol{K}_2$ are the constant gain matrices of the regulators that must be designed appropriately, see [36] and [41] for details.

Concerning METALINE versus ALINEA efficiency, field implementations and simulations results have concluded in the following [35]:

- ALINEA does not require any design effort, whereas METALINE application needs a sophisticated design procedure that relies on advanced control-theoretic methods (LQ optimal control).
- In case of urban freeways with high on-ramps density, METALINE did not provide advantages over ALINEA (ALINEA applied independently at each controllable on-ramp) under recurrent congestion.
- In case of non-recurrent congestion (e.g. because of an incident), METALINE prevails compared to ALINEA, because of more completed measurement information.

There is reluctance in some system operators, to implement RM due to the possibility that congestion will be moved from the freeway to the adjacent street network. In reality, an RM application could create both positive and negative effects on the adjacent road network traffic. It is obvious, based on previously notions and declarations, that in case an efficient control strategy is applied for RM, the throughput of the freeway will be increased. In more details, when the rush hour begins, queues at the on-ramps may be created, so as to avoid congestion forming on the freeway, that may results in the temporarily diversion toward the urban network. Nevertheless, because of the avoidance or reduction of the congestion, the freeway will be able to accommodate a higher throughput, so the drivers from urban network will be attracted and the overall network performance will be improved. In the Corridor Peripherique in Paris, France, a specially designed field evaluation had been done, that depicted the positive effect of RM on the freeway and the adjacent road network traffic conditions [42].

Nonlinear optimal RM strategies prevail, compared to reactive RM strategies, because the latter may assist until a certain extent, but they require suitable set values, and also they have more or less a local character. A superior coordination level is needed for freeway networks or long stretches, that calculates in real time optimal and equitable set of values from a preventive, strategic point of view. An optimal control strategy like this, should clearly take into consideration:

- the prevailing traffic state, both on the freeway and on the on-ramps
- predictions of the demand over a sufficiently long time period
- the restricted on-ramps storage capacity
- constraints of RM mentioned above
- nonlinear traffic flow dynamics, comprising the restricted capacity of the infrastructure
- the incidents currently appearing in the freeway network.

The control strategy, relying on this complete information, should deliver set values for the whole freeway network, over a time period, in order to respect all present constraints and minimize an objective criterion, like the total time spent in the overall network (comprising the on-ramps).

Appropriate dynamic models, can express the nonlinear traffic dynamics, in the form

$$\boldsymbol{x}(k+1) = \boldsymbol{f}\left[\boldsymbol{x}(k), \boldsymbol{r}(k), \boldsymbol{d}(k)\right] \tag{12}$$

where the state vector $\boldsymbol{x}$ includes all traffic densities and mean speeds of freeway sections, as well as all the ramp queues; the control vector $\boldsymbol{r}$ includes all controllable ramp volumes; and the disturbance vector $\boldsymbol{d}$ includes all on-ramp demands and turning rates (at network bifurcations or at off-ramps). The constraints of RM are given by **(3)** and the constraints of the queue read

$$l_i(k) \le l_{i,\max} \tag{13}$$

where l_i queue length (in vehicles) for each on-ramp i. The total time spent in the whole system over a time period K may be defined as

$$T_s = T\sum_{k=0}^{K}\left[\sum_{i=1}^{n}\rho_i(k)\cdot\Delta_i + \sum_{i=1}^{m} l_i(k)\right] \tag{14}$$

where $\rho_i(k)$ is the traffic density (in veh/km) in segment i at time $k\cdot T$.

So, for a given current (initial) state $\boldsymbol{x}(0)$ from corresponding measurements or estimates, and given demand predictions $\boldsymbol{d}(k)$, $k = 0,\ldots,K-1$, the problem composes in defining the ramp flows $\boldsymbol{r}(k)$, $k = 0,\ldots,K-1$, in order to minimize the total time spent **(14)** subject to the nonlinear traffic flow dynamics **(12)** and the constraints **(3)** and **(13).**

Plenty of reported works have considered and solved the previously described problem, and variations of it [19], [20], [43]-[44].

Integrated freeway network traffic control comprises different types of control measures which coordinate. The corresponding control strategies, ordinarily are designed and applied independently, thus failing to take advantage of the synergistic impacts that may be obtained from coordination of the respective control actions. An appropriate extension of the optimal control approach described above, gives a sophisticated concept for integrated freeway network traffic control. More specifically, the dynamic model **(12)** of freeway traffic flow could be extended, so as to be able to incorporate more control measures, in addition to RM rates $\boldsymbol{r}(k)$. Then

$\boldsymbol{r}(k)$ is replaced in (12) by a general control input vector $\boldsymbol{u}(k)$ that includes all applied control measures of any type. An approach like this was implemented in the integrated, generic freeway network control tool AMOC [45], [46], where RM and route guidance, as also RM and VSL are considered simultaneously.

2.2.2. Ramp Metering Effects

In Figure 3 there are two instances for a motorway on-ramp, the first instance (a) is without RM, and the second (b) is with RM, where q_{in} is the upstream motorway flow, d is the ramp demand, q_{con} is the mainstream outflow in presence of congestion, and q_{cap} is the freeway capacity. If congestion exists, it is known that q_{con}

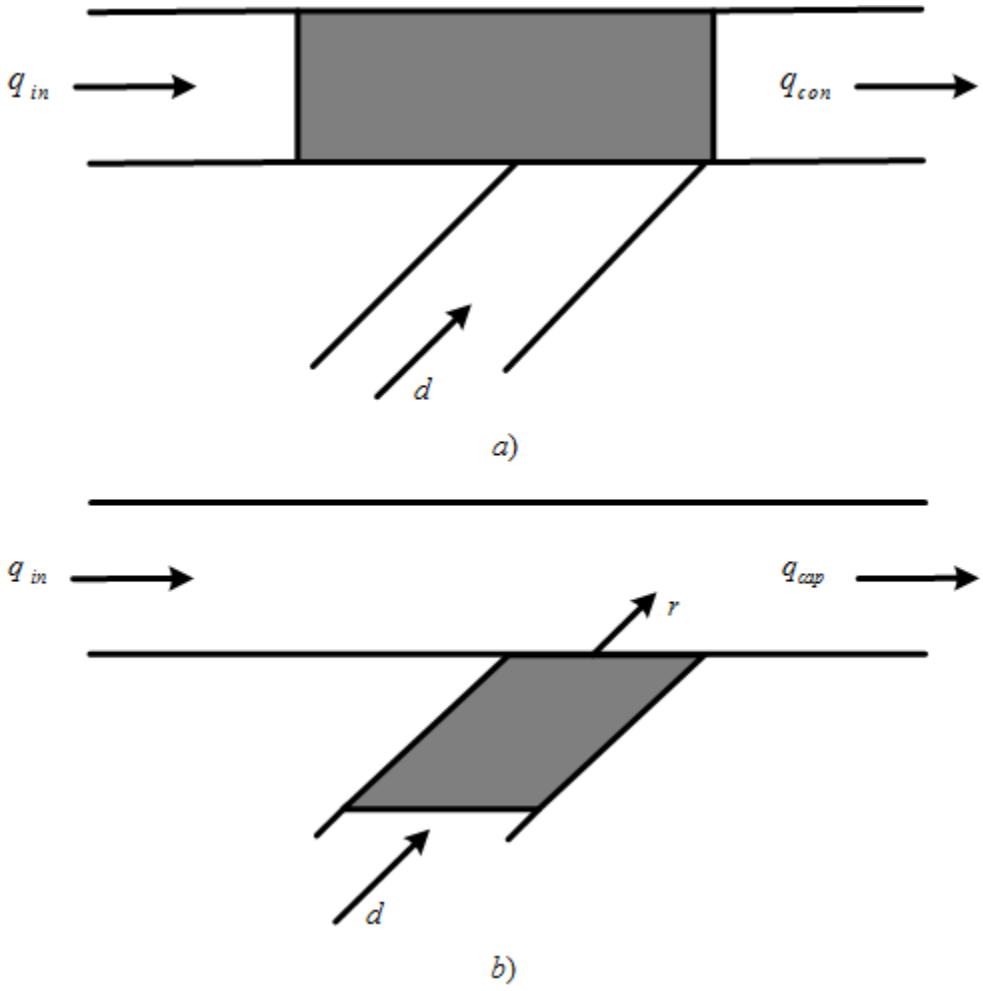

Figure 3: Two instances: (a) without RM and (b) with RM, the grey areas show congestion zones.

is lower by some 5%-10%, compared to the motorway capacity [1]. In Figure 3(b), RM is implemented, so as to keep capacity flow on the mainstream. An impact of RM is the creation of a queue at the on-ramp, but because q_{cap} is higher than q_{con}, RM has as a result the reduction of the total time spent (incorporating the waiting time at ramps). The total time spent improvement on ΔT_s (in %) [47], is given by

$$\Delta T_s = \frac{q_{cap} - q_{con}}{q_{in} + d - q_{con}} 100 \tag{15}$$

In Figure 4 there are two instances of a motorway stretch that contains an on-ramp and an off-ramp, the first instance (a) is without RM control, and the second (b) is with RM control. The assumption that $q_{con} = q_{cap}$ is considered, i.e., no capacity drop phenomenon because of congestion, so as to clearly distinguish the different impacts of RM. The part of the upstream flow that leaves at the off-ramp, is defined with the exit rate $\gamma(0 \le \gamma \le 1)$, so it can be easily proved [47] that the exit flow with no control

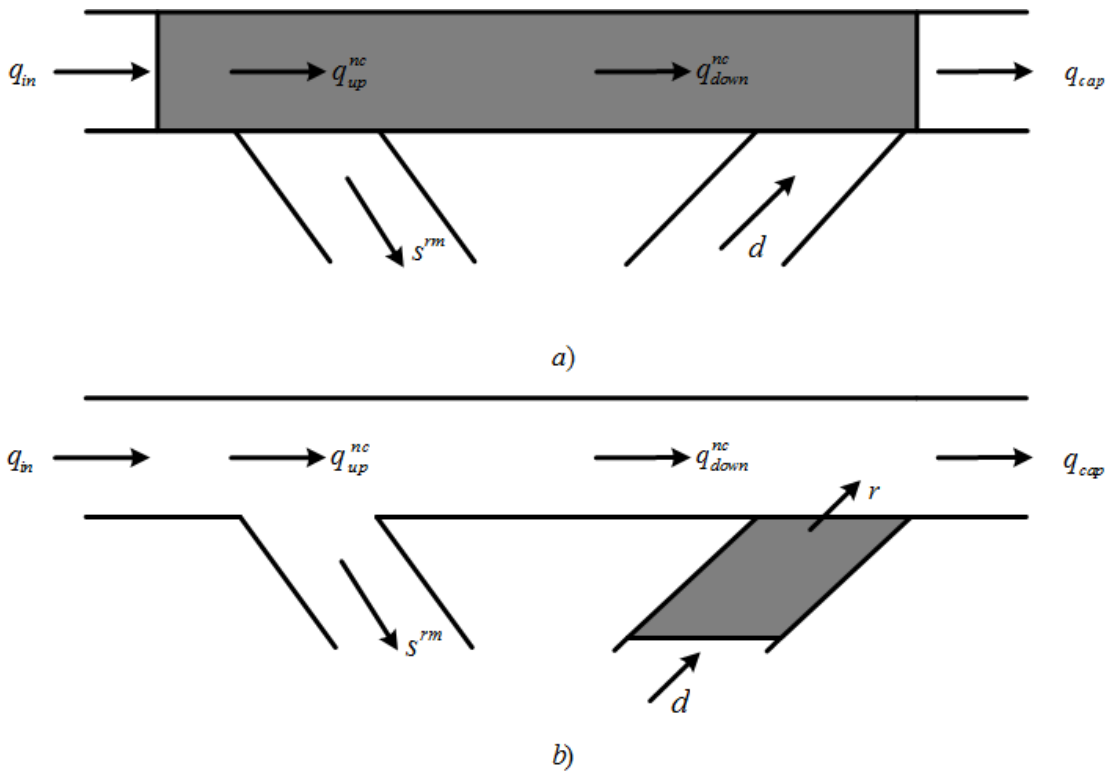

Figure 4: Two instances: (a) without RM and (b) with RM.

is given by

$$s^{nc} = \frac{\gamma}{1-\gamma}\left(q_{cap} - d\right) \tag{16}$$

whereas with metering control

$$s^{rm} = \gamma \cdot q_{in} \ . \tag{17}$$

For the reason that $(1-\gamma)q_{in} + d > q_{cap}$ holds (otherwise no congestion would have been created), it follows that s^{nc} is less than s^{rm}, so RM increases the outflow thus decreasing the total time spent in the system. The improvement of the total time spent in this case can be easily shown that is equal to

$$\Delta T_s = \gamma \cdot 100 \ . \tag{18}$$

The selection of the routes, from the road users, toward their destinations, has as a goal, the minimization of their individual travel times. In case a control measure (e.g., RM) is introduced that might affect the delay experienced in some network links (e.g., on-ramps), a part of the drivers consequently will change their ordinary route, so as to take advantage from, or avoid disadvantages due to the new network conditions. In Figure 3(b), for example, the upstream flow q_{in} will possibly increase, whereas the ramp demand d will decrease as compared to Figure 3(a). The route choice behavior of drivers is predictable to a large extent, so RM could be used also in order to enforce an operationally desired traffic flow distribution in the whole network, e.g., rat-running phenomenon avoidance, increased or decreased use of underutilized or overloaded, respectively, parallel arterials, etc. For sure, the amended routing behavior of drivers should be taken into consideration in the design and evaluation procedures of RM control strategies.

Many field evaluation conclusions (see e.g., [48]) depict that RM enhances the merging behavior of traffic flow at freeway intersections which possibly have an important positive effect on traffic safety because of less lane changes and decreased driver stress. In addition, the network efficiency increase is anticipated to lead to accordingly improved network traffic safety and reduced pollutant emissions to the environment.

2.3. Variable Speed Limits

This section includes an overview of VSL strategies, a description for the fundamental diagram and the VSL impacts.

2.3.1. Overview of VSL strategies

Field implementations of VSLs are numerous; the oldest one dates from the 1960s, in USA [49]. Some of the countries that have implemented VSL are Australia, Canada, England, Germany, Greece, USA, Sweden [50], [51], [52], [53]. The scope of this section is not to describe in detail the VSL implementations, but to present the main common characteristics of VSL strategies.

VSLs are used to define the maximum or in some cases the minimum speed, at which vehicles can legally drive on specific motorway stretches, and they are changed, taking into account the current traffic and weather conditions. The speed limit in each location is indicated with VMS on appropriate gantries, or on roadside poles, that are usually spaced 0.5-2 km. Across work zones, VSLs are displayed on portable trailers, that allow drivers to keep the most efficient and safe speeds, so as not to put themselves, the others drivers and the workers in danger. VSLs can be used either as a standalone system or as a part of a managed motorway, such as a congestion management system.

The biggest part of the mentioned VSL systems, have a common structure. Firstly, they have the data collection and processing, then the control algorithm or decision logic, and at the end the display. Among these elements there is a communication system, for data exchange.

The data collection procedure is performed with various types of detectors and sensors, according to the application. Normally, traffic detectors are measuring, speed, flow and occupancy. Many applications use as inputs weather conditions, thus measurements like temperature, rain and snow intensity, humidity, wind speed and direction, are important. Surface measurements, i.e., measurements of pavement conditions, like wet, salted, snowy, are sometimes helpful, for some applications. After the data collection, data are processed and then transmitted to the control logic of the system. At this stage, faulty or missing data is treated. Events like, congestion, incidents, constructions are affecting the system operation.

The control logic, of a VSL system is the most important part of the system, and makes it responsive to dynamic conditions. The biggest limitation in most of VSL systems is the control logic. The problem is, that VSL systems rely on simple threshold based control logics, but these thresholds are very sensitive to day-to-day stochastic variations, requiring tuning of the parameters, which is a painful procedure. In addition, most of the strategies were designed with a lack of a proper knowledge on VSL traffic impacts and on the operation which lead to successful objectives. The control logic contains the application of coordination between displayed VSL, and the application of standard rules and restrictions. VSL systems are usually automated, but nevertheless, manual operators are also considered, together with a closed-circuit television system.

Display updating is common every one or two minutes. The display of the speed limits is in speed intervals of multiples of 5 and 10, for miles per hour and kilometers per hour, respectively. In some implementations, not only the VSL is displayed, but also the minimum admissible speed limit. Sometimes there are auxiliary signs which provide the users with real time information, like advices, warning messages and others, that seem to improve the operation and the users' perception about the system. Variable speed limits can be either advisory or mandatory, but usually are mandatory, i.e., they must be obeyed by the law, and in some cases they are automatically enforced, for better driver compliance.

There are numerous VSL objectives that can be categorized in three groups: improve safety, more efficient use of highway and environmental effect reduction. It is usual for some systems to have objectives that belong to more than one of the three aforementioned groups. Most of the systems are focusing on safety improvement, so they are displaying appropriate speed limits, encountering traffic conditions. For the cases of safety improvement, some specific objectives are: reduction of driver error, maximization of driver compliance, infractions reduction. The systems that aim to highway efficiency improvement, aim to reduce the travel times by stabilizing the traffic flow, because of more uniform speeds. So far, very few applications indicate a traffic flow efficiency improvement, and if there is an improvement, it is very small. Concerning the reduction of the environmental effect, very few systems are considered it. There are some applications that reported a very small reduction on

pollutants emissions. Further research based on environmental aspects is needed, see, e.g., [54], [55].

The conclusions from existing VSL applications are not clear. Concerning the accidents rates reduction, in some applications it reaches the 20-30 percent, whereas on others applications the reduction is minor. In most applications the average speed and lane changing is reduced, the drivers' compliance is increased and the distribution of lane use is better. Improvements on the traffic flow throughput and reductions of travel times are, if any, minor.

2.3.2. Fundamental Diagram Description

The fundamental diagram represents approximately the traffic flow states in the steady state (i.e., traffic conditions changes are minor in space and time), in a given road-section, and can be a flow-occupancy or flow-density diagram, which have an inverse U shape, or a speed-flow diagram, which has a left-turned U shape, (see Figure 5). For the case of a flow-occupancy diagram (see Figure 5(a)), the mean speed of a particular traffic state is proportional to the slope of the line that connects the particular traffic state point with the origin. For the production of a fundamental diagram for a specific motorway location, traffic variables measurements (flow,

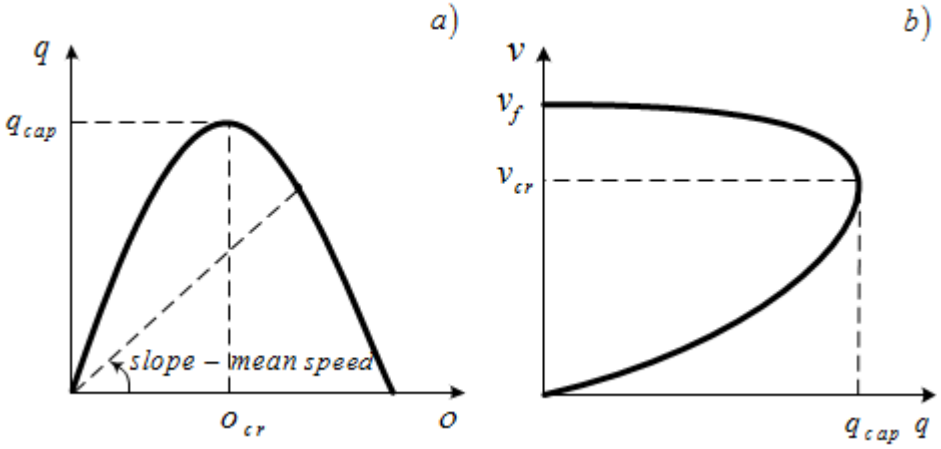

Figure 5: α) Flow-occupancy and, b) speed-flow diagrams, where: q is flow (veh/h), o is occupancy (%), v is mean speed, q_{cap} is capacity flow, o_{cr} is critical occupancy, v_f is free speed, and v_{cr} is critical mean speed.

occupancy, mean speed) are needed, that are fitted in an appropriate mathematical function. However, if the underlying spatiotemporal traffic flow phenomena are not considered appropriately, the procedure of the fundamental diagram production may lead to faulty results. More specifically, in order to be able to recognize the area around the critical occupancy (capacity flow), using real data, measurements from active bottleneck locations are required (see [15] for more details).

2.3.3. VSL Effects

This section contains a description of the preliminary results and the potential VSL impact.

2.3.3.1. Preliminary Results

In the past, there were very few investigations reporting the exact impact of VSL on aggregate traffic flow behaviour, e.g., on the fundamental diagram (flow-density curve). Zackor [56], summarized some early investigations, using traffic data with and without VSL, for a two-lane German motorway. A speed homogenization is observed, for individual vehicles, and for motorway lanes, under the impact of VSL. Zackor's results are helpful for a better VSL impact comprehension on individual vehicle speed distribution, whereas the impact of VSL on aggregate traffic flow behavior is not of the critical occupancy or critical density, under the impact of VSL.indicated. This notification was reported also in [56], but somehow in a qualitative way. In Figure 6(a) it is depicted that "at lower or mean traffic volumes, the mean speed is lower due to the reduction effect whereas, at higher volumes, an increase is detected due to the stabilizing effect. Thus, both capacity and speed rise by about 5 to 10 percent at the same time" [56]. Zackor [56] did not address the potential increase of the critical occupancy or critical density, under the impact of VSL.

Figure 6(b) depicts the VSL-induced fundamental diagram change, using a quantitative model proposed in [57]. The results addressed in [56] were the basis for this proposal. b in this figure is the ratio of the applied VSL divided by the free speed v_f without considering VSL; b equal to 1, by convention, corresponds to no-VSL case. The depicted capacity flow increase is rather excessive. Some later Dutch investigations did not observe any capacity increase, that could be assigned to VSL[58], although under advisory and not mandatory VSL.

In more lately research concerning VSL control, the considered VSL impact was to

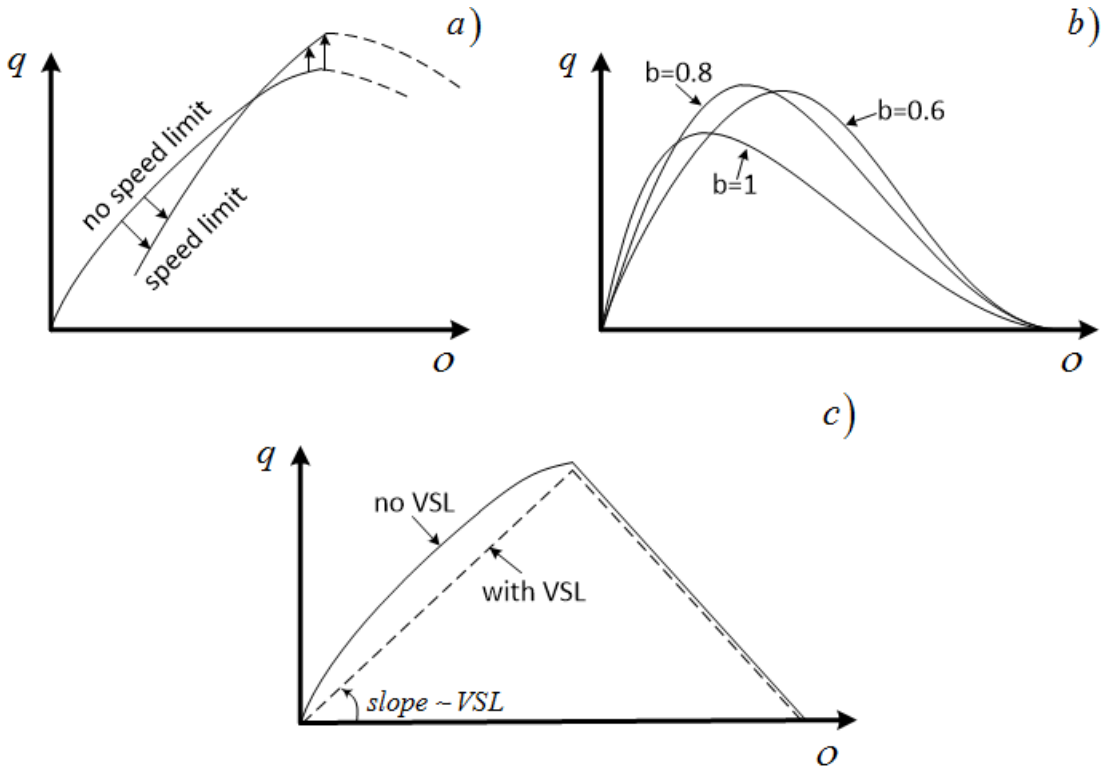

Figure 6: α) Fundamental diagram change due to speed limits [22]; b) Cremer model for VSL impact [23], where b = 1, corresponds to no speed limit, b = 0.8 corresponds to VSL = 0.8 , and b = 0.6 corresponds to VSL = 0.6 ; c) Hegyi model for VSL impact.

just replace the left part of the flow-occupancy curve by a straight line with slope corresponding to the displayed VSL, (see Figure 6(c)) [59]. Various subjects concerning the impact of VSL in modelling and control have been explored recently, see, e.g. [50], [60], [61], [62].

Concluding, the empirical proof seems to be very limited, and in fact no real consensus on the possible impact of VSL on aggregate traffic flow behavior. The prospects of VSL impact, along with a summary of the basic findings by [10], are following below.

2.3.3.2. Mean Speed Reduction at Undercritical Occupancies

The assumption that a VSL, displayed at undercritical occupancies, will decrease (with the appropriate driver compliance) the (otherwise higher) mean speed (Figure 7(a)), seems to be reasonable enough. The displayed VSL and the driver compliance are likely to be responsible for the range of this effect. The VSL-affected states serve the same flow but at lower speed and higher occupancy, that entails increased travel

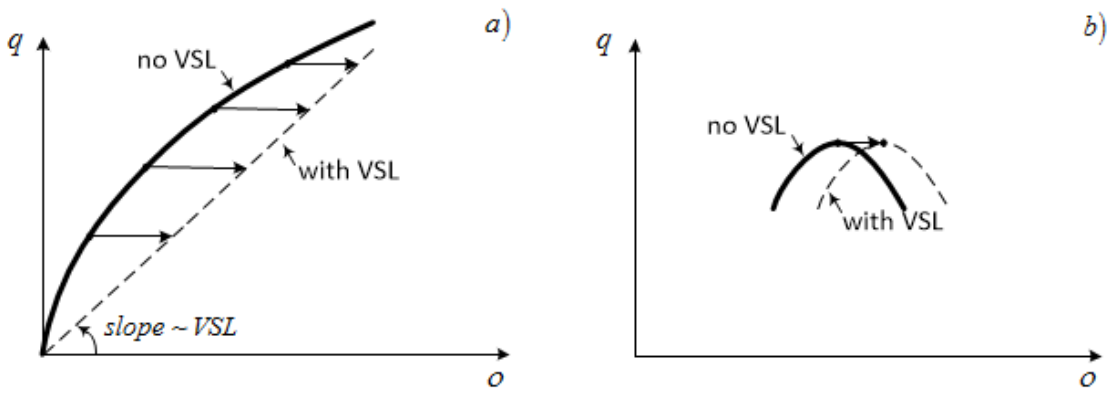

Figure 7: α) Potential impact of VSL on undercritical mean speeds; b) diagrams cross-point, with and without VSL.

times accordingly. The implementation of VSL at undercritical occupancies is very possible to increase travel times and as a result exacerbate traffic flow efficiency. The state transition described above, when implementing VSLs at undercritical occupancies, could be exploited in a different way. When VSLs are implemented upstream of a bottleneck, that soon is going to become active, provisionally (for the traffic state transition period, triggered by the VSL) will decrease the mainstream flow arriving in the bottleneck area, and as a result, the bottleneck activation and the resulting congestion are retarded. The reason of the temporary flow decrease, during the VSL-triggered traffic state transition, is that in the VSL state, the occupancy (and density) is higher than in the non-VSL state. Thereby, the flow is temporarilydecreased during the transition, to create the higher traffic density of the VSL state. This VSL impact is the main one exploited by Hegyi [59].

A significant issue is that the impact of VSL activation upstream of a possible bottleneck at undercritical occupancies, as depicted in the state transition of Figure 7(a), presents many similarities with the impact of local RM, in the case of limited ramp storage space. The similarities are following: The mainstream flow arriving at the bottleneck area, during the state transition of Figure 7(a), is decreased similarly to RM, with the difference that the ramp flow is decreased instead of the mainstream flow, so as to avoid or delay the congestion onset of the merge area. The mainstream density during the state transition of Figure 7(a), is increased similarly to RM, with the difference that in RM the vehicles are stored in the ramp and not in the mainstream. When the state transition has been finished, the mainstream flow returns to its pre-transition values (nearly equal to the upstream arriving mainstream

demand), similarly to RM where the ramp queue is released, when it covers the whole ramp, so as to avoid interference with the adjacent motorway. In more details, when the free storage space is about to be exhausted, RM actions are stopped and queue control is activated, trying to keep a maximum admissible ramp queue (see [4], [63], [64]), where the ramp outflow becomes significantly equal to the arriving ramp demand. When the state transition has been finished, the mainstream density keeps its increased value similarly to RM, where the queue of the on-ramp remains full until the arriving demand falls to sufficiently low values. In case of VSL activation upstream of a mainstream bottleneck in undercritical conditions, delays are produced for the concerned vehicles, due to the lower speed, similarly to RM, where the vehicles queueing on the ramp are delayed, but this impact is compensated by the avoidance or retarding of the bottleneck congestion and its related vehicle delays.

A theoretical analysis of VSL and RM application, and their similarities, are referred in [65], [66].

2.3.3.3. Throughput Increase and Congestion Delay at Overcritical Occupancies

The flow-density curves (for VSL and non-VSL), according to the Hegyi model [59] (Figure 6(c)), meet but do not cross, whereas according to Zackor [56], there is a genuine cross-point of both curves, somewhere near the critical occupancy (Figure 6(a)). If indeed there are cross-points, they likely lie at increasing occupancy values for decreasing VSL, due to the accordingly decreasing slope of the undercritical VSL-affected curves. It is possible that there is no cross-point for very low VSL. Since the VSL impact is assessed at occupancies near or higher than the cross-point, the following, partly overlapping questions are essential.

- Where is the cross-point (if any) located with respect to the non-VSL critical occupancy?
- Are VSL-induced critical occupancies higher than their non-VSL counterparts?
- Are VSL-induced flows higher at overcritical occupancies than their non-VSL counterparts?
- Is there a flow capacity increase for some VSL?

The above aspects are partly examined by [10], and the corresponding conclusions are summarized in the following section.

2.3.4. Variable Speed Limit Impact on Aggregate Traffic Flow Behaviour

The variable speed limits impact on aggregate traffic flow behaviour, in the form of the flow-occupancy diagram, was studied in [10], using traffic data from a European motorway, where a flow/speed threshold-based VSL control algorithm was implemented. The basic interest of this work was to confirm some long-held conjectures (Section 2.3.3) concerning the VSL impact on the shape of the flow-occupancy diagram.

Some conclusions of the work reported previously, are summarized in what follows.

- When speed limits are implemented at undercritical occupancies, they have the effect of decreasing the slope of the flow-occupancy diagram. In addition, the smaller the imposed speed limit, the larger the decrease in the slope of the flow-occupancy diagram. This impact may be helpful to keep the traffic flow back and retard the congestion onset at downstream bottlenecks, as discussed in Section 2.3.3.2.
- The VSL-affected flow-occupancy curve intersects (at least for some VSL) the non-VSL curve, and the critical occupancy is moving to higher values in the flow-occupancy diagram. The most cross-points were depicted to lie around or beyond the non-VSL critical occupancy. This impact could be used to keep more vehicles in the motorway, without getting into the congestion. These cross-points reveal that the mean speed at overcritical densities is higher when a speed limit is imposed than in no-VSL cases, and a reason of that may be due to the homogenization effects, that were discussed earlier.
- Concerning the possible capacity flow increase, the data investigation was rather inconclusive, because while a slight increase is visible for some VSL at some locations, at other locations there is no observable capacity increase for any VSL value. An appropriately designed control strategy, for throughput maximization could be used at the locations where VSL yield a capacity increase, as practiced in [67].

- For sufficiently low VSL values, capacity flow is lower in the fundamental diagram compared with the non-VSL cases, independently of the fact that capacity flow may increase for some VSL values or not. This effect has as a result the deliberately creation of a controllable mainstream congestion with benefit upstream of an uncontrolled potential bottleneck, to avoid its activation and the throughput reduction, because of the capacity drop.

2.4. Mainstream Traffic Flow Control

This section consists of the basic MTFC concept, its implementation as well as its integration with RM.

2.4.1. MTFC Basic Concept

The basic idea of MTFC is to regulate the mainstream traffic flow upstream of a bottleneck location with a suitable control strategy, so as to avoid capacity drop. In Figure 8, a local aspect of this basic idea is illustrated. As long as $q_{in} \le q_{cap}^{down}$ holds, the bottleneck in Figure 8 is not activated (and no MTFC is needed): In that case $q_{out} \approx q_{in}$. The bottleneck would be activated in case q_{in} becomes higher than q_{cap}^{down}, causing the reduction of q_{out}. Using MTFC, a controlled outflow q_c is implemented equal to the bottleneck capacity (or less if the bottleneck is because of a merging on-ramp). Clearly, MTFC cannot avoid mainstream congestion completely because $q_{in} > q_{cap}^{down}$; nevertheless it has some benefits as follows.

- The congestion outflow in the MTFC case is bigger than in the no-control case, because the capacity drop is avoided; thus the harmful effects of capacity drop are reduced.
- Because of the bigger outflow with MTFC, the created congestion in the MTFC case (a) has higher internal speed and (b) compared to the no-control case is space-time shorter. As a result the harmful effects of BOR are potentially improved, because of blocked off-ramps reduction.

The effects of BOR cannot completely avoided with MTFC, because anyway the MTFC strategy leads to a (controlled) mainstream congestion. Hence, MTFC is less efficient than RM (with sufficient on-ramp storage space), which is able to completely

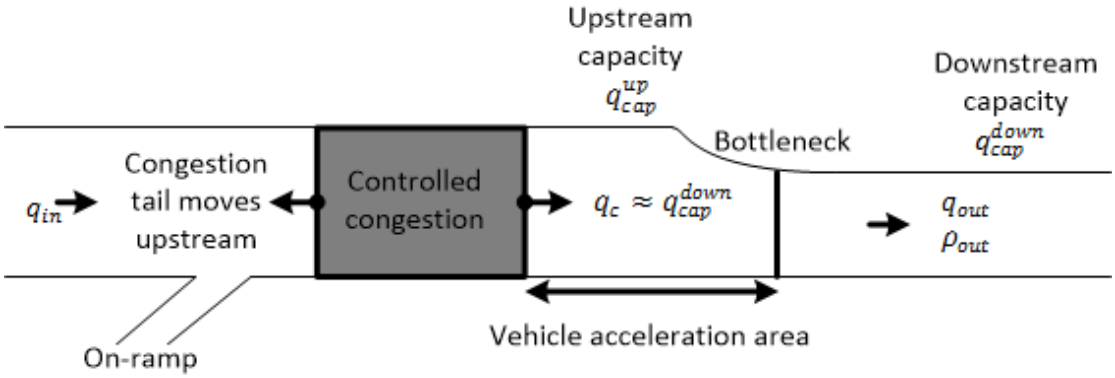

Figure 8: MTFC local aspect.

avoid congestion. However, the simulation results reported in [68] and [69] indicate that the difference in terms of cost function may be small.

The reason of the capacity drop phenomenon occurrence, seems to be the need for vehicles to accelerate from low speeds within the bottleneck congestion to higher as they reach the congestion head [70]. As shown in Figure 8, a mainstream controlled congestion will be created upstream of the flow control location and vehicles exiting this controlled congestion area will have a relatively low speed, that is possible to be lower than the critical speed v_{cr} leading to bottleneck capacity flow q_{cap}^{down} (Figure 5(b)). In order to achieve to have capacity flow q_{cap}^{down} at the downstream bottleneck area Figure 8, vehicles must be permitted to accelerate to the critical speed v_{cr} (around 70 km/h), i.e., the speed that succeed capacity flow, and get in the downstream bottleneck area with a critical speed around v_{cr}. Considering all the above, the head of the intentionally produced mainstream congestion (i.e., the location of the mainstream flow control) should be located sufficiently upstream of the addressed bottleneck. A distance of about 700 m should be appropriate for vehicles to accelerate from low speeds to 70 km/h, taking into account Figure 2.7 of [71]. In case of absence of acceleration area, the capacity drop might not be avoided, and this is the most possible reason for having very small improvements in [72].

The meaning of the acceleration area proposed length, is to consider it as an upper limit, which should suffice even if the appropriate controlled mainstream flow q_c is

very low. In case q_c is lower but close to q_{cap}^{down}, the vehicles leaving the controlled congestion area will have a relatively increased speed, thus it would not be necessary to accelerate significantly. In order to have an efficient traffic flow through the bottleneck for any value of q_c, i.e., for any vehicle speed, it is prudent to post a suitable speed limit for vehicles getting in the acceleration area, as also for vehicles in the downstream bottleneck area.

The MTFC concept can be applied to many types of motorway bottlenecks. An on-ramp merge bottleneck is one of the most ordinary types of bottleneck. Other types of bottlenecks are: merging motorways, lane-drop areas and work zones, strong curvatures, tunnels, bridges, strong grades, and strong weaving sections.

2.4.2. Implementation of MTFC

The controlled outflow q_c, delivered by the MTFC strategy, can be implemented on the motorway mainstream using three alternative actuators.

The first alternative actuator is special green-red traffic signals, one for each mainstream lane. They are located on suitable gantries above the motorway and are operated with asynchronous phasing as proposed and tested in [70]. The traffic cycles of the signals are very short, with green phases permitting just 1-2 vehicles to pass at a time, in contrast to urban junctions traffic lights. Thus, vehicles coming from upstream may not have to stop (as in front of urban traffic lights), but instead just slow down as needed so as to create the ordered controlled mainstream flow q_c. Traffic lights is the most immediate and direct actuator for MTFC, however appropriate campaigns may be needed in order to make the new control measure familiar to the road users, and also some related traffic regulations might have to be changed.

The second alternative actuator is variable speed limits, which is the actuator used in this thesis. This actuator could be used to slow down the motorway traffic flow sufficiently so as to produce the ordered controlled mainstream flow q_c. The VSL range utilized in the majority of current installations does not overcome $[60,120]$ km/h, which is not enough for the implementation of low values of controlled flows q_c. Consequently, the lower admissible bound of the usual range of implemented

VSL should be decreased for MTFC, and appropriate campaigns should inform the road users on the rationale and usefulness of MTFC system, so as to make the compliance to the displayed VSL higher.

Emerging vehicle-infrastructure integration systems is the third alternative actuator. Also this actuator could be used, in the same way as VSL above, to slow down equipped vehicles, so as to produce the ordered controlled mainstream flow q_c. An issue that appears for this actuator is that an investigation must be done concerning the required penetration level of equipped vehicles for appropriate operation.

The actuators described above should not be activated for the period that MTFC is not needed (e.g., in the free flow period). Upstream of the MTFC area suitable pre-signals and messages on variable message signs should be applied before the actual activation of the actuator (e.g., in the beginning of the peak period), so as to warn arriving drivers for the impending activation. However, the same pre-signals and messages should not stop warning arriving drivers, also during MTFC, concerning the applied mainstream traffic control further downstream.

2.4.3. Integrate MTFC with RM

MTFC can be suitably developed as a stand-alone control measure, but also as a combination with available or new RM. In case of combination with available or new RM, the developed integrated control strategy should be capable to:

- combine MTFC with every available control measure, in order to succeed efficiency and equity; and
- permit for easy implementation of several operational policies or specifications. Some examples of this could be that MTFC will be active only if RM is close to become inactive because of full on-ramps; or MTFC will be active only if the on-ramp waiting time, due to RM, overcomes a pre-specified threshold; or MTFC and RM might enable a predefined division of capacity; or a predefined distribution of delays among mainstream and on-ramps vehicles (see also [73]).

2.5. Traffic Flow Model Formulation

METANET [74], the traffic flow model used in this thesis, is a macroscopic second-order traffic flow model. It has been validated for numerous networks (see e.g. [18],

[75]) using real traffic data and it was found to reproduce the real traffic conditions (free flow, critical, congested) with notable accuracy. METANET comprises the model without taking into account VSL control measures, whereas in this thesis VSL control measures are also incorporated in the model as suggested by [76].

2.5.1. Introduction

A directed graph represents the motorway network. The graph nodes are placed at locations where a major change in road geometry occurs, as well at junctions, bifurcations, on-ramps and off-ramps. The graph links represent motorway stretches. Each link has homogeneous geometric characteristics, i.e., number of lanes, curvature, no on-ramps or off-ramps.

The aggregate traffic behavior at certain times and locations is defined by appropriate variables. The time and space arguments are discretized. The discrete time step is defined by T (typically $T = 10\,\text{s}$). A motorway link m is divided into N_m segments of equal length L_m (typically $L_m = 500$ m). The traffic in each segment i of link m at discrete time $t = kT, k = 0,1,\ldots,K$, is macroscopically characterized via the following variables: The traffic density $\rho_{m,i}(k)$ (in $veh/km/lane$) is the number of vehicles traveling along segment i of link m at time $t = kT$, divided by L_m and by the number of lanes λ_m; the mean speed $v_{m,i}(k)$ (in km/h) is the mean speed of the

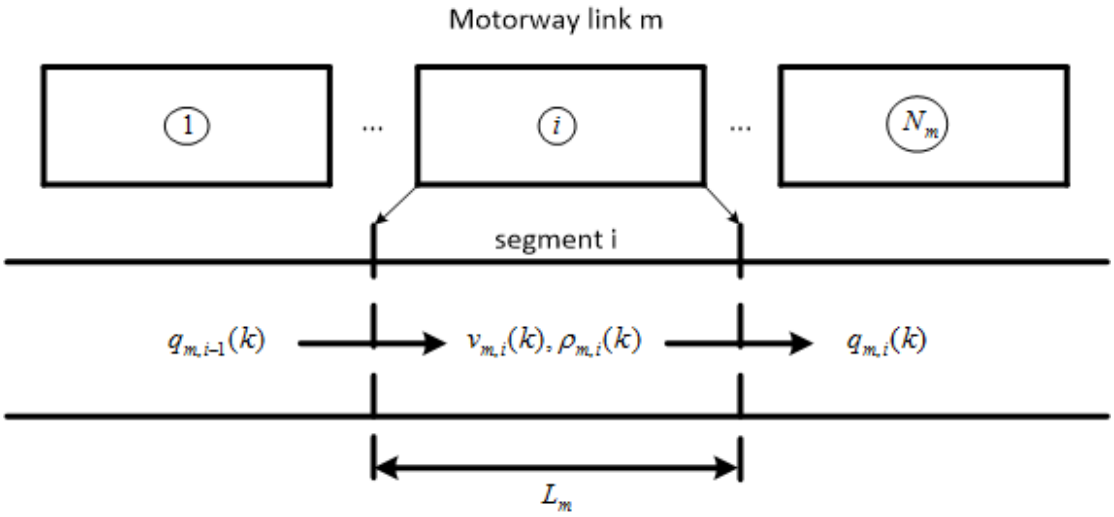

Figure 9: Discretized motorway link.

vehicles included in segment i of link m at time $t = kT$; and the traffic volume or flow $q_{m,i}(k)$ (in veh/h) is the number of vehicles leaving segment i of link m during time period $[kT, (k+1)T]$, divided by T.

2.5.2. Modeling of the Motorway Link

The variables previously defined are calculated for each segment i of link m at each time step k by the following equations:

$$\rho_{m,i}(k+1) = \rho_{m,i}(k) + \frac{T}{L_m \lambda_m}\left[q_{m,i-1}(k) - q_{m,i}(k)\right] \quad (19)$$

$$q_{m,i}(k) = \rho_{m,i}(k) v_{m,i}(k) \lambda_m \quad (20)$$

$$\begin{aligned} v_{m,i}(k+1) &= v_{m,i}(k) + \frac{T}{\tau}\left\{V\left[\rho_{m,i}(k)\right] - v_{m,i}(k)\right\} \\ &+ \frac{T}{L_m}\left[v_{m,i-1}(k) - v_{m,i}(k)\right]v_{m,i}(k) - \frac{\nu T}{\tau L_m}\frac{\rho_{m,i+1}(k) - \rho_{m,i}(k)}{\rho_{m,i}(k) + \kappa} \end{aligned} \quad (21)$$

$$V\left[\rho_{m,i}(k)\right] = v_{f,m} \exp\left[-\frac{1}{\alpha_m}\left(\frac{\rho_{m,i}(k)}{\rho_{cr,m}}\right)^{\alpha_m}\right] \quad (22)$$

where equation (19) reflects the conservation equation, equation (20) reflects the flows to be replaced in (19), equation (21) reflects an empirical dynamic mean speed equation, where equation (22) must be replaced, equation (22) reflects the static speed-density relationship corresponding to the fundamental diagram, and τ which is a time constant, ν which is an anticipation constant and κ, are model parameters that are the same for all the network links. For increased accuracy under certain conditions, two additional terms could be added to equation (21) [18].

The model without the VSL impact comprises three link-specific constant parameters in the speed-density curve (22). These parameters are the free speed $v_{f,m}$ encountered at zero density ($\rho_{m,i} = 0$), the critical density $\rho_{cr,m}$ at which traffic flow is near to capacity $q_{cap,m}$, and α_m. The capacity of the fundamental diagram (flow-density curve) is given by combining equations (20)-(22) under stationary (i.e., $v_{m,i}(k+1) = v_{m,i}(k)$) and spatially homogeneous (i.e., $v_{m,i-1} = v_{m,i}$ and $\rho_{m,i+1} = \rho_{m,i}$) conditions for $\rho_{m,i} = \rho_{cr,m}$ (i.e., the critical density).

$$q_{cap,m} = v_{f,m} \cdot \rho_{cr,m} \exp\left(-1/\alpha_m\right) \quad (23)$$

2.5.3. Modeling VSL impact

The impact of VSL can be incorporated in the previously described link model, taking into account the assumption that in each link, a single VSL value (if any) is displayed. In real field implementations, the number of the VMS depends on the link length. Nevertheless the above assumption is not very restrictive, for two reasons. The first reason is that if a higher spatial resolution of VSL is desired, then maybe the links should be selected accordingly short. The second reason is that if a lower resolution of VSL values is desired it is likely for the user of the related software tools METANET (simulator) and AMOC (optimal control, see Section 2.6) to produce clusters of links, where each cluster will have the same VSL value.

The VSL values are reflected to the link-specific VSL rates $b_m(k)$, that prevail, by definition, during $\left[kT,(k+1)T\right)$. The VSL rates correspond to naturally control variables, with an allowable value range $b_m(k) \in \left[b_{\min,m},1\right]$, where $b_{\min} \in (0,1)$ is a lower allowable bound for VSL rates. Below, the appropriate incorporation of the previously defined VSL rates into the link model (19)-(22), is executed. Following the lines of previous works ([59], [77], [78]), the incorporation is implemented by rendering the static speed-density relationship (22), b_m-dependent. Based on available real data, evidenced in [10], [76] this is enabled by rendering the three parameters included in (22) b_m-dependent using the following linear function:

$$v_{f,m}^*\left[b_m(k)\right] = v_{f,m} b_m(k), \quad (24)$$

and using the following affine functions:

$$\rho_{cr,m}^*\left[b_m(k)\right] = \rho_{cr,m}\left\{1+A_m\left[1-b_m(k)\right]\right\} \quad (25)$$

$$\alpha_m^*\left[b_m(k)\right] = \alpha_m\left[E_m-(E_m-1)b_m(k)\right], \quad (26)$$

where $v_{f,m}$, $\rho_{cr,m}$, α_m denote the specific non-VSL values for these parameters as in (22), whereas A_m and E_m are constant parameters that are estimated based on real data.

Equation (24) disclose that b_m is equal to the VSL-induced $v^*_{f,m}$ divided by the non-VSL $v_{f,m}$ or approximately equal to the displayed VSL divided by the legal speed limit without VSL. As a result, if $b_m(k)=1$, there is no VSL applied , else $b_m(k)<1$. The parameters of equations (24)-(26), are taking their non-VSL values, if $b_m(k)=1$. For $A_m>0$ and $E_m>1$, equations (25) and (26) propose that $\rho^*_{cr,m}$ and α^*_m are affine increasing functions for decreasing b_m beginning from their non-VSL values for $b_m(k)=1$.

Equation (22) is replaced for the extended model by:

$$V\left[\rho_{m,i}(k),b_m(k)\right]=v^*_{f,m}\left[b_m(k)\right]\exp\left[-\frac{1}{a^*_m\left[b_m(k)\right]}\left(\frac{\rho_{m,i}(k)}{\rho^*_{cr,m}\left[b_m(k)\right]}\right)^{a^*_m\left[b_m(k)\right]}\right], \quad (27)$$

and the VSL-induced capacity flow is defined by:

$$q^*_{cap,m}\left[b_m(k)\right]=v^*_{f,m}\left[b_m(k)\right]\cdot\rho^*_{cr,m}\left[b_m(k)\right]\cdot\exp\left(-\frac{1}{a^*_m\left[b_m(k)\right]}\right). \quad (28)$$

Traffic data from a European VSL-equipped motorway location, were used to calibrate the extended speed-density curve (27) that includes equations (24)-(26) [76]. The legal speed limit in this motorway is 70 mph and the applied VSL values are 60 mph, 50 mph, and 40 mph, that correspond to VSL rates $b_m\in\{1,0.86,0.71,0.57\}$. The calibration procedure gave different parameter values for different motorway locations; namely for some locations the real data and resulting values of A and E show a VSL-induced capacity $q^*_{cap,m}(b_m)$ increase, for some VSL values, compared to the non-VSL capacity $q_{cap,m}$, whereas at other locations, capacity increase was not notified for any VSL.

2.5.4. Modeling of the Origin Link

Origin links are links that receive traffic demand d_o and forward it into the motorway network. For the modelling of these links, a simple queue model is used (Figure 10).

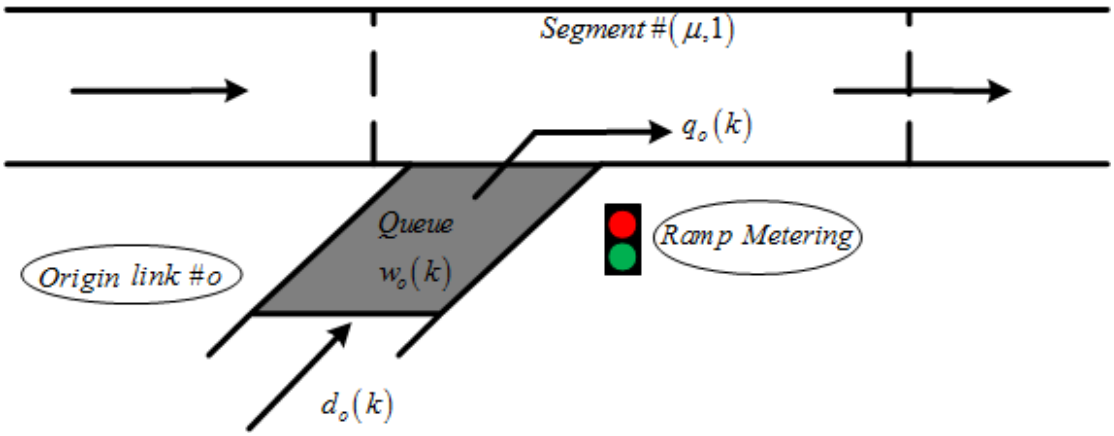

Figure 10: The queue model of the origin link.

The origin link outflow q_o depends on the arriving demand, on the corresponding mainstream segment $(\mu,1)$ traffic conditions, and on the RM control measures existence. In case RM is applied, the outflow $q_o(k)$ leaving origin o during period k is a portion $r_o(k)$ of the outflow $\hat{q}_o(k)$ that would leave without RM. The value of $r_o(k) \in \left[r_{\min,o},1\right]$, is the metering rate for the origin link o, i.e., a control variable where $r_{\min,o}$ is a minimum admissible value. No RM is applied if $r_o(k)=1$, else $r_o(k)<1$. The following conservation equation describes the queuing model:

$$w_o(k+1) = w_o(k) + T\left[d_o(k) - q_o(k)\right], \tag{29}$$

where $w_o(k)(veh)$ is the queue length in origin o at time kT, and $d_o(k)(veh/h)$ is the demand flow at o. The outflow $q_o(k)$ is defined with the following equation:

$$q_o(k) = r_o(k)\hat{q}_o(k) \tag{30}$$

with

$$\hat{q}_o(k) = \min\left\{\hat{q}_{o,1}(k), \hat{q}_{o,2}(k)\right\} \tag{31}$$

and

$$\hat{q}_{o,1}(k) = d_o(k) + w_o(k)/T \tag{32}$$

$$\hat{q}_{o,2}(k) = Q_o \min\left\{1, \frac{\rho_{\max} - \rho_{\mu,1}(k)}{\rho_{\max} - \rho_{cr,\mu}}\right\}, \tag{33}$$

where $Q_o\left(veh/h\right)$ is the capacity flow of the on-ramp, i.e., the maximum possible outflow of the on-ramp, under free-flow traffic conditions in the mainstream, and $\rho_{\max}\left(veh/km/lane\right)$ is the maximum density in the network. The uncontrolled outflow $\hat{q}_o\left(k\right)$, according to equations **(31)-(33)**, is defined by the current origin demand if $\hat{q}_{o,1}\left(k\right) < \hat{q}_{o,2}\left(k\right)$, else it is defined by the geometrical capacity Q_o (if the mainstream density is undercritical, i.e., $\rho_{\mu,1}\left(k\right) < \rho_{cr,\mu}$), or by the decreased capacity because of mainstream congestion (if $\rho_{\mu,1}\left(k\right) > \rho_{cr,\mu}$).

2.5.5. Modeling of the Node

Nodes represent motorway junctions and bifurcations, encompassing on-ramps and off-ramps. Through a number of input links, traffic flow enters a node n and is distributed to the output links according the equations below:

$$Q_n\left(k\right) = \sum_{\mu \in I_n} q_{\mu,\mathrm{N}_\mu}\left(k\right) \tag{34}$$

$$q_{m,0}\left(k\right) = \beta_n^m(k) Q_n\left(k\right) \quad \forall m \in O_n \tag{35}$$

where I_n is the set of links entering node n, O_n is the set of links leaving n, $Q_n\left(k\right)$ is the total traffic volume entering node n at period k, $q_{m,0}\left(k\right)$ is the traffic volume that leaves n via out-link m, and $\beta_n^m(k) \in \left[0,1\right]$ is the portion of $Q_n\left(k\right)$ that leaves n through link m (turning rates).

The upstream impact of the downstream link density, at a network node n (e.g. in case of congestion spillback), has to be taken into consideration in the last segment of the incoming links. This is done via:

$$\rho_{m,N_m+1}\left(k\right) = \frac{\sum_{\mu \in O_n} \rho_{\mu,1}^2\left(k\right)}{\sum_{\mu \in O_n} \rho_{\mu,1}\left(k\right)}, \tag{36}$$

where $\rho_{m,N_m+1}\left(k\right)$ is the virtual density downstream of any entering link m to be used in equation **(21)** for $i = N_m$ and $\rho_{\mu,1}\left(k\right)$ is the density of the first segment of the leaving link μ. The reason of the quadratic form existence is to account for the fact

that congestion on one leaving link may spill back into the entering link even for the case that there is free flow in the other leaving links.

In addition, the downstream impact of the upstream link speed, at a network node n, has to be also taken into account, according equation **(21)** for $i=1$. The flow-weighted average is used to calculate the required upstream mean speed value:

$$v_{m,0}(k)=\frac{\sum_{\mu\in I_n} v_{\mu,\mathrm{N}_\mu}(k)q_{\mu,\mathrm{N}_\mu}(k)}{\sum_{\mu\in I_n} q_{\mu,\mathrm{N}_\mu}(k)}, \qquad \textbf{(37)}$$

where $v_{m,0}(k)$ is the virtual speed upstream of any leaving link m that is needed in equation **(21)** for $i=1$.

2.5.6. General Dynamic Model

A non-linear macroscopic discrete-time state-space model is obtained, by connecting the above developed equations, for the whole network

$$x(k+1)=f\left[x(k),u(k),d(k)\right],\ x(0)=x_0, \qquad \textbf{(38)}$$

where x is the state vector, u is the control vector, and d is the disturbance (external variable) vector. The state vector comprises of the densities $\rho_{m,i}$, the mean speeds $v_{m,i}$ of every segment i of every link m, and the queues w_o of every origin o. The control vector comprises of the VSL rates b_m of every link m where VSL are applied, and of the RM rates r_o of every origin o that is metered. The disturbance vector comprises of the demand d_o at every origin o and the turning rates β_n^m at every bifurcation node n.

2.6. Optimal Control Problem Formulation

The integrated motorway network traffic control problem is formulated as a discrete-time dynamic optimal control problem, over a given optimization horizon K_p, with constrained control variables. A feasible-direction algorithm is used for solving this problem, which is efficient even for large-scale networks ([79], [22]). An extended version of the open-loop optimal control tool AMOC, incorporates the above extended formulation (to incorporate the VSL impact) and the numerical solution algorithm [20]. This tool can consider, coordinated RM, system optimum route

guidance, variable speed limits (using the extension introduced via equations **(24)-(27)**, and integrated control combining simultaneously all the control measures.

The general discrete-time formulation of the optimal control problem is given as follows:

Given disturbance predictions $d(k), k = 0,1,\ldots,K_p - 1$ and the initial state $x_0 = x(0)$, minimize

$$J = \theta\left[x\left(K_p\right)\right] + \sum_{k=0}^{k_p-1} \varphi\left[x(k), u(k), d(k)\right] \quad \textbf{(39)}$$

subject to equation **(38)** and the inequality constraints imposed on the ramp metering rates $r_{\min,o} \le r_0(k) \le 1$ and the VSL rates $b_{\min,m} \le b_m(k) \le 1$.

The total time spent (TTS) by all vehicles in the network (including the waiting time experienced at the ramp queues) can be selected as the cost criterion, which is a natural objective for the traffic system. A penalty term may be used to suppress high-frequency oscillations of the optimal control trajectories.

The exact cost criterion used as **(39)** is written below:

$$J = T\sum_{k=1}^{K_p-1}\sum_{m}\sum_{i}\rho_{m,i}(k)L_m\lambda_\mu + T\sum_{k=1}^{K_p-1}\sum_{o}w_o(k) + T\sum_{k=1}^{K_p-1}\sum_{m}a_b\left[b_m(k) - b_m(k-1)\right]^2 + T\sum_{k=1}^{K_p-1}\sum_{m}a_r\left[r_o(k) - r_o(k-1)\right]^2, \quad \textbf{(40)}$$

where a_b and a_r are weighting factors for the corresponding penalty terms.

As the problem is nonconvex, the solution determined by AMOC may reflect a local minimum; however, previous works demonstrate that good solutions (from an application point of view) can be found. The solution of AMOC consists of the optimal VSL rate trajectories as also the corresponding optimal state trajectory. The solution algorithm can account for control variables that change less frequently their values than the state variables. Moreover, to speed up the solving algorithm convergence, links can be grouped in clusters that have the same VSL rates.

The AMOC extension [68], that takes into account VSL rates, necessitated the specialized and rigorous incorporation of accordingly extended generic necessary

optimality conditions along with the corresponding Jacobian matrices, etc. Thus AMOC has become a universal optimal control tool that is immediately applicable to any, (even large-scale) motorway network to deliver optimal control results with low enough computational effort, that would allow for a real-time application of the tool. The above tool characteristics show a clear progress, compared to most previous works concerning optimal VSL control.

In conclusion, the reasons of taking ideal solutions in a simulation environment, by the optimal control problem are the "perfect" model, the exact knowledge of (future) disturbances (demands and turning rates) and the lack of some VSL constraints. The optimal control solutions for sure cannot be outperformed (in simulation) by any other control strategy, however can be used to evaluate the efficiency of other strategies, under different scenarios.

2.7. Model Parameter Calibration

The procedure of model parameter calibration aims at enabling the macroscopic model of the motorway network to represent traffic conditions with sufficient accuracy. The unknown parameters estimation, contained in equation **(38)** is not a trivial task, because the system equations are highly nonlinear in both the parameters and the state variables. The most ordinary approach is the minimization of the discrepancy between the model calculations and the real process data in the sense of a quadratic output error functional [18], [75], [80]. The measurable output vector of the non-linear system **(38)**, is defined by y (typically consisting of flows and mean speeds at various internal motorway locations), which may be related to the model state via

$$y(k) = g\left[x(k), p\right]. \tag{41}$$

So, the parameter estimation problem can be formulated as a least-squares output error problem:

Given the disturbance and control vector trajectories, the measured process output $y^m(k)$ for some $k \in M \subseteq \{1, 2, \ldots, K\}$, and the initial state x_0; find the set of parameters p minimizing the cost functional

$$J(p)=\sum_{k\in M}\left\|y(k)-y^{m}(k)\right\|_{Q}^{2} \tag{42}$$

subject to **(38)** and **(41)**, where Q is a positive definite, diagonal matrix.

The value of M may be a subset of the simulated points because usually the simulation step is set to a value, e.g. $T=5\,\mathrm{s}$, that is less than the utilized measurement interval, i.e., $60s$. The selected region of the model parameters is a closed allowable region of the parameter space that may be defined on the basis of physical considerations. An appropriate nonlinear programming routine named Nelder-Mead is used to determine the optimal parameter set, which is a derivative free method and is suitable for the calibration of macroscopic traffic flow models [81]. For each selection of a new parameter vector p, the corresponding value of the performance criterion **(42)** is calculated by a simulation run of the model equations as depicted in Figure 11. In [18] it has been shown that the model is most sensitive with respect to the values of the parameters used in the fundamental diagram **(22)**. The parameter vector consists of the free speed $v_{f,m}$, the critical density $\rho_{cr,m}$ and a_m, for every link m, and the parameters τ, ν, κ, δ, φ, $v_{\min}$ and $\rho_{\max}$ that are common for all the network links. Sometimes, for the reduction of the parameter vector dimension, some of the common parameters are taking constant values based on previous experience. The calibration tool that was used is CALISTO [82].

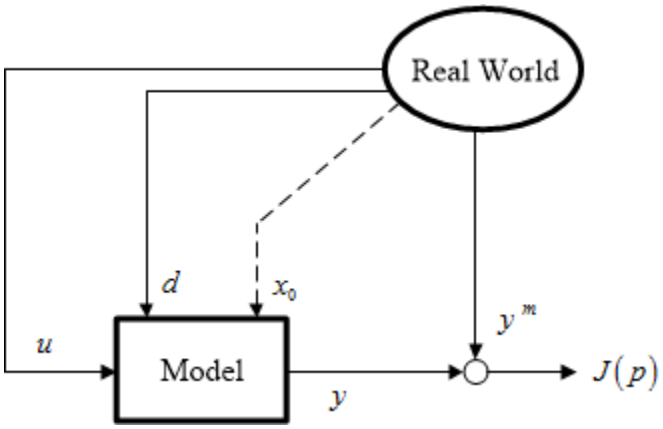

Figure 11: Performance criterion calculation.

3. MTFC for Multiple Bottlenecks

In this Chapter, a new control concept with MTFC for multiple bottlenecks is proposed. Section 3.1 contains a literature overview of MTFC like strategies, whereas section 3.2 refers to multiple bottleneck inspiration and causes. Section 3.3 presents the necessity of the approach. The implementation aspects of MTFC are included in section 3.4, the feedback controller is described in section 3.5, and section 3.6 summarizes the practical application aspects.

3.1. Literature Overview

The goal of MTFC is to regulate the flow upstream of a bottleneck location, in order to maximize throughput at the specific location. MTFC has been proposed in [68] and [69] and was demonstrated to lead to substantial improvement in traffic flow efficiency indicators. In initial investigations, a traffic-responsive control system for tunnel flow control had been proposed in [83]. A sag bottleneck exists in the tunnel (Holland tunnel between New Jersey and New York City), and it was noticed that the throughput was decreasing strongly, when a speed breakdown occurred. This corresponds to capacity drop phenomenon, in recent terminology. The reason for the throughput reduction was the need for vehicles to accelerate at the head of the forming congestion. In order to make decisions on the tunnel's flow control, real time measurements from the bottleneck location were used by the system, so as to avoid the speed breakdown occurrence within the tunnel. The first simple feedback control strategies of the bang-bang type were proposed in [84] and [85], whereas more sophisticated heuristic feedback algorithms were suggested in [83] and [86]. The traffic-light-based entrance control system of the San Francisco-Oakland Bay Bridge, is another MTFC-like system, which was demonstrated in [87] and has been in operation for more than 35 years. The algorithm used [88] seems to have substantial similarities with the algorithm suggested in [89]. Finally, some fixed-time mainline control actions have been investigated in [90] and [91], using traffic lights as a new traffic management tool for motorways.

Due to the fact that traffic lights on the motorway mainstream are not prevalent so far, researchers have considered MTFC enabled using VSL with various control strategies

and traffic application settings. The similarities of MTFC with RM applied to bottleneck locations, are described in [68]; while [69] refers to the valuable potential benefits of MTFC, based on suitable use of VSL, via an optimal control approach applied to a large-scale motorway infrastructure. The conclusion from these papers is that if MTFC is applied upstream of an active bottleneck location, it is able to avoid the capacity drop. Nevertheless, for use in real field implementation, the optimal control approach utilised may be cumbersome. An ALINEA-like feedback controller for MTFC, based on VSL, was proposed and tested via microscopic simulation in [72]. The improvements achieved in this work were marginal, and a possible reason is the absence of an acceleration area, as proposed in [68] and [69]. A feedback motorway traffic control using VSL, was proposed in [92], by using H-infinity control theory, although without clearly addressing throughput maximization at bottleneck locations. In [93], further suggestions have been made, but with a switching I-type regulator that seems to be inappropriate in delivering stable control actions. Finally, [24] demonstrated in a field experiment the feasibility of mainstream flow control by use of VSL, albeit for a different task, namely the dissolution of moving jams.

3.2. Multiple Bottleneck Inspiration and Causes

The idea of this control concept is based on an extension of the feedback controller presented in [6], taking into account the MTFC application concept presented in [68] and [69] in case of multiple bottlenecks. The issue of multiple bottlenecks for the case of RM control was addressed in [94]. More precisely in [94], a control strategy applicable to local RM in presence of random-location bottlenecks downstream of a metered on-ramp was studied. This strategy uses a number of PI-ALINEAs, each provided with the measurements from a separate detector site. A suitable decision device is designed, in order to determine the overall RM action from all PI-ALINEAs outputs. The logic of this system architecture is that the random bottleneck can be defined by at least one detector, a little after it emerges and the PI-ALINEA output corresponding to the bottleneck location should become sovereign for defining the overall RM action. So, the new proposal combines the concepts developed in [6] and [94] to derive a feedback law for VSL-based MTFC addressing multiple downstream bottlenecks.

There are various causes of multiple bottlenecks. Some of the causes are high demand

of consecutive uncontrolled on-ramps, bad weather, strong lane changing, lane drops, speed limit changes, etc.

3.3. The Proposed Approach

It has been assumed, in earlier works with MTFC, that feedback control actions taken for treating different bottleneck locations do not interfere with each other and can be handled separately. Nevertheless, sometimes this is not possible, e.g., when potentially active bottlenecks are in close proximity or interact with each other or are uncertain because of a number of possible reasons. In cases like these, the same MTFC system (i.e., single controlled area) should be enabled to handle multiple downstream bottlenecks. As a result, this calls for a suitable extension of the basic MTFC concept. The approach that is proposed is to specify the outflow q_c to be equal to the smallest outflow computed for the different bottlenecks. An important issue that must be addressed is the identification of the bottleneck locations. For this scope, availability of sufficiently dense measurements from the mainstream is required.

3.4. MTFC Implementation Aspects

VSL are utilized in this work as an MTFC actuator. Mainstream congestion will be formed upstream of the MTFC location. The vehicles exiting the congested area will be characterized by a speed lower than the critical speed that is needed to achieve bottleneck capacity flow q_{cap}^{down}. In order to avoid that, vehicles should be allowed (and encouraged) to accelerate to the critical speed v_{cr} (about 70 km/h), i.e., the bottleneck location speed at which capacity flow occurs. This is realized by placing the head of the created mainstream congestion upstream enough of the addressed bottleneck so that the vehicles have the possibility to accelerate from low speeds to 70 km/h. In absence of an acceleration area, the capacity drop phenomenon may not be avoided [68], [69]. At this point, it is interesting to refer to the empirical field experiments carried out in Japanese motorways at sag bottlenecks, where drivers were alerted to accelerate promptly at the head of the congestion so as to reduce the capacity drop level [95].

Acceleration of vehicles in the acceleration area and in the downstream bottleneck area depends on q_c and on their individual speed. The goal would be to have vehicles

that adopt the critical speed which leads to q_{cap}^{down}. Thus, it is advisable to post an appropriate speed limit for vehicles entering the acceleration area and the downstream bottleneck area. In the mainstream controlled section, vehicles move slower than in the upstream sections; this may bear some risks for vehicles approaching with a high speed the tail of the congested area. VSL may therefore be used also to gradually reduce the speed of arriving vehicles to the level prevailing at the controlled congestion, resulting in a reduction of the safety risk for arriving vehicles.

When using VSL as an MTFC actuator, some restrictions apply to the posted speed limits. The first restriction has to do with the VSL values: speed limits can take only discrete values within the range of permitted VSL (e.g., multiples of 10 km/h). A second issue is the speed limit difference between two consecutive posted VSL at the same gantry that is not allowed to be bigger than a predefined value (e.g., 20 km/h). Moreover, the difference of speed limits between two consecutive VSL-posting gantries must be considered. Finally, speed limits are not permitted to change their values more frequently than a predefined time interval (e.g., 1 min). This time interval could be used as the control period of the control strategy.

3.5. Feedback Controller

In the current section, the cascade feedback MTFC controller that was introduced in [6] is presented and an appropriate extension of this controller for the case of multiple bottlenecks is described.

The feedback controller developed in [6] regulates the traffic density ρ_{out} (see Figure 8) via appropriate real-time changes of the mainstream flow q_c. This is performed via appropriate VSL actions upstream of the bottleneck location. The flow q_{out} is maximized when ρ_{out} equals the critical density ρ_{cr}, thus the density set-point $\hat{\rho}_{out}$ of the control loop has to be set equal to ρ_{cr}. The process to be controlled is represented by a single-input-single-output (SISO) system with the VSL rate b as the control input and ρ_{out} as the control output.

Figure 12 depicts the MTFC cascade feedback controller structure designed in [87]. An integral (I) controller is included in the secondary loop while a Proportional-

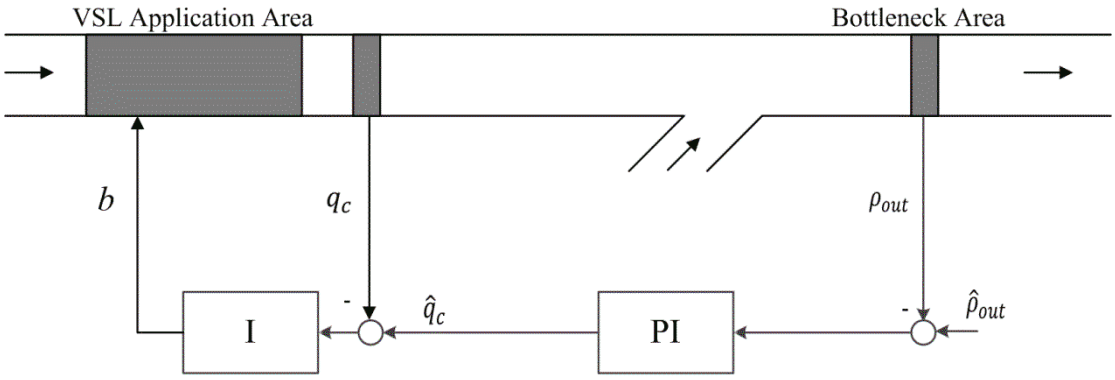

Figure 12: MTFC feedback controller structure using VSL as actuator.

Integral (PI) controller is included in the primary loop. The secondary loop is affected by the VSL rate b delivered by the secondary controller that will determine the outflow q_c. Downstream of the VSL application area, q_c is measured, fed back, and compared to the desired flow $\hat{q}_c$ delivered by the primary controller. The measured density (or occupancy) at the bottleneck area is used by the primary loop that compares it with the set-point $\hat{\rho}_{out}$. The I controller for the secondary loop is described by

$$b(k) = b(k-1) + K_I\left(\hat{q}_c(k) - q_c(k)\right) \tag{43}$$

with K_I as the integral gain of the controller. The PI controller for the primary loop is described by

$$\hat{q}_c(k) = \hat{q}_c(k-1) + K_I'\left(\hat{\rho}_{out}(k) - \rho_{out}(k)\right) + K_P'\left(\rho_{out}(k-1) - \rho_{out}(k)\right) \tag{44}$$

where K_I' and K_P' are the integral and proportional gains of the controller, respectively.

In case of multiple-bottleneck locations, a set of PI controllers (see Figure 13) are now used in the control strategy. Each controller takes measurements from a separate detector site, downstream of the acceleration area. An appropriately designed decision device, determines the overall MTFC action from all PI controllers' outputs. This strategy is similar to the strategy used in [94] for RM. The currently most critical bottleneck is identified by the decision logic, and the output of the PI controller corresponding to the identified bottleneck location becomes dominating for determining the overall MTFC action. The candidate bottleneck locations may be as

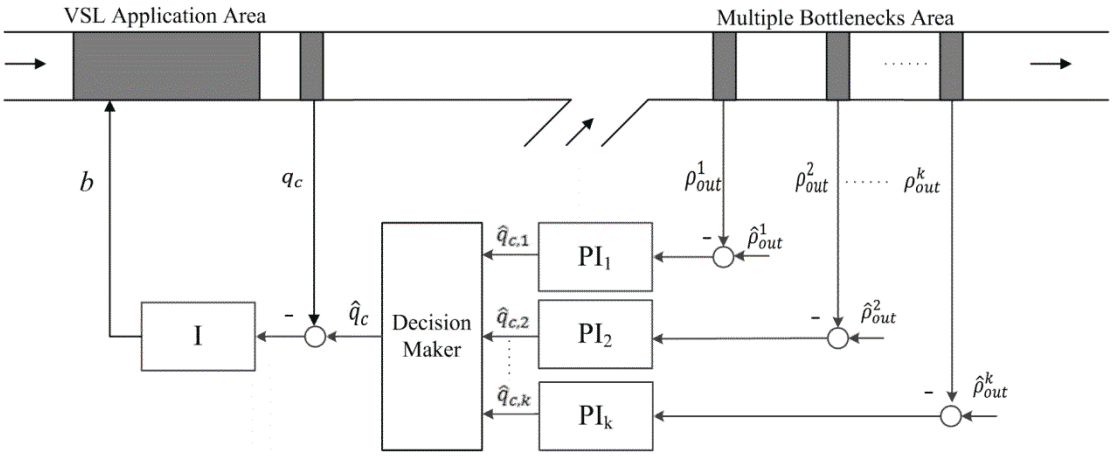

Figure 13: MTFC feedback controller structure for multiple bottlenecks using VSL as actuator.

many as the available downstream measurements.

The equation for the primary controller **(44)** is now replaced by

$$\hat{q}_{c,i}(k) = \hat{q}_{c,i}(k-1) + K_{I,i}^{'}\left(\hat{\rho}_{out,i}(k) - \rho_{out,i}(k)\right) + K_{P,i}^{'}\left(\rho_{out,i}(k-1) - \rho_{out,i}(k)\right), i = 1,\ldots,n \quad \textbf{(45)}$$

where $\hat{q}_{c,i}(k)$ represents the output of the i-th PI controller at time instant k.

The chosen decision policy is simply

$$\hat{q}_c(k) = \hat{q}_{cj}(k) \quad \textbf{(46)}$$

with

$$j = \arg\min_{i=1,\ldots,n}\left\{\hat{q}_{c,i}^{sm}(k)\right\} \quad \textbf{(47)}$$

$$\hat{q}_{c,i}^{sm}(k) = \alpha_{sm}\cdot\hat{q}_{c,i}(k) + (1-\alpha_{sm})\cdot\hat{q}_{c,i}^{sm}(k-1), i = 1,\ldots,n \quad \textbf{(48)}$$

where $\hat{q}_{c,i}^{sm}(k)$ represents the exponential smoothing of $\hat{q}_{c,i}(k)$ computed through equation **(48)** in which α_{sm} is a parameter within $[0,1]$. The tuning of parameter α_{sm} was based on a trade-off of actually addressing the relevant bottleneck at each point in time versus too frequent switching between different bottlenecks. It turned out to be a relatively easy task. A value equal to 0.7 has emerged and was used in the simulations. The controller that corresponds to the smallest (smoothed) flow value is

selected as the most critical. A smoothed flow is used for avoiding fast switching among the controllers.

The methodologies applied for tuning the controllers are the zone-based procedure, as described in [96], and the SIMC PID tuning method [97], for the secondary and primary controller, respectively. The gain values that emerged and were used in the simulations in Chapter 5 are the following: $K_I = 0.0006 \quad h \cdot lane/veh$ for the secondary controller, and $K_I^{'} = 1.5 \quad km/h/lane$ and $K_P^{'} = 13.0 \quad km/h/lane$ for the primary controllers.

3.6. Practical Application Aspects

The current section summarizes some practical VSL implementation aspects detailed in [6]. Firstly, VSL rates are supposed to use only predefined discrete values. In particular $b \in \{0.2, 0.3, \ldots, 1.0\}$. The VSL rate $b(k)$ delivered by the control strategy is rounded off to obtain the applied VSL rate.

In addition, as mentioned in section 3.4, VSL is applied also upstream of the controlled congestion. In particular, the difference between two consecutive posted VSL rates at the same gantry is limited to 0.2, and the same limit applies also to the difference between the posted VSL rates at two consecutive gantries. A last constraint is that a constant VSL rate equal to 0.9 is applied in the acceleration and bottleneck areas, whenever MTFC is active. All the above constraints are considered only in feedback control, not in the optimal control case.

The application of the VSL rate $b(k)$ delivered by the control strategy begins when the measured density $\rho_{out,i}$ at a bottleneck area i becomes higher than an activation threshold and ends when the measured density $\rho_{out,i}$ at all bottleneck areas becomes lower than a deactivation threshold (lower than the activation one). The control period is set to $T_c = 60 \quad s$. This value is appropriate for practical purposes, as it is used in current VSL installations in various countries.

4. Integrated Traffic Flow Control for Multiple Bottlenecks with Delay Balancing

In Chapter 4, a new integrated traffic flow control concept for multiple bottlenecks with delay balancing is proposed. The first section of this chapter 4.1 contains a literature overview of integrated control concepts, whereas the second section 4.2 refers to the feedback control structure. The next section 4.3, presents RM actuator characteristics, and section 4.4, MTFC actuator characteristics. The delay estimation is included in section 4.5, and the flow distribution for delay balancing is described in section 4.6.

4.1. Literature Overview

Various traffic management measures have been suggested and implemented to extenuate motorway traffic congestion, but in case each one of them is considered independently, a lot of benefits that would result from synergy (integration) of different control measures are missed. Also the respective control actions by individually designed control algorithms may even be contradictory under certain conditions. For instance, RM is the most direct and efficient measure for motorway traffic flow control, but the metered flow may be actually released whenever queue management strategies are activated with the aim of avoiding the creation of over-long on-ramp queues that spill over to the adjacent network [2]. On the other hand, VSL can be utilized to enable MTFC [69], [68], but very small VSL values may not considered acceptable for long time periods by the responsible road authority or the drivers.

In order to overcome some of the above restrictions, the integration of control actions has been considered in previous works. For instance, RM was integrated either with route guidance [20], [98] or with VSL [68], [69], [72], [93], [99]-[100]. Nevertheless, most of these approaches are based on sophisticated methods, e.g., nonlinear optimal control approaches, that may be cumbersome in field implementations due to their black-box character and their requirement for many more measurements, demand prediction and model validation. There was an effort lately for the design of feedback control approaches that integrate RM and VSL, and are more suitable for field

implementations. However, for example in [73], a quite specific layout is taken into account without considering the delays experienced by drivers, while the algorithm in [101] can be applied only in case of limited-length moving jams.

The proposed feedback-based integrated motorway traffic flow control strategy for multiple bottlenecks with delay balancing, is based on previous concepts for multiple bottlenecks, developed either for RM [94] or for MTFC (see Chapter 3); and it generalizes the delay balancing idea of [102], [103] to apply to an arbitrary number of RM or VSL actuators via suitable definition of a knapsack optimization problem. The new concept aims at throughput maximization while integrating an arbitrary number and type of such actuators, located upstream of all bottlenecks, through an optimization algorithm that balances the delays experienced by drivers behind each actuator in a desired pre-specified way. The concept is simple and robust. Many practical aspects, related to ramp metering and VSL implementation aspects, have been considered, and simulation results are presented for a real motorway stretch in the United Kingdom using a validated second-order macroscopic traffic flow model and real demands.

4.2. Feedback Control Structure

The proposed feedback control structure is shown in Figure 14. A set of n Proportional-Integral (PI) controllers is utilized, each fed with a corresponding measurement from a potential bottleneck site, downstream of all actuators. The measured density $\rho_{out,i}$ at the bottleneck location i at time instant k is compared with the set-point $\hat{\rho}_{out,i}$, usually set around the critical density value, at which capacity flow is achieved at that location. The PI-type regulator for the bottleneck location i is given by:

$$\begin{aligned}\hat{q}_{t,i}(k) &= \hat{q}_{t,i}(k-1) + \hat{K}_{I,i}\left(\hat{\rho}_{out,i} - \rho_{out,i}(k)\right) \\ &+ \hat{K}_{P,i}\left(\rho_{out,i}(k-1) - \rho_{out,i}(k)\right), i = 1,\ldots,n\end{aligned} \quad (49)$$

where $\hat{q}_{t,i}(k)$ demonstrates the output of the i-th regulator (i.e., the flow to be implemented by all actuators), $\hat{K}_{I,i}$ and $\hat{K}_{P,i}$ are the integral and proportional gains, respectively. Each regulator output is truncated in order to remain within a range of flow values $\left[\hat{q}_{t,\min}(k), \hat{q}_{t,\max}(k)\right]$. The determination of these time-varying bounds is

explained later. At the next time-period, the truncated values are used, as the $k-1$ values in (49) to avoid the well-known windup phenomenon for PI regulators.

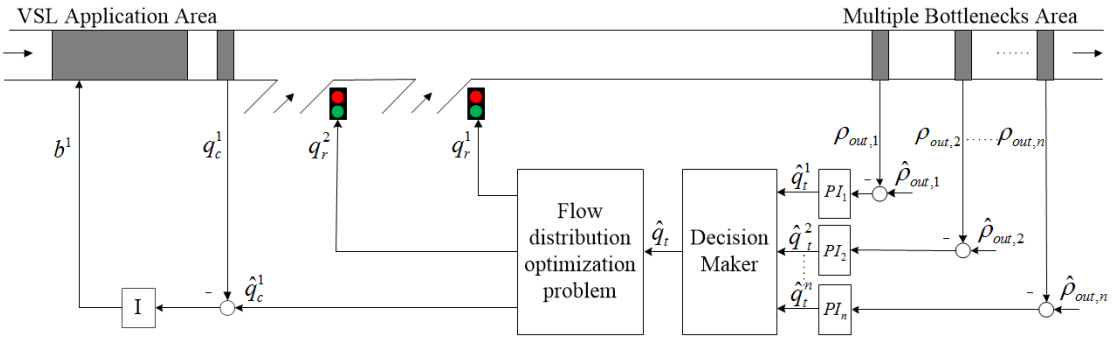

Figure 14: Integrated control structure for multiple bottlenecks and balanced delays.

It should be noted that the stability of the closed-loop ramp metering system with a PI-type regulator has been rigorously proved by Lyapunov stability arguments [104]. In fact, with its control parameters appropriately tuned, the regulator was found to be universally applicable to a range of distances between the on-ramp and downstream bottlenecks. The latter was also an empirical observation found in [105].

An appropriately designed decision algorithm determines the overall action from all PI controller outputs. More precisely, the currently active bottleneck is specified, and the output of the corresponding PI controller is selected for implementation. This is realized in the following way (see Chapter 3):

$$\hat{q}_t(k) = \hat{q}_{tj}(k) \tag{50}$$

with

$$j = \arg\min_{i=1,\ldots,n} \left\{ \hat{q}_{t,i}^{sm}(k) \right\} \tag{51}$$

$$\hat{q}_{t,i}^{sm}(k) = \alpha_{sm} \cdot \hat{q}_{t,i}(k) + (1-\alpha_{sm}) \cdot \hat{q}_{t,i}^{sm}(k-1), i = 1,\ldots,n \tag{52}$$

where $\hat{q}_{t,i}^{sm}(k)$ in (52) demonstrates the exponential smoothing of $\hat{q}_{t,i}(k)$ with α_{sm} a parameter within $[0,1]$. The controller that corresponds to the smallest (smoothed) flow value is chosen and is implemented in the time interval $\left(kT,(k+1)T\right]$, where T

is the control period. The smoothed flow is used, in order to avoid frequent switching to different controllers, which may be caused by measurement noise.

The determined total flow $\hat{q}_t(k)$ must then be distributed to the available actuators so that the bounds of each actuator flow are respected. Such bounds exist because of operational and policy-related issues. For instance, in case of RM a queue management policy may create lower bounds for the actuator, while in case of MTFC specific VSL lower and upper bounds are present. If $q_r^i(k)$ is the flow to be implemented by the i-th RM system and $q_c^i(k)$ is the flow to be implemented by the i-th MTFC system, both at time period k, then the total flow distribution should satisfy

$$\hat{q}_t(k) = \sum_{i=1}^{n_r} q_r^i(k) + \sum_{i=1}^{n_c} \hat{q}_c^i(k) \tag{53}$$

$$q_{r,\min}^i(k) \le q_r^i(k) \le q_{r,\max}^i(k), i = 1,\ldots,n_r \tag{54}$$

$$\hat{q}_{c,\min}^i(k) \le \hat{q}_c^i(k) \le \hat{q}_{c,\max}^i(k), i = 1,\ldots,n_c \tag{55}$$

where n_r and n_c are the numbers of RM and MTFC actuators available, respectively. In Figure 14 example, two RM and a single MTFC actuator are utilized. More than one MTFC system could be present, in case of merging motorways (as in [106]). However, in order to avoid cases where drivers experience more than one piece of queue/delay, it is assumed that an MTFC system, if any, is always located upstream of all metered on-ramps that feed the mainstream section which includes the bottleneck locations. At the multiple bottleneck area, uncontrolled on ramps and off ramps can be present in between the bottleneck locations without affecting the implementation of the control concept. The bounds of the actuators **(54)** and **(55)** will be determined in sections 4.3 and 4.4. Considering **(53)**-**(55)** the following can derived:

$$\sum_{i=1}^{n_r} q_{r,\min}^i(k) + \sum_{i=1}^{n_c} \hat{q}_{c,\min}^i(k) \le \hat{q}_t(k) \le \sum_{i=1}^{n_r} q_{r,\max}^i(k) + \sum_{i=1}^{n_c} \hat{q}_{c,\max}^i(k) \tag{56}$$

and, as a consequence, the bounds used to truncate the outputs of **(49)** are given by:

$$\hat{q}_{t,\min} = \sum_{i=1}^{n_r} q_{r,\min}^i(k) + \sum_{i=1}^{n_c} \hat{q}_{c,\min}^i(k) \tag{57}$$

$$\hat{q}_{t,\max} = \sum_{i=1}^{n_r} q_{r,\max}^{i}(k) + \sum_{i=1}^{n_c} \hat{q}_{c,\max}^{i}(k). \quad (58)$$

Generally, there may be an infinite number of flow distributions that satisfy (53)-(55). An approach that leads to a desired delay balancing across the involved actuators is presented in section 4.6.

In case one of the PI controllers fails, a possible reaction would be to just ignore this specific PI controller and the corresponding bottleneck location measurement. The system will recognize a possible congestion at this location from the next upstream measurement within the bottleneck location area. If the failure is on the detector measurement, pre-designed observers could be in place to use the other available measurements to estimate the missing detector. Depending on the number of detectors and bottlenecks, it may not be practical to design one observer for each and every combination of failure, but some main cases could be considered.

4.3. Ramp Metering

For the RM case, the ramp flows determined by the flow distribution algorithm can be implemented directly using traffic lights. The lower bounds required in inequality (54) can be specified by the queue management policy applied. A Proportional (P) controller with feed forward on-ramp demand may be utilized [4] to limit the on-ramp queue:

$$q_{qm}^{i}(k) = \left(w_i(k) - \hat{w}_i\right)/T + d_i^{sm}(k-1) \quad (59)$$

where $w_i(k)$ is an estimate of the queue on the on-ramp i at time instant k, $\hat{w}_i$ is the utilized set point, which is usually the maximum admissible on-ramp queue length for the on-ramp i, and $d_i^{sm}(k-1)$ is an exponentially smoothed value of the past demand measurements, which is used as an estimate of the demand for the next time period. The values obtained from (59) should be truncated in order to respect an infrastructure-related upper bound $\bar{q}_{r,\max}^{i}$ and a policy-related lower bound $\bar{q}_{r,\min}^{i}$. In the field, an estimated of the on-ramp queue can be obtained using a Kalman filter estimator [107].

Concerning the upper bounds required by inequality (54), it can be specified by the available demand:

$$q_d^i(k) = w_i(k)/T + d_i^{sm}(k-1) + c_i, c_i \geq 0 \quad \textbf{(60)}$$

where the constant c_i is used to ensure that the bound is not conservative in case of an underestimation of demand through smoothing. The truncation of the values obtained by **(60)** is applied utilizing the bounds defined earlier for **(59).**

4.4. MTFC via VSL

When MTFC is enabled by VSL, a secondary loop with an Integral (I) controller is used for each MTFC system (see [6] for details). This secondary loop compares the flow measurement q_c^i, which is collected downstream of VSL's i application area, with the corresponding desired flow $\hat{q}_c^i(k)$, delivered by the flow distribution algorithm, to calculate the VSL rate b^i. The VSL rate b^i is defined as the VSL-induced free speed divided by the non-VSL free speed and is approximately equal to the displayed VSL divided by the legal speed limit without VSL. The I-controller reads

$$b^i(k) = b^i(k-1) + K_I^i\left[\hat{q}_c^i(k) - q_c^i(k)\right] \quad \textbf{(61)}$$

where K_I^i is the integral gain of the controller. This I-type controller could be replaced by a look-up table of VSL rates versus desired flows, as an alternative [108].

Some practical VSL implementation aspects are considered. Posted VSL rates can only take predefined discrete values. As a consequence, the VSL rates delivered by **(61)** are rounded to the closest discrete value to obtain the corresponding posted VSL rates $\bar{b}^i(k) \in \left\{\hat{b}_i^{\min}, \hat{b}_i^{\min} + \Delta b, \ldots, \hat{b}_i^{\max}\right\}$ where Δb is the practiced discrete VSL increment, e.g. $\Delta b = 0.1$. In addition, the difference between two consecutively posted VSL rates at the same gantry is limited to $\Delta b_{\max}$, as often required in practice. Thus, the lower bound for the VSL rate that can be implemented is given by:

$$b_{\min}^i(k) = \max\left\{\hat{b}_{\min}^i, \bar{b}^i(k-1) - \Delta b_{\max}\right\} \quad \textbf{(62)}$$

and the upper bound is given by:

$$b_{\max}^i(k) = \min\left\{\hat{b}_{\max}^i, \bar{b}^i(k-1) + \Delta b_{\max}\right\} \quad \textbf{(63)}$$

Applying these bounds to (61), one can specify the bounds required by inequality (55) as:

$$\hat{q}^i_{c,\min}(k) = q^i_c(k) + \left[b^i_{\min}(k) - b^i(k-1) \right] / K^i_I \quad (64)$$

$$\hat{q}^i_{c,\max}(k) = q^i_c(k) + \left[b^i_{\max}(k) - b^i(k-1) \right] / K^i_I. \quad (65)$$

Furthermore, due to safety reasons, VSL may also be applied upstream of the controlled congestion. In that case, the difference between the posted VSL rate at two consecutive gantries is limited to $\delta b_{\max}$. Lastly, a constant VSL rate equal to 0.9 is applied in the acceleration area whenever MTFC is active [6].

4.5. Delay Estimation

As stated previously, the flow distribution to the available actuators will be determined in order to balance the delays experienced by the respective groups of drivers upstream of each actuator. An estimation of these delays is necessary, in order to achieve this goal.

For the case of vehicles queueing on an on-ramp i because of RM actions, $\delta^i_r(k+1)$ denotes the estimated delay to be experienced by drivers exiting the ramp at the next time period if a ramp flow $q^i_r(k)$ is implemented. Assuming no internal vehicle sinks and sources, and that vehicles enter and exit according to the first–in-first-out rule, an estimate of the delay is [103]:

$$\delta^i_r(k+1) = A^i_r - B^i_r q^i_r(k) \quad (66)$$

where $A^i_r = w_i(k) / d^{sm}_i(k-1) + T$ and $B^i_r = T / d^{sm}_i(k-1)$.

For the case of vehicles delayed by the controlled congestion due to MTFC actions, the delay can be estimated if the travel time under free flow conditions is subtracted from the currently experienced travel time for all the freeway segment located upstream of the control point that experience a speed smaller than the free flow speed v_f. This delay can be considered as having two components. The first component is the delay experienced within the most downstream part of the controlled congestion, where no on-/off-ramps are present, hence there are no internal sinks and sources, and vehicles enter and exit according to the first-in-first-out rule as at on-ramp queues;

whereas the second component considers the delay experienced farther upstream and is estimated by use of available speed measurements. Thus, the estimate of the delay due to the i-th MTFC system is given by:

$$\delta_c^i(k+1) = A_c^i - B_c^i q_c^i(k) \tag{67}$$

where $A_c^i = A_c^{i*} + N_i(k)/q_{in,i}^{sm}(k-1) + T - L_i/v_f$, $N_i(k)$ is an estimate of the number of vehicles within the most downstream (ramp free) motorway segment at time k, L_i is the length of that segment, $q_{in,i}^{sm}(k-1)$ is an exponentially smoothed value of the past inflow measurements at the entrance of this motorway segment, and $B_c^i = T/q_{in,i}^{sm}(k-1)$. Finally, A_c^{i*} is the second component of the delay that can be calculated based on speed measurements for all the segments that experience a speed smaller than the free flow speed v_f and are located further upstream.

4.6. Flow Distribution for Delay Balancing

The following knapsack optimization problem solution, delivers the flows to be applied for each actuator:

$$\min \sum_{i=1}^{n_r} \frac{\left(A_r^i - B_r^i q_r^i(k)\right)^2}{B_r^i} + \sum_{i=1}^{n_c} \frac{\left(A_c^i - B_c^i \hat{q}_c^i(k)\right)^2}{B_c^i} \tag{68}$$

subject to the linear equality **(53)** and the bounds on the decision variables **(54)** and **(55)**.

This problem is a convex optimization problem that is always feasible due to the fact that the bounds defined by **(56)** are considered for the truncation of the values calculated by **(49)**. It can be easily seen that by applying the first-order optimality conditions, delay equalization is achieved as long as none of the bounds is active. In case some bounds are active (for some actuators) then delay equalization is achieved for the rest of the actuators. The solution of this knapsack problem can be calculated using the computationally efficient algorithm developed by Brucker [109] within a finite number of iterations. An important notification is that the cost criterion **(68)** can be readily extended with additional weights so as to lead to any desired linear relations among the delays of different actuators, i.e., other than delay equalization.

Different actuators may feature different control periods. For instance, RM may be most efficient with a period of 20 *s*, whereas VSL cannot switch more frequently than each 1 min, in order to avoid driver irritation. In cases like this, the different control periods must be multiples of an equal-smaller common divisor, which is the period employed for the controller **(49)**. Then, at the time periods that it is not necessary to update the flow to be implemented by some actuator, its two bounds required by **(54)** and **(55)** are both set equal to the last decided flow value for the same actuator. The corresponding inequalities are acting as equalities, since both bounds are set equal to the same value. As a consequence, the solution of the knapsack problem is such that the flows of all actuators that are not updated remain indeed the same as in the last controller period, while all other flows are decided in order to guarantee delay equalization for all other actuators.

The application of the ramp flows and the VSL rates delivered by the control strategy starts when the measured density $\rho_{out,i}$ at a bottleneck location i becomes higher than an activation threshold, and ends when the measured densities at all bottleneck areas become lower than a deactivation threshold (which is lower than the activation threshold).

5. Feedback and Optimal Control Results for MTFC with multiple bottlenecks

In this chapter the approach presented in chapter 3 is applied to a model of a motorway in Perth, Australia. The network model and model calibration are described in section 5.1. Section 5.2 investigates the no control case, whereas section 5.3 investigates the optimal control approach, where VSL are used as an MTFC actuator. The examined motorway includes two bottleneck locations. Section 5.4 describes the case of taking into account only the one bottleneck, and section 5.5 the case of taking into account both bottlenecks.

5.1. Network Model and Model Calibration

A stretch of the Kwinana Freeway in Perth, Australia is considered for the simulations. The length of the considered stretch is about 19.8 *km* and extends from Leach Highway to Anketell Road. Figure 15 depicts the considered stretch. Arrows demonstrate links divided into a number of segments, indicated by vertical lines, whereas circles represent nodes. The nodes are positioned mainly at locations where on-ramps and off-ramps are connected to the mainstream. The METANET model has been validated for the stretch under consideration using real 2012 data. Figure 16 depicts the motorway stretch along with the positions of the detectors

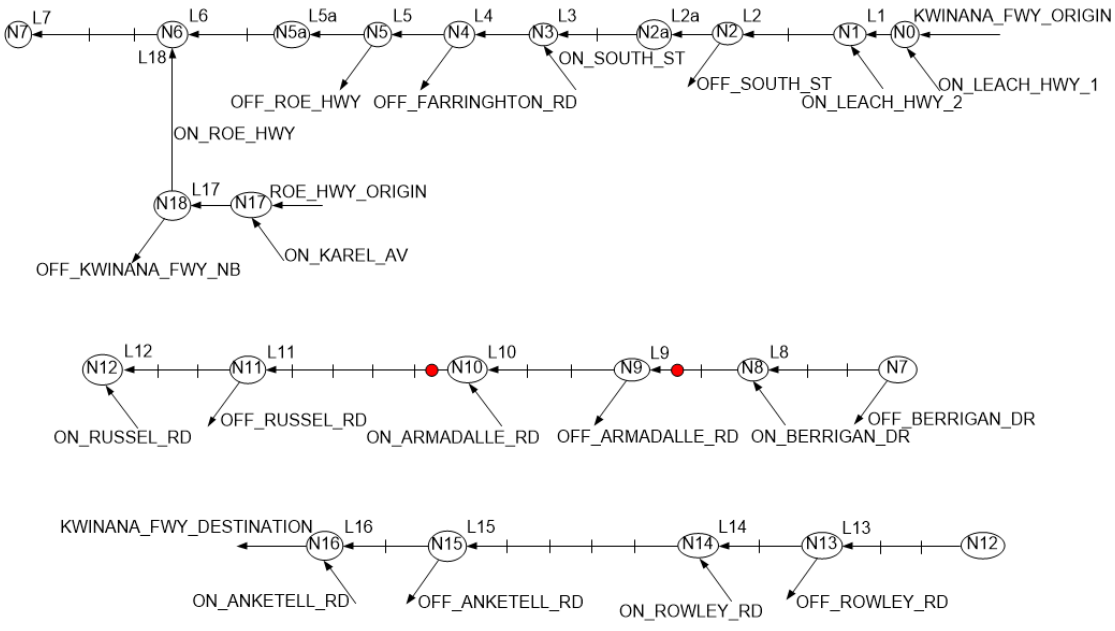

Figure 15: Motorway stretch with the two bottleneck areas marked with dots.

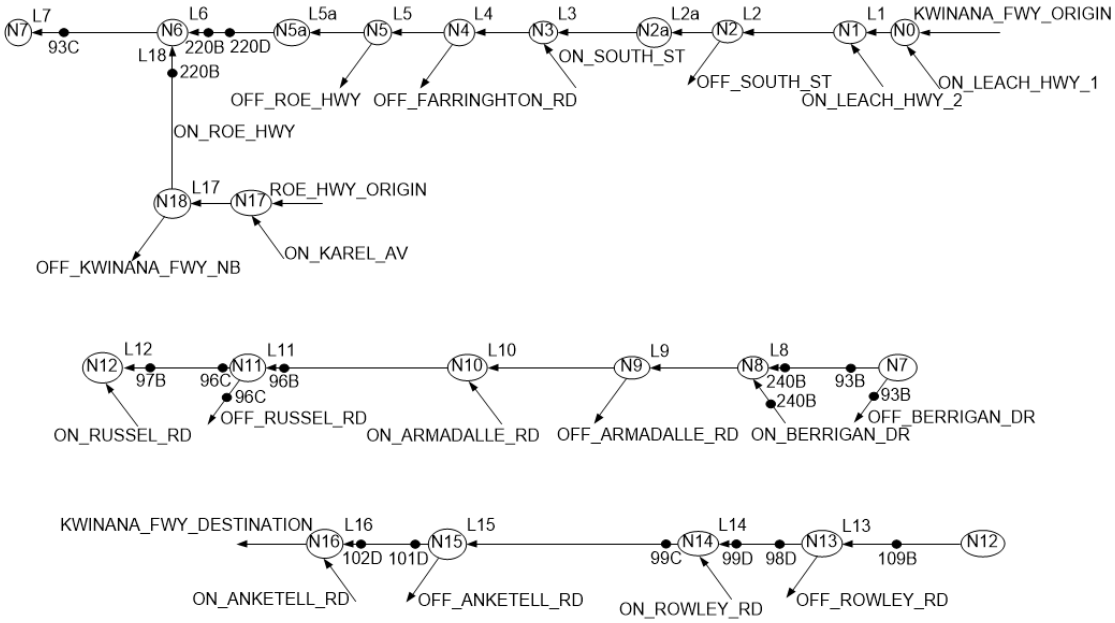

Figure 16: Motorway stretch along with the positions of the detector stations (bullet points with associated ids).

stations (bullet points with associated ids). The exclusive exit lane (Armadale Road off-ramp), corresponding to a lane drop in this section activates a bottleneck at L9 or/and the merge bottleneck at L11 (Armadale Road on-ramp), so there are two potentially active bottlenecks that are shown in Figure 15 with red bullets. The calibration procedure described in section 2.7, gave the following model parameters: $\tau = 22.8 \quad s, \quad \nu = 24.0 \; km^2 / h, \quad \delta = 2.3 \quad h / km, \quad v_{f,m} = 109.7 \; km / h, \quad \rho_{cr,m} = 27.6$ $veh / km / lane, \; \alpha_m = 2.3$. Figure 17 and Figure 18 depict measured versus predicted speeds and flows, respectively, from 13:00 to 20:00 on 15.10.2012, using the parameters mentioned above. It is observed that the model creates the congestion at the right place and for the right period of time (Figure 17), while it produces at parts a higher capacity drop downstream of the merge area compared to the real flow measurements (Figure 18). A reason could be that in the real measurements (and traffic conditions), the real capacity of the infrastructure is under-represented. This is, firstly, because the 220B and Roe Highway flows are upstream of the merge area (there is no measurement in the merge area); secondly, and most importantly, the merging of two quite high flows at high speeds is likely to lead to premature breakdown, i.e., to congestion forming before the infrastructure capacity is actually

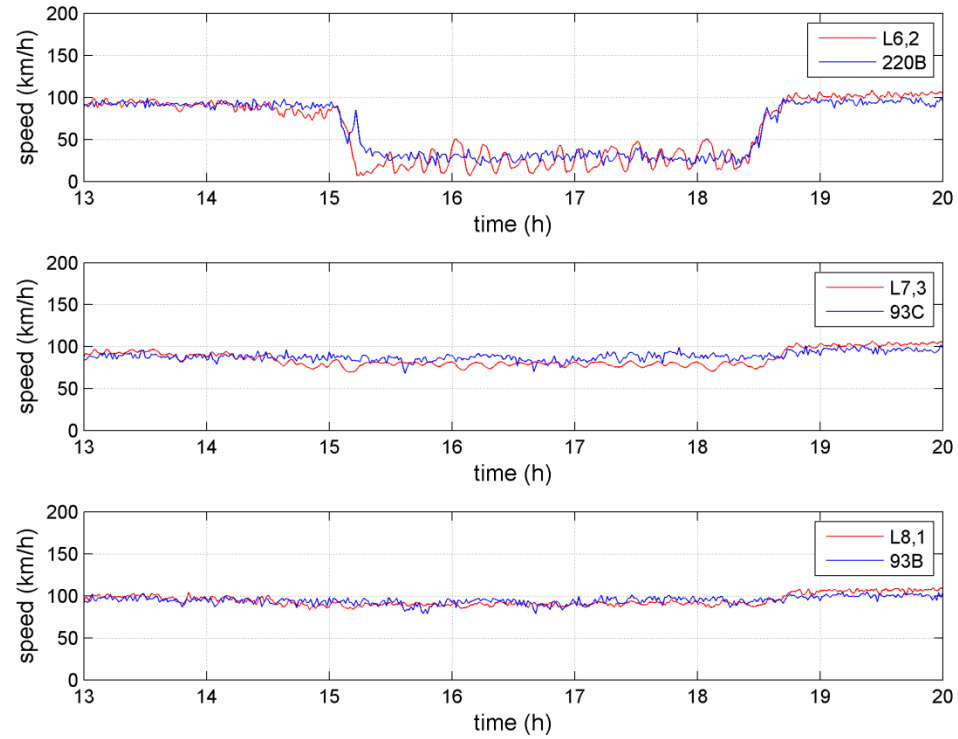

Figure 17: Measured versus predicted speed from 13:00 to 20:00 on 15.10.2012.

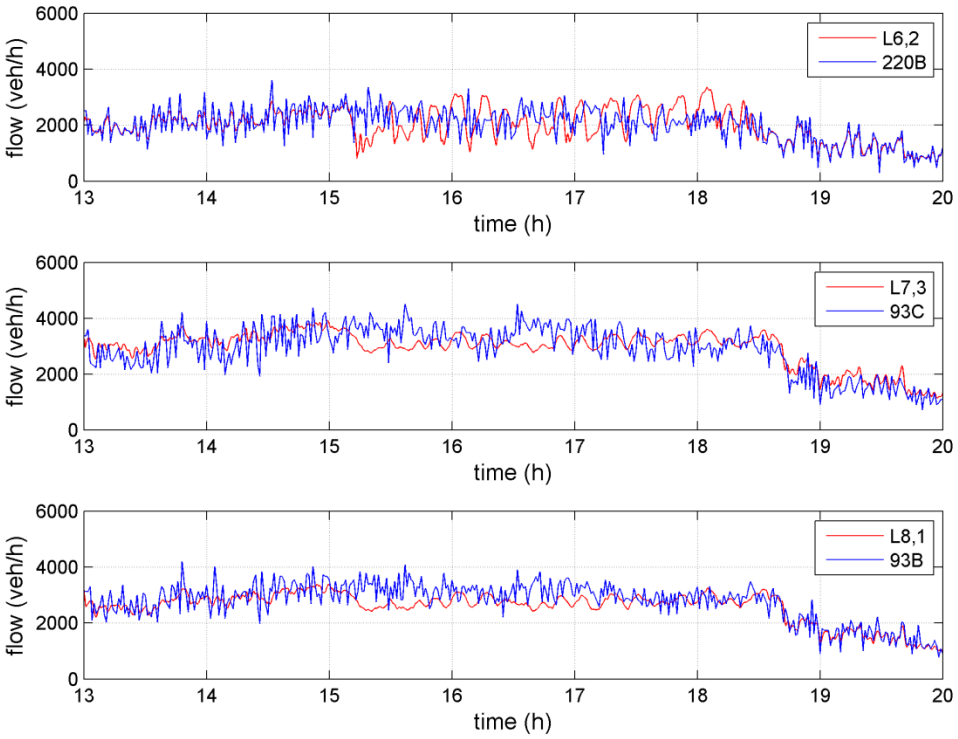

Figure 18: Measured versus predicted flow from 13:00 to 20:00 on 15.10.2012.

TABLE 1: SUMMARY OF SIMULATED SCENARIOS

Strategy	Description	TTS (veh·h)	%
No control	No-Control case.	7,145.4	-
Scenario 1	Optimal MTFC with VSL (AMOC) applied at L8.	5,391.3	24.55
Scenario 2	Feedback MTFC with VSL applied at L8 for the bottleneck at L11.	6,141.4	14.05
Scenario 3	Feedback MTFC for multiple bottlenecks with VSL applied at L8 for the bottlenecks at L9 and L11.	5,924.0	17.09

reached; this is supported by the relatively low total flow (less than 2000 $veh/h/lane$) observed just before the breakdown.

In TABLE 1 the strategies and the TTS results are summarized, and in Figure 19 the speed contour plots for all the strategies considered, are demonstrated.

5.2. No-Control Case

The no-control case is the case where no control strategies are applied, and is the base case that is going to be used to quantify any improvements arising from the use of control. Figure 20 shows the density, speed, and flow profiles for the two bottleneck areas. The demonstrated traffic situation is complex enough and can be explained as follows. At $t=15.6$ h the merge area of the ON_ARMADALE_RD on-ramp (L11) reaches the factual capacity of 4000 veh/h. Mainstream congestion is created after $t=15.7$ h, as the flow arriving at L11 continues to increase. As a result, the mainstream flow decreases due to the capacity drop phenomenon. This congestion propagates upstream but lasts only for 10 minutes, as it is clearly visible in Figure 19. Immediately after, congestion is created at L9 at around $t=15.8$ h which propagates upstream over 6.9 km and lasts for about 3.5 h (see Figure 19). The onset of this

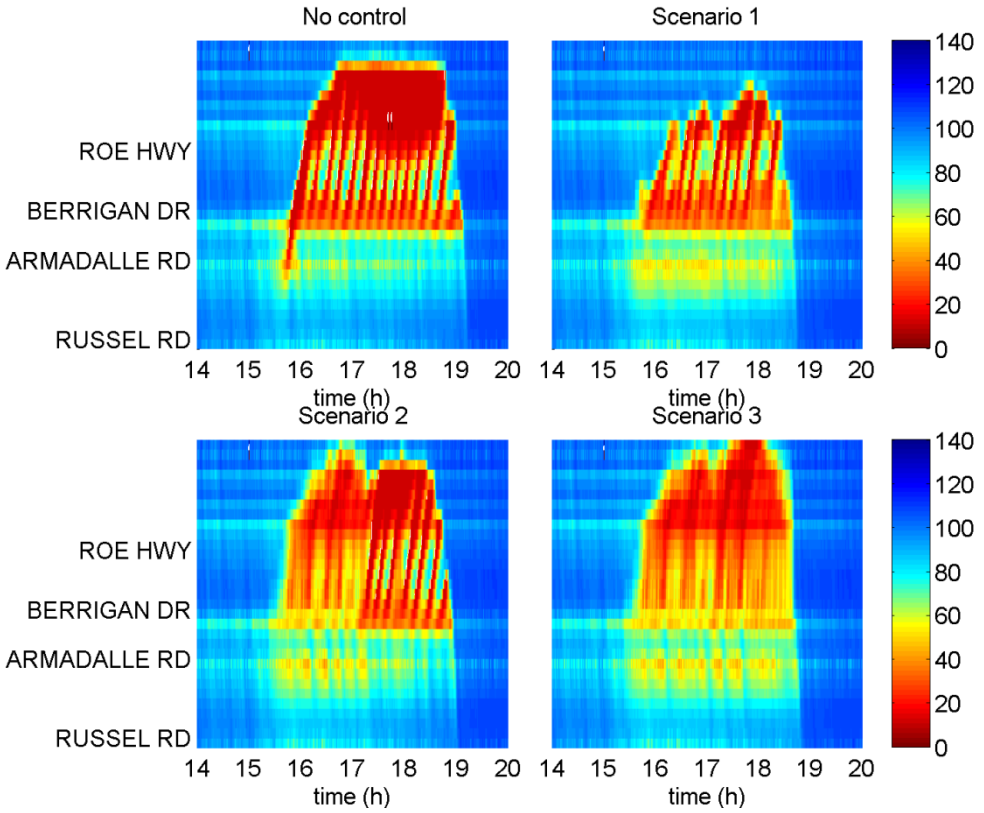

Figure 19: Speed (km/h) contour plots for the No-control case, Scenario 1, Scenario 2, and Scenario 3.

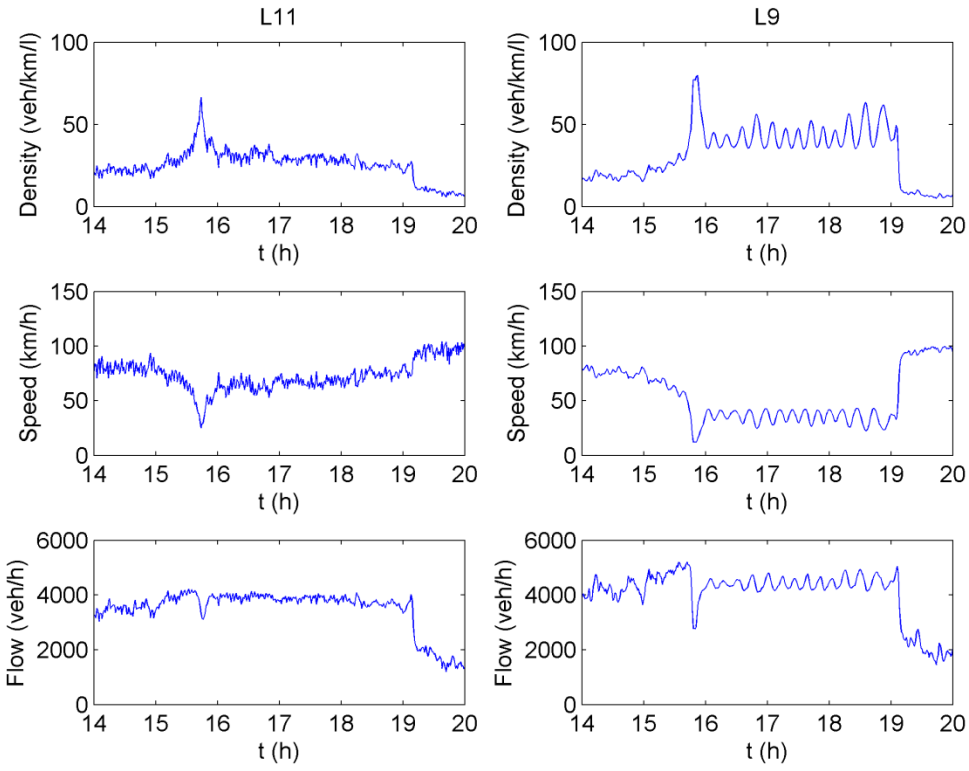

Figure 20: Traffic conditions at the two bottleneck areas, for the no-control case.

second congestion is due to the lane drop at node N9, from three lanes on link L9 down to two lanes on link L10, while the trigger is the spillback of congestion from L11. When congestion is created at L9, the flow feeding L11 is reduced, causing resolution of congestion at links L10 and L11. The resulting TTS is equal to 7,145.4 *veh·h*.

5.3. Scenario 1

In optimal MTFC (via AMOC), disturbances (demand and turning rate profiles) are assumed to be known for the whole period. Thus, optimal MTFC results may be viewed as an upper bound of achievable performance. This upper bound is used for the assessment of feedback control cases.

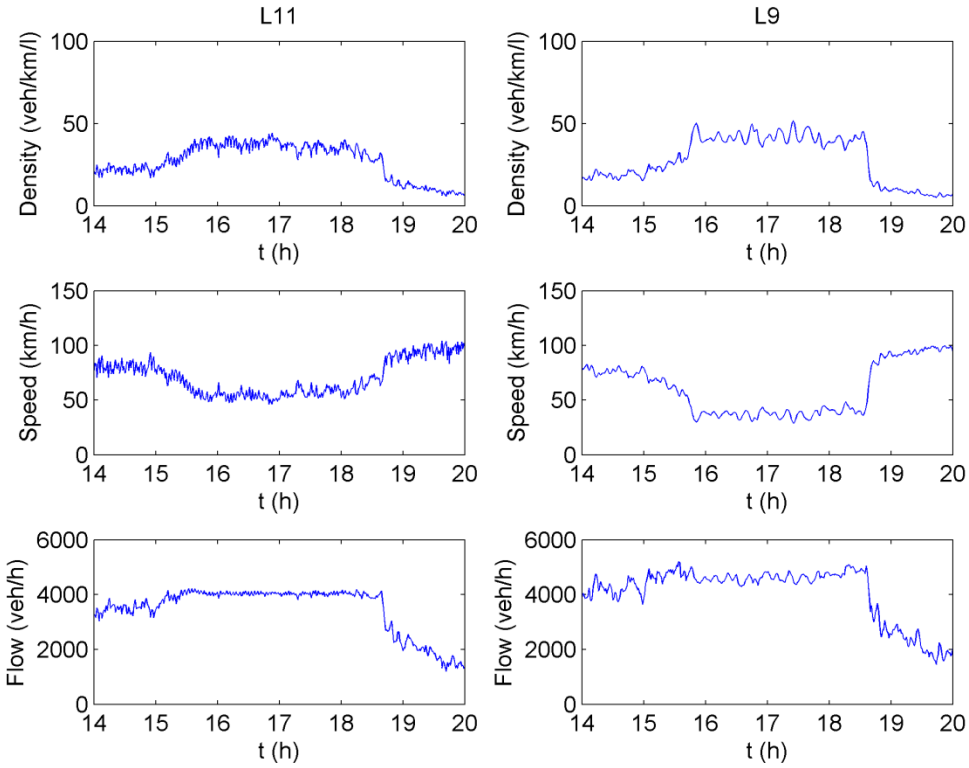

Figure 21: Traffic conditions at the bottleneck areas for scenario 1.

The controlled link is L8 (see Figure 15), while all other links are uncontrolled. A change can be made in VSL rates every 60 *s* as in the feedback control case. The lowest admissible value for the VSL rates is $b_{\min,m} = 0.2$, as higher values may limit the MTFC performance. The resulting TTS is 5,391.3 *veh·h*, which is a 24.55% improvement, compared to the no-control case. Figure 21 displays the density, speed

and flow profiles for both bottleneck areas. The optimal VSL rate trajectories are shown in Figure 22(a).

Due to the VSL actions, capacity flow is achieved at L11 (see Figure 21) which leads to minimization of TTS. The VSL rate for L8 departs from 1.0 and is varied appropriately, at times approaching almost 0.2 (the lowest limit for VSL). This results in the creation of a controlled mainstream congestion upstream of the acceleration area. This congestion extends over some 5.4 *km* for 2.8 *h*, which is smaller (in space and time) than in the no-control case and has higher internal speed (see Figure 19). After the end of the congestion period, the VSL rate is gradually increased back to 1.0 (see Figure 22(a)).

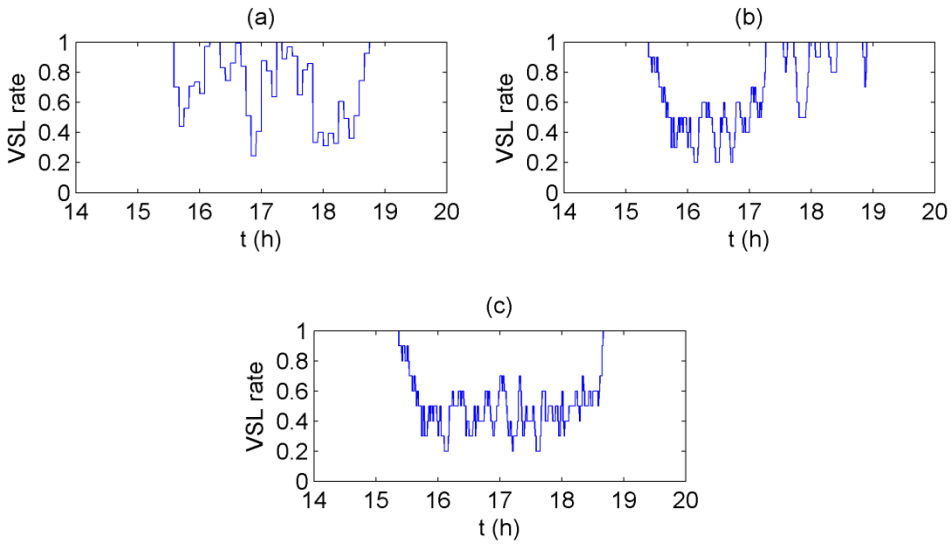

Figure 22: (a) Scenario 1: Optimal VSL rates. (b) Scenario 2: VSL rates given by the feedback controller for a single bottleneck. (c) Scenario 3: VSL rates given by the feedback controller for multiple bottlenecks.

5.4. Scenario 2

Feedback MTFC via VSL is applied here, with the constraints described in section 3.4. The VSL application area is link L8, whereas upstream of L8 there are safety limits, while downstream of L8, up to L11, a constant VSL rate $b = 0.9$ is applied whenever MTFC is active. Density measurements are taken from the first segment of L11, while flow measurements are taken from the first segment of L9; thus a single bottleneck is addressed in this scenario. The set-point for the primary controller is set

to $\hat{\rho}_{out} = 38$ $veh/km/lane$. The activation threshold is set equal to 28 $veh/km/lane$, while the deactivation threshold is set equal to 21 $veh/h/lane$.

The resulting TTS is 6,141.4 *veh·h*, which is a 14.05% improvement compared to the no-control case. The density, speed and flow profiles for both bottleneck locations are shown in Figure 23. The dashed line shows the density set-point utilized by the primary controller for L11, while the red line shows the period during which MTFC is active according to the activation/deactivation thresholds used. The feedback VSL rate trajectories are shown in Figure 22(b) and the feedback flow trajectories are shown in Figure 24(a).

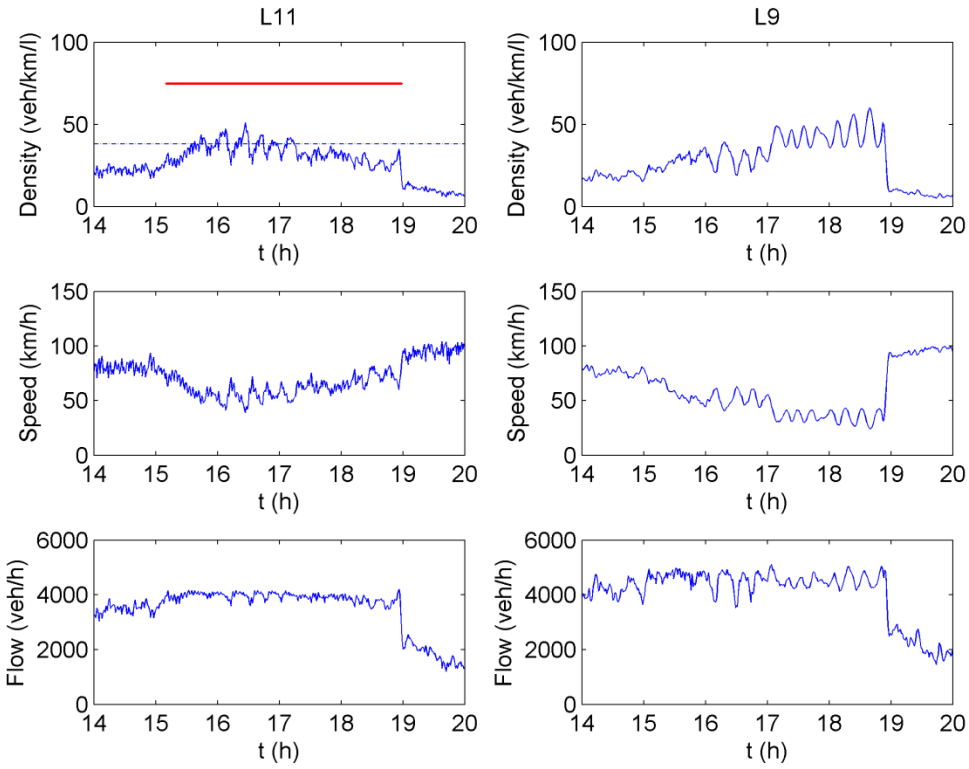

Figure 23: Traffic conditions at the bottleneck areas for scenario 2.

The VSL rate at L8 [see Figure 22(b)] is seen to behave similarly as in the optimal control case, but it reaches lower VSL rates. The flow decrease at the first segment of L9 [see Figure 24(a)] shows the impact of VSL on the controlled variable q_c. On the same diagram, the dashed line shows the output $\hat{q}_c$ of the primary controller. This flow is a reference for the secondary controller and, at most times, it is closely followed by the controlled variable. However, the flow at the bottleneck area (second

segment) of L9 is higher than what can be accommodated by L10, and, as a result, congestion is created there (Figure 19) which is not visible at the density measurement location farther downstream and can therefore not be addressed by the single-bottleneck feedback MTFC scheme. In other words, congestion is created between the VSL and the addressed L11 bottleneck, and this result indicates the necessity for a logic that can treat multiple bottlenecks.

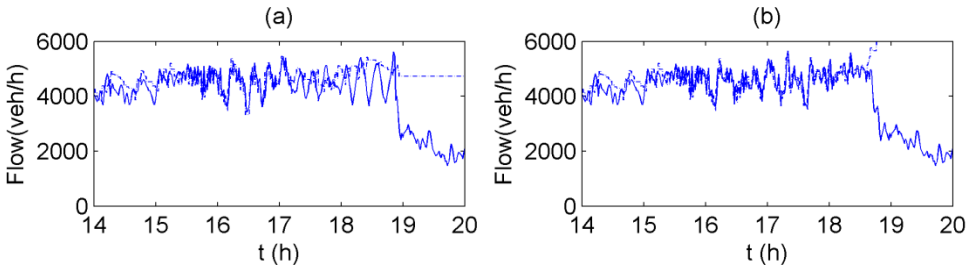

Figure 24: (a) Scenario 2: Flow (solid line) at the first segment of L9 and set-point (dashed line) given as an output of the primary loop of the feedback controller for a single bottleneck. (b) Scenario 3: Flow (solid line) at the first segment of L9 and set-point (dashed line) given as an output of the primary loop of the feedback controller for multiple bottlenecks.

5.5. Scenario 3

The proposed feedback MTFC for multiple bottlenecks is applied here to control both known bottleneck locations (L9 and L11); of course, additional downstream measurements, corresponding to farther potential bottlenecks, could be used, however this would not alter the presented results as no other bottlenecks are actually present.

VSL is applied at L8, whereas upstream of L8 there are safety limits, while downstream of L8, up to L11, there is a constant VSL rate $b = 0.9$ whenever MTFC is active. The set-point for the primary controller of L9 is set to $\hat{\rho}_{out} = 36$ $veh/km/lane$ and for the primary controller of L11 is set to $\hat{\rho}_{out} = 38$ $veh/km/lane$. Again, the activation threshold is set equal to 28 $veh/km/lane$ while the deactivation threshold is set equal to 21 $veh/h/lane$.

The resulting TTS is 5,924 $veh \cdot h$, which is a 17.09% improvement compared with the no-control case, clearly better than scenario 2. The density, speed, and flow profiles for both bottleneck locations are shown in Figure 25. The dashed lines show

the density set-points utilized by the primary controllers while the red lines show the periods for which each one of the primary controllers is selected by the decision policy defined by the equations **(46), (47), (48)** during the active period.

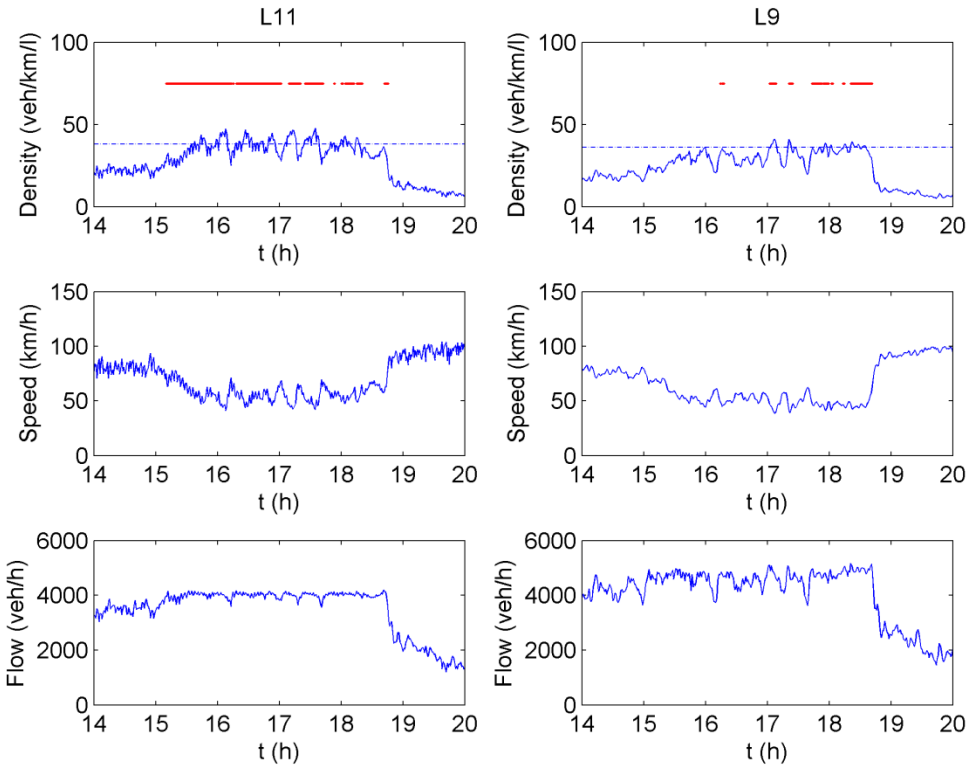

Figure 25: Traffic conditions at the bottleneck areas for scenario 3.

The feedback VSL rate trajectories are shown in Figure 22(c) and the feedback flow trajectories are shown in Figure 24(b).

Up to $t = 16.23h$, the primary controller that receives measurements from the bottleneck area at L11 is active, which means that up to this point the situation is identical to Scenario 2; after this, the decision logic switches to the upstream bottleneck whenever necessary to avoid its activation. The VSL rate at L8 [see Figure 22(c)] is gradually decreasing from 1 to 0.2 (the lowest limit for VSL). Compared to the previous scenarios, the VSL rates are less varying with time, and they take more moderate values. The flow decrease at the first segment of L9 [see Figure 24(b)] shows the impact of VSL on the controlled variable q_c. This decrease leads to avoidance of the bottleneck activation at both L11 and the second segment of L9, whereas a controlled mainstream congestion upstream of the acceleration area is created. This congestion extends over some 6 km for 2.5 h, which is smaller (in space

and time) than in the no-control case and Scenario 2, having also a higher internal speed (see Figure 19).

6. Feedback Integrated Control Results with Delay Balancing

In the current chapter the approach presented in chapter 4 is applied to a motorway in the United Kingdom. Section 6.1 includes the network model and model calibration description. The no control case is investigated in Section 6.2, whereas sections 6.3 and 6.4 investigate cases with local RM without queue constraints and local RM with queue constraints, respectively. Section 6.5 describes a case with coordinated RM, and section 6.6 a case with feedback MTFC. Finally, section 6.7 examines a case with integrated control.

6.1. Network Model and Model Calibration

A network in United Kingdom is considered for the simulations. The length of this stretch is 11.3 km. The graph of the motorway stretch is depicted in Figure 26. Arrows represent links divided into a number of segments, indicated by vertical lines. Links ON2 and ON3 are in fact motorway-to-motorway connections, modelled here as on-ramps. The METANET model has been calibrated for the stretch under consideration using MIDAS data [110] for the AM peak of September 9, 2014.

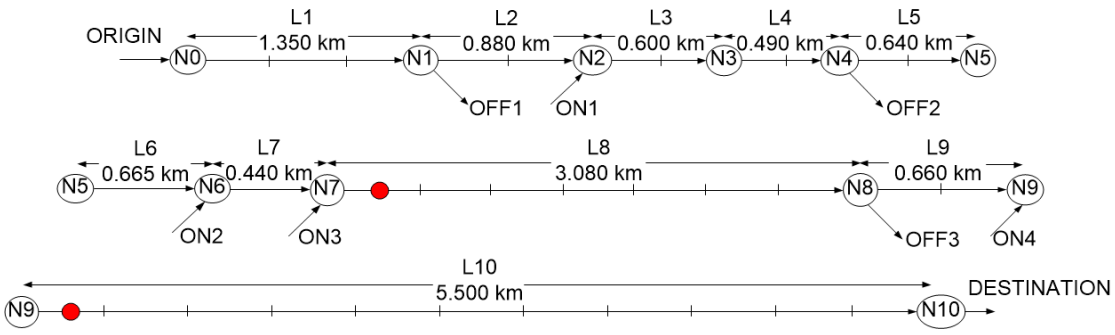

Figure 26: The motorway stretch considered. The two bottleneck areas are marked with red dots

Figure 27 depicts the motorway stretch along with the positions of the detectors stations (bullet points with associated ids). The active bottlenecks are located at links L8 and L10, i.e., a multiple bottleneck case exists if the on-ramp ON4 is not controlled. The calibration procedure described in section 2.7, gave the following model parameters: $\tau = 18.34 s$, $\nu = 19.84\ km^2/h$, $\delta = 1.97\ \ h/km$ for all the links and for L1 until L9 $v_{f,m} = 106.63\ km/h$, $\rho_{cr,m} = 28.94\ veh/km/lane$, $\alpha_m = 2.61$, and for

L10 $v_{f,m} = 104.22\ km/h$, $\rho_{cr,m} = 28.05\ veh/km/lane$ $\alpha_m = 2.76$. Figure 28 and Figure 29 depict measured and predicted speeds, respectively, from 5:00 to 11:00 on 9.9.2014, using the parameters mentioned above. It is observed that the model creates the congestion at the right place and for the right period of time.

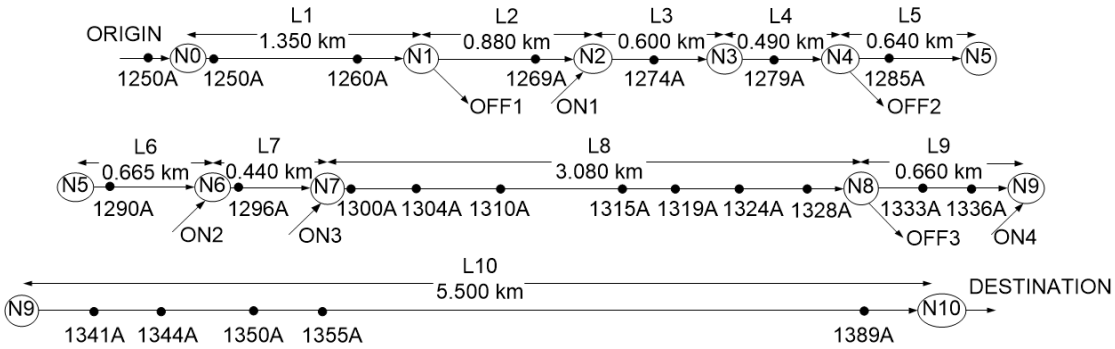

Figure 27: Motorway stretch along with the positions of the detector stations (bullet points with associated ids).

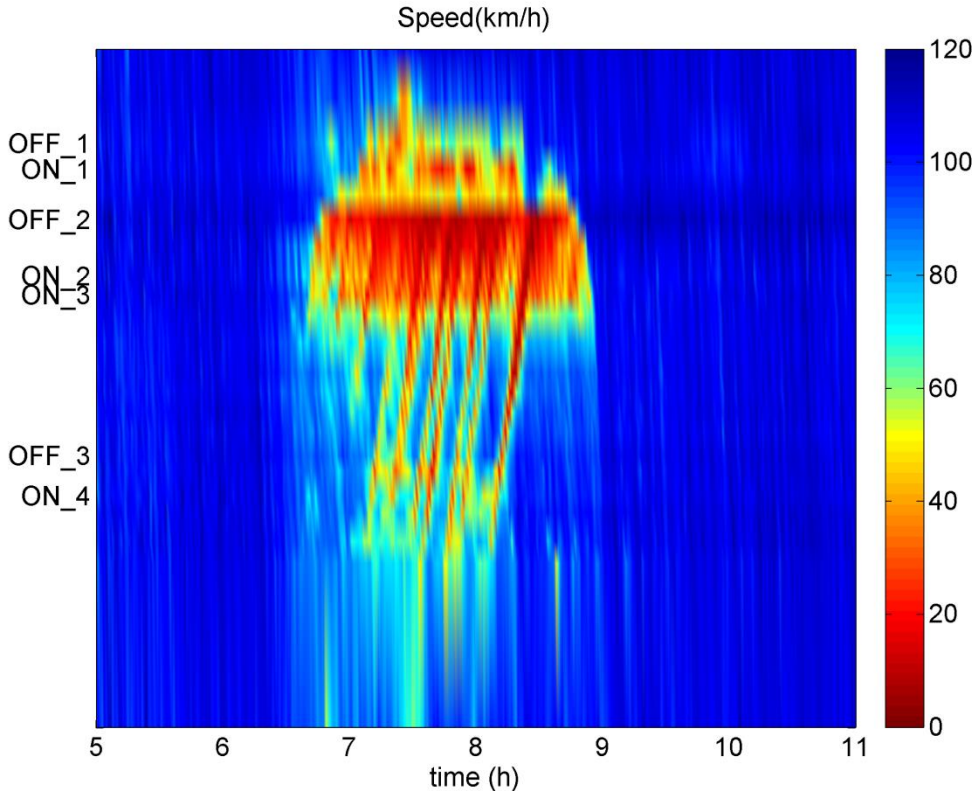

Figure 28: Speed (km/h) contour plot with the real data.

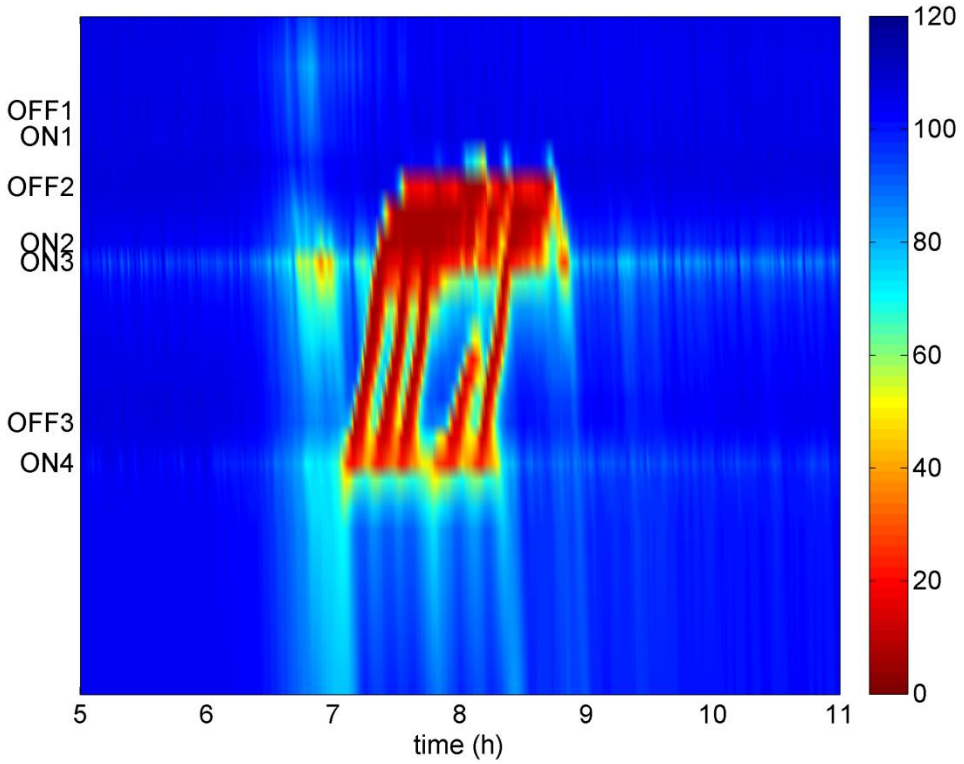

Figure 29: Speed (km/h) contour plot for the no-control case.

TABLE 2: DESCRIPTION OF SCENARIOS AND ACHIEVED RESULTS

	Description	TTS (*veh·h*)	Improvement (%)	TD (*veh·h*)	Improvement (%)
No-Control	Calibrated no-control case	3949	-	1178	-
Scenario 1	Local ramp metering with two separate controllers (no queue constraints)	3133	20.7	361	69.4
Scenario 2	Local ramp metering with two separate controllers (queue constraints)	3437	13.0	665	43.5
Scenario 3	Coordinated ramp metering for multiple bottlenecks	3539	10.4	767	34.9
Scenario 4	MTFC enabled via VSL for multiple bottlenecks	3408	13.7	540	54.2
Scenario 5	Integrated control for multiple bottlenecks	3139	20.5	340	71.1

A set of control scenarios are explored (see TABLE 2), each for a time horizon of 6 hours (5-11 AM). The utilized model time step is set to 5 *s*. The demand profiles for all the origins of the network are presented in Figure 30. A different control structure is used per scenario in order to show the increased efficiency achieved by the proposed approach. The efficiency measures considered (see TABLE 2) are the Total Time Spent (TTS), in the network, which is the sum of the Total Travel Time (TTT) and the Total Waiting Time (TWT) at the origins, as also the Total Delay (TD), which is given as the sum of the Mainstream Delay and TWT at the origins.

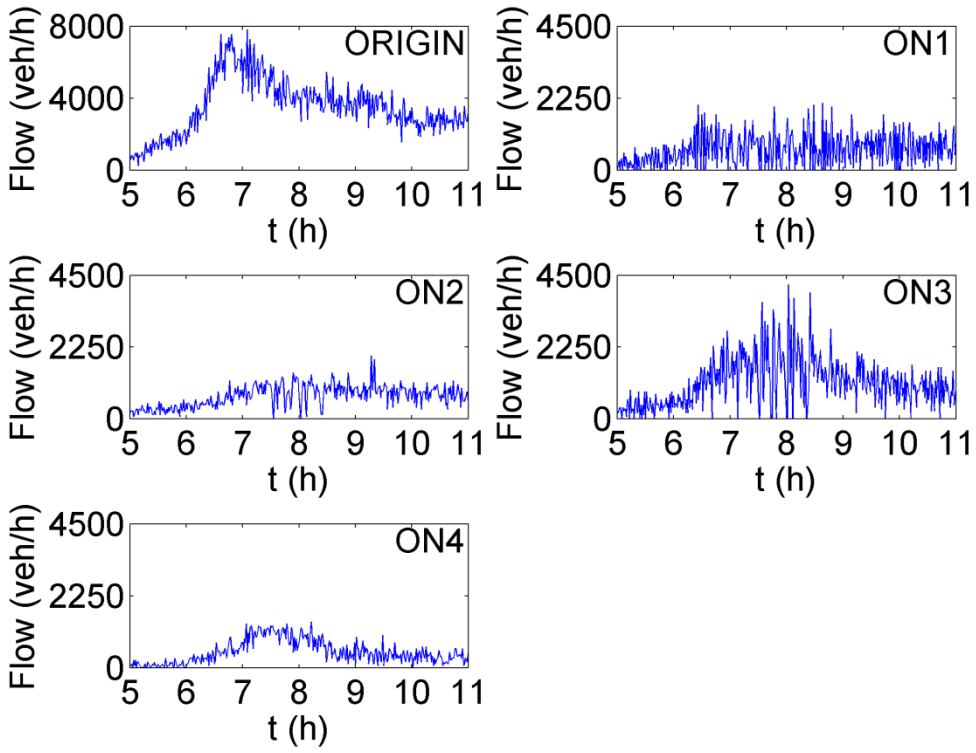

Figure 30: Demand profiles for all the origins of the network.

6.2. No-Control Case

No-control is the base case that will be used to quantify any efficiency improvements arising from the use of control actions. Figure 29 demonstrates the no-control speed contour plot for the time horizon under consideration. At $t = 6.75\ h$, the merge area of the ON3 on-ramp reaches its factual capacity of about 6000 veh/h. A short-lived congestion is created, lasting for about 15 *min*, without any major propagation of the phenomenon further upstream. At $t = 7\ h$, congestion is created at the merge area of the ON4 on-ramp because the demand is exceeding capacity (around 6200 veh/h) at the specific area. A capacity drop of around 15% is created and congestion propagates

upstream over 6.6 km triggering more severe congestion phenomena at the merge area of ON3 that last till about $t = 9\,h$. The resulting TTS in the network is equal to 3949 $veh{\cdot}h$, while TD is equal to 1178 $veh{\cdot}h$.

6.3. Scenario 1

Scenario 1 applies local RM actions using two separate controllers; a first controller receives measurements from the first segment of link L8 and acts using RM at on-ramp ON3; a second controller receives measurements from the first segment of link L10 and acts using RM at on-ramp ON4. In both cases the control period was set to 20 s while the P-term gain value in **(49)** was set to zero and the I-term gain value was set to 90 $km{\cdot}lane/h$ (which corresponds to the well-known ALINEA regulator [34]). The set-points of the controllers are set equal to the respective factual critical densities, namely 35 $veh/km/lane$ and 29 $veh/km/lane$. No queue management actions are considered in order to investigate what is the upper bound of efficiency that can be achieved for the case of the most direct control measure, i.e., local RM. This is implemented by setting a very high admissible value for the queue in **(59)**.

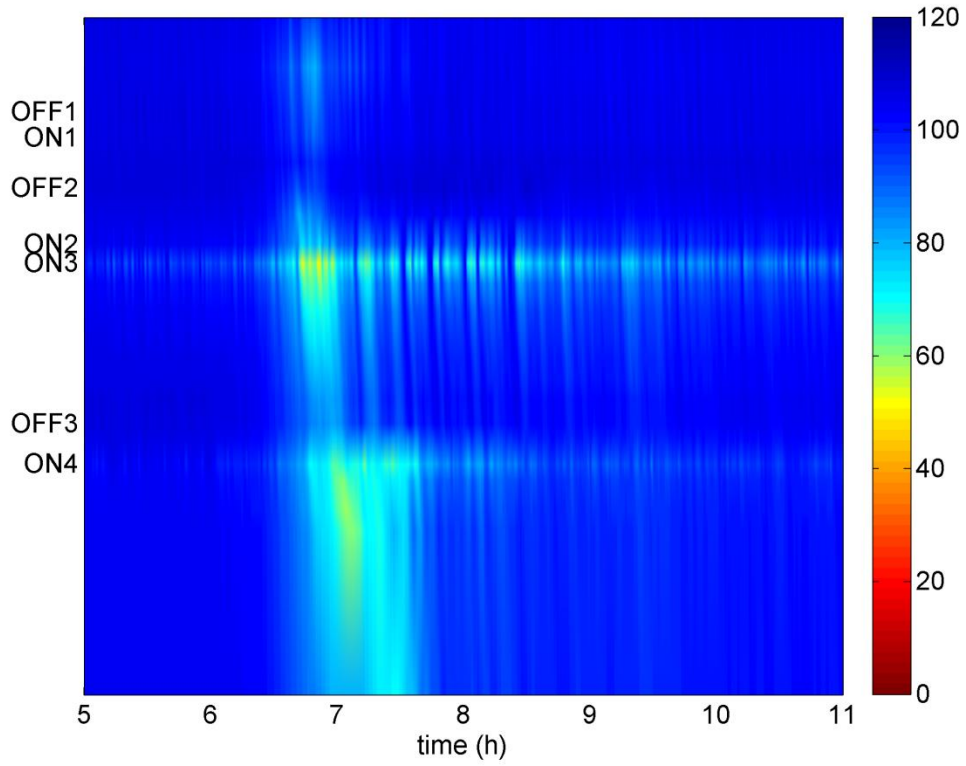

Figure 31: Speed (km/h) contour plot for the Scenario 1.

Compared to the no-control case, the resulting TTS is reduced by 20.7% while the resulting TD is reduced by a remarkable 69.4%. The speed contour plot for Scenario 1 is presented in Figure 31, while the queues created on the on-ramps due to RM actions are shown in Figure 32. At both bottlenecks, density values are maintained around the corresponding set-points; thus capacity flow is achieved at L8 and L10, which leads to minimization of TTS. The real delays experienced by drivers queueing at on-ramps ON3 and ON4 are displayed in Figure 33(a) and are as expected, completely unbalanced.

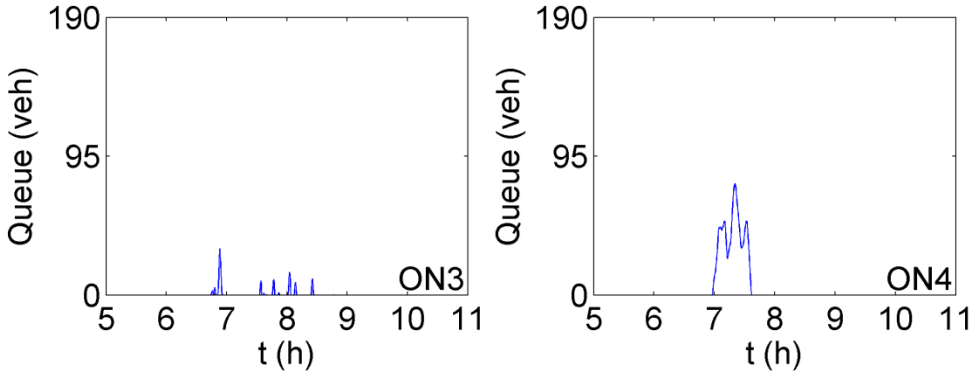

Figure 32: Queue profiles for Scenario 1.

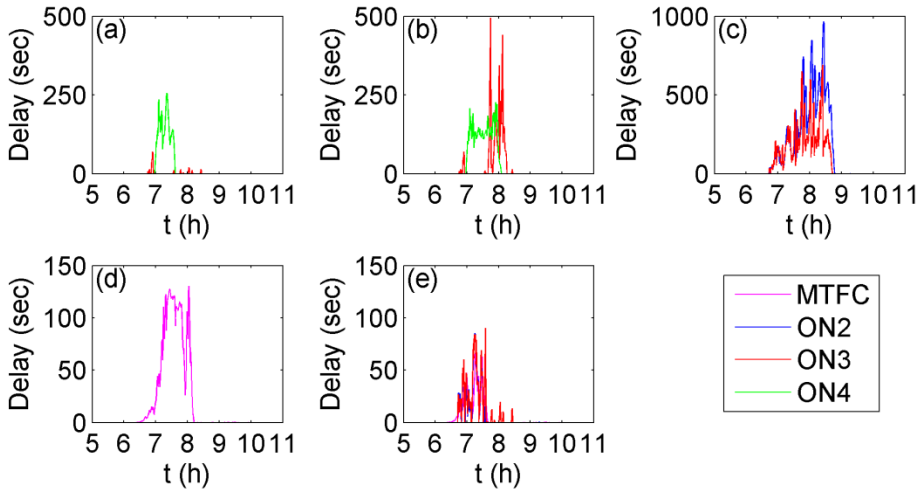

Figure 33: Delay profiles for (a) Scenario 1; (b) Scenario 2; (c) Scenario 3; (d) Scenario 4; Scenario 5.

6.4. Scenario 2

Scenario 2 applies local RM actions as in Scenario 1. The only difference is that queue management actions are now considered. The maximum admissible queues (based on an estimate of the real storage space of the infrastructure) are 92 *veh* for ON3 and 40 *veh* for ON4. Compared to the no-control case, the resulting TTS is now

reduced by 13.0%, while the resulting TD is reduced by 43.5%. The speed contour plot for Scenario 2 is presented in Figure 34, while the queues created on the on-ramps due to RM actions are shown in Figure 35.

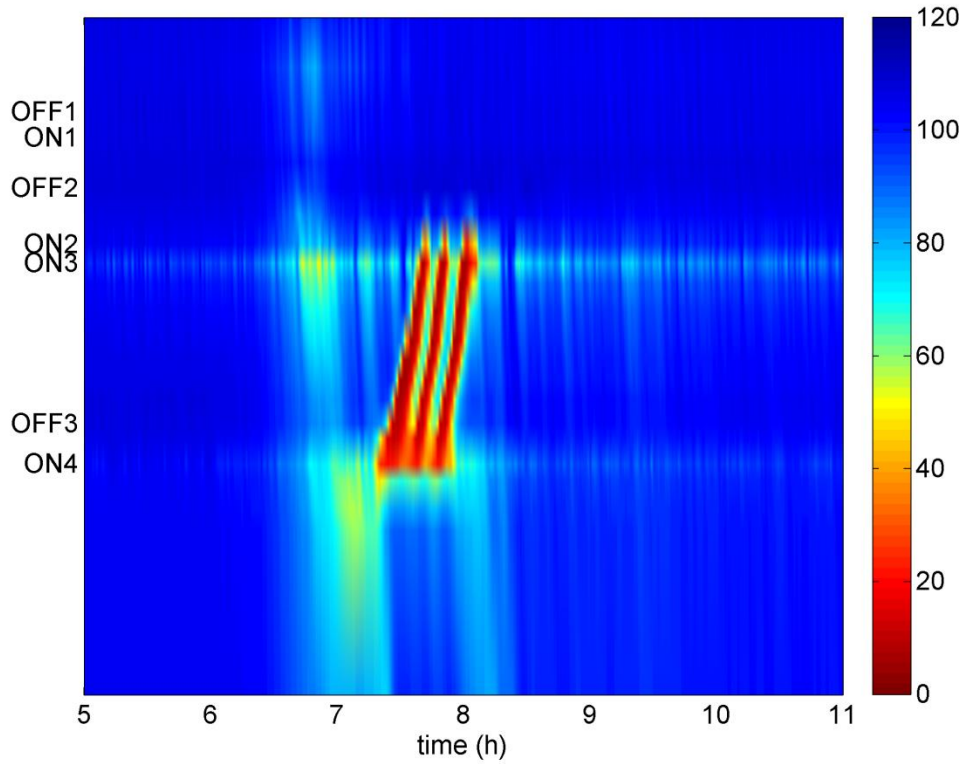

Figure 34: Speed (km/h) contour plot for Scenario 2.

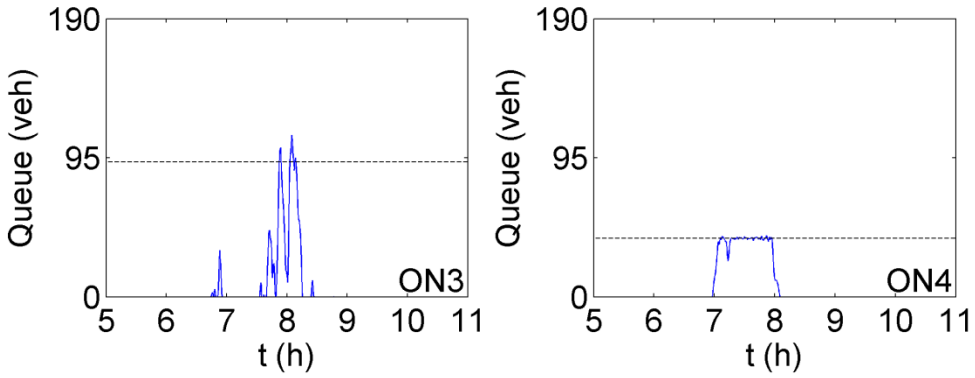

Figure 35: Queue profiles for Scenario 2.

As expected, RM actions are now just delaying the onset of congestion because the applied queue controller releases the ON4 on-ramp flow in order to maintain the queue around its maximum admissible value. Congestion propagates upstream causing further RM actions at the ON3 on-ramp. However, due to queue management actions there, the ON3 on-ramp flow is also released leading to congestion

propagation up to the area of the OFF2 off-ramp, which is a major connection to another motorway. The quite unbalanced on-ramp queue delays are displayed in Figure 33(b).

6.5. Scenario 3

Coordinated RM is applied in Scenario 3 utilizing the proposed new approach. Both bottleneck locations, L8 and L10, are considered by a single control structure, and RM is applied at on-ramps ON2 and ON3, i.e., the two on-ramps that are situated upstream of both bottlenecks, (with maximum admissible queues of 180 *veh* for ON2 and 92 *veh* for ON3) with a control period of 20 *s*. The utilized density set-points remain the same as in the previous scenarios, i.e., $\hat{\rho}_{out,1} = 35\ veh/km/lane$ for L8 and $\hat{\rho}_{out,2} = 29\ veh/km/lane$ for L10, while the gain values are $\hat{K}_{I,1} = \hat{K}_{I,2} = 5\ km \cdot lane/h$ and $\hat{K}_{P,1} = \hat{K}_{P,2} = 30\ km \cdot lane/h$.

Compared to the no-control case, the resulting TTS is reduced by 10.4% while the resulting TD is reduced by a remarkable 34.9%. The speed contour plot for Scenario 3 is presented in Figure 36, while the queues created on the on-ramps due to RM actions are shown in Figure 37.

At both bottlenecks, density values are maintained around the corresponding set-points up to 7:45 AM, i.e., up to the point that queue management actions are applied at on-ramp ON3. The delays experienced by drivers queueing at on-ramps ON2 and ON3 are displayed in Figure 33(c) and are, as expected, balanced up to 7:45 AM. Later on, the optimizer asks for stronger RM actions to be applied at on-ramp ON2, since metering at on-ramp ON3 is practically inactive. However, because the arriving demand at ON2 has meanwhile fallen to low levels (lower than the lower bound applied on the RM flow), the ordered metering does not materialize, and congestion at L10 cannot be avoided. This low demand after 7:45 AM is the reason why the queue created at on-ramp ON2 is never reaching the maximum storage capacity.

It is interesting to note that this coordinated RM scenario can be readily modified to act towards balancing of the relative on-ramp queues (as in the well-known HERO system [103]), rather than balancing of the respective time-delays. This would enable a better exploitation of the available storage space in both on-ramps, before queue management actions are activated. For the present infrastructure and demand

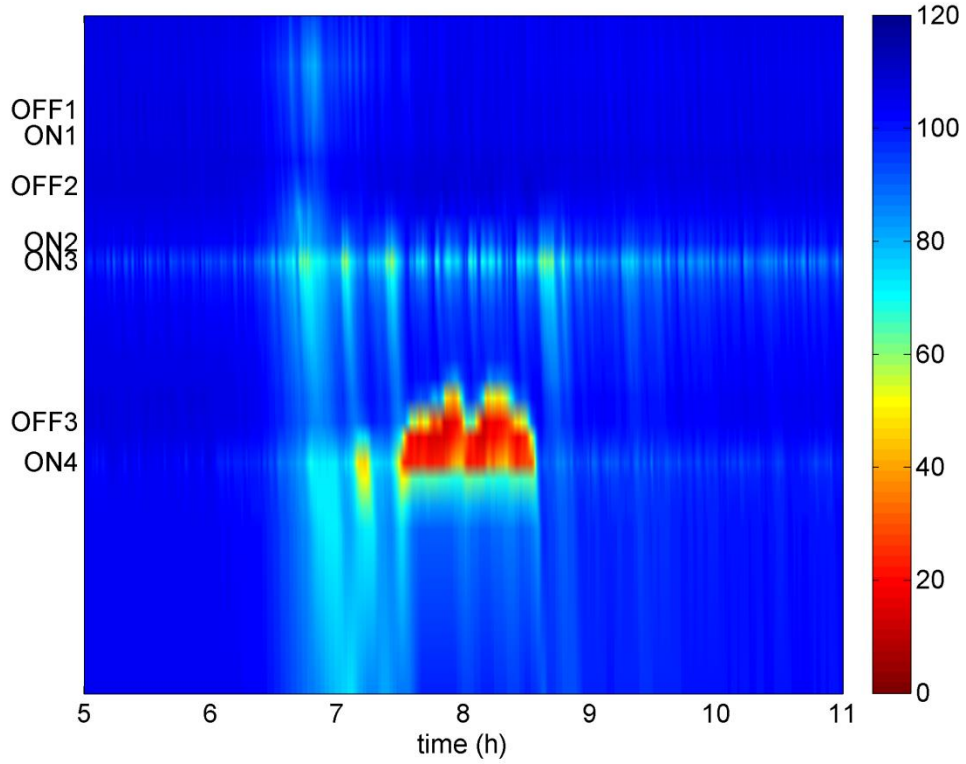

Figure 36: Speed (km/h) contour plot for Scenario 3.

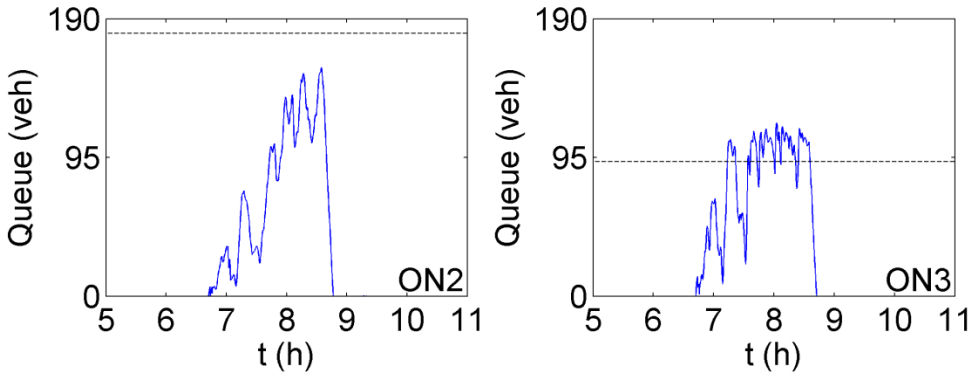

Figure 37: Queue profiles for Scenario 3.

configuration, this approach leads to an improvement in the TTS value of 18.9% and to an accordingly smaller mainstream congestion.

6.6. Scenario 4

Scenario 4 applies feedback MTFC for two bottleneck locations, L8 and L10. The VSL application area comprises links L4 and L5, whereas upstream of L4 there are safety-related VSL; the acceleration area comprises links L6 and L7. The control period was set to 60 *s*. The utilized density set-points as well as the gain values are the

same with those used for Scenario 3. The following values are used for various parameters required by the secondary controller (the $i=1$ index has been dropped for simplicity): $\hat{b}_{\min}=0.2$, $\hat{b}_{\max}=1.0$, $\Delta b=0.1$, $\Delta b_{\max}=0.1$, $\delta b_{\max}=0.3$ and $K_I=0.0015\ h\cdot lane/veh$.

Compared to the no-control case, the resulting TTS is reduced by 13.7% while the resulting TD is reduced by 54.2%. The speed contour plot for Scenario 4 is presented in Figure 38 while the VSL rate trajectory is shown in Figure 39. Note that no queues are created as no RM is applied.

The VSL rate is gradually decreased from 1 (no speed limit) to 0.2 (the lowest admissible limit for VSL), and a controlled congestion is created at the VSL application area. The onset of congestion at the merging area of the ON4 on-ramp is delayed up to a few minutes after 7 AM, i.e., up to the point at which the secondary I-regulator is saturated due to reaching the lower admissible VSL rate bound of 0.2. The delay experienced by drivers within the controlled congestion is displayed in Figure 33(d).

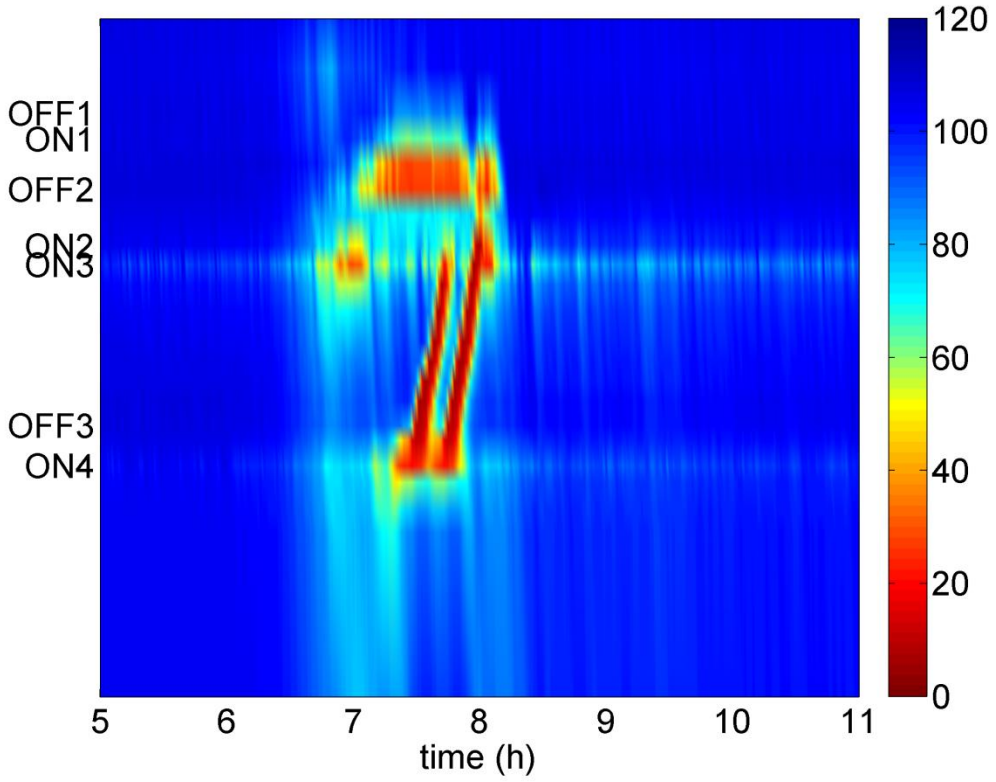

Figure 38: Speed (km/h) contour plot for Scenario 4.

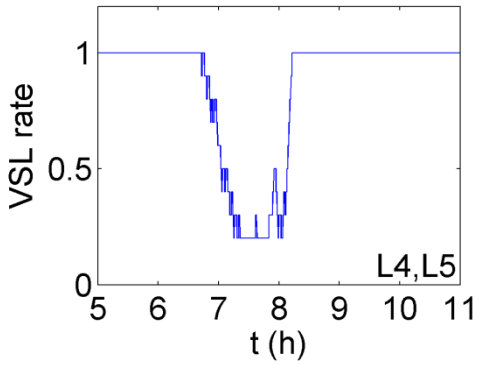

Figure 39: VSL rate for Scenario 4.

6.7. Scenario 5

Integrated control is applied in Scenario 5 using three actuators, i.e., two ramp meters applied at on-ramps ON2 and ON3 (with maximum admissible queues of 180 *veh* for ON2 and 92 *veh* for ON3) with a control period of 20 *s*; and a VSL-enabled MTFC with a control period of 60 *s* as in Scenario 4. The same gains and the settings as in Scenario 4 are used. Both bottleneck locations are considered using the integrated concept presented in Chapter 4 aiming at delay balancing for the three actuators.

Compared to the no-control case, the resulting TTS is now reduced by 20.5% while the resulting TD is reduced by a remarkable 71.1%. The speed contour plot for Scenario 5 is presented in Figure 40, while the VSL rate trajectory due to MTFC actions, as well as the queues created on the on-ramps due to RM actions, are shown in Figure 41. At both bottlenecks, density values are maintained around the corresponding set-points, thus capacity flow is achieved at L8 and L10. This is done without any queue saturation for the two on-ramps and without any saturation of the VSL rates. The created mainstream controlled congestion is much smaller (in space and time) than in the no control case, having also higher internal speed.

Finally, the delays experienced by drivers are displayed in Figure 33(e). It can be concluded that the (highest) efficiency of Scenario 1 is virtually reached, while delay balancing is achieved for the utilized actuators. As no ramp meter is applied at on-ramp ON4, there are no queues created there (Figure 41), and as a result there is no delay experienced by the drivers on this ramp.

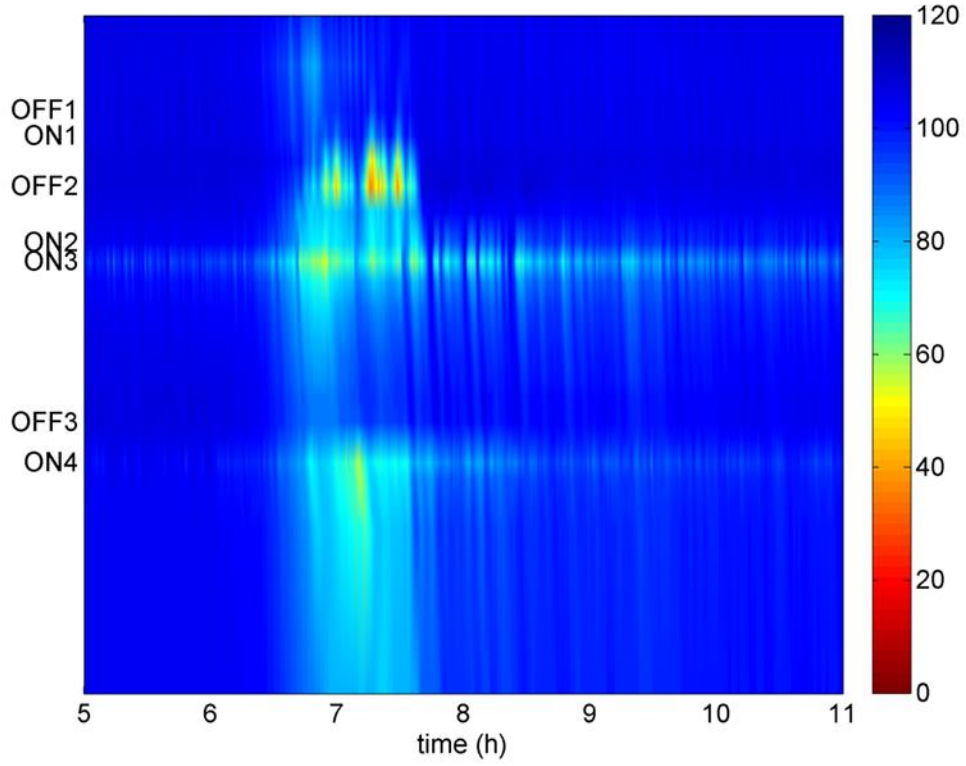

Figure 40: Speed (km/h) contour plot for Scenario 5.

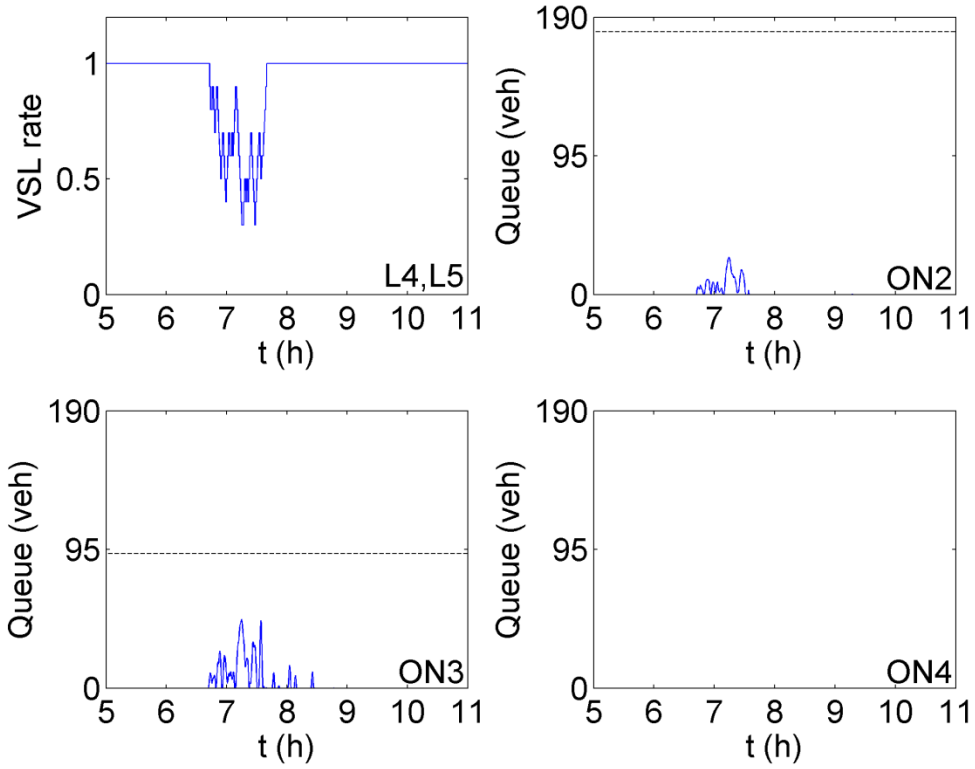

Figure 41: VSL rate and queue profiles for Scenario 5.

7. Conclusions and Perspectives

Chapter 7 summarizes the findings and conclusions of this thesis, together with some comments on potential future research. A summary of the overall conclusions is included in section 7.1, whereas in section 7.2 the main contributions of this thesis are referred. At the end, in section 7.3 some issues about potential future research are discussed.

7.1. Overall Conclusions

Feedback MTFC enabled via VSL for multiple bottlenecks on motorways has been proposed in this thesis. The evaluation of the proposed control strategy, using the METANET macroscopic traffic simulator, and its comparison with the optimal control approach, for a real network, demonstrates its efficiency. The feedback concept is approaching the performance of optimal control, is more robust (as no model or demand predictions are needed), and can be immediately implemented in the field as it considers practical and safety constraints.

A feedback-based integrated motorway traffic flow control concept for multiple bottlenecks is also proposed in this thesis. Integration is achieved subject to balancing of delays experienced by drivers. The suggested concept has been evaluated using the validated METANET macroscopic traffic flow simulator for a real infrastructure and has been compared to other control structures. The integrated controller presented in Scenario 5 is shown to be superior as it takes advantage of all the available storage capacity required for queueing upstream of the bottlenecks. The feedback controller is robust as there is no need, neither for any predictions of the demand nor for any model calibration or parameter identification. Practical and safety constraints have been considered, and, as a result, the concept is appropriate for field implementations.

7.2. Thesis Contribution

In the following bullets, the main contributions of this thesis are summarized.

- A new control concept for feedback MTFC enabled via VSL for multiple bottlenecks on motorways can be immediately implemented in the field, improving the current traffic conditions on motorways.

- A new control concept for feedback-based integrated motorway traffic flow control concept with delay balancing for multiple bottlenecks can be also immediately implemented in the field, ameliorating the present traffic conditions on motorways.

7.3. Potential Research

Future research activities could focus on further extensions of the proposed concepts at a network level so as to apply coordination between different integrated controllers. Safety impact of the control strategies under various traffic scenarios could also be considered using microscopic simulation. Another issue is the calibration of the traffic flow model using real VSL data. Unfortunately no current system is operating such a broad range of VSL rates [0.2,1] at the moment. VII systems are continuously becoming more and more part of the daily drivers' life worldwide. As a result it is almost compulsory to take them into consideration and incorporate them in the proposed control concepts. Thus, a good idea would be to use VII systems instead of VSL announced on VMS as an MTFC actuator. A very useful extension of the proposed integrated control concept is the possibility of having MTFC actuator between metered ramps. This will complicate the situation, as the queue created by an MTFC system may affect any upstream metered ramps used by the system, i.e., a different speed limit applied because of MTFC actions may lead to different critical points for the upstream merge areas and also the drivers leaving a ramp, will face two delays, the delay of the ramp metering actuator and the delay of the MTFC actuator which is not fair.

A very essential issue is the field implementation of the proposed control concepts but before this, it would be reasonable to apply the proposed control concepts to other networks and observe the results in simulation. This procedure will confirm the validity of the concepts and may help to the concepts extension.

References

[1] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 4, pp. 271–281, 2002.

[2] I. Papamichail, A. Kotsialos, I. Margonis, and M. Papageorgiou, "Coordinated ramp metering for freeway networks – A model-predictive hierarchical control approach," *Transp. Res. Part C Emerg. Technol.*, vol. 18, no. 3, pp. 311–331, 2010.

[3] Centre for Economics and Business Research, "The future economic and environmental costs of gridlock in 2030," *WHITE Pap. Cent. Econ. Bus. Res. London*, 2014.

[4] E. Smaragdis and M. Papageorgiou, "Series of new local ramp metering strategies," *Transp. Res. Rec.*, vol. 1856, pp. 74–86, 2003.

[5] M. Davarynejad, A. Hegyi, J. Vrancken, and van den B. Van, "Motorway ramp-metering control with queuing consideration using Q-learning," *14th Int. IEEE Conf. Intell. Transp. Syst. Washington, DC, USA*, pp. 1652–1658, 2011.

[6] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Local feedback-based mainstream traffic flow control on motorways using variable speed limits," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1261–1276, 2011.

[7] G. Shao-long, M. Jun, W. Jun-li, S. Xiao-qing, and L. Yan, "Methodology for variable speed limit activation in active traffic management," *13th COTA Int. Conf. Transp. Prof.*, vol. 96, pp. 2129–2137, 2013.

[8] A. Kotsialos and M. Papageorgiou, "Efficiency and equity properties of freeway network-wide ramp metering with AMOC," *Transp. Res. Part C Emerg. Technol.*, vol. 12, no. 6, pp. 401–420, 2004.

[9] M. J. Cassidy and J. Rudjanakanoknad, "Increasing the capacity of an isolated merge by metering its on-ramp," *Transp. Res. Part B Methodol.*, vol. 39, no. 10, pp. 896–913, Dec. 2005.

[10] M. Papageorgiou, E. Kosmatopoulos, and I. Papamichail, "Effects of variable speed limits on motorway traffic flow," *Transp. Res. Rec.*, vol. 2047, pp. 37–48, 2008.

[11] Y. Wang, M. Papageorgiou, G. Sarros, and W. Knibbe, "Feedback route guidance applied to a large-scale express ring-road," *Transp. Res. Rec.*, vol.

1965, pp. 79–88, 2006.

[12] "European Transport Safety Council, 10th Annual Road Safety Performance Index (PIN) Report," 2016.

[13] Z. Wang and C. M. Walton, "An investigation on the environmental benefits of a variable speed control strategy," *Res. Rep. SWUTC/06/473700-00072-1*, 2006.

[14] P. Mitchell, "Speed and road traffic noise," *A Rep. Comm. by UK Noise Assoc.*, 2009.

[15] M. Papageorgiou, E. Kosmatopoulos, M. Protopappas, and I. Papamichail, "Effects of variable speed limits (VSL) on motorway traffic," *Intern. Rep. 2006-25. Dyn. Syst. Simul. Lab. Tech. Univ. Crete, Chania, Greece*, 2006.

[16] A. Srivastava and N. Geroliminis, "Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model," *Transp. Res. Part C*, vol. 30, pp. 161–177, 2013.

[17] E. D. Arnold, "Ramp metering: a review of the literature," *Tech. Assist. Rep. Virginia Transp. Res. Counc.*, 1998.

[18] M. Papageorgiou, J. M. Blosseville, and H. Haj-Salem, "Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris – Part I: Modelling," *Transp. Res. Part A, Gen.*, vol. 24, no. 5, pp. 345–359, 1990.

[19] M. Papageorgiou and R. Mayr, "Optimal decomposition methods applied to motorway traffic control," *Int. J. Control*, vol. 35, no. 2, pp. 269–280, 1982.

[20] A. Kotsialos, M. Papageorgiou, M. Magneas, and H. Haj-Salem, "Coordinated and integrated control of motorway networks via non-linear optimal control," *Transp. Res. Part C*, vol. 10, no. 1, pp. 65–84, 2002.

[21] G. Gomes and R. Horowitz, "Optimal freeway ramp metering using the asymmetric cell transmission model," *Transp. Res. Part C Emerg. Technol.*, vol. 14, no. 4, pp. 244–262, 2006.

[22] M. Papageorgiou and M. Marinaki, "A feasible direction algorithm for the numerical solution of optimal control problems," *Intern. Rep. 1995-4, Dyn. Syst. Simul. Lab. Tech. Univ. Crete , Chania, Greece*, 1995.

[23] J. Hourdakis and P. G. Michalopoulos, "Evaluation of ramp control effectiveness in two twin cities freeways," *Transp. Res. Board 81st Annu. Meet. Washington,D.C.,USA*, 2002.

[24] A. Hegyi, S. P. Hoogendoorn, M. Schreuder, H. Stoelhorst, and F. Viti, "SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory," *2008 11th Int. IEEE Conf. Intell. Transp. Syst.*, pp. 827–832, Oct. 2008.

[25] A. Hegyi and S. P. Hoogendoorn, "Dynamic speed limit control to resolve shock waves on freeways-Field test results of the SPECIALIST algorithm," *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst., Funchal, Port.*, pp. 519–524, 2010.

[26] Y. Wang, M. Papageorgiou, G. Sarros, and W. J. Knibbe, "Real time route guidance for large-scale express ring-roads," *Intell. Transp. Syst. Conf.*, pp. 224–229, 2006.

[27] P. Papadimitratos, A. de la Fortelle, K. Evenssen, R. Brignolo, and S. Cosenza, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 84–95, 2009.

[28] J. A. Wattleworth, "Peak-period analysis and control of a freeway system," *Highw. Res. Rec.*, vol. 157, pp. 1–21, 1965.

[29] L. S. Yuan and J. B. Kreer, "Adjustment of freeway ramp metering rates to balance entrance ramp queues," *Transp. rRsearch*, vol. 5, pp. 127–133, 1971.

[30] D. Tabac, "A linear programming model of highway traffic control," *Proc. 6th Annu. Princet. Conf. Inf. Sci. Syst. Princeton, NJ*, pp. 568–570, 1972.

[31] C. F. Wang, "On a ramp-flow assignment problem," *Transp. Sci.*, vol. 6, pp. 114–130, 1972.

[32] M. Papageorgiou, "A new approach to time-of-day control based on a dynamic freeway traffic model," *Transp. Res.*, vol. 14B, pp. 349–360, 1980.

[33] D. P. Masher, D. W. Ross, P. J. Wong, P. L. Tuan, P. L. Zeidler, and S. Peracek, "Guidelines for design and operating of ramp control systems," *SRI, Menid Park. CA, Stanford Res. Inst. Rep. NCHRP 3-22, SRI Proj. 3340*, 1975.

[34] M. Papageorgiou, H. Haj-Salem, and J. M. Blosseville, "ALINEA: A local feedback control law for on-ramp metering," *Transp. Res. Rec.*, vol. 1320, pp. 58–64, 1991.

[35] M. Papageorgiou, H. Hadj-Salem, and F. Middelham, "ALINEA local ramp metering: Summary of field results," *Transp. Res. Rec.*, vol. 1603, pp. 90–98, 1998.

[36] M. Papageorgiou, J.-M. Blosseville, and H. Hadj-Salem, "Modeling and realtime controlof traffic flow on the southern part of Boulevard Peripherique in Paris-Part II: Coordinated on-ramp metering," *Transp. Res.*, vol. 24A, pp. 361–370, 1990.

[37] L. S. Yuan and J. B. Kreer, "An optimal control algorithm for ramp metering of urban freeways," *Proc. 6th IEEE Annu. Allert. Conf. Circuit Syst. Theory. Allerton, Illinois*, 1968.

[38] A. Kaya, "Computer and optimization techniques for efficient utilization of urban freeway systems," *Proc. 5th IFAC World Congr. Paris, Fr.*, vol. 12, no. 1, 1972.

[39] M. Athans, P. K. Houpt, D. Looze, D. Orlhac, S. B. Gershwin, and J. L. Speyer, "Stochastic control of freeway corridor systems," *Proc. 1975 IEEE Conf. Decis. Control*, pp. 676–685, 1975.

[40] L. Benmohamed and S. M. Meerkov, "Feedback control of highway congestion by a fair on-ramp metering," *Proc. 33rd IEEE Conf. Decis. Control. lake Buena Vista, FL*, vol. 3, pp. 2437–2442, 1994.

[41] C. Diakaki and M. Papageorgiou, "Design and simulation test of integrated corridor control for M8 eastbound corridor in Glasgow," *Dyn. Syst. Simul. Lab., Tech. Univ. Crete, Chania, Greece, Intern. Rep. 1995-3*, 1995.

[42] H. Haj-Salem and M. Papageorgiou, "Ramp metering impact on urban corridor traffic: field results," *Transp. Res.*, vol. 29A, pp. 303–319, 1995.

[43] M. Papageorgiou, "Applications of automatic control concepts to traffic flow modeling and control," *Berlin, Ger. Springer-Verlag*, 1983.

[44] N. Bhouri, M. Papageorgiou, and J. M. Blosseville, "Optimal control of traffic flow on periurban ringways with application to the Boulevard Peripherique in

Paris," *Prepr. 11th IFAC World Congr. Tallinn, Est.*, vol. 10, pp. 236–243, 1990.

[45] A. Kotsialos, M. Papageorgiou, and A. Messmer, "Optimal coordinated and integrated motorway network traffic control," *14th Int. Symp. Transp. Traffic Theory, Jerusalem, Isr.*, pp. 621–644, 1999.

[46] R. C. Carlson, D. Manolis, and I. Papamichail, "Integrated ramp metering and mainstream traffic flow control on freeways using variable speed limits," vol. 00, no. 2011, 2012.

[47] M. Papageorgiou, J. M. Blosseville, and H. Hadj-Salem, "La fluidification des rocades de l'Ile de France: Un project d'importance," *Dyn. Syst. Simul. Lab., Tech. Univ. Crete, Chania, Greece, Intern. Rep. 1998-17*, 1998.

[48] T. McLean, C. Brader, S. Hangleiter, M. Tsavachidis, C. Damas, B. Maxwell, and P. Barber, "Urban integrated traffic control evaluation results," *Deliv. 8.3, Eur. Transp. Telemat. Proj. TABASCO, Brussels, Belgium*, 1998.

[49] H. Greenberg and A. Daou, "The control of traffic flow to increase flow," *Oper. Res.*, vol. 8, no. 4, pp. 524–532, 1960.

[50] L. Kianfar, P. Edara, and C. Sun, "Operational analysis of a freeway variable speed limit system in St. Louis, Missouri," *J. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 385–398, 2015.

[51] J. Geistefeldt, "Capacity effects of variable speed limits on German freeways," *Procedia Soc. Behav. Sci.*, vol. 16, pp. 48–56, 2011.

[52] M. Robinson, "Examples of variable speed limits applications," 2000.

[53] M. Mirshahi, J. Obenberger, C. A. Fuhs, C. E. Howard, R. A. Krammes, B. T. Kuhn, R. M. Mayhew, M. A. Moore, K. Sahebjam, and C. J. Stone, "Active traffic management: The next step in congestion management," *Tech. Rep. FHWA-PL-07-012,Federal Highw. Adm.*, 2007.

[54] B. Liu, D. Ghosal, C. N. Chuah, and H. M. Zhang, "Reducing greenhouse effects via fuel consumption-aware variable speed limit (FC-VSL)," *IEEE Trans. Veh. Technol.*, vol. 61, no. 1, pp. 111–122, 2012.

[55] S. K. Zegeye, B. De Schutter, J. Hellendoorn, and E. A. Breunesse, "Variable speed limits for green mobility," *Proc. 14th Int. IEEE Conf. Intell. Transp.*

Syst. (ITSC 2011), Washington, DC, pp. 2174–2179, 2011.

[56] H. Zackor, "Speed limitation on freeways: traffic-responsive strategies," *Concise Encycl. traffic tansportation Syst. (M. Papageorgiou, ed.), Pergamon Press. Oxford, United Kingdom*, pp. 507–511, 1991.

[57] M. Cremer, "Der verkehrsflub auf schnellstrassen," *Springer Verlag, Berlin, Ger.*, 1979.

[58] S. Smulders, "Control of freeway traffic flow by variable speed signs," *Transp. Res.*, vol. 24B, pp. 111–132, 1990.

[59] A. Hegyi, "Model predictive control for integrating traffic control measures," *Ph.D Diss. Delft Univ. Technol., Delft, Netherlands*, 2004.

[60] B. G. Heydecker and J. D. Addison, "Analysis and modelling of traffic flow under variable speed limits," *Transp. Res. Part C Emerg. Technol.*, vol. 19, no. 2, pp. 206–217, 2011.

[61] V. L. Knoop, A. Duret, C. Buisson, and B. van Arem, "Lane distribution of traffic near merging zones: influence of variable speed limits," *13th Int. IEEE Conf. Intell. Transp. Syst. Funchal, Madeira Island, Port.*, pp. 485–490, 2010.

[62] W. Liu, Y. Yin, and H. Yang, "Effectiveness of variable speed limits considering commuters' long-term response," *Transp. Res. Part B Methodol.*, vol. 81, no. 2, pp. 498–519, 2015.

[63] S. Xiaotian and R. Horowitz, "A localized switching ramp-metering controller with a queue regulator for congested freeways," *Proc. 2005 , Am. Control Conf. , Portl. , OR, USA*, pp. 2141–2146, 2005.

[64] R. Jiang, E. Chung, and B. J. Lee, "Local on-ramp queue management strategy with mainline speed recovery," *8th Int. Conf. Traffic Transp. Stud. Chang. China*, vol. 43, pp. 201–209, 2012.

[65] J. P. Lebacque and H. Haj-Salem, "Ramp metering, speed management, and braess-like paradoxes," *Proc. 5th Trienn. Symp. Transp. Anal. Le Gosier, Guadaloupe, French West Indies*, 2004.

[66] J. P. Lebacque and H. Haj-Salem, "Speed limit control: a problem formulation and theoretical discussions," *Proc. 4th Trienn. Symp. Transp. Anal. Sao Miguel, Azores*, pp. 421–426, 2001.

[67] I. Papamichail, K. Kampitaki, M. Papageorgiou, and A. Messmer, "Integrated ramp metering and variable speed limit control of motorway traffic flow," *Proc 17th IFAC World Congr., Seoul, Korea*, p. 14 084–14089, 2008.

[68] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal motorway traffic flow control involving variable speed limits and ramp metering," *Transp. Sci.*, vol. 44, no. 2, pp. 238–253, 2010.

[69] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, "Optimal mainstream traffic flow control of large-scale motorway networks," *Transp. Res. Part C Emerg. Technol.*, vol. 18, no. 2, pp. 193–212, 2010.

[70] M. Papageorgiou, I. Papamichail, A. Spiliopoulou, and A. Lentzakis, "Real-time merging traffic control with applications to toll plaza and work zone management," *Transp. Res. Part C Emerg. Technol.*, vol. 16, no. 5, pp. 535–553, 2008.

[71] F. L. Hall, "Chapter 2: Traffic stream characteristics," *Traffic Flow Theory State-of-the-Art Report, N. Gartner, C.J. Messmer, A.K. Rathi, Eds. Washington, DC FHWA/TRB/ORNL*, 2001.

[72] J. Zhang, H. Chang, and P. A. Ioannou, "A simple roadway control system for freeway traffic," *Proc. Amer. Control Conf., Minneapolis, MN*, pp. 4900–4905, 2006.

[73] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Integrated feedback ramp metering and mainstream traffic flow control on motorways using variable speed limits," *Transp. Res. Part C*, vol. 46, pp. 209–221, 2014.

[74] A. Messmer and M. Papageorgiou, "METANET: A macroscopic simulation program for motorway networks," *Traffic Eng. Control*, vol. 31, no. 8, pp. 466–470, 1990.

[75] 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 Trans. Intell. Transp. Syst.*, vol. 3, pp. 282–292, 2002.

[76] K. Kampitaki, "Integrated control of traffic flow using ramp metering and variable speed limits," *M.S. thesis, Dept. Prod. Eng. Manag. Tech. Univ. Crete, Chania, Greece*, 2008.

[77] A. Alessandri, A. Di Febbraro, A. Ferrara, and E. Punta, "Optimal control on freeways via speed signalling and ramp metering," *Control Eng. Pract.*, vol. 6, no. 6, pp. 771–780, 1998.

[78] "European Comission. Roadmap to a single european transport area - towards a competitive and resource efficient transport system," *WHITE Pap. Eur. Com. Brussels, Belgium*, 2011.

[79] M. Papageorgiou, "Optimierung," *Oldenbourg,Munchen, 2nd Ed.*, 1996.

[80] M. Cremer and M. Papageorgiou, "Parameter identification for a traffic flow model," *Automatica*, vol. 17, pp. 282–292, 1981.

[81] J. Spiliopoulou, A., Papamichail, I., Papageorgiou, M., Tyrinopoulos, Y., Chrysoulakis, "Macroscopic traffic flow calibration using different optimization algorithms," *Proc. Int. Symp. Transp. Simulation, Ajaccio, Corsica, Fr. June 1−4*, 2014.

[82] J. Spiliopoulou, A., Papamichail, I., Papageorgiou, & Chrysoulakis, "A Calibration tool for macroscopic traffic flow models," *Proc. 3rd Int. Symp. 25th Natl. Conf. Oper. Res. Volos, Greece, June 26−28*, pp. 215–225, 2014.

[83] D. C. Gazis and R. S. Foote, "Surveillance and control of tunnel traffic by an on-line digital computer," *Transp. Sci.*, vol. 3, no. 3, pp. 255–275, 1969.

[84] K. W. Crowley and H. Greenberg, "Holland tunnel study aids efficient increase of tube's use," *Traffic Eng.*, vol. 35, pp. 20–41, 1965.

[85] R. S. Foote and K. W. Crowley, "Developing density controls for improved traffic operations," *Highw. Res. Board Rec.*, vol. 154, no. 1, pp. 24–37, 1967.

[86] R. S. Foote, "Instruments for road transportation," *High Speed Gr. Transp. J.*, vol. 18, no. 15, pp. 260–281, 1968.

[87] M. S. McCalden and Jr, "A traffic management system for the San Fransisco-Oakland bay bridge," *ITE J*, vol. 54, no. 5, pp. 46–51, 1984.

[88] L. Chen, A. M. May, and D. M. Auslander, "Freeway ramp control using fuzzy set theory for inexact reasoning," *Transp. Res. Part A*, vol. 24, no. 1, pp. 15–25, 1990.

[89] A. Spiliopoulou, I. Papamichail, and M. Papageorgiou, "Toll plaza merging

trafffic control for throughput maximization," *J. Transp. Eng.*, vol. 136, no. 1, pp. 67–76, 2010.

[90] E. Jacobson and J. Landsman, "Case studies of US freeway-to-freeway ramp and mainline metering and suggested policies for Washington State," *Transp. Res. Rec.*, vol. 1446, pp. 48–55, 1994.

[91] K. A. Haboian, "A case for freeway mainline metering," *Transp. Res. Rec.*, vol. 1494, pp. 11–20, 1995.

[92] Y. H. Chiang and J. C. Juang, "Control of freeway traffic flow in unstable phase by H-infinity theory," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 193–208, 2008.

[93] X. Y. Lu, P. Varaiya, R. Horowitz, D. Su, and S. E. Shladover, "A new approach for combined freeway Variable Speed Limits and Coordinated Ramp Metering," *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst., Funchal, Port.*, pp. 491–498, 2010.

[94] Y. Wang, M. Papageorgiou, G. John, I. Papamichail, R. Geoffrey, and Y. William, "Local ramp metering in random-location bottlenecks downstream of metered on-ramp," *Transp. Res. Rec.*, no. 2178, pp. 90–100, 2010.

[95] Y. Murashige, "Countermeasure strategies against traffic congestion on motorways in Japan.," *Procedia Soc. Behav. Sci. 16*, pp. 110–119, 2011.

[96] G. Ellis, "Control System Design Guide- A Practical Guide," *Amsterdam. The Netherlands: Elsevier*, 2004.

[97] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *J. Process Control*, vol. 13, no. 4, pp. 291–309, 2003.

[98] A. Karimi, A. Hegyi, B. De Schutter, J. Hellendoorn, and F. Middelham, "Integrated model predictive control of dynamic route guidance information systems and ramp metering," *7th Int. IEEE Conf. Intell. Transp. Syst. 2004. IEEE, Washington, DC, USA*, pp. 491–496, 2004.

[99] A. Hegyi, B. De Schutter, and H. Hellendoorn, "Model predictive control for optimal coordination of ramp metering and variable speed limits," *Transp. Res. Part C Emerg. Technol.*, vol. 13, no. 3, pp. 185–209, 2005.

[100] A. Muralidharan and R. Horowitz, "Computationally efficient model predictive

control of freeway networks," *Transp. Res. C, Emerg. Tchnol.*, vol. 58, pp. 532–553, 2015.

[101] N. Mahajan, A. Hegyi, G. S. van de Weg, and S. P. Hoogendoorn, "Integrated variable speed limit and ramp metering control against jam waves - a COSCAL v2 based approach," *IEEE 18th Int. Conf. Intell. Transp. Syst.*, 2015.

[102] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang, "Review of road traffic control strategies," *Proc. IEEE*, vol. 91, no. 12, pp. 2043–2067, 2003.

[103] I. Papamichail and M. Papageorgiou, "Balancing of Queues or Waiting Times on Metered Dual-Branch On-Ramps," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 438–452, Jun. 2011.

[104] Y. Wang, E. Kosmatopoulos, M. Papageorgiou, and I. Papamichail, "Local ramp metering in the presense of a distant downstream bottleneck: Theoretical analysis and simulation study," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, pp. 2024–2039, 2014.

[105] Y. Kan, Y. Wang, M. Papageorgiou, and I. Papamichail, "Local ramp metering with distant downstream bottlenecks: A comparative study," *Transp. Res. C, Emerg. Technol*, vol. 62, pp. 149–170, 2016.

[106] R. C. Carlson, A. Ragias, I. Papamichail, and M. Papageorgiou, "Mainstream traffic flow control of merging motorways using variable speed limits," *2011 19th Mediterr. Conf. Control Autom.*, pp. 674–681, 2011.

[107] G. Vigos, M. Papageorgiou, and Y. Wang, "Real-time estimation of vehicle count within signalized links," *Transp. Res. C, Emerg. Technol.*, vol. 16, pp. 18–35, 2008.

[108] R. C. Carlson, I. Papamichail, and M. Papageorgiou, "Comparison of local feedback controllers for the mainstream traffic flow on freeways using variable speed limits," *J. Intell. Transp. Syst.*, vol. 17, pp. 268–281, 2013.

[109] P. Brucker, "An O(n) algorithm for quadratic knapsack problems," *Oper. Res. Lett.*, vol. 3, no. 3, pp. 163–166, 1984.

[110] H. Agency, "Motorway Incident Detection and Automatic Signalling (MIDAS) Design Standard," *1st ed., available online www.midas-data.org.uk.Bristol, UK*, 2007.