PhD Thesis

Active Traffic Management Strategies In Metropolitan Freeways: Modeling and Empirical Assessment of Dynamic Speeds Limits

Author:

Josep Maria Torné

PhD Supervisors:

Dr. Francesc Soriguera
Dr. Francesc Robusté

PhD Program in Transport Infrastructures and Engineering
Civil Engineering School of Barcelona
Technical University of Catalonia– UPC – Barcelona**Tech**

Barcelona, October 2013

Fortunate is he, who is able to know the causes of things

Publius Vergilius Maro

ACTIVE TRAFFIC MANAGEMENT IN METROPOLITAN FREEWAYS:
MODELING AND ASSESSING DYNAMIC SPEED LIMIT STRATEGIES
Josep Maria Torné

Abstract

Traffic congestion is a major source of inefficiencies in metropolitan freeways (e.g. productivity loss, energy waste, pollutant emissions, among other externalities). Active traffic management (ATM) strategies are used in order to face the problem and alleviate its effects by managing recurrent and non-recurrent congestion with a combination of real-time and predictive operational strategies.

Dynamic speed limit management (DSL) is an ATM tool believed to bring some benefits to freeway traffic, like the increase in capacity and/or the vehicles' speed homogenization. Currently it is being implemented in many metropolitan freeways worldwide. This thesis is devoted to assess such conceptual benefits with empirical data and robust traffic flow models.

Empirical data are obtained from the C-32 freeway in Barcelona, the first DSL implementation in Spain. The detailed analysis of these data allows characterizing the effects of DSL management on traffic behavior. It is found that, under the right conditions, DSL can increase capacity due to a reduction in the speed dispersion. This finding allows defining new ATM policies.

A coordinated strategy for traffic management considering both ramp metering and DSL is proposed in order to reduce the capacity drop phenomenon in the vicinity of a freeway on-ramp. The effectiveness of this control method is tested using simulation with an extension of the traditional cell transmission model, which incorporates the ability to reproduce DSL, together with the capacity drop. Results show amelioration in the performance indicators of the system, highlighting the equity-friendly component of this coordinated strategy.

Finally, results from a cost-benefit assessment of DSL considering the main externalities suggest that the social profitability of DSL management in metropolitan freeways is limited when applied alone. The potential synergies of applying a "pack" of different ATM strategies in a coordinated way, define challenging issues for further research.

Key Words: *Dynamic speed limits, ramp metering, capacity drop, freeway merge modeling, traffic environmental assessment, traffic safety, traffic fundamental diagram calibration, cell transmission model, freeway traffic..*

Dr. Francesc Soriguera   Dr. Francesc Robusté

Acknowledgements

Fer un doctorat deixa petjada, almenys en el meu cas. A mi m'ha permès créixer tant professional com personalment. El camí del doctorat és un camí ple de reptes i de dificultats. Afrontar-les i tractar de superar-les és una escola que no té preu, o millor dit, potser el té massa elevat. Malgrat tot, crec que ha valgut la pena En aquestes línies intentaré compensar d'alguna manera l'ajut que tantes persones m'han lliurat al llarg d'aquests anys de doctorat. Espero no deixar-me a ningú.

És indubtable que aquesta tesi no hagués arribat a bon terme sense la incansable ajuda dels meus directors de tesi: en Francesc Soriguera i en Francesc Robusté. Cadascú en una vessant diferent, han estat claus en aquest procés. Us agraeixo les hores dedicades a ajudar-me a fer una bona feina, tot i que de vegades voldries deixar-ho estar o conformar-te amb menys. Crec que aquesta és una de les millors lliçons que em porto. Com deia un conegut eslògan publicitari: *La feina ben feta no té fronteres.* Així doncs, gràcies per haver-me ensenyat, guiat i modelat per tal d'intentar fer recerca de qualitat, amb el caire de repte (i d'explorador solitari, no ens enganyem...) que aquesta tasca comporta tantes vegades.

Molt especialment, voldria agrair el suport que m'han lliurat els meus pares. Els hi dec moltes coses i també una part de la tesi. El seu gran bagatge científic (i que tot just ara arriba a la seva fi...) ha estat una ajuda inestimable durant aquests darrers anys. Gràcies pels vostres consells i ànims, crec que sóc un gran afortunat per tenir-vos! També vull donar les gràcies a les meves germanes Anna, Núria i Maria, al meu germà Pau i als meus cunyats i familiars per haver-me recolzat sempre, tot fent pinya.

Vull donar les gràcies a molts amics que m'han donat suport al llarg d'aquests anys: a en Jaume que em va engrescar amb el doctorat i a qui li dec tants bons consells; a en Vicente, Lluís, Ferran, Miguel A. i tants d'altres amics de Tarragona (*vamos!*), L'Hospitalet i Pearson pels ànims que sempre m'heu transmès i perquè heu hagut d'aguantar-me en els moments de més neguit, ben propis d'aquests darrers mesos; a en Thomas i l'Eduard per

haver pogut compartir la il·lusió (i la incertesa) de la recerca i, sobretot, per la vostra amistat; a en Miguel per haver-me regalat una guspira del seu gran talent artístic en forma de portada de tesi i *last but not least*, als amics de Suïssa (Jean David, Edgar, Nerio, David, Toni Z...) que tant m'han ensenyat i ajudat, *merci à tous!* Tampoc vull oblidar-me de tants amics companys de les sortides a la muntanya, en bicicleta, prenent un beure... que m'he anat trobant pel camí durant tots aquests anys, ja em perdonareu que no us mencioni individualment, però us tinc a tots presents.

Gràcies a tots els que heu passat pel CENIT i hem compartit tantes jornades: Dulce, Alberto, Javi, Quique, Marta, Xisco, Carles, Albert, Mireia, Miquel, Leif, Magín, Sergi, Pau, Óscar, Deme, Nati, Pilar, Anna, Patty...

Thanks to Nikolas Gerolimins, to give me the opportunity to work with him during my stage in the École Polytechnique Fédérale de Lausanne, mainly for transmitting me his enthusiasm for the research, for the support you have provided me during different stages of the thesis and, last but not least, for the good time spend together. You have played a crucial role in my doctorate. I also want to thank all the members of LUTS (Jack, Mohsen, Mehmed, Nan, Yuxuan, Burak and Christine) for your hospitality during my stay, your unselfish assistance and the good times we have shared together (what amazing football matches played!). Thanks for those unforgettable months.

To Thomas Schreiter from the Delft University of Technology, who provided me helpful references and unselfish assistance for the modeling issue and to Mónica Menéndez, for your hospitality during my visit to ETH in Zurich and your encouraging comments about my research.

Vull agraïr al Servei Català del Trànsit, i en particular a l'Òscar Llatje i en Jordi Galindo: sense les dades que em vàreu facilitar aquesta tesi hagués quedat coixa.

Molt especialment, vull donar les gràcies a la Universitat Politècnica de Catalunya i en particular, al personal administratiu (Sandra, Mónica, Sílvia i Lorena) pel vostre recolzament al llarg d'aquest procés.

Finalmente, también quiero agradecer al Ministerio de Economía y Competitividad y a la Cátedra abertis que me haya facilitado el apoyo económico necesario para desarrollar esta tesis.

To all of you, thank you very much.

Table of contents

List of figures

List of tables

Part I

Introduction and objectives

CHAPTER 1

Thesis overview

Traffic monitoring and control systems have benefited from intensive research and development during the last decades. Considerable advances in real-time traffic detection, data processing and control algorithms have enabled a deeper understanding of vehicular traffic, as well as novel ways to manage transportation facilities. Traffic management, based on the spontaneous infrastructure utilization, is approaching its end. Unfortunately, drivers still show significant levels of resistance against traffic control innovations. This will evolve, like in many other aspects of life, where society shows an extraordinary commitment in order to improve global efficiency.

In this context, different Active Traffic Management (ATM) strategies appear. The objective is to dynamically manage and control both traffic demand and available capacity in order to maximize efficiency indicators. The election of such indicators depends on the pursued objectives, but generally includes:

- **Environmental friendliness**. There is a great concern about the environmental impact of traffic, as reflected by the fact that cars around the world will become appreciably cleaner in the years to come. Low fuel consumptions, low air and noise pollution are the goals. In large urban areas the environmental effects of traffic are considered critical. Actually, air pollutants in many cities repeatedly reach concentrations far above acceptable limits. Hence, governments are urged to take the necessary steps to improve air quality, and traffic is the main source of air pollution and greenhouse gas emissions [Keuken, et al., 2005].

- **Operational efficiency**. Society needs to get the most from infrastructure investments. So, efficient operation is a requirement. However, the objectives differ depending on the perspective considered. Agencies look for the network optimum, minimizing the total travel time. Drivers pursue the user optimum, in order to minimize their individual travel time. Information is needed in order to fit together both perspectives and achieve drivers' commitment for the overall benefit.

- **A satisfactory level of safety**. Traffic accidents are a scourge for modern societies. Huge efforts are made in order to reduce the number of fatalities and injuries on the roads. Any traffic management strategy has to bear that in mind, and promote safe operations.

- **Network reliability.** Predictable travel times and precise arrival time estimations are appreciated by road users as they play a fundamental role in the traveler planning process, e.g. make departure time choices easier. In addition, network unreliability implies significant scheduling costs. This means that an efficient transportation system needs to be reliable. Since a substantial part of road network unreliability is caused by accidents, safer traffic systems are also translated into more reliable networks.

This thesis quantifies the effects on these objectives resulting from ATM strategies on freeways. Several ATM strategies have recently addressed freeway traffic sustainability (i.e. saving time and energy, decreasing accident risk and diminishing pollutant emissions). The common approach involves control strategies over the traffic stream by using variable traffic signs (e.g. ramp control, dynamic route guidance or dynamic speed limits - DSL). Among them, DSL is one of the most attractive policies, because of its apparent simplicity and the relative reduced implementation costs. Many worldwide transportation authorities have recently spend large amounts of money in extensive installation of dynamic speed limit systems. Unfortunately, the investment in research has not been proportional. In addition, drivers fail to see the usefulness of the DSL due to inconsistent or incomplete communication regarding its value and functionality. This discourages drivers from strict speed limit compliance.

The scope of the thesis includes, but is not limited to, the dynamic speed limit strategies. In many cases several ATM strategies are complementary and should be combined in order to provide greater benefits. Given the predominance of the DSL systems and the data availability, focusing on such strategy is a reasonable way to address the ATM strategies.

1.1 Research objectives

The present thesis uses the vast amount of freeway traffic data available at traffic management centers in order to improve our knowledge regarding DSL strategies. This is achieved by using data measured on a metropolitan freeway in the vicinity of Barcelona (Spain). The C32 is one of the most demanded freeways towards Barcelona from the South

(Figure 1.1). It links the city to its airport and with many satellite towns. The selected stretch is 14.5 km long, mainly with 3-lanes, equipped with loop detectors every 500 m, measuring the average speed, flow and occupancy. In addition, 10 Variable Message Signs (VMS) displaying dynamic speed limits are installed. Recurrent congestion episodes grow in the morning and evening rush periods.

Figure 1.1 *Test site location.*
Source: © OpenStreetMap contributors, CC BY-SA

Two main research directions are addressed in the thesis: a modeling and an empirical approach. Each one constitutes a differentiated part of the thesis dissertation. The modeling approach is the most common way to deal with active traffic management strategies. This is mainly due to the reticence of many traffic authorities to improve their traffic monitoring proceedings, and test ATM strategies on real highway labs. The popularity of traffic flow models (either microscopic or macroscopic), capable of incorporating these strategies also favors the modeling approach. For instance, most of DSL research is based on a particular second order model, i.e. METANET [Messmer **and Papageorgiou**, 1990] or on its further extensions. However, almost any first order model has provided neither a trustful approach to the DSL modeling nor the capacity drop phenomenon reproduction. In order to fill this gap, a simple approach, fully consistent with the Lighthill-Whitham [Lighthill and Whitham, 1955] and Richards [Richards, 1956], LWR model, is presented (see Chapter 4).

In another order of things, the modification of traditional macroscopic models in order to reproduce active control strategies has always implied the introduction of new parameters. The simplistic (but usual) calibration of these parameters may lead to unfeasible merging behaviors when uncontrolled merges are modeled. This is the issue addressed in Chapter 3, where a simple calibration procedure is proposed in order to recover the full consistency of the modeling of the merging mechanism. This plays a crucial role for merging bottlenecks.

The empirical part of the thesis aims to characterize the drivers' behavior when facing different speed limits. The effect of this behavior on stationary traffic states, via the fundamental diagrams representation, constitutes one of the contributions of this empirical analysis (see Chapter 5).

Real empirical data also allows a quantitative analysis in terms of congestion reduction, pollutant emissions and traffic safety on a particular freeway corridor (see Chapter 6). Thus, the social profitability of the policy can be assessed. This assessment also compares empirical with simulation results, validating the proposed traffic flow model.

The thesis has a clear practical application, as it directly emerged from the traffic administration intention of improving traffic management on DSL corridors. The thesis outcomes contribute in the efficiency improvement of freeway management, allows the definition of new traffic management strategies and facilitates the decision making process at traffic management centers.

1.2 Scientific contributions

This thesis contributes to science in the following ways:

- *A novel methodology for calibrating the two-parameters involved in the Asymmetric Cell Transmission Model (ACTM) merge model.* [Gomes and Horowitz, 2006] The standard simplistic parameter calibration fails on the attempt to provide an accurate prediction of the queue length in an uncontrolled merge junction. The proposed method recovers the physical consistency of the dynamics involved. (Chapter 3)

- *Development of an innovative extension to the original Cell Transmission Model* [Daganzo, 1995] *capable of reproducing traffic phenomena not yet modeled with first order models.* Includes the ability to incorporate DSL strategies by introducing an inverse-lambda fundamental diagram together with an approach to the capacity drop effect. (Chapter 4)

- *An innovative coordinated ramp metering and DSL freeway traffic management strategy.* A strategy fully based in a novel model of endogenous merge capacity is proposed. The bottleneck discharge flow is maximized taking advantage of the merge capacity mechanism. Both mobility and equity indicators present significant improvements. (Chapter 4)

- *ATM strategies show synergistic relationships.* They should be implemented together whenever possible. (Chapter 2)

- *Insights into the effect of DSL strategies in aggregated traffic flow.* DSL strategy is capable of improving the freeway performance by increasing the queue discharge rates when suitable speed limits are set. Significant variations on the fundamental diagrams appear (e.g. lowering of the slope of the free-flow branch or the shift of the critical occupancy to higher values). (Chapter 5)

- *The application of DSL policies should not be indiscriminate but respond to a detailed analysis, as it may result in negative social profits.* Results are highly sensitive to the different monetary values given to each term of the objective function. An increase in the cost of traffic externalities or alternatively, the application of marginal costs to the short additional delays caused by the DSL control (instead of the higher average costs) may imply a switch from profitable to non-profitable strategy. (Chapter 5)

- *Macroscopic models based on the CTM are suitable for reproducing DSL strategies.* It is found that the simulation model accuracy when considering the average travel time in congested conditions is more than acceptable. (Chapter 5)

1.3 Publications from this thesis outline

The results of this thesis have already been published in a high impact SCI Journal:

- Soriguera, F., J.M. Torné and D. Rosas. (2013) Assessment of Dynamic Speed Limit Management on Metropolitan Freeways. *Journal of Intelligent Transportation Systems*, 17, (1), 78-90.

Accepted for presentation in the following international conferences:

- Torné, J.M., F. Soriguera and N. Geroliminis. (2014) Coordinated active traffic management freeway strategies using the capacity-lagged cell transmission model. *Proceedings of the 93rd Transportation Research Board Annual Meeting.* Washington, D. C.

- Torné, J.M. and F. Soriguera. (2014) On the Consistency of Freeway Macroscopic Merging Models. *Proceedings of the 93rd Transportation Research Board Annual Meeting.* Washington, D. C.

- Torné, J.M., F. Soriguera and N. Geroliminis. (2012) Modifications of the asymmetric cell transmission model for modeling variable speed limit strategies. *Proceedings of the 1st European Symposium on Quantitative Methods in Transportation Systems.* Lausanne, Switzerland.

- Torné, J.M., D. Rosas and F. Soriguera. (2011) Evaluation of speed limit management on C-32 highway access to Barcelona. *Proceedings of the 90th Transportation Research Board Annual Meeting.* Washington, D. C.

Whereas the following communications are submitted:

- Torné, J.M., F. Soriguera and D. Ramoneda. (2014) Empirical evidences of dynamic speed limits impact on a metropolitan motorway. *Submitted to the 18th Pan-American Conference of Traffic and Transportation Engineering and Logistics.*

1.4 Dissertation outline

In essence, the thesis is divided in four chapters, plus this introduction and the literature review. The relationship between the chapters and its main content is schematically outlined in Figure 1.2. The figure clearly shows that the research focuses on two main issues (traffic flow modeling –Chapters 3 and 4– and empirical assessment –Chapter 5 and 6–). The literature review (Chapter 2) sets a common framework for the rest of the chapters.
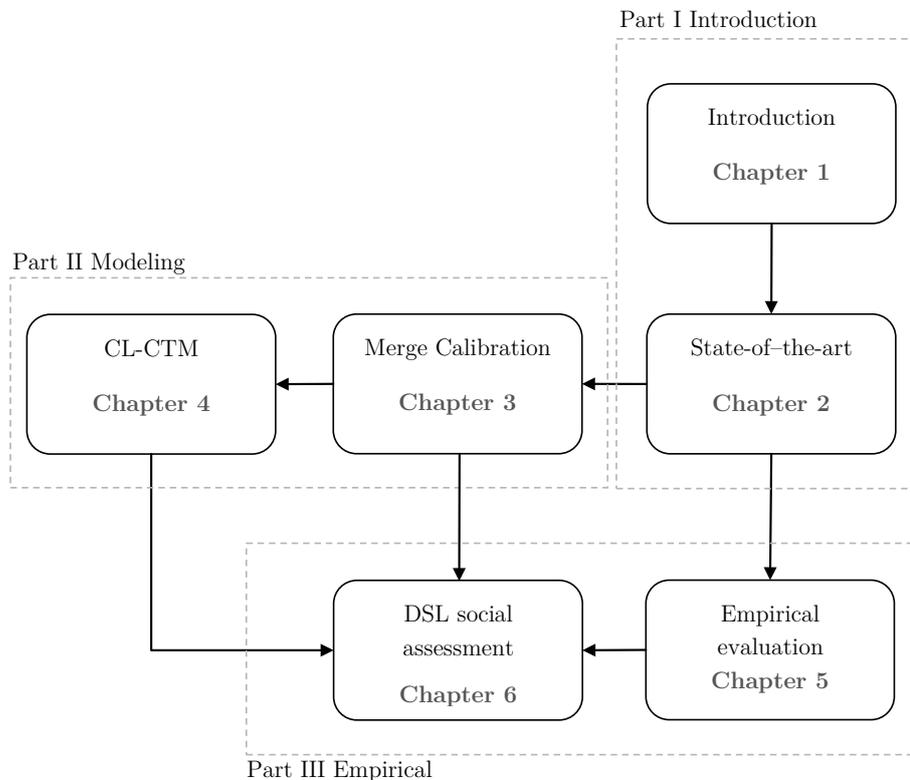
Figure 1.2 *Overview of the content of this dissertation.*

Specifically, Part I contains the introductory chapter (Chapter 1) and the state-of-the-art (Chapter 2), where the main ATM strategies are revisited. Common practices around the world are described and the main controversial topics involving these policies are sketched. Part II addresses the modeling issues. This includes Chapter 3 proposing a calibration method for the ACTM in non-metered merges, and Chapter 4 presenting a coordinated ATM strategy using the capacity-lagged CTM extension. The proposed strategy combines ramp metering together with DSL in order to reduce the capacity drop in a merging bottleneck. Part III presents the empirical approach. Chapter 5 provides an empirical evaluation of the DSL policy analyzing the aggregated traffic flow behavior from loop detector measurements. Chapter 6 presents a cost - benefit assessment where different concepts are quantified via an objective function which provides an estimation of the social profitability of the DSL policy. This chapter also shows a real application of the previous models. Finally, Chapter 7 provides some conclusions and suggests future works.

A state-of-the-art for active traffic management strategies

2.1 Introduction

An increasing consensus exists among practitioners that a good way to increase the efficiency (e.g. productivity, equity and sustainability) in road networks is to introduce some kind of control over the traffic flow by means of active management strategies. ATM is defined as a continuous process which dynamically manage recurrent and non-recurrent traffic congestion in real time (measuring and analyzing traffic data for implementing the most promising control strategies) in order to maximize the efficiency of highway facilities [Kurzhanskiy and Varaiya, 2010a].

There is a lack of scientific literature describing the main available ATM strategies in an integrated manner and from a general point of view. Some provide brief summaries of the most remarkable techniques, but without going inside the most controversial topics [Brinckerhoff, 2010; Lennie, et al., 2009]. While others include partial literature reviews in their publications' introduction [Carlson, et al., 2010; Papageorgiou, et al., 2008].

The present chapter tries to fill this gap by describing three main, mostly already applied, ATM strategies: two focused on the supply side (dynamic speed limits and managed lanes, Section 2 and 3) and one on the demand side (ramp metering, Section 4). Each strategy will be presented with a conceptual description, a scientific literature review and references to international experiences as well as their main results.

2.2 Dynamic speed limits

Speed limits are usually set to attempt to limit road traffic speed for several reasons (e.g. safety or pollution). When maximum allowed speed in a particular spot varies (manually or automatically) depending on traffic conditions or weather they are usually known as variable or dynamic speed limit (DSL) strategies.

Control by means of variable speed signs was first introduced more than three decades ago in Germany [Zackor, 1972] and in early 1980s in the Netherlands [Remeijn, 1982]. Because of its apparent simplicity and reduced implementation cost, nowadays, it is one of the most attractive policies and many cities around the world have introduced DSL systems, isolated or coordinated with other control measures. However, the effectiveness of the policy is still a controversial issue.

The claimed benefits of these actuations imply a reduction of traffic related emissions and accident rate, as well as an improvement in congestion reduction (e.g. reduction of stop&go traffic episodes or capacity increase). These claims results from studies mainly focused on control algorithms evaluated using second order traffic flow models [Carlson, et al., 2010; Hegyi, et al., 2005; Lu, et al., 2010]. However, few evaluations of the DSL with real data can be found in the literature. Some exceptions include [Hegyi and Hoogendoorn, 2010; Heydecker and Addison, 2011; Papageorgiou, et al., 2008; Soriguera, et al., 2013; Torné, et al., 2011; Torné, et al., 2013]. Real implementations of these DSL systems use somehow complex speed limit change algorithms, using different types of real time information. However, most of them are not sufficiently based on traffic flow behavior and lead to blind and inefficient algorithms.

2.2.1   The impact of DSL in aggregate traffic flow

The main claimed effect of DSL is the speed homogenization (i.e. speed variance reduction). This leads to positive impacts on traffic safety [Lee, et al., 2006], pollutant emissions [Baldasano, et al., 2010; Soriguera, et al., 2013; Stoelhorst, 2008; Torné, et al., 2011] and highways capacity [Papageorgiou, et al., 2008; Zackor and Papageorgiou, 1991]. Non-conclusive empirical evidences presented in the latter study seem to suggest a certain shift of the critical occupancy to higher values in the flow-occupancy diagram due to DSL strategy, and possibly caused by the speed harmonization. However, to the author's knowledge, there is no literature analyzing speed homogenization based on individual vehicle speed data. Despite that, from the traffic engineering point of view, the impact of DSL on aggregate traffic flow behavior is what matters, this empirical study may be necessary for attempting a conclusive cause and effect analysis of the expected reduction of stop&go episodes, lane-changing maneuvers and capacity increase due to DSL strategies.

In spite of this, different approaches have been proposed since the early investigation of [Zackor, 1972]. Most of them [Carlson, et al., 2010; Cremer, 1979; Hegyi, et al., 2005] are

based in few empirical evidences and not always tested with real data. Certain Fundamental Diagram (FD) shapes are suitable to justify an increase in capacity due to the stabilizing effect produced by the DSL when it is introduced in lower than critical densities (Figure 2.1(a) and (b)) [Cremer, 1979; Zackor and Papageorgiou, 1991]. Others are less optimistic in this point, though (Figure 2.1(c) and (d)) [Carlson, et al., 2010; Hegyi, et al., 2005].

(a)   (b)

(c)   (d)
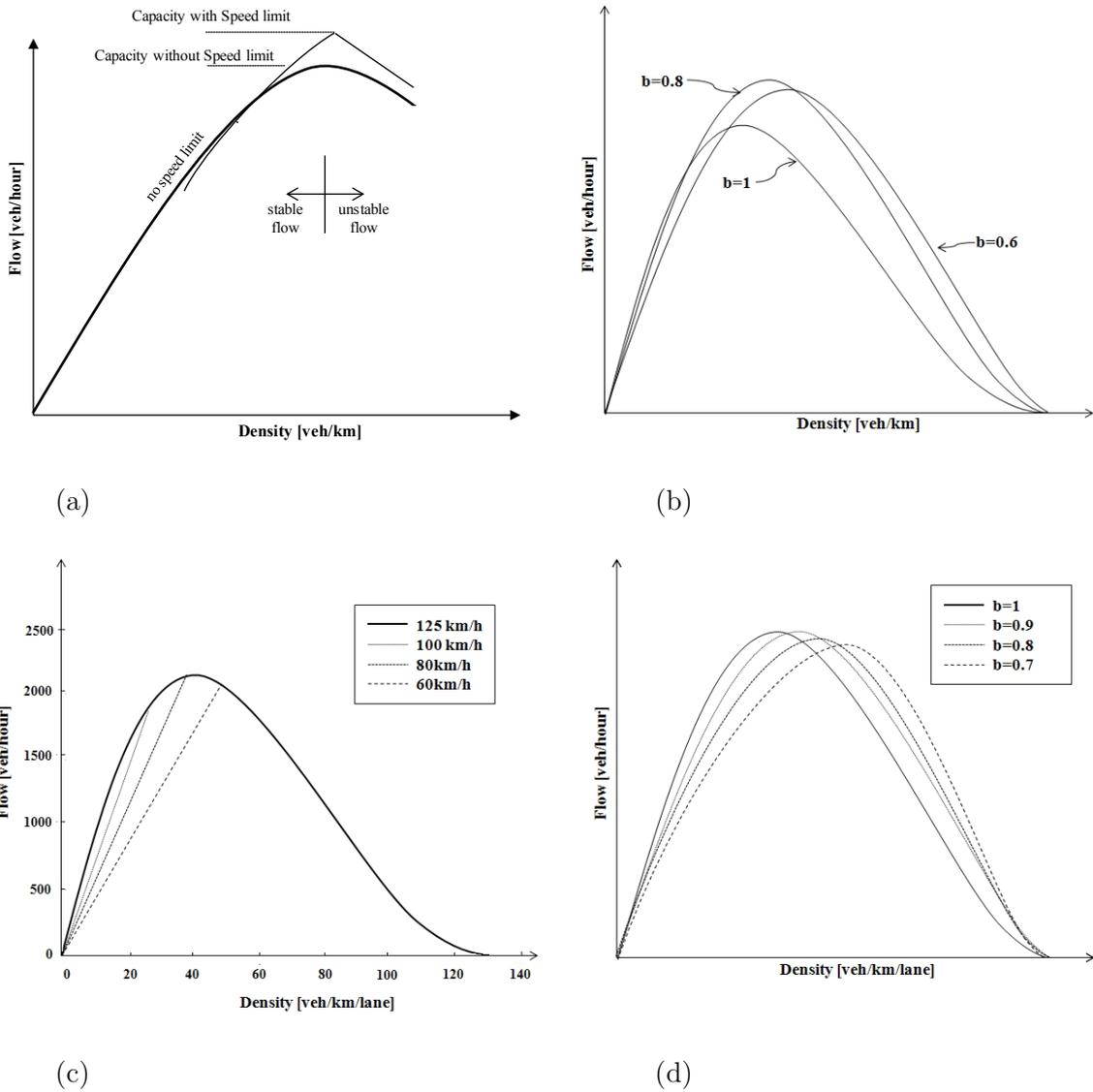
Figure 2.1 *Impact of DSL in the FD. (a)* [Zackor and Papageorgiou, 1991] *(b) speed limit=b*Vfree* [Cremer, 1979] *(c)* [Hegyi, et al., 2005]*(d) speed limit=b*Vfree* [Carlson, et al., 2010]

Figure 2.1(a) represents FD modification under low to medium traffic volume, where average speed was lower due to the speed limit reduction. For higher volumes, there was a theoretically increase up to 10% in both capacity and optimal speed, due to the stabilizing

effect achieved, based on analytical results. In fact, Zackor's results followed the conclusions presented by [Cremer, 1979] analytical models, which are summarized in Figure 2.1(b), and are quite similar to the analytical ones presented by [Carlson, et al., 2010] (Figure 2.1(d)). Indeed, the capacity increases referenced by Zackor were a bit exaggerated, as could be observed in field studies from the Netherlands [Hegyi and Hoogendoorn, 2010; van den Hoogen and Smulders, 1994], Great Britain [Papageorgiou, et al., 2008] and Spain [Soriguera, et al., 2013; Torné, et al., 2010; Torné, et al., 2011]. The Dutch cases concluded that DSL measures did not substantially contribute to improve capacity, whilst the English and -Spanish considered the data analysis were rather inconclusive. In fact, [Hegyi, et al., 2005] considered that the only effect of these measures is the emergence of a straight line on the left side (free flow) of the flow-occupancy diagrams corresponding to the maximum allowed speed (Figure 2.1 (c)) without mentioning any capacity increase.

In conclusion, there seems to be no agreement on the effect of DSL in aggregated traffic flow behavior. As many of the DSL benefits are originated in the microscopic traffic flow dynamics, only individual vehicle data or a smart use of aggregations allows observing certain phenomena. Unfortunately, micro data is not usually available, and results using aggregations are still non-conclusive. So, there is need to more empirical studies in order to obtain robust empirical facts which are expected to clear up the question. Moreover, there is an inherent difficulty when analyzing DSL effect for certain speed limitations. As data collected during specific speed limit values may not cover the whole range of possible occupancies, conclusions related with capacity increase or critical occupancy/density shift could be difficult to draw.

2.2.2   Mainstream metering and other features of DSL strategies

It is reasonable to assume that when DSL measures are applied in situations below the critical occupation (and assuming a certain degree of acceptance by the drivers), would reduce the mean speed observed (i.e. decrease the slope of the free flow branch of the FD). The effect of this measure is to move horizontally flow-density plot points, i.e. holding the same flow and increasing density, compared to the original state. This way, the activation of speed limit below critical speed upstream of a mainstream bottleneck in undercritical conditions will increase travel time and hence produce an active restriction to flow [Carlson, et al., 2010; Hegyi, et al., 2005; Soriguera, 2010]. In that case, this effect may be similar to a mainstream ramp metering.

The control strategy may define a certain set of speed limits upstream of a bottleneck such that downstream flow does not trigger the bottleneck activation. In fact, most of the reported studies focus on optimal coordination between DSL and ramp metering (e.g. [Carlson, et al., 2010; Hegyi, et al., 2005] who modeled hypothetical highway networks and [Lu, et al., 2010] a real stretch in San Francisco). In certain configurations, the final outflow will be maximized, leading to a corresponding decrease of the total time spent in

the system at the downstream bottleneck. It should be noted that this would only happen if capacity drop in the bottleneck section is avoided. Indeed, it is essential to select suitable time varying speed limits in order to achieve the desired onset congestion delay. Thresholds usually work with pre-selected values according to the measured occupation or flow on the road. Yet occupancy is a better explanatory of capacity [Papageorgiou, et al., 2008].

2.2.3   International experiences

Three main purposes motivate to set DSL applications: safety, congestion or sustainability. Some of these motivations arise when a reduction of the supply (e.g. bad weather or work zones) or an increase in demand occurs (e.g. special events or rush-hour in a metropolitan freeway). The early motivation for promoting DSL projects was to increase safety or reduce congestion. In the United States, examples can be found since the 1960s in many states, such as, New Jersey, Washington State, Nevada, Arizona or New Mexico [Lennie, et al., 2009; Zarean, et al., 2000]. However, the safety benefits of DSL have already partially been contrasted whilst promising results have been found in different countries. Close to London, the M25 ring motorway is one of the most publicized examples. Since 1995, DSL has been operational on the western quadrant of this highway. In 2004, a business case for 'before and after' data was run between Junctions 15 and 16 resulting on a 10% injury accidents reduction [Highways Agency, 2004]. Similar results were found for M42 [Sultan, et al., 2008]. Other simulation-based studies analyze the effect on the crash potential reduction in an American highway section resulting on a 5–17% reduction [Lee, et al., 2006]. For interested readers, additional information about reported safety benefits and an overview of Australasian DSL practices can be found in [Lennie, et al., 2009].

Taking into account that traffic is considered the main source of air pollution and greenhouse gases emissions, for many European countries DSL is a useful tool to achieve the European Union sustainability requirements in terms of air quality [European Parliament Council, 2008]. Accordingly, DSL projects have been developed in Germany, Netherlands, Austria, Great Britain, France, Switzerland, Denmark and Spain. A number of European countries have documented a reduction in vehicle emissions after the implementation of DSL [Baldasano, et al., 2010; Keller, et al., 2008; Thomas, 2006], whilst only simulation studies can be found for U.S. [Wang and Walton, 2006]. In the Netherlands, local traffic NO_x emission reductions were reported by 20 and 30 percent range and PM_{10} was reduced by approximately 10 percent in four test locations [Olde, et al., 2005]. The Great Britain experience presents initial results for vehicle emissions in M42 between 4 and 10 percent reductions depending on the pollutant [Brinckerhoff, 2010; Thomas, 2006].

2.3 Managed lanes

Managed Lanes (ML) can be described as strategies which package lane(s) depending on certain criteria (e.g. vehicle destination, occupancy or qualification) for different purposes (e.g. reduce lane changes, optimize the use of the lanes, improve the level of service for carpool vehicles). ML can facilitate traffic flow in response to changing conditions (i.e. variation on the on/off ramp demand, special events, work zones or incidents) by controlling user eligibility, access (by pricing or destination discrimination) and lane assignment.

Traditional lane management strategies include high occupancy vehicle (HOV) lanes, high occupancy toll (HOT) lanes and dynamic lane assignment (DLA). Of the listed managed lane strategies, the first two are commonly implemented in U.S. (and rarely in Europe) and the third one currently only implemented in some European countries. However, the addition of new ML strategies both Europe and U.S. may be easily supported since most corridors are heavily instrumented and users are steadily familiar with dynamic operating conditions that reflect different traffic states [Brinckerhoff, 2010; Mirshahi, et al., 2007].

2.3.1   HOV and HOT lanes

HOV lanes are highway lanes intended to carry only vehicles with more than a predetermined number of occupants (e.g. 2+ or 3+). The primary purpose of an HOV lane is to increase the total number of people moved through a congested corridor by offering two kinds of incentives to the users: travel time savings and reliability and predictability travel time. So, this strategy can be consider as a tool for actuating in the demand side, more than optimizing the available capacity of an infrastructure (without considering the mentioned below smoothing effect), i.e. focusing on people mobility. However, financial sustainability and better capacity utilization are the main motivations which have led to switch some of the existing HOV facilities to value pricing schemes (i.e. HOT lanes). HOT lanes allow single or lower occupancy vehicles to use an HOV lane for a fee, while maintaining free travel for qualifying HOVs and green vehicles. The need to expand financing options has made HOT lanes a more publicly and politically acceptable option, whilst the availability of electronic toll collection has made it technically viable. Because they are often constructed within the existing road space they are criticized as being an environmental tax or perk for the rich.

2.3.1.1   Strategy definitions

Examples of HOV facilities include bus-only roadways, freeway lanes reserved for HOVs, bypass lanes at metered freeway entrance ramps, and special lanes on arterial streets. On the highways, there are usually two paradigmatic options. (i) Non-separated HOV facilities

share the right of way with general purpose lanes (GPL). Access is allowed continuously throughout the length of the lane. The configuration –widely introduced in the U.S.– offers flexibility although they are hard to be enforced. HOV lanes can be easily converted to GPL during off-peak hours. (ii) Separated HOV facilities are separated from GPL by physical barriers or a buffer space. The barriers might be temporary or permanent. Temporary barriers are mostly used with contraflow lanes. Permanent barriers reduce greatly the time and space flexibility of the system. Both are more expensive than the first-mentioned configuration. From the operational point of view, two relaxations have been introduced: (i) switching from the 7/24 system to only certain hours and days of week, and (ii) allow exceptions to the types of vehicles that use these lanes (such as users of HOT lanes, or low emission and energy efficient vehicles).

2.3.1.2   *Literature review*

However, despite the efforts of both engineers and researchers, the impacts of non-separated HOV lanes are not yet fully understood. In fact, an understanding of the relation between HOV lanes and system capacity is necessary. A central issue in this debate is whether an HOV lane affects the capacity of the GPL passing through a bottleneck. A number of studies report that carpool lanes unduly penalize LOVs by creating congestion in non-carpool lanes [Chen, et al., 2005; Kwon and Varaiya, 2008; Laval and Daganzo, 2006] and reducing the capacity of bottlenecks. The reason is based on the increase in lane-changing activity, especially when the GPL are congested near an active bottleneck. Indeed, there is also wastage in freeway's queue storage space due to underutilized HOV lanes which extends the queue length in adjacent lanes. On the contrary, empirical evidences support that capacity may actually increase because disruptive vehicle lane changing diminishes close to the bottleneck [Cassidy, et al., 2010; Menendez and Daganzo, 2007]. Both authors refer the existence of a smoothing effect that is characterized by significantly higher bottleneck discharge flows in adjacent lanes. The effect was consistently reproduced across days and sites; and was so pronounced that even an underutilized carpool lane could increase a bottleneck's total discharge flow. In addition, the findings suggest that highway congestion could also be reduced by inducing the smoothing effect through other means, whether or not it includes a carpool lane (e.g. dynamic lane assignment).

2.3.1.3   *International experiences*

The opening of the bus-only lane on the Shirley Highway in Northern Virginia/Washington, D.C. in 1969 represent the first freeway HOV application in the U.S. In contrast, few HOV lanes are currently operating in Europe and Australia. Madrid, Linz, Leeds, Amsterdam, Stockholm, Oslo and –currently under development– the C58 in the north access to Barcelona are the main projects to date. As none of them are non-separated HOV lanes, the aforementioned capacity discussion can't be fully applied to

these cases. However, the main feature for separated HOV lanes can be applied: carry vehicles with a higher number of occupants considerably increasing the number of people moved during congested periods. This quality remains even when the number of vehicles that use the HOV lane is lower than on the adjoining GPL. Apart from that, enhancing bus transit operations is an additional objective to be focused on. As a result of an HOV facility (i) bus travel times, (ii) schedule adherence and (iii) vehicle plus labor productivity will be improved. In addition, this policy will attract new bus riders, enhance transit cost effectiveness and, as some drivers are induced to share rides, environmental benefits can accrue [Turnbull, et al., 2006]. Focusing on Madrid case (operational since 1994), the results showed. public transport mode share grew from 23% to 35% after implementation, while single occupancy vehicle use dropped from 70% to 48% [Schijns and Eng, 2006].

Due to the U.S. Department of Transportation's strong endorsement for using HOT lanes (generally by consider converting HOV to HOT lanes) as an effective strategy to address congestion, many such projects have arisen. Pricing differential by occupancy is most commonly applied as a HOT lane. However, HOV-vehicles continue to represent the majority of vehicles on existing HOT facilities. Most of the toll schemes are based on a simple principle: the more congested general-purpose lanes (GPL) are the higher HOT fares would be charged for prevailing free flow in HOT lanes. However, the lack of widespread toll models consistent with traffic flow theory results on toll schemes based on practitioners experience and non clear methodologies [Chung and Recker, 2011]. Multiple variations in pricing application (dynamic, e.g. I-95 Miami; flat rate, e.g. I-15 Salt Lake City; or time-of-day rating, e.g. I-25 Denver) and access (occupancy, e.g. SR-167 Seattle; green-vehicles class, e.g. I-15 San Diego; or midway accessibility, e.g. I-10 W Houston) are found throughout the U.S. Meanwhile, almost any previously defined HOT lane scheme has been found in Europe, though they could provide the much-needed funds in times of fiscal constraint. However, other value pricing techniques are more frequent in Europe (such as area licensing scheme in London or cordon in Stockholm). For interested readers, additional information can be obtained in [Federal Highway Administration, 2008; Richardson and Bae, 2008; Turnbull, et al., 2006].

2.3.2   Dynamic lane assignment

DLA strategies try to address one of the most disrupting activities of freeway traffic: lane changes and the indiscriminate use of each one of the lanes based only on the drivers self criteria, with neither information nor control. Commonly, drivers try to maximize their own benefit without considering the efficient performance of the whole system. Therefore, big traffic flow management inefficiencies caused by first-in-first-out (FIFO) queues generation on highways may arise.

As far as author's knowledge is concern, little literature applying lane-changing phenomenon knowledge to specific lane-use operations has been reported. Most of them focus on evaluations of present facilities which simultaneously use additional ATM strategies: e.g. M42 in the Great Britain [Sultan, et al., 2008] with hard-shoulder plus DSL

or A4-A86 in France [Aron, et al., 2013] with hard-shoulder running. Others analyze the highway capacity increase thanks to *platooning* (vehicles travelling as clusters) by automated vehicles strategies, leading to lane-changing and spacing (i.e. gaps) reduction [Hall and Caliskan, 1999]. However, we'll focus on traffic operations assuming non-automatic vehicles, which is the common case in any highway. Among the little methodological literature found, two references must be remarked: [Jella, et al., 1993] focus on urban intersections and [Daganzo, et al., 2002] on freeways. The next section will summarize the DLA strategies presented in the latter reference.

2.3.2.1 *Daganzo et al.'s guidelines for DLA*

The goal of these strategies is to avoid that the lane changing friction and O/D weaving creates problems even to other drivers not involved in these actions (e.g. blocking parts of the freeway).

Off-ramp

Possible strategies to avoid spill-back generated by exit-traffic queues lead to:

- **Avoid total trunk jam** by forming a made up queue of exit-traffic only, in order to avoid the creation of FIFO queues which will affect through vehicles. DLA strategies include: (i) increase the off-ramp capacity by using the hard shoulder, (ii) queue the vehicles that want to exit in a specifically reserved lane in order to avoid last-minute incorporations, (iii) ban lane changing immediately before and after the problematic off-ramp.
- **Avoid underuse of alternative off-ramp(s).** Due to off ramp low capacity, undesirable spill-back of the queues are generated, although few vehicles want to exit. Then, the problematic off-ramp can be closed (before the spill-back is generated) and increase storage capacity of an alternative exit(s) by using the hard shoulder (if it exists) or other lanes. Suitable information about alternative exits must be given on advance to motorists.

On-ramps

In order to avoid total freeway blockage due to a demanded entrance, implemented strategies may assure free flow traffic in the entrances by lane-changing prohibition during peak hours for right lane(s) upstream problematic entrances. Again, suitable information must be given on advance to motorists.

2.3.2.2 *European dynamic lane assignment experiences*

Among the strategies described before, the most extended is the use of hard-shoulder as a driving lane for certain periods of time and locations. Multiple examples can be found in Europe and rarely in the rest of the world. The first case dates back to 1996 in a short

stretch near Cologne (Germany). Only experiences from France and Great Britain will be tackled, giving up other interesting cases (e.g. the Netherlands').

In France, hard shoulder running was introduced in 2000 near Paris, where two urban freeways (A4 and A86) share a 2.3 km-long weaving section. As the traffic flows of the two freeways are added, traffic is particularly dense at some hours on the weaving section. In order to reduce the congestion, an auxiliary lane was opened to the traffic 4 hours per day on average (Figure 2.2(a) and (b)). This lane is taken from the hard shoulder. At peak hours, a moveable barrier is removed, opening part of the hard shoulder to the traffic. The impact on congestion is quite obvious due to the enlarged offer. In fact, almost 10% capacity increasing as well as level of service improvements has been reported [Desnouailles and Cohen, 2007]. Similarly, the impact on safety has been positive by decreasing the number of accidents due to the reduction in traffic density [Aron, et al., 2013].

(a)   (b)

(c)
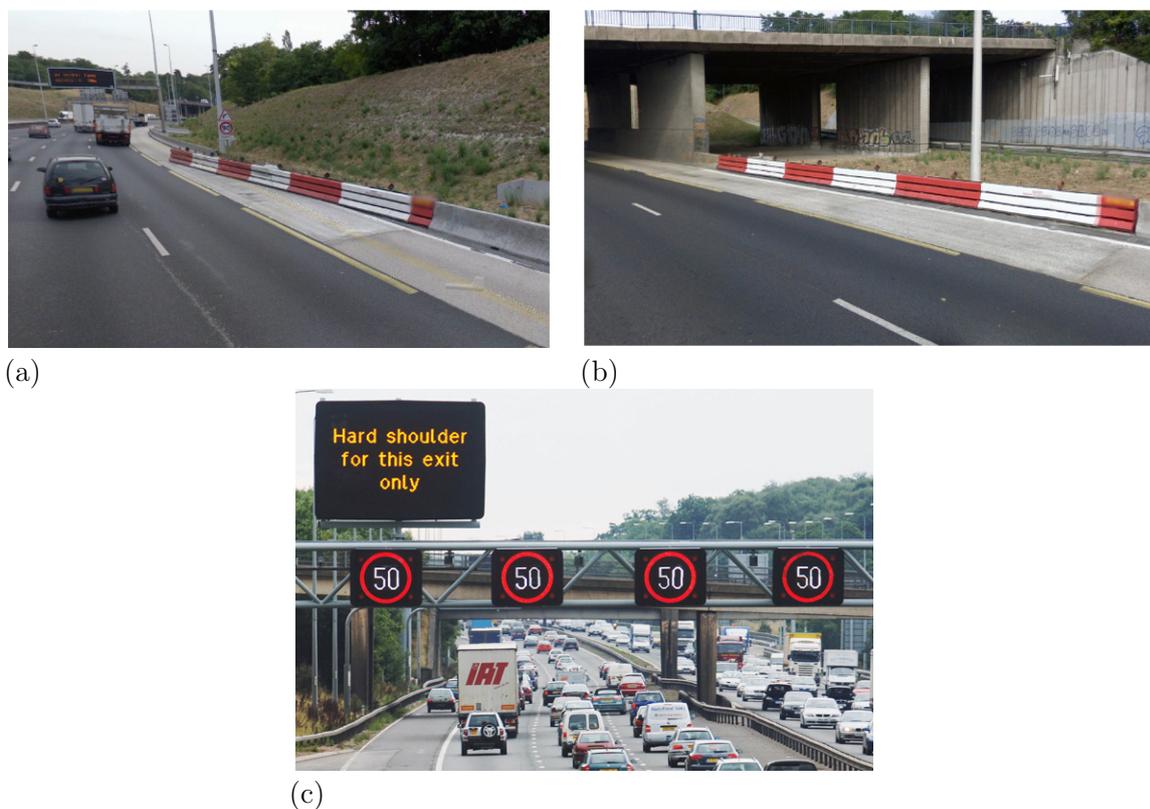
Figure 2.2 (a) Closed and (b) opened to traffic hard shoulder lane in the A4-A86 concurrent freeway (France), and (c) hard shoulder together with DSL operational scheme in the M42 freeway (Great Britain)
Source: © Google Maps

In Great Britain, an operational scheme using the hard shoulder along 17 km on the M42 three-lanes freeway close to Birmingham became operational in 2006. It works together with DSL during periods of peak demand. The operation of three lanes plus hard-shoulder sections (4L-DSL, see Figure 2.2(c)) has increased the observed freeway capacity by an

average of 7% (compared to non-DSL case) and 10% (compared to three lanes plus DSL, i.e. 3L-DSL) [Sultan, et al., 2008]. Apart from that, 4L-DSL has reduced average journey times during periods of recurrent severe congestion, compared to the 3L-DSL case. It was also demonstrated that on average over all weekdays the variability of journey times has been reduced by 27% and 34%, when compared to non-DSL and 3L-DSL respectively. Although there is no definitive quantitative conclusion on safety, the Great Britain authorities consider that this operation is ''Globally At Least Equivalent'' (GALE, i.e. introducing this new strategy will not increase the total risk compared to the current case) and propose its generalization [Halbert and Orme, 2008].

2.4 Ramp metering

Ramp metering (RM) is the most widely employed strategy to control freeway traffic in the U.S. and less expanded in the rest of the world. It is recognized that RM can directly control the flow into the freeway. The principle is straightforward: short delays at on-ramps is the (relatively low) price to pay for maintaining capacity flow on the freeway itself, leading to substantial savings for each individual road user.

The effectiveness of RM for improving some performance indicators (e.g. average travel time or total time spent) resides in its ability to avoid capacity drop performance of the main-line flow thanks to the metered entrances. As a result, a suitable control strategy is capable of actually maintaining the maximum throughput, around the capacity flow value. Thus, a pertinent RM strategy should maintain critical density and capacity-flow conditions for full exploitation of the potential RM benefit. Unfortunately, a fact that is very frequently neglected in RM designs and installations [Papageorgiou and Papamichail, 2007]. A consequence of this kind of strategies is the wastage of the freeway's ability to store traffic. Other approaches may involve (i) allowing the creation of queues but avoid the blockage of much demanded sections. In that case, equity requirements arise and should be faced via linking control actions at multiple ramps [Geroliminis, et al., 2011]. (ii) Manage entering traffic according to its destination in order to avoid spill-back from downstream off-ramps. (iii) Allow the entering traffic that does not want to exit at the next off-ramp to be incorporated in the trunk lanes.

2.4.1   Mainly ramp metering strategies

RM strategies may be classified according to various characteristics. The most basic being the usage or not of (i) real-time measurements and (ii) the geographical extent of the strategy's operation:

- **Real time strategies** base their decisions on collected real-time measurements from the freeway mainstream and the on-ramps; in contrast, fixed-time strategies are designed on the basis of historical data and are applied daily in the same

(clock-based) way. Obviously, fixed-time strategies are blind to the prevailing traffic conditions and may therefore either underload or overload the freeway.

- **Local strategies** control individual on-ramps based on real-time measurements collected in the vicinity of the on-ramp. In its turn, they may be applied independently to several individual metered ramps of the freeway.

- **Coordinated strategies** control all metered ramps of a freeway (network) based on all available real-time measurements. Coordinated strategies may be further subdivided into optimal control strategies, hierarchical strategies or rule-based strategies according to the basic design approach and their functional architecture. Next sections will describe both local and coordinate strategies.

It must be noted that any RM strategy needs to address the main RM side effect: the temporary formation of vehicle queues at on-ramps. Consequently, suitable ramp queue management strategies (requiring precise vehicle count estimations) may avoid queued vehicles to interfere with adjacent street traffic.

2.4.1.1   Local RM strategies

The strategies are evaluated, although traffic signals are not necessarily modified, at each control time interval (e.g. 20s…60s). At the end of each running period, time-averaged measurements of traffic flow or occupancy from the ending period are used to calculate (via the corresponding strategy) the ramp flow to be implemented in the next period. The most popular local RM strategies are: (i) the demand capacity strategy, (ii) ALINEA [Papageorgiou, et al., 1991] (both described in next paragraphs) and their variations (e.g. real-time estimation of the critical occupancy (AD-ALINEA)); and other approaches based on neural network or fuzzy logic.

- The **demand-capacity strategy** [Masher, et al., 1975] attempts to add to the last measured upstream flow as much ramp flow as necessary to reach the known downstream freeway capacity. The DC strategy is not a feedback but a feedforward (open loop) disturbance rejection scheme, which is generally known to be sensitive to various further non-measurable disturbances (e.g. a slow vehicle, a short shock wave, merging difficulties, etc.). A remarkable drawback of this strategy is that targets a pre-specified flow capacity value which may lead to further efficiency degradation due to its inherent uncertainty.

- **ALINEA** is a feedback (closed loop) RM strategy. Its scheme is called an I-type (integral) regulator in the classical Automatic Control Theory. It is well-known that this regulator leads automatically to a certain objective occupancy, under stationary average conditions. This is a particularly attractive feature of ALINEA as it automatically rejects possible disturbances (e.g. upstream flow or a biased implementation of the ordered ramp flow rate).

2.4.1.2   *Coordinate ramp metering strategies*

The main inefficiencies of local ramp algorithms are manifested in cases of (i) high demand, as travelers at the on-ramps may experience significantly higher delays than those in the mainline or adjacent ramps, leading to inequity issues, (ii) short on-ramps, where long queues can spillback to the arterial network and create significant delays in adjacent streets and (iii) multiple active bottlenecks. Consequently, co-operation of multiple ramps to avoid congestion must be considered. Some selected coordinate RM strategies are briefly described:

- An open-loop **optimal control** can be applied to coordinate RM strategies. Many different models have arisen (see [Papageorgiou and Kotsialos, 2002] for a wide literature review). Founded on different information (e.g. current and predicted traffic state/demand and physical features of the network) and subjected to certain constraints (e.g. minimize an objective criterion such as the total time spent in the whole network), the control strategy delivers both traffic evolution and optimal controls over the optimization time horizon.

- **Hierarchical control** structure decomposes the overall freeway control into several concatenated components or layers and a feedback link with measurements information, defining a close-loop structure. The main objective of hierarchical control is to achieve computational feasibility and robustness of the control solutions by determining the system-wide, nominal metering rates first and then adjusting them according to real traffic conditions. Previous efforts have employed the hierarchical control concept based on two, three (i.e. estimation, optimization and direct control [Papageorgiou, 1983]) or four (i.e. initialization, estimation, optimization and tactics processors) [May, 1975] layers.

2.4.2   International experiences

The success of early RM applications in U.S. cities (e.g. Chicago, Los Angeles, Minneapolis or Seattle) led to the expansion of RM systems in many other metropolitan areas as well as in other worldwide cities (e.g. Sydney, Toronto or Tel Aviv). Among them, several coordinated RM strategies are or were in operation in U.S. [Bogenberger and May, 1999; Geroliminis, et al., 2011; Hadi, 2005; Piotrowicz and Robinson, 1995].

The development of RM systems in Europe has a different level of implementation across countries. Leuven (Belgium) has the only system currently in operation across the country. In Great Britain, the firsts systems were introduced in the junction 10 of the M6 (in Walsall, Birmingham) in 1986. More than 30 locations (in Leeds, Manchester and Birmingham) with ALINEA's RM strategy are currently operational. This strategy is also operative in several locations in France (i.e. Paris, Lyon and Bordeaux). Similar approach is currently operating in more than 100 locations around Germany, as well as the

Netherlands which run a wide network of RM facilities [Papageorgiou and Papamichail, 2007].

2.5 Conclusions

ATM techniques are generally less expensive and more feasibly (due to recurrent space limitations) than the physical expansion of facilities. So, they can commonly be used for dealing with traffic congestion. The presented literature review may be useful for having an inspection of the rising ATM strategies. Among them, hard shoulder running is one of the most effective traffic management techniques because it can easily provide additional capacity. Besides, it should be taken into account that most of the presented strategies show synergistic relationships between them and should be implemented together. Actually, research has shown that these strategies are rarely implemented in isolation, showing practitioners' awareness of these features.

PART II
Modeling

Overview

Many different traffic flow models have been developed since the first attempts in the 1930's (e.g. [Greenshields, et al., 1934]). From a historical point of view four big families may be distinguished, ranging from the simple characterization of the fundamental relation to computer intensive microscopic simulations. Macroscopic and mesoscopic model families are in between [van Wageningen-Kessels, 2013]. In essence, the fundamental relation describes the static relation between the speed of a vehicle and the distance to its leader. The rest of the families describe how this relation evolves over time and space. The description of the traffic dynamics and the level of detail differ in each family of models. Microscopic models trace the movement of individual vehicles, while macroscopic models aggregate many vehicles and describe their 'average' behavior (e.g. flow, density and avg. speed). Mesoscopic models imply an aggregation level halfway between the microscopic and macroscopic families.

This modeling part of the thesis deals mainly with the fundamental relation and with macroscopic models. Mesoscopic level is addressed when a greater level of detail is required. The main objective consists in presenting an extension to a classical first order macroscopic traffic flow model (i.e. Cell Transmission Model, [Daganzo, 1994]) including a new characterization of the fundamental diagram that allows the modeling of DSL strategies. This was motivated by the fact that most of freeway traffic models dealing with DSL rely on second order models [Carlson, et al., 2010; Hegyi, et al., 2005; Lu, et al., 2010], although most of them still present inconsistencies when modeling real traffic flow behavior and may lead to effects contradicting the physics of traffic. Until now, first order models have not been used for this purpose due to their hypothetical inability to reproduce the capacity drop phenomenon which is crucial for reproducing congested periods in bottleneck sections.

On the way to construct this new model, the Asymmetric Cell Transmission Model (ACTM, [Gomes and Horowitz, 2006]) was explored in order to assess its suitability.

Surprisingly, an inconsistency was observed when simulating uncontrolled merges. Chapter 3 deals with an approach for recovering the physical feasibility of the ACTM's merging behavior, by proposing a simple calibration procedure of the model parameters in uncontrolled merges. A comprehensive justification of the proposed method and its full consistency with Newell-Daganzo merge model is also provided.

Chapter 4 constitutes one of the main contributions of the whole thesis. The CTM model extension proposes a representation of the DSL effect on the fundamental diagram considering the capacity drop effect. The CTM is extended by introducing some simple rules which govern the capacity drop mechanism. In addition, a new traffic control strategy is proposed. The coordinated (i.e. ramp metering and DSL) strategy takes advantage of the total merge capacity. The strategy proposes suitable speed limits and metered flows for maximizing the available capacity. Results obtained from simulation using this innovative control strategy are promising and may motivate further research efforts on this direction.

CHAPTER 3

Freeway merging models

3.1 Introduction

Many freeway traffic macroscopic models have arisen (e.g. [Daganzo, 2005; Kuhne and Beckschule, 1993; Laval and Leclercq, 2008; Lebacque, 2003; Messmer and Papageorgiou, 1990; Papageorgiou, et al., 1990; Payne, 1971]) since the seminal work by Lighthill and Whitham [Lighthill and Whitham, 1955] and Richards [Richards, 1956] where the well-known LWR continuous traffic model was developed by solving a partial differential equation (i.e. the vehicle conservation equation) assuming an equation of state (i.e. the flow as a function of the density: the fundamental diagram). Models based on this LWR assumption are called 1st order models. Variants exist, depending on the method employed to solve the equations. Notably, the LWR model captures the main phenomena of freeway traffic: the different behavior between free-flow and congested regimes and the formation and dissipation of queues upstream of a bottleneck. However, important limitations have also been noted (e.g. [Kerner, 2013; Kerner and Rehborn, 1996]). Criticisms are related to the central hypothesis of a static fundamental diagram, pointing out, for example, the inability to reproduce the hysteresis phenomenon or the sudden capacity drop due to congestion. It should be pointed out that recently, [Torné, et al., 2012] have proposed a first order model including the capacity drop phenomenon. In spite of this, these limitations gave rise to second or higher order models, which retain the principle of vehicle conservation from the LWR model, but replace the static fundamental diagram with a dynamic function aiming to reproduce some of these traffic flow features, such as the asymmetric driving behavior resulting in a hysteresis loop. From a mathematical point of

29

view, it would be more accurate to name these models as "two/multiple-equation models", rather than "second/higher order models" [Aw and Rascle, 2000], since most of the proposals in the literature involve only first order partial differential equations.

The simplicity of first order models makes them appealing. Their finite difference approach known as the Cell Transmission Model [Daganzo, 1994 and 1995] makes them suitable for computer simulation and widely used. The Asymmetric Cell Transmission Model (ACTM) [Gomes, 2004; Gomes and Horowitz, 2006] is an evolution of the previous to account for the restricted merging dynamics at controlled ramps. The ACTM is considered to be an adequate model for assessing real time control strategies in road networks with RM installations [Gomes, 2004; Gomes and Horowitz, 2006; Kurzhanskiy, et al., 2009; Kurzhanskiy and Varaiya, 2010b] and its use has been promoted in any type of freeways network, including those without ramp control. However, it is shown in the present chapter that the common simplistic calibration of the two merging parameters involved in the ACTM in non-controlled on-ramps, leads to an unrealistic merging model. The goal of the research is to analyze the merging model presented in the ACTM and propose a calibration methodology to preserve its desirable physical coherence in non-controlled junctions.

The chapter is structured as follows. Section 3.2 provides a brief overview of the state-of-the-art on macroscopic merge models. Section 3.3 focuses on the classic Newell-Daganzo merge model, while ACTM merge model is described in Section 3.4. In Section 3.5 a theoretical calibration procedure for the ACTM merge model is presented, accompanied by a feasibility analysis of the parameter values. Next, Section 3.6 presents an example application. Finally, conclusions are outlined in Section 3.7.

3.2 Freeway merge models revisited

Traffic flow behavior at freeway merges has been extensively analyzed in the literature because of its crucial role in traffic flow modeling, dynamic network loading and dynamic traffic assignment. The merging at junctions is one of the most disrupting operations in a freeway network, causing queues that may affect all the incoming links. Two situations result in queues at merges: (i) an exogenous queue from downstream spills back and affects both merging approaches and (ii) the merge itself becomes an active bottleneck due to the restrictions imposed by the traffic blending mechanism. This last cause is analyzed here.

Consider a merge with two incoming links (i.e. the upstream mainline and an onramp) which blend in a single branch (i.e. the downstream mainline). Among the different macroscopic merge models [Banks, 2000; Daganzo, 1995; Gomes, 2004; Gomes and Horowitz, 2006; Jacobson, et al., 1989; Jin and Zhang, 2003; Kotsialos, et al., 1999; Lebacque, 1996; Newell, 1982; Ni and Leonard II, 2005; Papageorgiou, et al., 1991; Papageorgiou, et al., 1990]*,* the most simple and robust approach is proposed by [Newell, 1982] and next reformulated in [Daganzo, 1995]. This is referred as Newell-Daganzo (ND) model in the present chapter. This merge model excels in being physically intuitive. At the same time, the stationary distribution patterns proposed have been observed for multiple

merging junctions in California [Bar-Gera and Ahn, 2010; Cassidy and Ahn, 2005], so that the results of the model have already been empirically validated.

The ND model assumes that the system maximizes the total merge outflow. This is achieved by introducing the non-concave/non-convex mid{} function in its formulation (see next section for the details on the ND model). A big drawback arises when dealing with optimization problems due to the non optimization-friendly nature of such function. To overcome this difficulty, the use of the concave min{} function and the linearization of the merging model is proposed (e.g. [Gomes, 2004; Gomes and Horowitz, 2006; Jin and Zhang, 2003; Kotsialos, et al., 2002; Ni and Leonard II, 2005]). However, the careless use of such artifacts may lead to the violation of the maximum outflow assumption [Tampère, et al., 2011]. This is the consequence of considering the unrealistic assumption of weak or no interactions between all the supply constraints, which is unacceptable for uncontrolled merges (e.g. [Gomes and Horowitz, 2006; Hegyi, et al., 2005; Lebacque, 1996; Papageorgiou, et al., 1990]). Remarkably, [Ni and Leonard II, 2005] have recently proposed a model whose solution space is equivalent to that of ND. In addition, the computational cost has been reduced thanks to having eliminated unnecessary and costly possibilities. Similarly, a generic node model [Tampère, et al., 2011] also satisfies the aforementioned solution space.

One controversial issue is observed among the former merge models: the definition of the blending parameter (i.e. the merge-ratio). No general consensus exists about its definition or calibration. This could be justified due to the fact that the understanding of the interaction between the merging streams for various traffic conditions still generates debate. Here, the merging ratio, 'β', is defined as the relative magnitude of both merging flows (i.e. $\beta = \frac{on-ramp\ flow}{upstream\ mainline\ flow}$). Table 3.1 groups the different models according to their definition of the blending parameter.

In all the previous models the merge capacity is exogenously determined by the downstream merging conditions. This capacity will decrease if a downstream queue spills back to the merge. However, the merge capacity could also be considered as an endogenous variable if the merge itself is considered to be an active bottleneck with capacity drop. In this case, the capacity of the merge might be influenced by the traffic states on one or both of its approaches and the ND theory may not hold [Banks, 1990; Cassidy and Bertini, 1999]. In this case, it would be reasonable to assume that the capacity is influenced by many factors like the merge demand, lane change maneuvers, vehicle accumulation in the queue, vehicles acceleration rate and length of the merging section, among others. Recently, [Leclercq, et al., 2011] has proposed an analytical model, based on the variational theory [Daganzo, 2005], that estimates the capacity drop with respect to the demand on the on-ramp and the different model parameters. Unfortunately, both incoming roads are assumed single-lane by the model, leading to a limited approach. Moreover, [Torné, et al., 2012] has also described an alternative approach to the problem, defining a simple methodology based on the ND priority-based merge queuing model.

Table 3.1 *Merge-ratio definition for different merge models.*

Author	Blending parameter definition
[Jin and Zhang, 2003]	The ratio between both link demands (Demand ratio)
[Cassidy and Ahn, 2005; Ni and Leonard II, 2005]	The ratio between both link capacities (Capacity ratio)
[Bar-Gera and Ahn, 2010]	The ratio between both link number of lanes (Lane ratio)
[Lebacque, 1996]	The ratio between the maximum possible density of each link. (Maximum density ratio)
[Gomes and Horowitz, 2006]	The relative position of the on-ramp with respect the downstream section boundary
[Hegyi, et al., 2005; Kotsialos, et al., 1999; Kotsialos, et al., 2002; Papageorgiou, et al., 1990]	A ratio of densities considering the actual density, the critical density and the maximum possible density in the mainstream link (Density ratio)
[Daganzo, 1995; Newell, 1982]	Priority distribution among both links (Branch priority)

3.3 The Newell-Daganzo merge model

The ND merge model assumes that, if there is enough demand, the macroscopic flow behavior at the merge is determined by the available downstream capacity and the driver priorities in turn-taking. This last parameter describes the interaction between the conflicting streams when both upstream branches are congested. With this assumption, the theory describes the possible traffic conditions upstream of a merge only considering the demand for each branch, the merge capacity and the turn-taking rates.

Consider a hypothetical junction with two links ('A' and 'B') competing to merge into a single link 'E' (Figure 3.1). The flow through each of the links, 'Y_A' and 'Y_B', is constrained by:

1. 'R_E', the maximum receiving flow, defined as the minimum between the capacity of link 'E' and its available number of slots for allocating the vehicles at the jam density. These are properties of the receiving cell.

2. 'S_A' and 'S_B', the maximum sending flows, defined as the minimum between the demand and the capacity of each merging link.
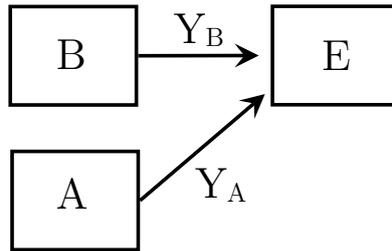
Figure 3.1 *Representation of a merge*

Consequently, the flow must satisfy:

$$Y_A + Y_B \leq R_E \tag{3.1}$$

$$Y_A \leq S_A \tag{3.2}$$

$$Y_B \leq S_B \tag{3.3}$$

All variables are expressed in vehicle units, and the time dimension (t) is omitted in the equations for notational convenience. All equations hold for one time step.

Figure 3.2 shows the feasible region for the merging flows. If $S_A + S_B \leq R_E$ all the demand can be accommodated and both branches are dictated by forward waves. The merging flows are then defined by the coordinates of point 'P' (i.e. $Y_A = S_A$; $Y_B = S_B$). When $S_A + S_B > R_E$ and not all the demand can be satisfied, it is necessary to include the priorities to solve for the merging flows. To do so, the priority of link 'i', 'p_i', is defined as the proportion of the total merging flow coming from link 'i' when both approaches are oversaturated (i.e. in queue). Therefore, $p_A + p_B = 1$). This is shown in Figure 3.2 as the 'Priority line' (i.e. the straight line from the origin with slope $\frac{Y_A}{Y_B} = \frac{p_A}{p_B} = \beta$). Because link 'B' is always as assigned to the highest priority link, the merge-ratio 'β' takes values between 0 and 1. [Bar-Gera and Ahn, 2010] suggest that merging-ratios can be reasonably estimated by the ratios between the numbers of lanes on the merging approaches, which is typically similar to the capacity-ratio (i.e. rate among capacities of the merging links).

Equations (3.4) and (3.5) describe the solution for the merging flows when $S_A + S_B > R_E$. More easily, the solution can be graphically obtained from Figure 3.4 as the middle point of the three intersections (i.e. M, M′ and M″) defined by the merge downsloping capacity line with the supply lines (i.e. S_A and S_B) and with the priority line.

$$Y_B = \text{mid}\{S_B, R_E - S_A, p_B \cdot R_E\} \tag{3.4}$$

$$Y_A = \mathrm{mid}\{S_A, R_E - S_B, p_A \cdot R_E\} \tag{3.5}$$

Three situations can be identified depending on the value of priorities, 'p_A' and 'p_B':

- Case I, when $p_B R_E \geq S_B$ and $p_A R_E < S_A$ leading to $Y_B = S_B$ and $Y_A = R_E - S_B$, corresponding to point M. Congestion in branch A.

- Case II, when $p_B R_E < S_B$ and $p_A R_E < S_A$ leading to $Y_B = p_B R_E$ and $Y_A = p_A R_E$, corresponding to point M''_{II}. Congestion in branch A and B.

- Case III, when $p_B R_E < S_B$ and $p_A R_E \geq S_A$ leading to $Y_B = R_E - S_A$ and $Y_A = S_A$, corresponding to point M'. Congestion in branch B.

Therefore, if $S_A + S_B > R_E$ not all the demand can be served, and queues are created in one, the other or both approaches. It is implicitly assumed that $S_A < R_E$ and $S_B < R_E$.
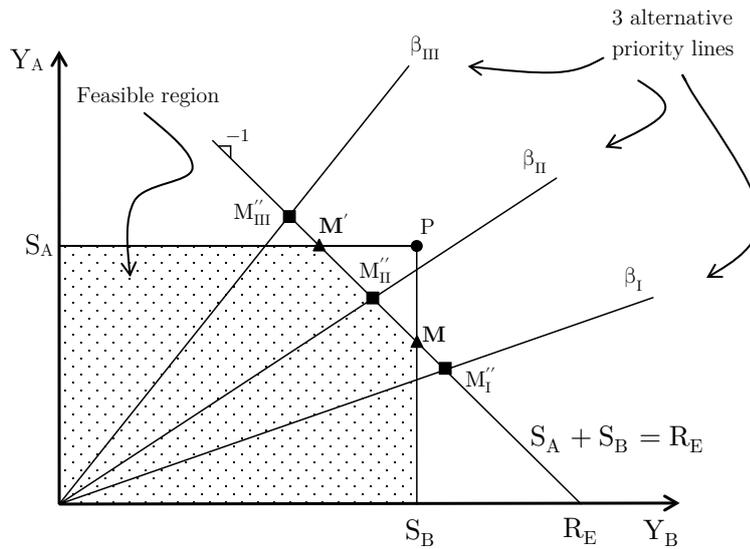
Figure 3.2 *Diagram of feasible merge flows considering three different priority ratios* (β_I, β_{II}, β_{III}).

3.4 The asymmetric cell transmission merge model

With the objective of formulating an optimal RM control scheme, [Gomes, 2004] developed the ACTM. The *asymmetric* tag was included in opposition to the CTM merge model symmetry (i.e. the model is not altered if labels 'A' and 'B' in Equations (3.4) through(3.5) are exchanged). This means that the ACTM makes independent allocations for each merging branch. [Gomes and Horowitz, 2006] simplified the original ACTM model by

eliminating one parameter (from three to two). This last version of the ACTM is considered here.
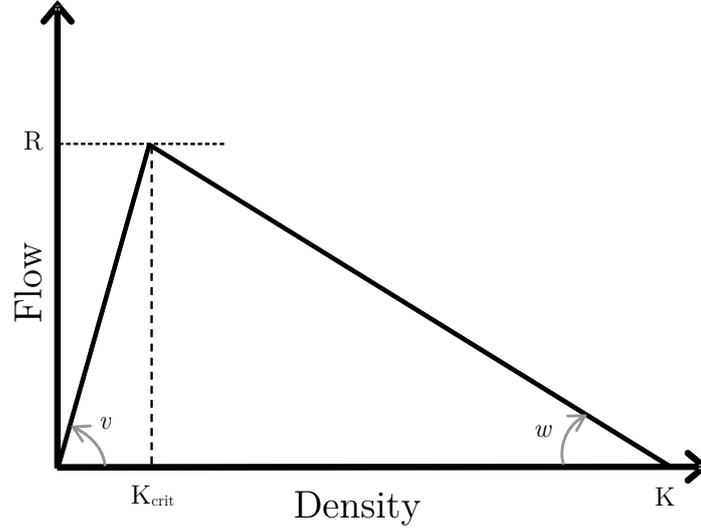
Figure 3.3 *A generic fundamental diagram.*

Consider the same hypothetical merge junction previously presented in Figure 3.1, and a generic triangular fundamental diagram (see Figure 3.3). ACTM computes flow on link 'A' (the on-ramp) as the minimum between the on-ramp demand, the on-ramp capacity and the mainline available space. Flow on link 'B' (the mainline) is computed as the minimum of what can be sent by the upstream section (i.e. the minimum between the demand and the upstream mainline capacity) and what can be received by the downstream section (i.e. the minimum between the downstream mainline capacity and the available space). Remarkably, any explicit constraint captures the fact that the blending of both flows must not exceed the downstream capacity, 'R_E'. All terms in Equations (3.6) and (3.7) are normalized to vehicle units.

$$Y_A = \min \left\{ \underbrace{S_A}_{\min(\text{on-ramp demand, on-ramp capacity})} ; \underbrace{\xi \, d_E(K_{jam,E} - k_E)}_{\text{mainline space}} \right\} \qquad (3.6)$$

$$
\begin{aligned}
Y_B = & \\
= \min & \left\{ \underbrace{S_B}_{\min(\text{demand, ups. capacity})} ; \underbrace{\frac{w_E \Delta t}{d_E}(K_E d_E - k_E d_E - \gamma Y_A)}_{\text{congestion term}} ; \underbrace{R_E}_{\text{mainline capacity}} \right\}
\end{aligned} \qquad (3.7)
$$

Being 'd_E' the length of cell, 'K_E' the jam density, 'v_E' the free-flow speed, 'w_E' the characteristic shock-wave speed and 'k_E' the density, all of them related to cell 'E'. 'Δt' is the time length of the simulation step. Finally, 'γ' and 'ξ' are the ACTM parameters.

In [Gomes, 2004; Gomes and Horowitz, 2006] 'ξ' is defined as an allocation parameter which captures the distribution of available space to vehicles entering from the on-ramp. No value was suggested. In practice, 'ξ' is usually assumed to be equivalent to the merging ratio (i.e. $\beta = \frac{Y_A}{Y_B} = \frac{p_A}{p_B}$ in the ND notation). This value is generally approximated by the ratio of capacities. Then, $\xi = \beta \approx \frac{on-ramp\ capacity}{mainline\ capacity}$, and $\xi \in [0,1]$ holds by definition. In the same original references, 'γ' is defined to be the flow blending coefficient that reflects the relative position of the on-ramp within the section. For simplicity, [Gomes, et al., 2008] suggest to use $\gamma = 0$, indicating that the on-ramp is at the beginning of each of the discretized sections. Note that, in this case, the on-ramp contribution is only considered in the mainline flow computation via 'k_E', updated every time step.

As it will be proved, this simplistic calibration leads to unfeasible merging flows in uncontrolled merges when both approaches are queued and congestion terms dominate.Note that with this calibration the expected behavior of the model should yield merging flows satisfying:

$$\frac{Y_A}{Y_B} = \xi \tag{3.8}$$

Instead, from Equations (3.6) and (3.7) we obtain:

$$\frac{Y_A}{Y_B} = \frac{\xi d_E}{w_E \Delta t} \tag{3.9}$$

Consequently Equation (3.8) holds only if $d/\Delta t = w_E$ hold. This means that results will only be consistent for a given relationship between the cell size 'd_E' and the duration of the time step, 'Δt'. However, this relationship turns out to be impossible given the $d_E/\Delta t \geq v_E$ restriction of the CTM, where 'v_E' is the free-flow speed of cell E. Note that values of $v_E \approx 5{\sim}6 w_E$ are generally measured on freeways.

Another consequence of this calibration is that yields $\frac{Y_A}{Y_B} \gg \xi$ given that $\frac{d_E}{w_E \Delta t} \gg 1$. This means that the portion of the total blended flow coming from 'A' (i.e. the on-ramp) is much greater than the ratio of capacities. This can be represented in the feasible flows diagram with a steeper priority line in relation to the ND model (see Figure 3.4).

In the diagram of feasible merge flows, when the solution is on the priority line indicates that both approaches are queued. This happens only if the intersection between the priority and the capacity lines falls inside the rectangular region defined by 'S_A' and 'S_B'. In Figure 3.4 it can be seen that for the simplistic calibration of the ACTM, this only happens for very reduced downstream capacities (i.e. smaller than $R_{E,II}$). These situations where the downstream merge capacity is smaller than the upstream supply (e.g. Case III in Figure 3.4) are not frequent. This means that in almost all the realistic situations, the ACTM merge solution will be on the $Y_A = S_A$ line (e.g. Case I in Figure 3.4). This results

in the on-ramp serving vehicles at capacity and all queues transferred to the upstream mainline. It is obvious that the behavior of the ACTM under this calibration is not realistic.
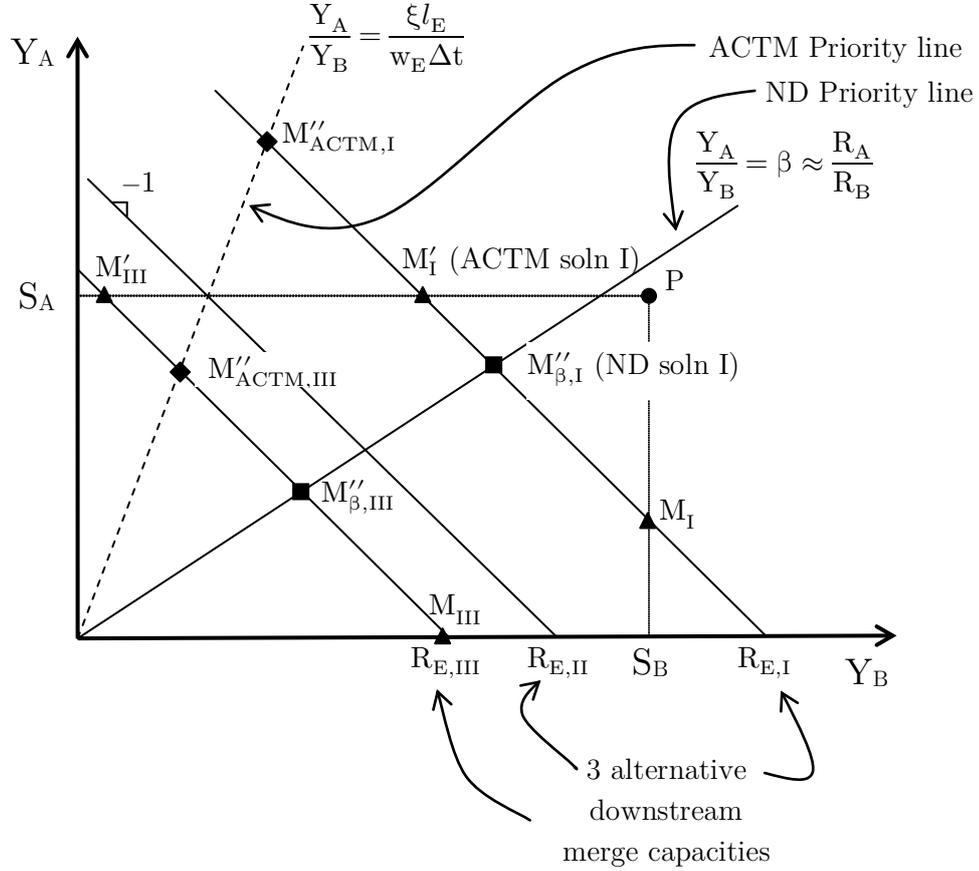
Figure 3.4 *ACTM vs ND diagram of feasible merge flows for three different merge capacities $\left(R_{E,I}, R_{E,II}, R_{E,III} \right)$ values.*

3.5 Recovering the physical coherent behavior of the asymmetric cell transmission merge model

Given the inconsistent behavior of the simplistic calibration of the two ACTM merging parameters in uncontrolled merging junctions, a new calibration procedure is needed. A simple expression is proposed here. The main assumption is that, in uncontrolled merging, the results obtained using the ACTM should be equivalent to those obtained according to the ND model. Because the ND merging model is globally accepted and empirically validated in uncontrolled merges [Bar-Gera and Ahn, 2010; Cassidy and Ahn, 2005] this assumption is consistent and eliminates the need for further empirical validation [.

Assume that $S_B + S_C > R_E$, $p_B R_E < S_B$ and $p_A R_E < S_A$ (Case II referred in Section 3.3). Consequently, it is expected that the dynamics of the model lead to $Y_B = p_B R_E$ and $Y_A = p_A R_E$, corresponding to point M''_{II} of Figure 3.2. Both branches are congested in this case. Note that this scenario is the same captured by [Bar-Gera and Ahn, 2010; Cassidy and Ahn, 2005] in their experiments. In this case, it can be assured that the congestion term in Equation (3.7) and the mainline space term in Equation (3.6) are dominant. In that case, Equations (3.6) and (3.7) are simplified to:

$$Y_A = \xi \ d_E (K_{jam,E} - k_E) \tag{3.10}$$

$$Y_B = \frac{w_E \Delta t}{d_E} (K_{jam,E} d_E - k_E d_E - \gamma Y_A) \tag{3.11}$$

In order to obtain a feasible calibration it is needed to fix one parameter and solve for the other consistently with the ND solution. Here, 'γ' is simply calibrated as the merge-ratio, according to [Bar-Gera and Ahn, 2010; Cassidy and Ahn, 2005] definition, (i.e. $\gamma \equiv \beta$). Thus;

$$\gamma = \frac{Y_A}{Y_B} = \frac{\xi \ d_E(K_{jam,E} - k_E)}{\frac{w_E \Delta t}{d_E} \left(K_{jam,E} d_E - k_E d_E - \gamma \left(\xi \ d_E(K_{jam,E} - k_E)\right)\right)} = \frac{\xi}{\frac{w_E \Delta t}{d}(1 - \gamma \xi)} \tag{3.12}$$

Considering that 'w_E' and 'd_E' are properties of the model so that the 'E' subscript can be eliminated and solving for 'ξ' we obtain:

$$\xi = \frac{w \Delta t \gamma}{(d + w \Delta t \gamma^2)} \tag{3.13}$$

Normalizing the congestion wave speed as $\overline{w} \triangleq \frac{w \Delta t}{d}$, Equation (3.13) is simplified to:

$$\xi = \frac{\overline{w} \gamma}{1 + \gamma^2 w} \tag{3.14}$$

Equation (3.14) provides the consistent value for 'ξ' when 'γ' is calibrated as the merge-ratio. It is only needed to check the consistency of the calibrated parameters with the range of possible values proposed by [Gomes, 2004; Gomes and Horowitz, 2006] for the ACTM. The only restrictive condition is:

$$\xi < \frac{1 - \overline{w}}{1 - \gamma \overline{w}} \tag{3.15}$$

And therefore, for the proposed calibration being consistent it is necessary to fulfill:

$$\frac{\overline{\gamma w}}{1 + \overline{\gamma^2 w}} < \frac{1 - \overline{w}}{1 - \gamma \overline{w}} \tag{3.16}$$

Because both γ and $\overline{w} \in [0,1]$, denominators at both sides of the inequality are positive. We can obtain:

$$\overline{w} < \frac{1}{(1 + \gamma - \gamma^2)} \tag{3.17}$$

The limiting case occurs for $\gamma = 1/2$ when $f(\gamma) = (1 + \gamma - \gamma^2)$ is maximum. It implies $\overline{w} < 0.8$ and this condition is fulfilled as '$\overline{w}$' is usually lower than 0.2.
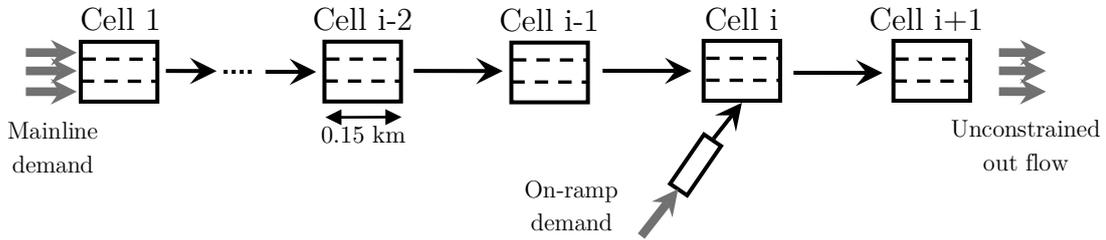
Figure 3.5 *Simulation test layout.*

3.6 Numerical experiment to illustrate the adequacy of the calibration

A simple simulation experiment is presented in order to compare the performance of the ACTM using the simplistic versus the modified calibration procedure (see Table **3.2**). The later will be referred as the 'modified ACTM' and the former simply as ACTM. The analytical results presented in the previous section are numerically confirmed. For uncontrolled merges, the ACTM shows an unfeasible behavior, while the results of the modified ACTM are consistent with those predicted by the ND model.

The simulation layout consists simply in a 3 km stretch of a 3 lane freeway with a one-lane on-ramp located at its downstream end. . The stretch is divided in a total of 20 cells of 0.15 km in length each (see Figure 3.5). The mainline fundamental diagram (for the 3 lanes) is defined by a capacity of 6000 veh/h, free-flow speed of 90 km/h and jam density of 400 veh/km, while the on-ramp capacity is set to 2000 veh/h. The merge-ratio, computed as the ratio of capacities, is $\beta = {}^{1}\!/_{3}$.

Table 3.2 *Merging model parameters*

Model	γ	ξ
ACTM (non-controlled ramps)	0	$\frac{1}{3}$ $(\equiv \beta)$
Modified ACTM	$\frac{1}{3}$ $(\equiv \beta)$	0.0652

The network is fed with a constant mainline demand of 5000 veh/h and a progressively increasing on-ramp demand. Demand profiles are shown in Figure 3.6 for the total simulation time (i.e. 60 min or 600 time steps). The simulation step is computed as the ratio between the cell length and the free-flow speed being $\frac{0.15}{90}$ hours (i.e. Δt=6 sec). This configuration allows observing multiple traffic states (on-ramp / mainline): uncongested / uncongested (U/U), congested / congested (C/C) and uncongested / congested (U/C).
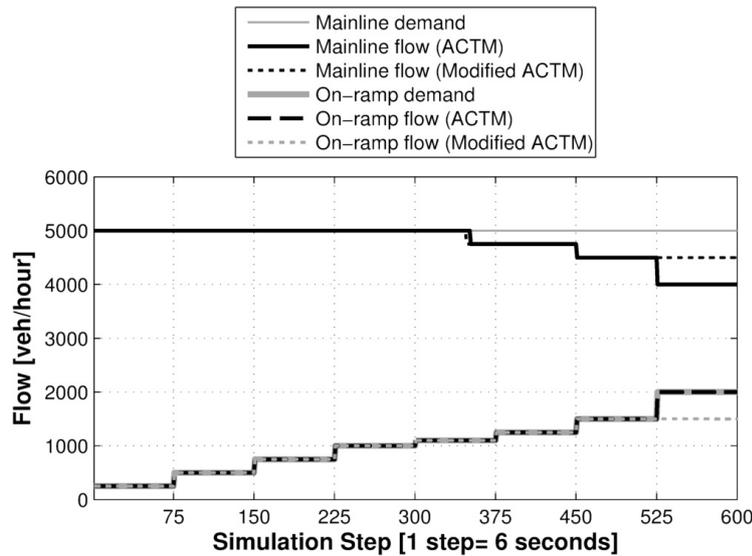
Figure 3.6 *Stationary merge flow simulation results.*

Both models behave identically while the total demand does not exceed the available capacity (i.e. until t=300, Figure 3.6). The U/U state holds and all the demand is served. After that, the aggregated demand exceeds the downstream mainline capacity. Because the relation between the ramp and the mainline demand is smaller than the merging ratio, the system switches to a U/C state, where the mainline demand cannot be fully satisfied but all the ramp demand is served. The mainline flow drops to a value equal to the difference between the mainline capacity and the on-ramp demand, corresponding to points along the feasible flows boundary satisfying $Y_A = S_A$ and $Y_B = R_E - S_A$ (see Figure 3.7). Still, this behavior is adequately modeled by both calibrations. At t=525, the increase in the on-ramp demand implies exceeding the merge-ratio. Therefore, according to the ND model and empirical observations, it should be expected for the system to switch to the C/C

state, with both approaches queued and merging flows given by the intersection between the priority and capacity lines in the feasible merge flows diagram. As it can be seen in Figure 3.7 and 3.8, this is adequately modeled by the modified ACTM. On the contrary the simplistic calibration of the ACTM does not predict queues on the ramp. The system remains in a U/C state still serving all the ramp demand.
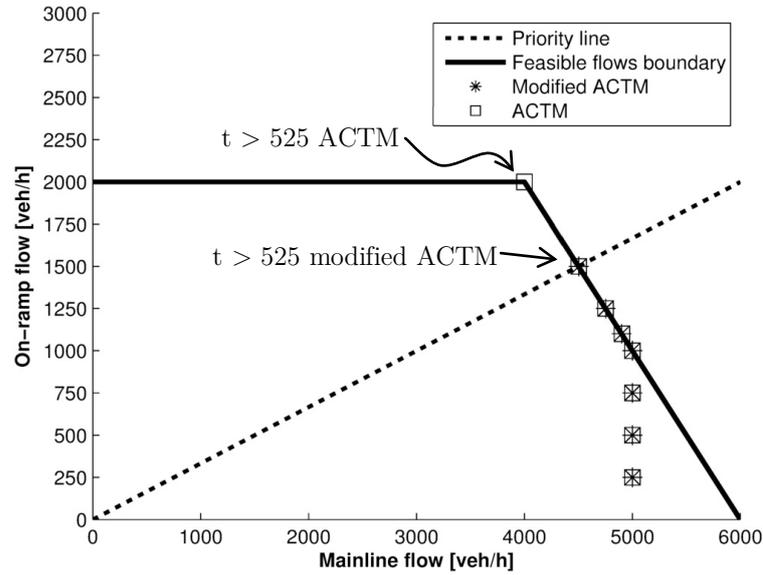
Figure 3.7 *Diagram of feasible flows for the simulation stationary results.*
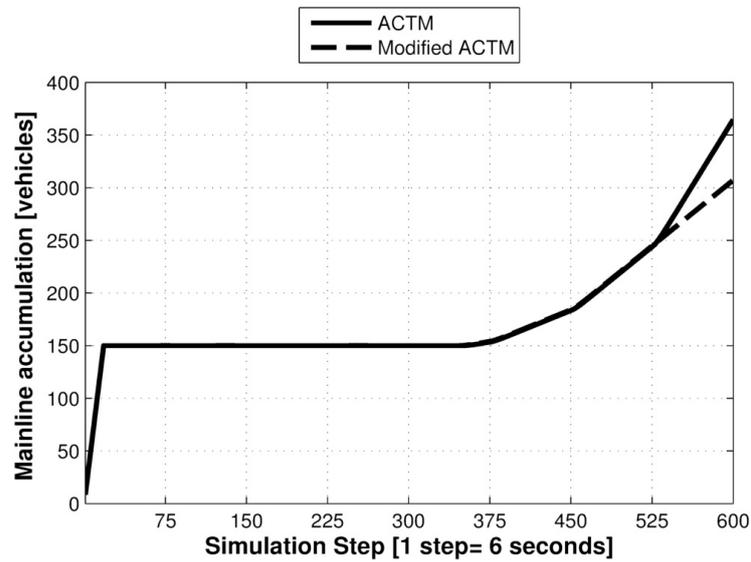
(a)
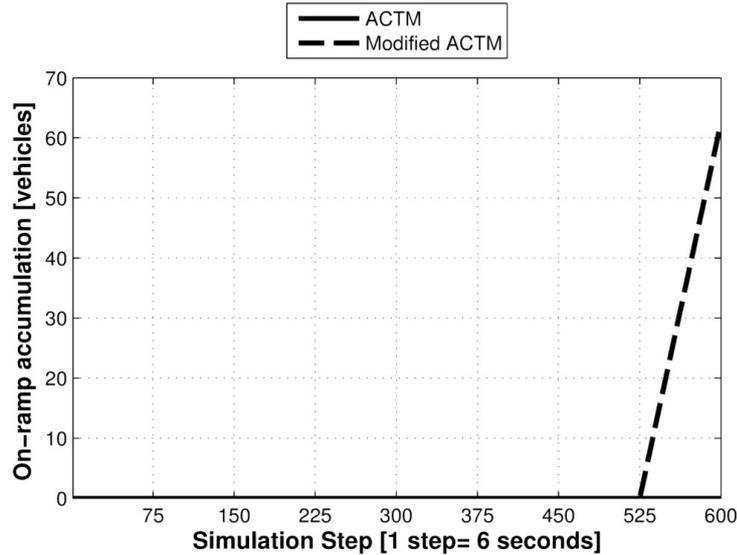
(b)

Figure 3.8 *Total amount of vehicles accumulated (a) in the mainline and (b) in the on-ramp.*

3.7 Conclusions

An accurate prediction of the queue length in a merge junction is an essential outcome of any traffic flow model. The standard simplistic calibration of the ACTM parameters in uncontrolled merge junctions fails on this attempt, because the obtained merging behavior is unfeasible. This probably happens because ACTM was developed focusing primarily on controlled merges. Uncontrolled merges were implicitly left out of the scope of the model, and not addressed.

The present paper proposes a methodology for calibrating the two-parameters involved in the ACTM model in uncontrolled merges, in order to recover its physical consistency. The method is condensed in one explicit analytical equation. Despite the introduction of an additional calibration expression, the simplicity of the model is preserved. Considering this calibration, the ACTM can continue being a suitable tool for operational planning and management of freeway corridors, preserving the optimization friendly formulation of the ACTM for controlled junctions and still resulting in a physically feasible behavior in uncontrolled merges.

Coordinated active traffic management freeway strategies using the capacity-lagged cell transmission model

4.1 Introduction

An increasing consensus exists among practitioners that a good way to increase the efficiency (e.g. productivity, equity and sustainability) in road networks is to introduce active management strategies (ATM) for traffic flow control. ATM is defined as a real-time process which dynamically manage recurrent and non-recurrent traffic congestion in real time (measuring and analyzing traffic data for implementing the most promising control strategies) in order to maximize the efficiency of highway facilities [Kurzhanskiy and Varaiya, 2010a]. Among all of them (e.g. DSL; RM, lane management or traveler information) the former mentioned policy is one of the most attractive ones, due to its apparent simplicity and reduced implementation cost. Many cities around the world have introduced DSL systems (e.g. [Papageorgiou, et al., 2008; Soriguera, et al., 2013]) isolated or coordinated with other control measures. The claimed benefits of these actuations imply a reduction of traffic related emissions and accident rates, as well as an improvement in congestion reduction (e.g. reduction of stop&go traffic episodes or capacity increase). It is reported [Smulders, 1990; Zackor, 1991] that congestion benefits are due to the

homogenization of traffic flow which allows for increased capacity. In fact, some recent data in European highways provides evidences to support this assertion [Ramoneda, 2013] while other studies result inconclusive in this point [Papageorgiou, et al., 2008]. Other contributions result from studies mainly focused on control algorithms evaluations using second order traffic flow models [Carlson, et al., 2010; Hadiuzzaman, et al., 2013; Hegyi, et al., 2005]. Anyway, the effectiveness of the policy is still a controversial issue, probably motivated for the few evaluations of the dynamic speed limit with real data that can be found in the literature, (e.g. [Heydecker and Addison, 2011; Papageorgiou, et al., 2008; Soriguera, et al., 2013]). Remarkably, most of them only analyzed aggregated data, while many of the DSL benefits are originated in the microscopic traffic flow dynamics. Thus, only individual vehicle data or a smart use of aggregations may allow to observe certain phenomena. Unfortunately, micro data is not usually available and aggregated results are still non-conclusive.

Still, one main drawback can be derived from DSL strategies by setting speed limits below free-flow speed. It has been frequently described how such strategies imply trigger an active bottleneck in the most downstream controlled section [Carlson, et al., 2010; Soriguera, 2010]. On addition, such active bottleneck may lead to the likely capacity drop occurrence, i.e. the observed difference between the freeway capacity and the queue discharge rate, due to inherent instabilities inside a queued bottleneck (e.g. lane changing maneuvers, sudden driver brakes, small gaps combined with high flows or short acceleration areas at the edge of a bottleneck). Consequently, any attempt to model such strategies requires incorporating the capacity drop phenomenon due to its significant role in the traffic flow behavior. Equally important, DSL algorithms must also consider the unnecessary drawbacks that will be induced if setting low speed limits during uncongested situations. In fact, real implementations of DSL systems support its strategy on processing different types of real time information. However, most of them are not sufficiently based on traffic flow behavior and lead to blind and inefficient algorithms [Soriguera, et al., 2013]. The research tries to fill this gap by presenting an extension of the original CTM [Daganzo, 1994 and 1995] to include capacity drop and also incorporating the ability to model DSL strategies.

The rest of the chapter is organized as follows. Next section briefly describes the background of the main components involved in the modeling issue. Section 4.3 presents a conceptual description of the extended CTM, i.e. the capacity-lagged cell transmission model (CL-CTM). The CTM's modifications proposed (both for a common and a merge section) are described in Section 4.4. Next, Section 4.5 presents some model applications to reproduce coordinated RM and DSL strategies. Then, Section 4.6 shows the performance of the extended model using a simple traffic network case study. Finally, a summary with conclusions and future research scope is given.

4.2 Background

Notably, most of the freeway traffic models dealing with DSL rely on the original METANET second order model [Messmer and Papageorgiou, 1990] or its further extensions [Carlson, et al., 2010; Hadiuzzaman, et al., 2013; Hegyi, et al., 2005]. An important property for such models is their ability to reproduce the capacity drop mechanism at active bottlenecks, being suitable for traffic flow control applications. Almost any first order model can be found in the literature modeling such phenomenon. Certainly, the original LWR first-order model [Lighthill and Whitham, 1955; Richards, 1956] does not reproduce the capacity drop phenomenon. Anyway, only two main attempts have been found [Edie, 1961; Lebacque, 2003], [Srivastava and Geroliminis, 2013] although none of them have fully succeeded on its purpose. The first one's contribution is almost reduced to a formal, but not validated, model. While, the second, although more complete than the other, still requires to be more extensively contrasted for a wider range of traffic scenarios.

Firstly, [Edie, 1961] observed a sharp speed drop in a small density range for some fundamental diagram (FD) plots and proposed a two-regime phase diagram to explicitly include the capacity drop. Commonly, this shape of the FD has been conceptualized as a 'reverse lambda', with linear branches. Two possible flow values result for a certain range of densities around the critical density value, depending on the traffic conditions, i.e. whether traffic is moving from the free-flow to the congested regime on the equilibrium curve or vice versa. For the first case, the maximum observed flow reaches the capacity flow, while going from congested regime to free-flow occurs via a maximum flow lower than the capacity flow. This value is also called the 'queue discharge capacity'. Certainly, such model behavior provides a simple formalization for the hysteresis phenomenon, first reported in [Treiterer and Myers, 1974]. Unfortunately, these types of models have a serious conceptual disadvantage: infinite speed shock waves can occur, leading to unrealistic results. The here proposed model tries to solve this problem.

More recently, other models assume a two-phase traffic flow model [Lebacque, 2003; Monamy, et al., 2012]: the former only consider simple links without nodes, while the latter also includes a node model. This model extends first-order traffic flow models by retaking the fact that traffic acceleration is bounded. One phase captures traffic equilibrium where flow and speed are functions of density whereas traffic acceleration is low. The second phase is characterized by constant acceleration. By building a modified sending or demand function (i.e. the characterization of the demanded flow to be sent from upstream to downstream sections), the capacity drop phenomenon is, in some way, reproduced. In essence, the demand at the head of a congested section is then not capacity, but a bit lower. In free-flow the demand function is unchanged, so still the demand is the current flow. The model has been partially validated with real data in a certain merging section part near Paris [Monamy, et al., 2012]. Although the model next proposed shares certain similarities with the two-phase traffic flow model, its Godunov scheme [Godunov, 1959] approach for solving the LWR's equations set up a great divergence. Recently, [Srivastava and Geroliminis, 2013] integrated capacity drop in an LWR model with intermediate (internal) freeway mainline boundaries. The authors utilize a fundamental diagram with

two values of capacity and they provide a memory-based methodology to choose the appropriate value in the numerical solution of the problem with a Godunov scheme. The model did not investigate any control methodologies to be developed.

4.2.1　　The cell transmission model and its main extensions

The CTM is one of the most referenced and contrasted first order models. It presents an approximation to the LWR's kinematic wave partial differential equation applying the Godunov's discrete approximation. Its simplicity and flexibility, has led many researchers to adapt the model to a vast range of applications, without (almost) altering the original modeling formulation and the FD. Both researchers and practitioners have employed CTM in multiple transportation applications such as dynamic traffic assignment, traffic prediction, signal control and RM [Alecsandru, et al., 2011]. The CTM poses a traffic flow mechanism based in a discretization of the network in a finite number of cells and certain rules for sending and receiving traffic flow.

For the macroscopic simulation's interest, existing literature shows how the original cell-transmission formulation has been adapted to achieve different purposes (e.g. [Daganzo, 1997; Laval and Daganzo, 2006]. Among all of them, two main contributions must be highlighted and also set the starting point for the model next proposed. (i) [Daganzo, 1999; Szeto, 2008] showed how the accuracy of the model will be enhanced if the downstream density, used to calculate the sending flows, was read a certain number of time intervals earlier than the current time. That way, the advancing of the queued-state shockwave follows a physical coherent behavior and also improves the accuracy results (being capable to exactly converge to the LWR model for certain parameter values, defined below). This model is also known as the lagged CTM, i.e. L-CTM. (ii)[Gomes and Horowitz, 2006] proposed the asymmetric cell transmission model (ACTM) which differs from the original merge model, because the merging flows are calculated separately via two-parameters which may be accurately calibrated to preserve the physical coherence of the ACTM merge model. Particularly, when uncontrolled merges are simulated, the simplistic calibration proposed in [Gomes and Horowitz, 2006] leads to unrealistic merge behavior. Chapter 3 has previously proposed a simple expression, Equation (3.13), for calibrating the two-parameters involved in the ACTM's merge model. With this calibration, queues can be developed both in the mainline and on-ramp when high demand occurs, in contrast with the original ACTM model.

4.3　The capacity-lagged cell transmission model

The proposed model almost follows the same CTM's assumption which considers the flow as a function of density (i.e. its FD) with a triangular or trapezoidal form. In fact, a

modification will be introduced in the triangular shape, in order to account for the capacity drop phenomenon. Thus, the proposed shape results on a mix between the triangular-shaped FD (for the free-flow branch) and the trapezoidal-shaped FD (for the congested branch), Figure 4.1(a). That way, the characteristic in common with respect the two-phase traffic flow model is sketched. At the same time, the aforementioned problems related with the inverse lambda-shaped FD are skipped. In fact, all the state transitions exhibit feasible values for the shock waves speeds, even when the sudden flow drop in the vicinity of the critical density occurs. Figure 4.1(a) depicts the main CL-CTM parameters involved in the FD, both considering a non-DSL case (where v is the free-flow speed) and a DSL-case (where V is the current speed limit value). Two different capacities may be associated to each case: 'Q^f' for the free-flow states and 'Q^d' for the states where the capacity drop has occurred, i.e. congested states. Both parameters satisfy $\frac{Q^f(v)}{1+\alpha} = Q^d(v)$, being $\alpha < 1$, and K corresponds to the jam density value.
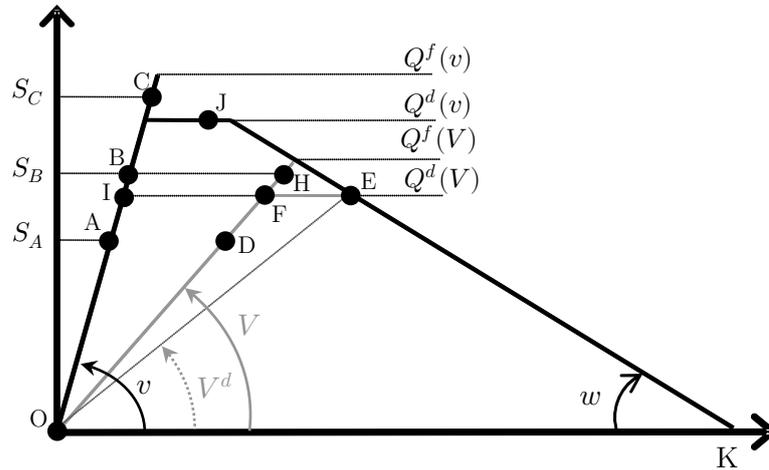
(a)
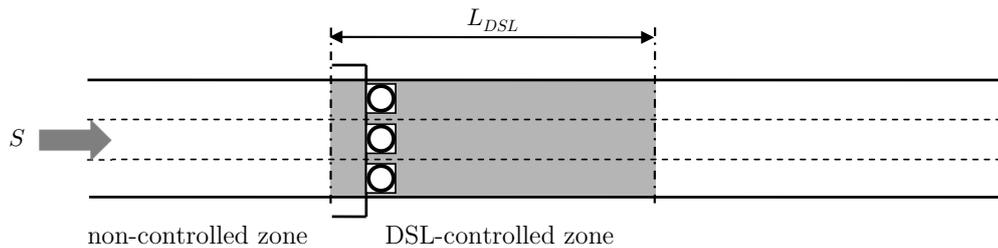
(b)

Figure 4.1 *(a) The equation of state of the capacity-lagged cell transmission model depicting different stationary traffic states and (b) an hypothetical three-lane stretch with a DSL-controlled area*

To illustrate the model behavior, consider traffic flows through a hypothetical three-lane stretch represented in Figure 4.1(b). At certain location a variable message sign indicates the speed limit has been lowered to 'V', being $v > V$. Obviously, at the same time the maximum outflowing value (i.e. capacity) for the DSL-controlled zone will be reduced. The DSL transitions from high to low speed values are expected to be done according to the

principle of flow conservation: i.e. the flow remains constant while the density suffers an increase. Unfortunately, such transition is not always observed in many proposed models (e.g. [Carlson, et al., 2010; Hegyi, et al., 2005]) temporary violating the principle. In those cases, a temporary flow decrease together with density conservation is observed during the DSL traffic state transition. It is commonly justified by considering that during the transition it is necessary the flow to be temporarily reduced to 'create' the higher traffic density of the DSL state. Anyway, it is clear that this behavior responds to a model artifact. To illustrate the model behavior different demand scenarios will be analyzed:

- **Demand A, 'S_A'.** Suppose that the upstream demand corresponds to the state A flow. As this value is lower than the capacity of the non-controlled sections, $S_A < Q^f(v)$, and no queues spill back from downstream, the upstream section will operate at state A. Once such flow crosses the DSL gantry, the vehicles will lower the speed, increase the density and maintain the same flow. Consequently, the DSL-controlled zone will operate at state D as the corresponding flow value doesn't exceed the feasible capacity of the zone, i.e. '$Q^f(V)$'. Once the vehicles cross the downstream edge of the DSL-controlled zone, they can speed up to recover speed 'v' while maintaining the same flow (i.e. state A).

- **Demand B, 'S_B'.** The demand increases further in such a value between both DSL-zone capacities, i.e. $S_B \in [Q^d(V), Q^f(V)]$. The non-controlled zone remains in free-flow state because $S_B < Q^d(v)$. However, this higher demand makes the DSL-controlled zone to enter in an instable state: the headways have become shorter, and it is challenging to maintain the same speed with such high density values. Any driver brake may destabilize the flow and activate a queue formation. If the upstream demand remains constant, the DSL-controlled zone will operate at state H. Downstream the DSL-controlled zone, the vehicles speed up to recover speed 'v' while maintaining the same flow (i.e. state B).

- **Demand C, 'S_C'.** As expected, the upstream demand increases above the DSL-zone capacity, being $S_C > Q^f(V)$. So, the DSL-controlled zone becomes an active bottleneck and a queue starts to grow upstream from the gantry location. Initially, the bottleneck discharges with the maximum capacity, '$Q^f(V)$', but when the queue reaches previous non-controlled cell, the capacity drop mechanism becomes active and the discharge flow drops to '$Q^d(V)$'. The stationary state for the DSL-controlled zone will correspond to the congested state E with a downgraded speed value, 'V^d'. In turn, the non-controlled sections will perform in state C until the growing queue doesn't reach them. Then, the flow will evolve to state E, with the correspondent shockwave. Finally, downstream the DSL-controlled zone, the vehicles will evolve to state I corresponding to the discharge flow during congestion (i.e. '$Q^d(V)$').

- **Demand C→B→A, $S_C \rightarrow S_B \rightarrow S_A$.** Consider the case where demand scenario 'S_C' has previously occurred and the stationary conditions are state F for controlled zones and state E for non-controlled. For a given time, e.g. at the end of the rush hour, the upstream demand falls to 'S_B'. The queue will not start to

dissolve, because $S_B > Q^d(V)$, although it will grow slower. Next, another demand decrease is observed corresponding to 'S_A'. In that case, the queue will start to decrease, because $S_A < Q^d(V)$. After a while, it is expected the queue will be dissolved. Note that the discharge flow will always be '$Q^d(V)$': no reason to recover the DSL freeway capacity '$Q^f(V)$' which will only happen when the whole zone will have recovered the free-flow state. So, while the queue remains in the no-control zone, it will perform in state E. Once the queue is dissolved, it will evolve to the free-flow state A. In turn, the final state for DSL-controlled zone will be D.

Among the traffic flow dynamics previously described, the capacity drop modeling and the ability to reproduce DSL strategies are not captured in the basic CTM and establish the main contribution of the research. The necessary extensions are easy to implement and fully consistent with the CTM. In fact, many previous studies have test the robustness of CTM for simulating real-world freeway segments [Gomes and Horowitz, 2006; Ishak, et al., 2006; Kurzhanskiy and Varaiya, 2010a] and for DSL-controlled freeway sections [Soriguera, et al., 2013].

4.4 Model formulation

The CL-CTM formulation is presented in two parts: Section 4.4.1 for the ordinary cells and Section 4.4.2 for the merging cells.

4.4.1 Ordinary cells

The model for ordinary cells differs from the original CTM by introducing (i) some rules to change the cell capacity depending on the traffic states governing the flow and (ii) the L-CTM. To apply the CL-CTM, the freeway is divided into N sections, with each section containing at most one DSL gantry varying the speed limit of the cell. Consequently, the discretization of real freeways with DSL installations must take into account such constraints. The model has two basic components: the mainline conservation equations and the capacity rules.

4.4.1.1 Mainline conservation equations

In order to better capture the behavior of the model, the original CTM sending and receiving functions, ([Daganzo, 1995], Equations 3a and 3b), has been split into a total of four expressions, Equations (4.1) to (4.4), one per each of the terms involved in the expressions. That way, it is easier to track which of the elements involved in the traffic dynamics govern the cells at any time.

$$T_{i,j}^1 \triangleq R_{i-1,j} \tag{4.1}$$

$$T_{i,j}^2 \triangleq V_{i-1,j-f_i} \cdot k_{i-1,j-f_i} \tag{4.2}$$

$$T_{i,j}^3 \triangleq R_{i,j} \tag{4.3}$$

$$T_{i,j}^4 \triangleq w_i\left(K_i - k_{i,j-l_i}\right) \tag{4.4}$$

Where '$T_{i,j}^1$' corresponds to the capacity, '$R_{i-1,j}$', of the sending cell '$i-1$' at time 'j', being $t = j \cdot \Delta t$ and 'Δt' the simulation period; '$T_{i,j}^2$' corresponds to what can be sent by the sending cell '$i-1$' at time '$j-f_i$' under the set DSL, i.e. '$V_{i-1,j-f_i}$'(note that if non active DSL, $V_{i-1,j-f_i} \equiv v_{i-1}$) via the density, '$k_{i-1,j-f_i}$'; '$T_{i,j}^3$' corresponds to the capacity, '$R_{i,j}$', of the receiving cell 'i', at time 'j' and '$T_{i,j}^4$' corresponds to what can be received by the receiving cell 'i' at time 'j', via the density, '$k_{i,j-l_i}$', in cell i at time $j - l_i$. Where 'f_i' and 'l_i' the two parameters introduced in the L-CTM: the lag 'f_j' is used to hinder the deterioration of accuracy when large cells are used, whereas the lag l_i is used to account for the difference in forward and backward shockwave speeds. Both parameters are expressed in terms of simulation step counts.

Finally, a reformulated expression of [Szeto, 2008], Equations 10 and 11, is posed:

$$q_{i,j} = min\{T_{i,j}^m\big|_{f_i,l_i=0}, T_{i,j}^m\} \ for \ m = \{1,2,3,4\} \tag{4.5}$$

Every time simulation step, one (or more) of the '$T_{i,j}^m$' terms may result the lowest (or dominant) one(s) in the min{} function of Equation (4.5) . Clear rules may be defined for setting which is the unique state governing every cell for a given time. To that end, a new variable is introduced to store such information: '$G_{i,j}$'. By definition, it can only exhibit four integer values: $G_{i,j} = \{1,2,3,4\}$. Such variable will play a crucial role in the election of the cell capacity. That way, the rules for updating '$G_{i,j}$' values must be carefully selected, making the election coherent with the model behavior feasibility. The following rules are stated:

$$G_{i,j} \triangleq 1 \leftrightarrow \{T_{i,j}^1 \mid (T_{i,j}^1 \leq T_{i,j}^2)\bigwedge(T_{i,j}^1 < T_{i,j}^3)\bigwedge(T_{i,j}^1 \leq T_{i,j}^4)\} \tag{4.6}$$

$$G_{i,j} \triangleq 2 \leftrightarrow \{T_{i,j}^2 \mid (T_{i,j}^2 < T_{i,j}^1)\bigwedge(T_{i,j}^2 < T_{i,j}^3)\bigwedge(T_{i,j}^2 \leq T_{i,j}^4)\} \tag{4.7}$$

$$G_{i,j} \triangleq 3 \leftrightarrow \{T_{i,j}^3 \mid (T_{i,j}^3 < T_{i,j}^1)\bigwedge(T_{i,j}^3 \leq T_{i,j}^2)\bigwedge(T_{i,j}^3 \leq T_{i,j}^4)\} \tag{4.8}$$

$$G_{i,j} \triangleq 4 \leftrightarrow \{T_{i,j}^4 \mid (T_{i,j}^4 < T_{i,j}^1)\bigwedge(T_{i,j}^4 < T_{i,j}^2)\bigwedge(T_{i,j}^4 < T_{i,j}^3)\} \tag{4.9}$$

4.4.1.2   Capacity rules

The next step consists on defining each cell capacity for the subsequent simulation step. It is considered that an upstream cell being in a congested state (i.e. $G_{i-1,j} = 4$) will make the immediately downstream one (i.e. cell i, which probably will also be in a queued state except if it is downstream the active bottleneck cell) to decrease its capacity. This behavior is consistent with the capacity drop mechanism within active bottleneck locations. When combining it, all together with lagged-density values (i.e. the density values from previous simulation steps via $l > 0$), the capacity drop phenomenon is well reproduced and the wave (or the information) with the new capacity travels together with the back of the queue growth. This hypothesis is reasonable for small simulation time steps. Anyway, it is stated that the available capacity for cell i at time '$j + 1$' depends on $G_{i-1,j}$ value:

$$R_{i,j+1} = Q^d(V) \leftrightarrow \{G_{i,j} \mid G_{i-1,j} = 4\} \text{ for } i \notin \mathcal{E} \tag{4.10}$$

$$R_{i,j+1} = Q^f(V) \leftrightarrow \{G_{i,j} \mid G_{i-1,j} \neq 4\} \text{ for } i \notin \mathcal{E} \tag{4.11}$$

where '$\mathcal{E}$' is the set of sections with on-ramps. This behavior can only be considered for ordinary links. In fact, a special case arises when cell 'i' is an on-ramp, which will be addressed in next section.

4.4.2   Merging cells

The merging cells model only differs from the CTM merge model ([Daganzo, 1995] Equations 7a-7b) by considering the merge capacity as an endogenous variable, next defined in Equation (4.12). Figure 4.2(a) presents the different variables involved. The capacity of the merge is considered to be influenced by the traffic states of its supplying branches and thus, cannot be considered as an exogenous variable [Cassidy and Bertini, 1999]. Thus, it's reasonable to assume that the capacity is influenced by many factors (e.g. merge demand, lane change maneuvers, vehicle accumulation in the queue, vehicles acceleration rate and length of the merging section, among others). Recently, [Leclercq, et al., 2011] proposed an analytical model, based on the variational theory [Daganzo, 2005], that estimates the capacity drop with respect to the demand on the on-ramp and the different model parameters. Unfortunately, both incoming roads are assumed single-lane by the model, leading to a partially unrealistic approach. In this chapter, an alternative approach to the problem is proposed by defining a simple methodology based on the Newell-Daganzo [Daganzo, 1995; Newell, 1982] priority-based merge model. The proposed model poses that in the vicinity of an area with mandatory lane changes (e.g. due to an on-ramp or a lane drop) two different mechanisms may trigger an additional restriction of the mainline capacity (i.e. capacity drop): instabilities inside a queue (type I) and the merging itself (type II). The first can arise along the length of the queue and the second in the

merging area. These capacity restrictions will not always become active and can fluctuate according to three different scenarios:

- Traffic flows through the bottleneck at freeway capacity and without capacity drop (i.e. '$Q^f(v)$').

- Queue is an active restriction limiting capacity to the queue discharge rate (i.e. '$Q^d(v)$') due to frequent lane changes or (de)accelerations or drivers' distractions.

- The merging is an active restriction limiting the capacity, 'R', being $R \in [Q^d(v), Q^f(v)]$

It should be noted that '$Q^f(v)$' capacity will only occur in areas with mandatory smooth merges under free-flow conditions. If demand increases and queues are created, the available freeway capacity may drop towards '$Q^d(v)$' value. Consequently, it is considered that when the bottleneck is performing at queue discharge rate both mechanisms (i.e. type I and II) are active. For the construction of the macroscopic model, it's essential to capture both capacity drop phenomena in the fundamental diagram calibration procedure. In fact, a recent study has partially succeeded in this issue, proposing useful fundamental diagram calibration methodologies for both triangular and inverse lambda shapes [Dervisoglu, et al., 2009].

An analytical model is proposed in Equation (4.12) for capturing the capacity dynamics of a section containing an on-ramp. It success to reproduce the capacity drop by considering two parameters, one per each capacity drop type: α for the type I and β for the type II, being $\alpha, \beta \in [0,1]$. In fact, 'α' is constant for a given fundamental diagram, while β depends on the present demand state, so '$R(\beta)$' holds. Note that the merge ratio γ is considered an exogenous model parameter, constant over time. This hypothesis is reasonable as it can be practically estimated by the ratios between the numbers of lanes on the merging approaches [Bar-Gera and Ahn, 2010], which are characteristic to every particular merge design.

$$R_{i,j} = Q_i^d(1 + \alpha\beta_{i,j}) \text{ for } i \in \mathcal{E} \tag{4.12}$$

where $\alpha = \left(\frac{Q^f(v)}{Q^d(v)} - 1\right)$, $\beta_{i,j} = 1 - \frac{|M|}{|P|}$, being '$P$' the point in the plane $q_{i,j} - r_{i,j}$ with coordinates $\left(Q_{i,j}^f \cdot \frac{Q_i^f}{Q_i^f + Q_i^{or}}, Q_{i,j}^f\left(1 - \frac{Q_i^f}{Q_i^f + Q_i^{or}}\right)\right)$ constant for a given FD network, Q_i^{or} corresponds to the on-ramp capacity and M depends on two different demand cases: if $S_{i,j} > S_{i,j}^{or}$ then $M = \left(\frac{S_{i,j}^{or}}{\gamma}, S_{i,j}^{or}\right)$, otherwise, $M = \left(S_{i,j}, S_{i,j} \cdot \gamma\right)$ (see Figure 4.2 (b)).

'Q_i^d' and $\alpha = \left(\frac{Q^f(v)}{Q^d(v)} - 1\right)$ may be calibrated from field measurements where capacity drop is observed. A particular case may be addressed if $\beta_{i,j} < 0$ (i.e. $\frac{|M|}{|P|} \geq 1$). Then, it means the merging is performing with 'Q_i^d' capacity and no additional capacity restriction may be introduced. In that case, '$\beta_{i,j}$' must be updated to $\beta_{i,j} = 0$ value. Next section will provide an example showing a congested merge where the merge capacity drops to '$Q^d(v)$' value.
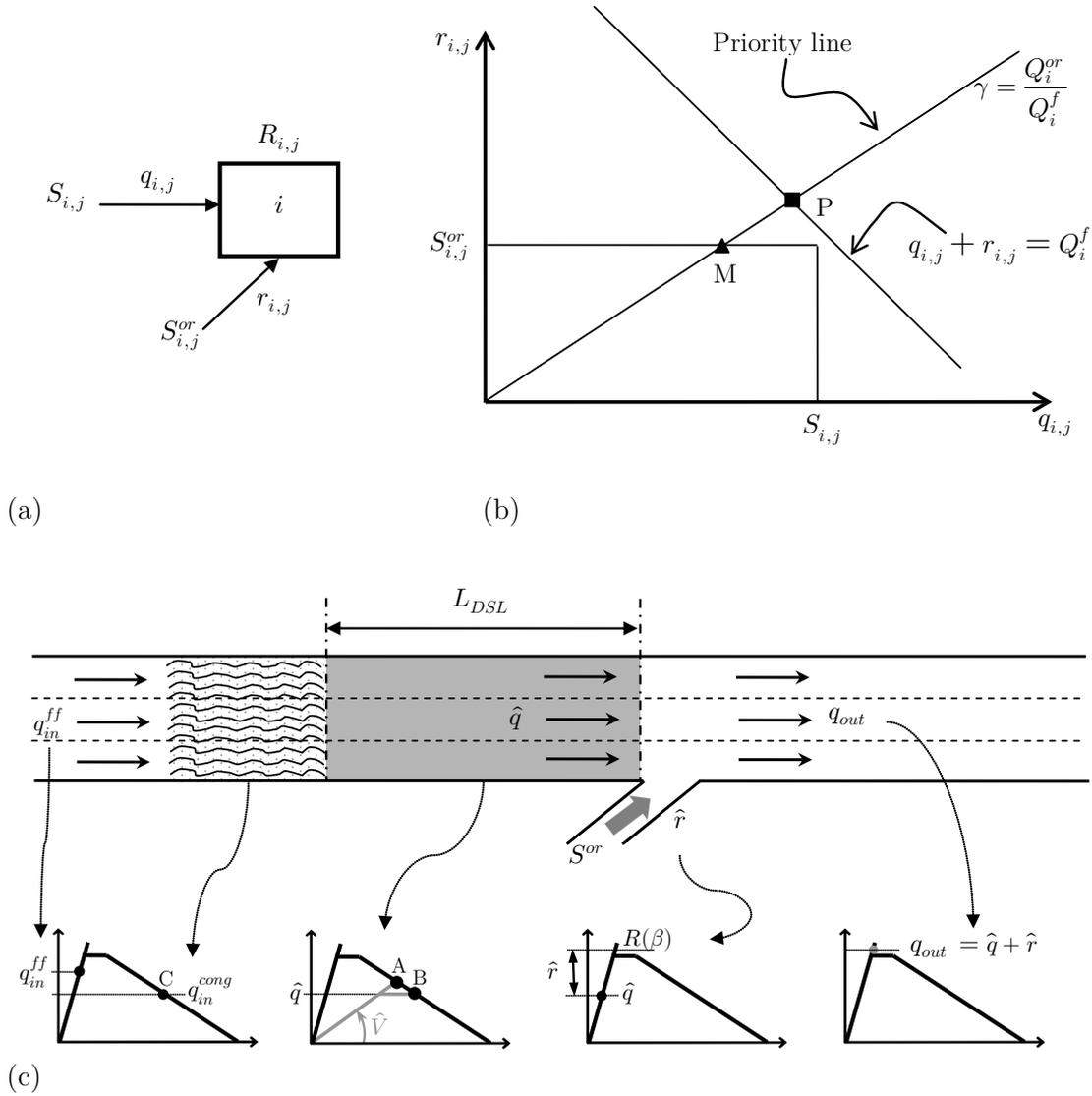
Figure 4.2 *(a) Representation of a merge (b) Diagram of feasible merging flows with points P and M involved in the merging capacity definition (c) Freeway operation under the proposed strategy*

4.5 Design of a ramp metering – dynamic speed limit coordinated strategy

The proposed strategy aims to improve drivers' equity in an active bottleneck induced by an entrance via introducing coordinated RM and DSL strategies to maximize the discharge flow. The proposed strategy explicitly considers and exploits the endogenous bottleneck capacity feature to obtain intelligent control schemes.

Consider certain demand scenario, being $S = q_{in}^{ff}$ the upstream mainline demand and S^{or} the on-ramp demand (Figure 4.2(c)), such that is creating an active bottleneck. When queues are starting to grow in both branches, the performance of the system varies in such a way that both mainline and on-ramp flows drop. Then, the system is operating at certain state corresponding to the P-point depicted in Figure 4.2(b). In this no-control involved scenario, the observed discharge flow exhibits $q_{out} = Q^d(v)$, i.e. a degradation of freeway operation has occurred due to the mainline throughput drop.

To improve this situation, it is stated that, if a suitable control scheme is introduced, capacity drop occurrence can be mitigated, then the bottleneck discharge flow, 'q_{out}', will increase, all leading to mobility performance indicators improvements (e.g. total time spent, TTS). Therefore, the objective is to maintain 'q_{out}' as far as possible from '$Q^d(v)$' capacity value, i.e. such that $q_{out} = R(\beta_{i,j}) > Q^d(v)$ for $i \in \mathcal{E}$. To this end, it can be stated that, for a given demand, exists certain $\hat{q}$-valued outflow coming from DSL-controlled sections (L_{DSL} lengthed) such that blended together with certain on-ramp rate, '$\hat{r}$', results in a final bottleneck outflow which holds Equation (4.13), corresponding to the mainline capacity for the on-ramp section (Figure 4.2(c)):

$$q_{out} = R(\beta_{i,j}) = \hat{q} + \hat{r} = Q^d(V) + \hat{r} \tag{4.13}$$

Note Equation (4.13) assumes that the final flow for the DSL-controlled stretch corresponds to the one corresponding to point B and C from Figure 4.2(c). Thus, it is assumed point A state cannot be hold due to the capacity drop phenomenon.

By considering the merge capacity expression, Equation (4.12), and assuming $S_{i,j} > S_{i,j}^{or}$ (which is common for freeway multiple lane cases) then, '$\beta_{i,j}$' is only function of $|M|$, being $M = \left(S_{i,j}^{or} \middle/ \gamma, S_{i,j}^{or} \right)$ its coordinates. In addition, it can also be considered $S_{i,j}^{or} \equiv \hat{r}$, due to the RM behavior. That way, Equation (4.13) can be rewritten as:

$$Q^d(v) \cdot \left(1 + \alpha \cdot \left(1 \pm \frac{\hat{r}\sqrt{\frac{1}{\gamma^2}+1}}{|P|} \right) \right) - Q^d(V) = \hat{r} \tag{4.14}$$

By isolating '$\hat{r}$' from Equation (4.14) and applying $Q^d(V) = \frac{\tilde{w}\hat{V}K}{\tilde{w}+V}$ where $\tilde{w} = \frac{Q^d(v)}{K-k^\lambda}$ and $k^\lambda = \frac{Q^d(v)}{v}$, an explicit expression in terms of 'V' is obtained:

$$\hat{r}(V) = \frac{|P|\left((1+\alpha) \cdot Q^d(v) - \frac{\tilde{w}VK}{\tilde{w}+V} \right)}{|P| + \alpha Q^d(v) \pm \sqrt{\frac{1}{\gamma^2}+1}} \tag{4.15}$$

The strategy will perform with a fixed DSL value (i.e. certain 'V') in a set of DSL-controlled sections, $\mathbb{C}$, such that $i \in \mathbb{C}$ and for any 'j' such that $t \in [t_{on}, t_{off}]$ (i.e. within the

control period). The effectiveness of such strategy, captured with the TTS indicator, is based on the suitable election of both time trigger parameters. 't_{on}' must be selected at the time the total demand (mainline plus on-ramp) exceeds the merge capacity. In turn, 't_{off}' must be never selected before the instant when the present speed is lower than 'V^d' value. If the condition for 't_{on}' activation stops to hold, the control can be deactivated. Within the period where the control is active and for sections $i \in \mathcal{E}$, it is assumed $\hat{r}(V) + S_{i,j} > R_{i,j}$ holds, implying the system operate following (4.16) and(4.17):

$$r_{i,j} = \hat{r}(V) \text{ for } i \in \mathcal{E} \tag{4.16}$$

$$f_{i,j} = R_{i,j} - \hat{r}(V) \text{ for } i \in \mathcal{E} \tag{4.17}$$

In turn, Equation (4.17) corresponds to '$Q^d(V)$' by the flow conservation principle. Next, on-ramp queue must be computed coherently:

$$\delta_{i,j} = max\{0, \delta_{i,j-1} + \overline{S_{i,j}^{or}} - \hat{r}\} \text{ for } i \in \mathcal{E} \tag{4.18}$$

where '$\overline{S_{i,j}^{or}}$' is the demand of vehicles aiming to enter the freeway, but not yet queued, at time 'j'.

Referring to the V-value election, it must be considered that it closely depends on the incoming demand. The suitable election of 'V' for obtaining a TTS amelioration in the whole system, results in a trade-off between the time spent in the mainline and in the on-ramp: the higher 'V' is set the lower the TTS in the mainline and the higher the TTS in the on-ramp. Next section provides an enlighten example with promising results.

4.6 Simulation results

The network consists of 33 equal sized cells (0.1 km each) with one-lane on-ramp located at kilometer post 2.9 (Figure 4.3(a)) preceded by a DSL-controlled section with length L_{DSL}= 0.1km. The parameters involved in the test are: v=90 km/h, w =18 km/h, K=420 veh/km, Δt=4 sec, $Q^f(v)$ =6300 veh/h, $Q^d(v)$ =5670 veh/h, l=2 simulation steps and f=0 simulation steps. For simplicity, the example will consider the RM strategy is always active, whilst the DSL strategy will operate for a set of 'V' such that $V = [40, 50, 60, 70, 80, 90]$ km/h for $t \in [0, 45]$. This election makes easier to obtain insights about the behavior of the proposed strategy. The demand profiles shown in Figure 4.3(b) are used for the mainline input and for the on-ramp. Different performance indicators are calculated over the 30 min of simulation whereby the last 18 min is a cool-down period with no inflows in all the links in order to have comparable traffic conditions (i.e. same total travelled distance) for all considered scenarios.

(a)

(b)   (c)
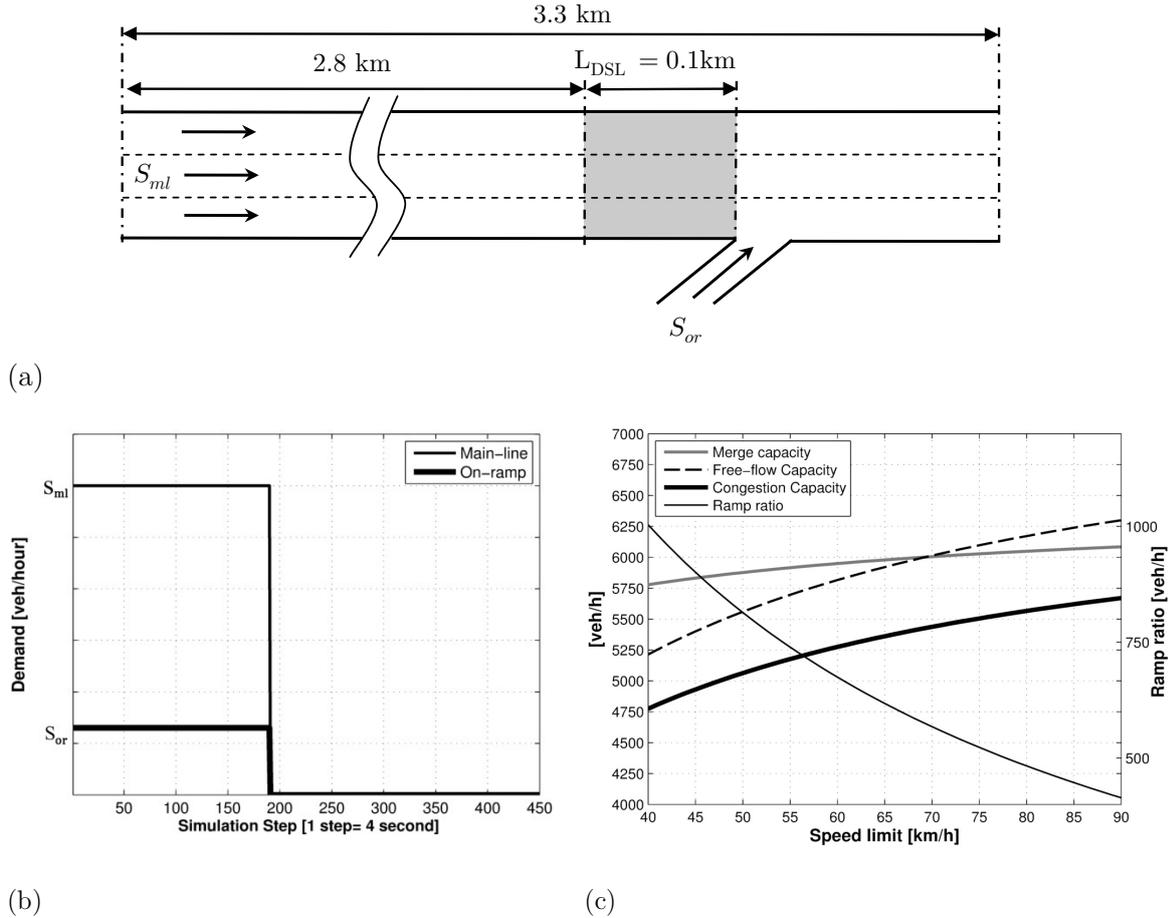
Figure 4.3 *(a) The three-lane simulation test freeway (b) Demand profile for the DSL-RM simulation case (c) Function plots for $R(V)$, labeled as 'Merge capacity'; $Q^d(V)$, labeled as 'Congestion capacity'; $Q^f(V) = \frac{wVK}{w+V}$, labeled as 'Free-flow capacity' and for the secondary ordinates axis, $\hat{r}(V)$, labeled as 'Ramp ratio'. The proposed parameters were here considered.*

Considering the previously defined parameters, the merge capacity, 'R', (by considering the square negative root value case in Equation (4.14) is computed for the proposed set of 'V' values (Figure 4.3(b)). That way, different pairs $(\hat{r}, V)$ are suggested for the strategy, all exhibiting reasonable values (i.e. $\hat{r} < 1500$ corresponding to the on-ramp capacity and $Q^d(v) < R < Q^f(v)$). The experiment consists on running the strategy considering different demand ratios S_{or}/S_{ml} and 'V' values. Main simulation results are gathered in Table 4.1 to support the former assertion.

The results show TTS which monotonically decrease with higher 'V' values. Thus, the higher the 'V' value is the higher the strategy benefits are (i.e. greater $\triangle$TTS reductions with respect the no-control case). However, for transportation applications an equity criterion may also be considered: the maximum acceptable waiting time in the on-ramps. It will be fixed in 4min, according to Minnesota's freeways thresholds [Geroliminis, et al., 2011]). That way, only the simulated scenarios exhibiting an average on-ramp delay under this threshold will be considered feasible. Consequently, the case presenting the closest (but

not higher) delay value will correspond to the optimal speed limit value for certain demand, i.e. 'V^*' underlined in Table 4.1. It can be observed how lower demand ratios correspond to lower 'V^*' values and, more significantly, to lower delay ratios, all leading to more equitable results as the network delays are definitely better balanced.

For the controlled situation, from t=0 a ramp meter is set to $\hat{r}$ =814 veh/h and DSL to V=50 km/h. Consequently, on-ramp queue growth is triggered and the mainline flow entering in the merge is constrained to the DSL cell capacity, corresponding to $Q^d(50)$ =5063 veh/h. That way, the strategy assures the merge capacity is fixed at 5877 veh/h, higher than the merge capacity observed for the non-controlled case, which leads to the observed time savings due to the capacity drop avoidance. The on-ramp queue grows with higher pace than for the no-control case, it stops to increase when the on-ramp demand shifts to the zero value at t=190 and finally, it is completely dissolved at t=269.

Table 4.1 *Simulation results and its variation with respect the non-control case, for different speed limit values, V, and S_{ml}/S_{or}, always considering $S_{ml} + S_{or}$ =6900 veh/h. (Legend: or≡on-ramp and ml≡main-line)*

	V [km/h]	No control	40	50	60	70	80	90
$S_{ml}/S_{or}=4$ TTS (ml+or)	[veh · h]	*695,39*	654,98	579,56	534,47	505,74	492,70	499,60
ΔTTS	[%]		-5,81%	-16,66%	-23,14%	-27,27%	-29,15%	-28,15%
or delay / ml delay	[-]	*1,86*	4,58	15,42	45,52	176,12	7467,01	6591,92
or delay	[min]	*1,20*	2,38	4,40	6,62	9,05	11,23	13,01
$S_{ml}/S_{or}=5$ TTS (ml+or)	[veh · h]	*758.29*	725.22	647.32	593.57	558.32	534.69	516.31
ΔTTS	[%]		-4.36%	-14.64%	-21.72%	-26.37%	-29.49%	-31.91%
or delay / ml delay	[-]	*0.31*	1.38	6.02	16.16	38.01	92.12	275.91
or delay	[min]	*0.24*	0.92	2.61	4.46	6.49	8.74	10.87
$S_{ml}/S_{or}=6$ TTS (ml+or)	[veh · h]	*748.43*	760.69	698.89	640.93	600.98	573.16	553.96
ΔTTS	[%]		1.64%	-6.62%	-14.36%	-19.70%	-23.42%	-25.98%
or delay / ml delay	[-]	*0.00*	0.00	2.47	7.78	17.79	36.57	72.70
or delay	[min]	*0.00*	0.00	1.34	2.92	4.67	6.59	8.74
$S_{ml}/S_{or}=7$ TTS (ml+or)	[veh · h]	*740.93*	761.86	732.66	679.29	635.85	604.97	583.30
ΔTTS	[%]		2.82%	-1.12%	-8.32%	-14.18%	-18.35%	-21.27%
or delay / ml delay	[-]	*0.00*	0.00	0.63	3.90	9.87	20.16	37.22
or delay	[min]	*0.00*	0.00	0.38	1.77	3.30	4.98	6.86

no-control case   V=50 km/h
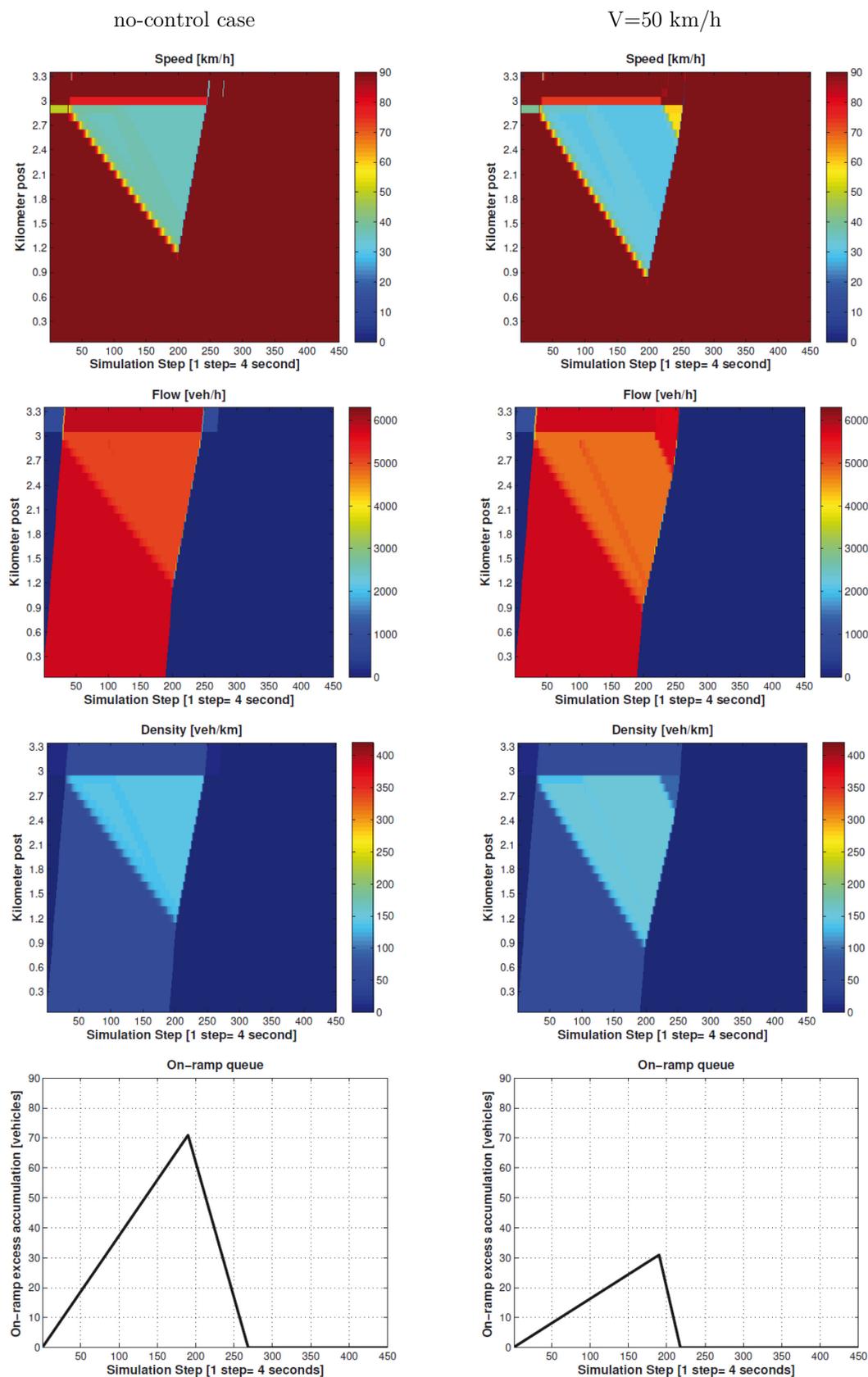
Figure 4.4 *Contour plots (1ˢᵗ-3ʳᵈ row) and (4ᵗʰ row) accumulated vehicles in the on-ramp of the simulation results, comparing the no-control case (first column) with the V=50 km/h case (second column) considering the demand case S_{ml}/S_{or} =5.*

4.7 Conclusions and further research

One of the main goals of this chapter has been the development of an innovative extension of CTM, i.e. CL-CTM, which has arisen capable to reproduce certain macroscopic phenomena not yet modeled with first order models. Among the included extensions, excels the ability to incorporate DSL strategies by introducing a modified inverse-lambda shaped fundamental diagram. However, for fully exploiting one of the main DSL benefits (i.e. to delay the onset of congestion by decreasing the mainstream flow) an approach to the capacity drop phenomenon is required. The presented model covers this requirement by presenting a CTM fully consistent approach. In fact, the CL-CTM differs only in the addition of some capacity rules (i.e. Equations (4.10) and (4.11)) and the L-CTM.

A new coordinated RM and DSL freeway traffic management strategy has also been proposed. Notably, it is mainly based in a novel endogenous merge capacity model here proposed. An illustrative example was presented under different control scenarios. It was shown in detail how the coordinated strategy succeeds in improving the discharge rate flow governing the merge which implies significant equity benefits for the freeway performance.

A main drawback of the posed strategy refers to capacity characterization which is assumed. It has been frequently described (e.g. [Geroliminis, et al., 2011]) how capacity is strongly influenced by many different factors affecting the fundamental diagram calibration accuracy (e.g. [Dervisoglu, et al., 2009]). Thereby, a pending issue may address the strategy considering the critical value of density at which capacity is observed because it is less sensitive and more stable, allowing the opportunity for more effective control.

Part III

Empirical assessment

Overview

Following the scientific method, any hypothesis constructed in order to answer a question (e.g. a traffic flow model), needs to be tested with an experiment. This is one of the main objectives of this empirical part of the thesis. In addition a cost-benefit assessment of DSL strategies, based on empirical measurements, is also provided. Data derive from C32 freeway, and captures different evolution stages along the progressive implementation of the first DSL installation promoted in Spain.

DSL are used on freeways to manage traffic flows with the intention of improving capacity. Chapter 5 is devoted to verify this assertion. The investigation focuses on the relationships between flow and occupancy (as an indicator of traffic flow density) under different speed limit scenarios. Freeway capacity and critical occupancy are estimated.

Chapter 6 presents a generic method for quantitatively evaluating the benefits and costs of ATM strategies. This is achieved by computing the performance of an objective function in before/after scenarios. The objective function considers delays, energy consumption, pollutant emissions and accident risk. Remarkably, all the required parameters can be easily calibrated from commonly available traffic data, only including aggregated data. Two main difficulties arise in the application of the method on the test corridor: (i) It is needed to obtain comparable scenarios by assuring the same traffic demand. As this is almost impossible to be accomplished with real data from different periods of time, the need for traffic simulation appears. (ii) Simulation implies to assume some drivers' behavior modification due to DSL. So, in both, empirical and simulation approaches, limitations arise and the results should be considered only as approximations. Considering these restrictions and in order to preserve the simplicity of the model validation, the capacity drop effect is neglected. In fact, the calibrated fundamental diagrams support this assumption on the test site.

Empirical evidence of dynamic speed limit effects on freeway traffic

5.1 Introduction

As has extensively been exposed in Chapter 2, no agreement exists on the effect of DSL in aggregated traffic flow behavior. Indeed, over time, two general views have evolved on the use of speed limits. The first emphasizes the homogenization effect (e.g. [Smulders, 1992; Zackor, 1979]) which may lead to capacity increments, whereas the second is more focused on avoiding or mitigating traffic flow breakdown by reducing the input flow at bottlenecks by means of speed limits (e.g. [Carlson, et al., 2010; Hegyi, et al., 2005]). The present chapter focuses on the first view. As previous works presenting empirical results were non-conclusive (e.g. [Papageorgiou, et al., 2008; Soriguera, et al., 2013]), there is need to more studies in order to obtain robust facts which are expected to clear up the question.

The present chapter tries to fill this gap by presenting an exhaustive empirical analysis of the DSL effects only considering aggregated traffic data and current speed limits for a significant number of rush hour periods. To address this evaluation, an empirical approach is proposed throughout the chapter. First, Section 5.2 presents the motorway test site located in Barcelona, together with the available data and the methods proposed for data validation. The proposed methodology for the fundamental diagram (FD) calibration is described in Section 5.3. Meanwhile, Section 5.4 is devoted to analyze the obtained results focusing on the drivers' DSL compliance and, while Section 5.5 studies its effect on the

traffic flow characterization. Finally, conclusions and avenues for further research are presented in Section 5.6.

5.2 The freeway site and available data

The freeway stretch selected for the test is part of the C32 freeway towards Barcelona with a length of approximately 5.7 km (Figure 5.1) containing two off-ramps and an on-ramp. The most downstream off-ramp (i.e. S2) triggers a bottleneck which creates recurrent congestion periods in the morning. Note that the mainline slow lane is devoted to S2 exiting ramp, producing a lane reduction in the trunk from three to two lanes. This fact justifies the test site location and sets a suitable frame for the proposed analysis. As a result, the fast lanes and the slow lane may be considered as belonging to different sections. Thus, they are named '54+490a' and '54+490b', respectively. For more details of the C32 freeway performance, it is recommended to check available literature (e.g. [Soriguera, et al., 2013]).

The stretch is equipped with 5 loop detector stations (measuring the average speed, flow and occupancy every $\Delta t =1$ min time intervals) and 4 Variable Message Signs (VMS) gantries displaying dynamic speed limits (between 40 and 80 km/h), two of them (i.e. R-VMS 2 and R-VMS 3) also contain a radar installation. Definitions for each of these quantities are given below.

$n^{\theta}_{(s,t)}$ Counting of vehicles crossing section 's', corresponding to each loop detector, at lane 'θ' during data interval 't' [vehs/min].

$n_{(s,t)}$ Aggregated counting of vehicles crossing section 's' during data interval 't' [vehs/min].

$o^{\theta}_{(s,t)}$ Occupancy of section 's'at lane 'θ' during data interval 't' [%].

$o_{(s,t)}$ Aggregated occupancy of section 's' during data interval 't' [%].

$v^{\theta}_{(s,t)}$ Time mean speed of vehicles crossing section 's' at lane 'θ' during data interval 't' [km/h].

$v_{(s,t)}$ Aggregated mean speed of vehicles crossing section 's' during data interval 't' [km/h].

$V_{(s,t)}$ Posted DSL governing section 's' during data interval 't' [km/h].

Two speed limit regulation measures in C32 highway are analyzed. Both were approved by the Autonomous Government of Catalonia on 2008 in order to improve air quality in the Barcelona metropolitan region. The first one came into force on January 1st 2008 limiting speeds to 80 km/h on major highways around Barcelona. The second one, applied one year later, was based on a dynamic speed limit management (DSL) maintaining the maximum speed limit of 80 km/h. Both considering 87 morning rush hour periods (from October

2008, February 2009, May and June 2010) and the presented speed limit regulations, three different scenarios are defined grouped by year criterion (i.e. 2008, 2009 and 2010).
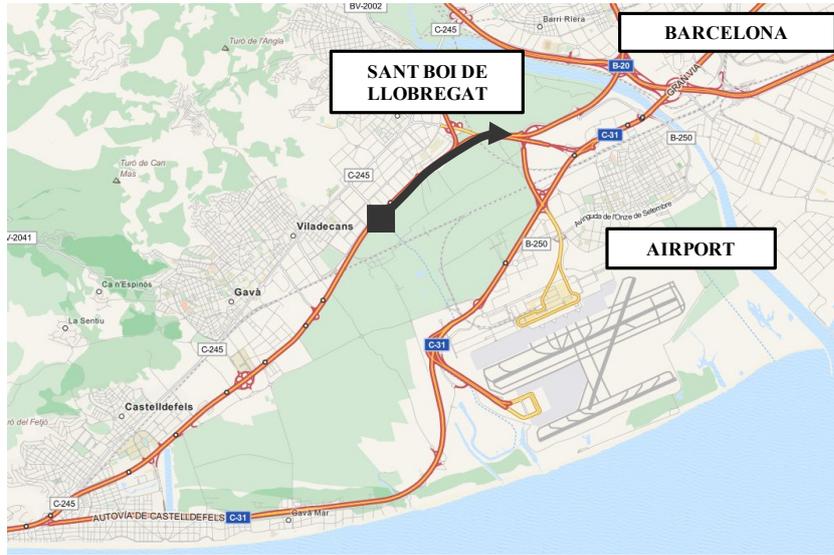
(a)
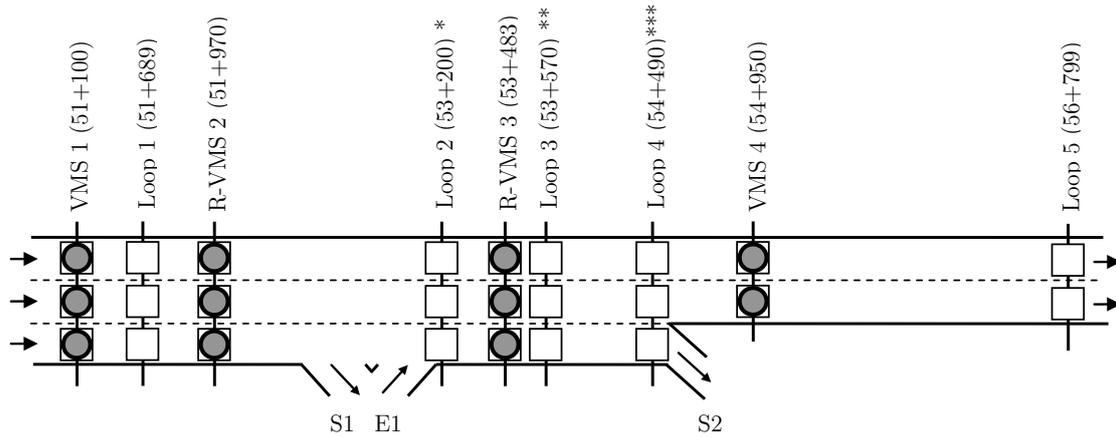
*   Error: drift in entering vehicles counts
**   Error: drift in vehicles counts for one lane in 2010 data
*** Section '54+490b' corresponds to the slow lane and '54+490a' to the rest of lanes

(b)

Figure 5.1 *Test site: (a) overview map, (b) stretch layout indicating the different installations with its respective kilometer points and errors observed in the available data. Font: Figure 5.2 (a) was obtained from © OpenStreetMap contributors, CC BY-SA.*

Taking the original loop detector data, data presenting very unreasonable values of speed, occupancy or flow (such as negative or very extreme values) are not considered for further computations. With this procedure, around 5-6% of the original data are removed for further analysis (see Table 5.1), but many loops present less than 1% values (i.e. Loop 1 and 3). However, other kind of errors may remain hidden, e.g. drifts in vehicles counts caused by loop detectors defects, although they can be easily detected with oblique curves of cumulative traffic flow (i.e. *N*-curves, [Cassidy and Windover, 1995]). In particular, such

errors were systematically observed in Loop 2 due to mainline congestion caused by E1 (Figure 5.3(a)) and also in Loop 3 for year 2010 data (Figure 5.3(b)). The affected loops present drifted curves falling below its corresponding position, i.e. Loop 2 must be the first beginning from the top, Loop 3 the second and so on. This kind of errors is inherent to the loop detectors technology. That is, during intense congestion episodes, vehicles are covering the magnetic loop for 'long time' leading to a biased counting of vehicles. Fortunately, speed and occupancy measurements are not affected.

(a)   (b)
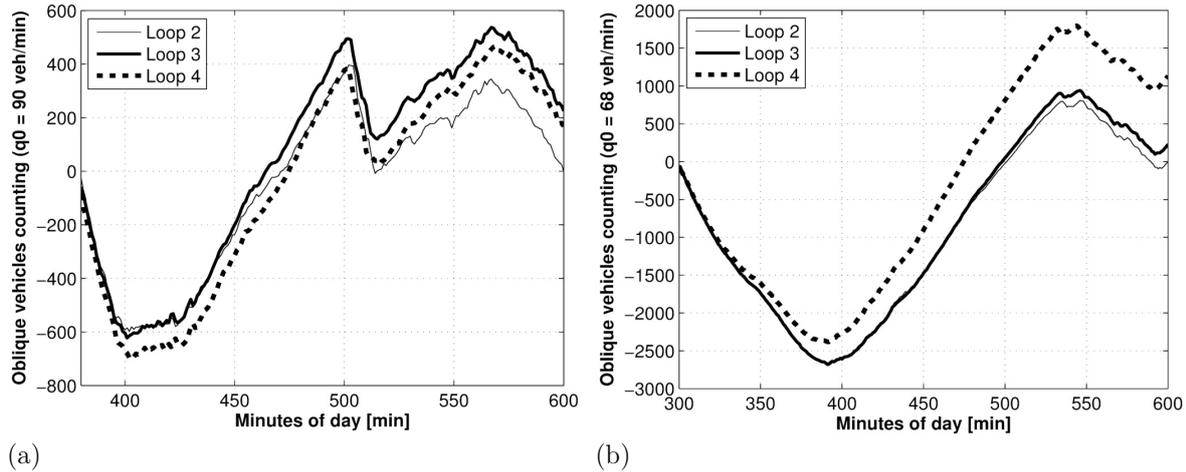
Figure 5.3 *Oblique plot with N-curves for Loops 2 to 4, summed over all lanes, corresponding to rush hour periods of (a) February 17th 2009 and (b) May 4th 2010.*

Table 5.1 *Definition of speed related categories*

Category	Description	Condition
A	DSL doesn't affect traffic performance. Too high DSL.	$v_{(s,t)} < V_{(s,t)} - \delta$
B	DSL affects traffic performance.	$\left\| v_{(s,t)} - V_{(s,t)} \right\| \le \delta$
C	Lack of DSL compliance. Too low DSL.	$v_{(s,t)} > V_{(s,t)} + \delta$
-	It is considered wrong data if the condition is not fulfilled	$0 < o^{\theta}_{(s,t)} \le 100$ $v^{\theta}_{(s,t)} \le 150 \text{ km/h}$ $n^{\theta}_{(s,t)} \le 44 \text{ veh/min}$

5.3 Data analysis methods

To obtain an insightful macroscopic analysis of DSL effects with empirical data, different methodologies have been sequentially applied. Every one of the proposed steps' aim is to eliminate, or at least substantially reduce, the noise in the measurements. The initial method step begins with very broad data filters and getting more specific with each following step. Briefly, the first step consists in a stratification of data available depending on its relative speed value with respect the current DSL. Next, the method tries to obtain the longest possible series of consecutive data intervals which verify some of the properties

expected for stationary traffic states. These intervals become 'candidates' for being stationary traffic states. Finally, the stationarity of these 'candidates' is tested to systematically construct bivariate diagrams under DSL conditions.

5.3.1 Data labeling and candidates' definition.

A prime task to be tackled is exploring the effectiveness of DSL system. Recently, [Long, et al., 2012] explore drivers perceptions in the implementation of a DSL system in St. Louis (USA) providing insightful facts which are coherent with the ones reported in the C32 case [Soriguera, et al., 2013; Torné, et al., 2011]. In fact, significant levels of dissatisfaction and resistance to innovation were observed, leading to lack of compliance with posted DSL speeds. Frequently, such resistance is based on perceptions of ineffectiveness and of high opportunity costs, strongly related with the drivers' perception of the value of time.

The next proposed methodology addresses these concepts in a simple way. To this end, data is stratified in three categories, '$\mathbb{C}$', labeled as 'A', 'B' or 'C' depending on its relative value by considering a tolerance parameter named 'δ', which is set to 10 km/h (see Table 5.1). The former stratification, combined with information about the posted DSL value and with the occupancy measurements, forms a three components frequency vector, i.e. (o, V, i) being $i \in \mathbb{C} = \{A, B, C\}$. The frequency of each vector occurrence allows obtaining many insights about drivers' behavior and the DSL governing algorithm. Section 5.4 analyzes the results for the different scenarios providing a full assessment for the DSL affectation and its compliance.

In addition, the latter step has also provided certain time-series data belonging to different 'i' categories, being $i \in \mathbb{C}$ (see Figure 5.4 top), the 'candidates' to which the below stationary criteria will be applied. Now, the expected goal is to obtain long data spans where all the constituting shorter periods (i.e. 1-min data) verify to hold the same DSL value and most of them the same 'i' category. To that end, short periods, named as 'holes' and exhibiting 'i' values different than the ones of its surrounding intervals, are accepted. So, the macroscopic perspective of the methodology is preserved, assimilating punctual data dysfunctions or oscillations. Figure 5.4 shows an example of two candidates of the same category with a hole between them which will become into a single candidate. '$T_{(i,j)}$' is defined as the time length of span 'j' and category 'i' computed as $T_{(i,j)} = \tau_{(f,j)} - \tau_{(o,j)}$, where '$\tau_{(o,j)}$' is the first time instant and '$\tau_{(f,j)}$' is the last time instant of span 'j', and $T_{(m,n)}$ is the duration of interval 'n' and category 'm' which is evaluated to be fused with span '$T_{(i,j)}$' and '$T_{(i,j+1)}$', such that $i, m \in \mathbb{C}$ and $i \neq m$.
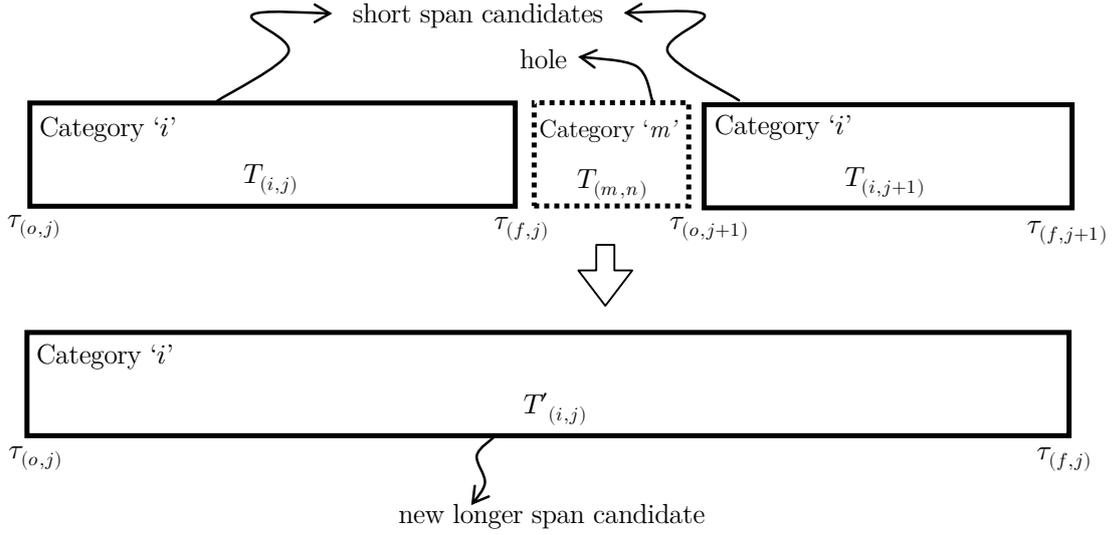
Figure 5.4 *Example of two candidates of the same category with a hole between them (top) next fused into one longer span candidate (bottom).*

Then, if '$T_{(m,n)}$' is sufficiently short in comparison with $T_{(i,j)} + T_{(i,j+1)}$ and $T_{(m,n)} \leq T_{max}$, the hole is shifted to category 'i', so $T_{(m,n)} \subset T'_{(i,j)}$. Equation (5.1) presents the condition to be verified:

$$T_{(i,j)} + T_{(m,n)} + T_{(i,j+1)} \leq (1 + \alpha)\big(T_{(i,j)} + T_{(i,j+1)}\big) \tag{5.1}$$

Being $i \neq m$, $m \neq$ D, $\alpha = 0.3$ and $T_{max} = 3$ min.

An special case arises when '$T_{(m,n)}$' matches with the first or the last interval of the rush hour time period. In that case, $T_{(i,j)} = 0$ and $T_{(i,j+1)} = 0$, respectively.

5.3.2   Construction of bivariate diagrams with near-stationary traffic data

It is known that, when traffic conditions are approximately stationary, reproducible bivariate relations exist, e.g. flow and occupancy. For this purpose, [Cassidy, 1998] defined the different steps involved in the construction of a bivariate diagram. However, its identification of nearly stationary periods by visual inspection of cumulative curves arises to be a non functional tool when big amounts of data are to be analyzed, as happens in the present case. The new approach is fully based on Cassidy's technique but differs on the way the stationary periods are detected by presenting a simple and robust method to replace the visual inspection procedure. Remarkably, it also differs from some FD calibration blind methods (e.g. [Dervisoglu, et al., 2009; Papageorgiou, et al., 2008]). In that cases, the methods do not consider that under non-stationary traffic conditions the

measurements obtained do not necessarily fall on the curve describing near-stationary traffic, hence the big dispersion observed in their bivariate diagrams. Actually, it must be noted that the main goal of the new proposed method is reduced to characterize two basic FD parameters which play a crucial role in the freeway performance: the maximum flow (i.e. the capacity) and the occupancy/density value where it happens (i.e. the critical occupancy/density). In the same way, the near-stationary condition helps to assure the obtained results mimic the representative behavior of a section, avoiding the non-stationary transition states.
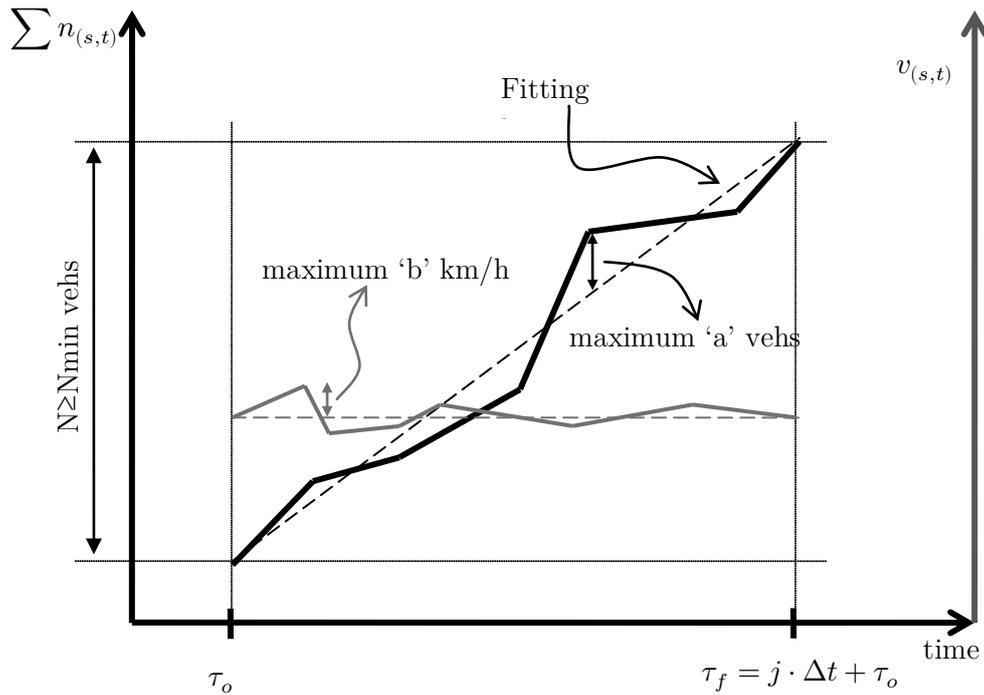
Legend:

τ_o First time period considered

τ_f Last time period considered

Figure 5.5 *Detecting stationarity from vehicle cumulative count.*

As previously referred to, in the detection of the stationary traffic states lies one of the key points for well reproducing the different bivariate relations. These stationary time intervals will be obtained from inspecting the multiple candidates obtained from the previous method step. This inspection consists of verifying two criteria, one in relation to flow and the other affecting speed. If almost constant speed and flow values were encountered within the span, approximate stationarity can be accepted. In particular, flow is considered to be almost constant if the measured cumulative count do not deviate more than 'a' vehicles from cumulative count related to the average flow within the period, Equation (5.2) and Figure 5.5.

$$\frac{\sum_{t=\tau_o}^{\tau_f} n_{(s,t)}}{\tau_f - \tau_o + 1} k\Delta t - a \leq \sum_{t=\tau_o}^{k\cdot\Delta t+\tau_o} n_{(s,t)} \leq \frac{\sum_{t=\tau_o}^{\tau_f} n_{(s,t)}}{\tau_f - \tau_o + 1} \cdot k\Delta t + a \qquad (5.2)$$

where 'a' is a parameter to be calibrated and $k = 1 \div \frac{\tau_f - \tau_o + 1}{\Delta t}$.

The second stationarity criterion deals with speed measurements. Being at least 'N_{min}' aggregated vehicles, speed was considered to be almost constant during the whole span considered, if the average speed of all the constituting shorter periods 't' (see Equation(5.3)) was contained within the confidence interval limits plus an additional value, 'b', of each particular measurement.

$$\frac{\sum_{t=\tau_o}^{\tau_f} n_{(s,t)} \cdot v_{(s,t)}}{\sum_{t=\tau_o}^{\tau_f} n_{(s,t)}} - v_{(s,t)} \cdot e_{V_{(s,t)}} - b \leq v_{(s,t)} \leq \frac{\sum_{t=\tau_o}^{\tau_f} n_{(s,t)} \cdot v_{(s,t)}}{\sum_{t=\tau_o}^{\tau_f} n_{(s,t)}} + v_{(s,t)} \cdot e_{V_{(s,t)}} + b \qquad (5.3)$$

where 'b' is a parameter to be calibrated and $e_{V_{(s,t)}} = \sqrt{\frac{CV_V^2}{n_{(s,t)}}}$ considering a 67% degree of confidence, being 'CV_V' the speed coefficient of variation taken as 0.1375 [Soriguera and Robusté, 2011].

Finally, for obtaining large stationary periods, the algorithm aggregates as much vehicles as possible while Equation (5.2) and (5.3) continue to hold. Once the process has finished, it is expected that multiple stationary states would have been obtained with reasonable time durations of at least 3-4 min, as proposed by [Cassidy, 1998]. Every stationary state lasts for 'T' minutes, aggregates $N \geq N_{min}$ vehicles, belongs to one particular '$\mathbb{C}$' category and exhibits a unique V-value. Once the aggregations have been computed, bivariate plots are depicted considering the mean occupancy and flow values of all the stationary spans considered. Additionally, (i) some error indicators are computed (i.e. '$\sigma(q)$' and 'FE'), whose information may be useful for validating the obtained results and (ii) a pair of mesoscopic variables (i.e. 'r_o' and 'r_v') are also computed for capturing the different behavior among the lanes, together with its corresponding error estimators (i.e. ' $\sigma(r_o)$' and ' $\sigma(r_v)$') (see Table 5.2). Note that all the error indicators are computed by considering the flow value among all the considered stationary spans of flow.

The next step pursues to identify representative points (i.e. average flow and occupancy stationary values) in order to easily characterize the two basic parameters concerning the FD diagrams calibration (i.e. capacity and critical occupancy). The proposed process divides the occupancy axis into intervals of equal length, i.e. Δo=5%, in the same way as was proposed by [Papageorgiou, et al., 2008]. For every range of occupancies, i.e. $[0, \Delta o]$, $[\Delta o, 2\Delta o]$, $[2\Delta o, 3\Delta o]$..., one single value represents each magnitude (i.e. '$\hat{o}$', '$\hat{q}$', '$\hat{r}_o$' and '$\hat{r}_v$'). All the depicted points are obtained via a weighted mean which considers the counting of vehicles, 'N_j', of every stationary 'j' state previously computed, e.g. the flow is computed as $\hat{q} = \frac{\sum_{j=1}^{n} \bar{q}_j \cdot N_j}{\sum_{j=1}^{n} N_j}$. Their corresponding error estimators are also computed, i.e. '$e(q)$' and '$\widehat{FE}$' (see Table 5.2). Note that, '$e(q)$' captures the temporal variation for all the stationary spans (i.e. for every constituting 1-min aggregated data) falling within certain

range of occupancies. That way, the error indicators are not dispersion estimators for the bivariate diagrams constructed with stationary states. Together, for the points encountering within every Δo-sized range of occupancies (i.e. before being aggregated in the latter process) an empirical first order curve is fitted and so, a calibrated FD is obtained.

Table 5.2 *Different variables associated with the bivariate diagrams*

Variable	Description	Expression	Units
$\sigma(q)$	Standard deviation of 'q' along the stationary period for a stationary span	$\sqrt{\dfrac{\sum_{t=\tau_o}^{\tau_f}(n_{(s,t)}-\bar{q})^2}{T}}$	veh/h, km/h
$\overline{r_o}, \overline{r_v}$	Mean 'o' and 'v' range among lanes for a stationary span	$\dfrac{\sum_{t=\tau_o}^{\tau_f}[\max(\cdot_{(s,t)})-\min(\cdot_{(s,t)})]}{T}$	%, km/h
$\sigma(\overline{r_o}), \sigma(\overline{r_v})$	Standard deviation of the mean 'o' and 'v' range among lanes for a stationary span	$\sqrt{\dfrac{\sum_{t=\tau_o}^{\tau_f}\left[\left[\max(\cdot_{(s,t)})-\min(\cdot_{(s,t)})\right]-r_{.}\right]^2}{T}}$	%, km/h
$\overline{FE}$	Fractional error for a stationary span [Cassidy and Bertini, 1999]	$\sqrt{\dfrac{\sigma(q)^2}{\bar{q}/N}}$	%
$e(q), e(r_o), e(r_v)$	Error estimation of 'q', 'r_o' and 'r_v' for every 'j' stationary state within the corresponding range of occupancies	$\pm\sqrt{\dfrac{\sum_{j=1}^{n}(N_j)^2\sigma(\cdot)_j^2}{\left(\sum_{j=1}^{n}N_j\right)^2}}$	veh/h, km/h, %
$\widehat{FE}$	Mean fractional error for every stationary 'j' state within the corresponding range of occupancies	$\dfrac{\sum_{j=1}^{n}FE_j\cdot N_j}{\sum_{j=1}^{n}N_j}$	%

5.4 Drivers' compliance and algorithm accuracy assessment

Remember that, for every section and year, a three component (i.e. (o, V, i) being $i \in \mathbb{C} = \{A, B, C\}$) frequency matrix is obtained. To simplify the results interpretation, different manipulations are done. First, the 3-dimensions matrix is reduced to 2-dimensions

by stratifying according to category, 'ℂ'. Second, data is manipulated in such a way that, for a given (o, V) pair of values, section and year, the addition of every normalized frequency rate across all 'ℂ' categories results equal to 100% (Equation (5.4)).

$$\sum_{i \in \mathbb{C}} (o, V, i) = 100\% \qquad (5.4)$$

In that way, different patterns can be observed depending on the category. For a given (o, V) pair, big percentage values in category 'A' mean malfunctioning in the DSL algorithm, which is greatly influenced by the relative position between the particular loop and the VMS gantry. If considering category 'B', it is desirable to find big percentage values. In fact, the higher the value, the better the algorithm efficiency and the drivers' DSL observance. However, category 'C' may shade interesting information about the lack of DSL compliance and its capability to act as a mainline metering. To this end, two effects will be studied (i) the effect of the distance between the loop localization and the radar installations and (ii) the trends for the region with low values of 'o' and 'V'. For visualization purposes, every particular cell is colored depending on the category presenting the higher frequency values for that particular (o, V) pair (see Figure 5.6). In that case, the categories considered are 'A', 'B', 'C' and a new specific category, named 'FF', corresponding to cases where $o \in [0,5]\%$ and $V=80$ km/h. It captures the free-flow state before the onset of congestion, no matter which is the corresponding category. In short, this method provides a useful snapshot of the drivers' behavior in a certain section under all traffic conditions. Results will be showed for three representative sections: one highly influenced by a radar installation (i.e. section 53+570 containing Loop 3 installation) and the other two are located far away from the radar influence (i.e. sections 54+490a and 54+490b). To easily compare results, cells exhibiting different categories across the two years available are highlighted with thick line borders in Figure 5.6.

The related numerical results are gathered in Table 5.3. The first row shows the total percentage of data contained in each one of the four zones depicted, while the others show how data included in a zone is distributed among the group 'ℂ' of categories. Obviously, given a section and a year, the first row sums a hundred and the aggregation within a column from second to fourth rows, results in the same value. It must be noted that the considered morning rush periods are differently defined in 2009 than in 2010. The latter year includes much more free-flow states than the first one. This difference must be considered when comparing both years' results and so, bigger proportions of data are expected in zone 'FF' for 2010 than for 2009.

A first glance at Figure 5.6 reveals a clear pattern for each of the depicted zones, no matter neither the section nor the year. Zone 'B', which mainly contains category 'B' states, sketches a linear tendency with negative slope of around 15%-occupancy wide values, covering from 0% up to 35-50% occupancy values, depending on the sections. In fact, this shape is coherent with the common empirical results obtained when plotting occupancy against speed. Zone 'C', mainly containing category 'C' states, takes the remaining zone below the green one, i.e. low DSL and occupancy values. Finally, zone 'A',

mainly containing category 'A' states, is irregularly spread above the zone 'B' and in general, is observed for low DSL and high occupancy values.
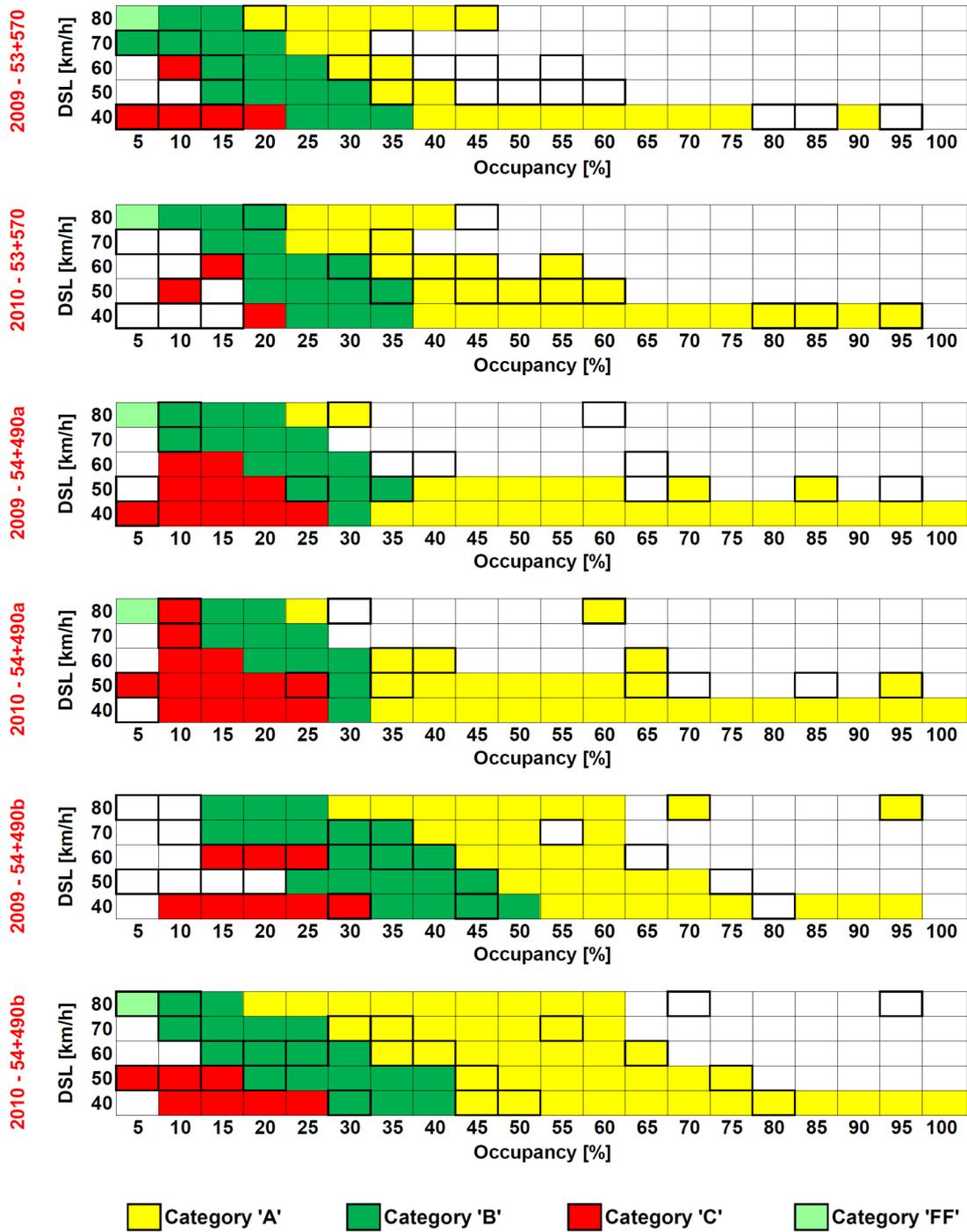
Figure 5.6 *Zone distribution of the four considered categories for sections 53+570, 54+490a and 54+490b, in years 2009 and 2010.*

Table 5.3 *Summary of data distribution among the different zones and categories considering sections 53+570, 54+490a and 54+490b, in years 2009 and 2010, expressed in percentage units [%].*

	Zones	year 2009				year 2010			
		A	**B**	**C**	**FF**	**A**	**B**	**C**	**FF**
53+570	Total data	*12.51*	*75.65*	*10.72*	*1.12*	*6.72*	*62.50*	*0.54*	*30.24*
	Category A	83.30	14.96	0.00	0.00	78.97	9.04	4.08	1.31
	Category B	16.70	80.12	21.30	100.00	20.28	88.23	13.27	95.37
	Category C	0.00	4.92	78.70	0.00	0.75	2.73	82.65	3.33
54+490a	Total data	*2.87*	*59.50*	*37.52*	*0.11*	*1.94*	*22.16*	*30.59*	*45.31*
	Category A	92.48	5.48	0.17	0.00	92.75	2.97	0.00	0.12
	Category B	7.52	85.40	12.10	40.00	7.25	91.43	0.00	18.09
	Category C	0.00	9.12	87.73	60.00	0.00	5.60	100.00	81.78
54+490b	Total data	*26.99*	*64.85*	*8.16*	*0.00*	*15.39*	*46.12*	*2.39*	*36.11*
	Category A	83.63	11.30	0.00	0.00	74.45	9.64	10.50	3.26
	Category B	16.29	82.83	18.23	0.00	25.44	88.17	18.25	84.91
	Category C	0.08	5.87	81.77	0.00	0.12	2.19	71.25	11.83

Focusing on section 53+570, it must be noted that most of data inside zone 'B' and 'FF' belongs to category 'B'. In absolute values, only a 2-4% of the total data belongs to 'C'. It shows the big effect of proximity to the radar installation (installed around 300m upstream of the section). Zone 'B' shape is quite homogenous with a maximum occupancy around 35% value for both years. However, on year 2010 a shift in the zone 'B' border to higher occupancy value is observed, probably induced by drivers' accommodation to the new technology, which leads to speed limit observance in more dense traffic conditions. Remarkably, a high rate of category 'A' data is also observed, i.e. around 20% of total data for 2009, within zone 'A' and 'B'. These values are two times higher than the ones observed in 54+490a section. This behavior may be imputed to the 2 km distance between the section and its DSL governing 'R-VMS 2' gantry. It means that the algorithm miscalculate the traffic flow evolution along this distance, leading to inappropriate DSL values. If considering section 54+490a, only up to 6% values are obtained. In that case, the section is governed by 'R-VMS 3', only 1 km upstream the referred section. This result well fits to the fact of being located immediately downstream of the bottleneck localization, far away from the closest radar installation. It means, lower percentage values for category 'B' together with higher category 'C' values. Moreover, zone 'B' band is quite narrow if compared to other sections. If focusing on zone 'FF', the lack of DSL attainment can be verified: data is distributed 40% for category 'B' and 60% for category 'C' on year 2009, whereas 20% and 80% on 2010, respectively. Finally, section 54+490b arises as a special section case, due to the fact of being a section devoted to exiting vehicles through 'S2'. It is well reflected in the high rate of speed limits fulfillment (i.e. high category 'B' and low category 'C' percentage values) and the lack of vehicles within zone 'FF'. In essence, the section only behaves in two different modes: (i) in free-flow conditions few vehicles cross the section (it corresponds to one slow lane located at the kilometer post 54+490), which

justifies the lack of data for occupancy up to 10% values on 2009; (ii) in congested conditions the vehicles exhibit low speed values, which increases the percentage of category 'A' vehicles for certain DSL values (i.e. around 18-30% of the total data, high if are compared with section 54+490a results exhibiting 3-6% values) or increase the percentage of category 'B' when the DSL value lowers, as can be observed in the zone 'B' band wide which considerably increases for low DSL cases.

5.5 Evaluation of DSL effects on freeway traffic

Certainly, one crucial issue remains to be addressed: to quantify the DSL impact in the freeway maximum flow characterization and its critical occupancy. The methodology described in Section 5.3.2 is applied to the different available data. All the test site sections are considered, except 53+200 one due to its systematic drift counting error. The DSL data will be compared with the non-DSL one and any FD modification among the scenarios will be attributed to the DSL strategy.

It should be noticed that the posed method only considers stationary states in the bivariate diagrams construction step, and they closely depend on three parameters: 'a', 'b' and 'N_{min}'. The values selection is based on three criteria: (i) low average fractional error values, i.e. '$\overline{\overline{FE}}$' which is defined as the weighted by vehicles count average of the 'FE' values for all the stationary states and sections within a particular scenario; (ii) assuring the existence of stationary states with high occupancy values and (iii) verifying 'b' holds $b < \delta$. Note that the first and the second condition define a trade-off: low 'a' and 'b' values combined high 'N_{min}' values lead to low $\overline{\overline{FE}}$ values but without capturing high occupancy states. Then, condition (ii) is not fulfilled. It cannot be assured that the calibrated parameters are the optimal one. However, the inspection of a vast number of parameter values together with criteria (ii) and (iii) procedure results a simple and effective approach for the FD characterization. Finally, the following parameters values were selected: a=10 veh, b =7 km/h and N_{min}=125 veh. The corresponding computed error is $\overline{\overline{FE}}$ = 5.31%. In this way, multiple bivariate diagrams (i.e. '$\bar{o}$' vs '$\bar{q}$') stratified by different criteria (i.e. *V*-value, category and year) were obtained (e.g. Figure 5.7 and Figure 5.8). In the light of the achieved results and only considering the obtained bivariate diagrams for solving the maximum flow characterization issue, it would be conclude that DSL has no affect. In particular, the scatter of points corresponding to category 'B' in DSL-controlled scenarios, frequently exhibit maximum flow values below the non-controlled ones. However, these particular results may hide some macroscopic effects of the DSL strategy which can only be observed when a suitable data aggregation eliminates transitory behaviors. So, the relevance of every individual point of the bivariate diagram is balanced thanks to the weighted mean process developed when aggregating the stationary points per every Δo-range of occupancy values (e.g. '$\hat{o}$' or '$\hat{q}$' variables). Moreover, when representing these variables, it is also depicted a vertical bar which corresponds to the flow error estimation, i.e. '$e(q)$'. This error captures the data dispersion of all the individual 1-min data aggregated within the Δo-width range of occupancy values. An enlighten example can be

found for a particular section in year 2010 (i.e. section 51+689 for the V=60 km/h case). In this situation, the bivariate diagram hints at no capacity increase because two points on year 2008 set are depicted above the year 2010 ones (Figure 5.7(a)). Nevertheless, when computing the least-squared best-fit line through the data and considering '$\hat{q}$' variables, a clear capacity increment is exhibited. Two comments must be given for the comprehension of the next results and figures: (i) note that the capacity increment is measured with respect to the mean value which represents the points within the corresponding Δo=5% range, and (ii) in the case that only one point falls inside the occupancy range, no fitting line but only the bar error is depicted.

Examining the different FD (e.g. Figure 5.7), it can be derived some general rules. (i) The free-flow branch presents the minimum error values, increasing for the congested branch. This result is consistent with the traffic flow behavior, which become more unstable for increasing occupancy values. (ii) The FD for year 2008 present little '$e(q)$' values within the free-flow branch. It is highly influenced by the big number of data available (remember only one speed category is defined) which drives down the error. (iii) It can be sketched a direct relation between the speed limit compliance and the homogeneity of the traffic flow, captured by the error estimators. It is verified when computing the mean 'FE' value (i.e. the average 'FE' value among all the computed stationary states aggregated per Δo-range for all the sections, stratified by speed category, '$\mathbb{C}$'). It results that in almost all the sections, the lowest value corresponds to the category 'B' data (see Table 5.4). In fact, Section 5.5.3. will tackle the question from another point of view, also obtaining similar results.

For the subsequent analysis only data belonging to category 'B' would be analyzed. As the scope of the present paper focus on the DSL effects, they can only be captured when certain degree of drivers' speed limit adherence exists.

Table 5.4 *Mean 'FE' estimation of the flow for different sections and categories, in terms of [%].*

Section / Category	A	B	C
51+689	5.2	3.7	3.2
53+200	3.5	3.4	4.1
53+570	4.9	3.1	3.8
54+490a	6.8	5.5	4.7

5.5.1   Effects on freeway capacity

For every year, section and speed limit, the maximum observed flow among the obtained '$\hat{q}$' values, defined as '$\hat{q}_{max}$', is computed.
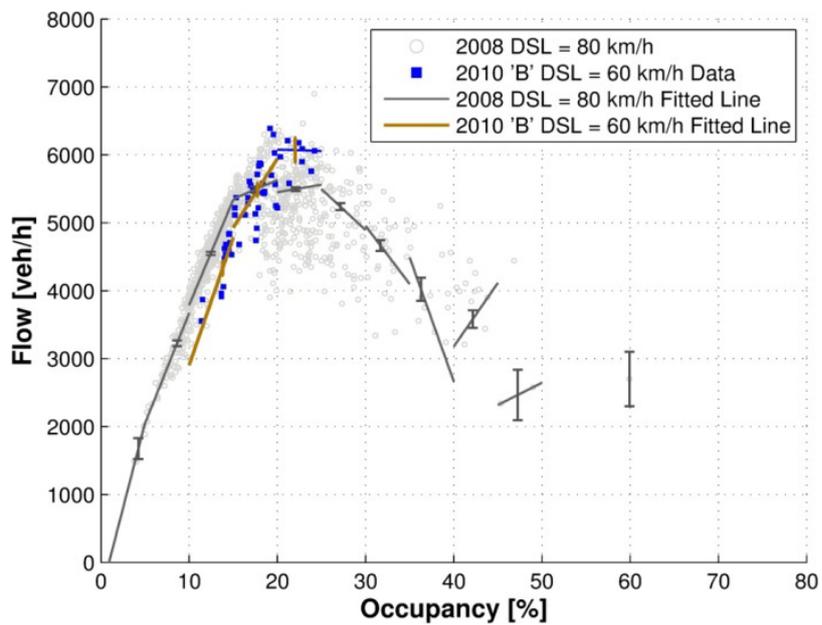
(a)
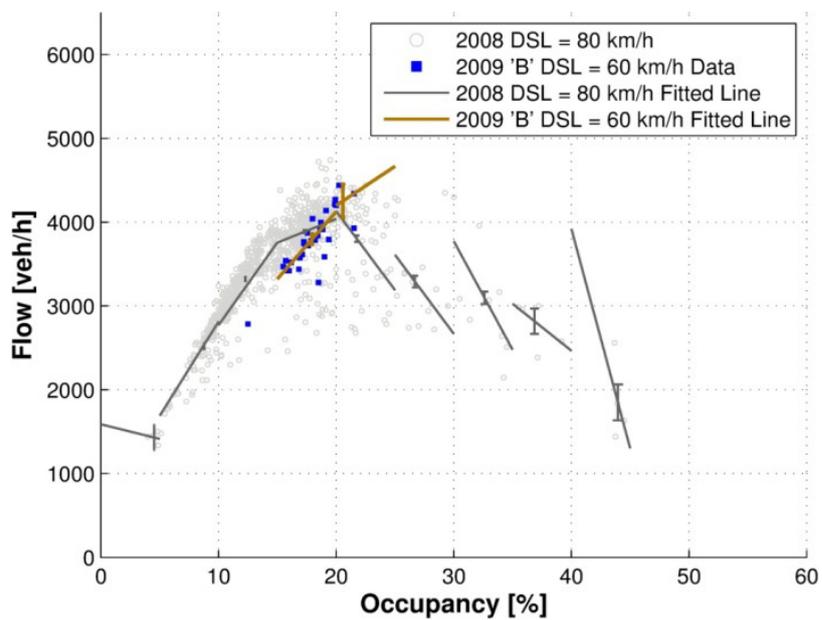
(b)
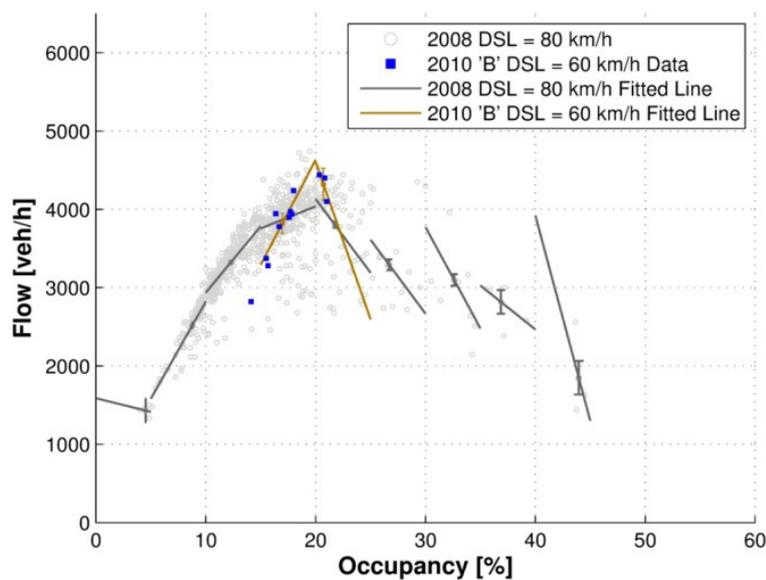
(c)
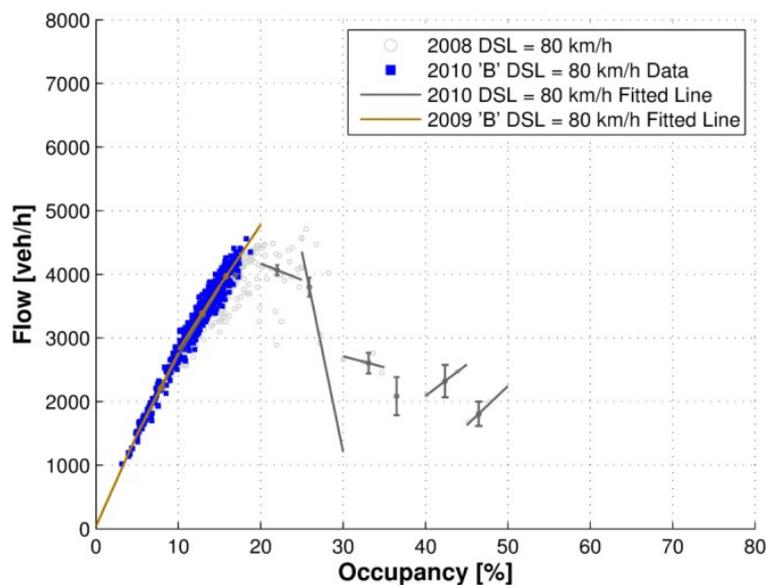
(d)

Figure 5.7 *Fundamental diagram calibration for 60 km/h speed limit (a) on year 2010 at section 51+689, (b) on year 2009 and (c) year 2010 at section 54+490a; and (d) on year 2010 at section 56+799. Each case is compared with respect the reference year and error estimations, 'e(q)', are depicted for every aggregated point with a vertical line.*

Table 5.5 presents the '$\hat{q}_{max}$' values, its related error named '$e(\hat{q}_{max})$', and the prevailing speed limit, 'V', together with information about the relative variation with respect the year 2008 data and the mean time length of the capacity state, i.e. '$\hat{T}$'. Note that the maximum flows for 54+490a/b sections are selected assuring the same DSL is set in both sections, as both sections share the same location.

Table 5.5 *Capacity value estimations for all the sections and only considering stationary states belonging to category 'B'.*

Zones / Year	Capacity estimation, $\hat{q}_{max} \pm e(\hat{q}_{max})$ [veh/h]			Variation with respect 2008 [%] (Active DSL value, [km/h])		Mean time length, $\hat{T}$ [min]		
	2008	2009	2010	2009	2010	2008	2009	2010
51+689	5496.47 ±27.22	5492.67 ± 230.51	6069.40 ± 186.18	-0.07 (60)	10.42 (60)	5.92	3.00	3.47
53+570	6060.22 ±19.73	6600.00 ±158.74	6628.79 ±113.26	8.91 (60)	9.38 (60)	7.01	3.00	4.56
54+490a	3888.36 ±18.70	4253.58 ±199.35	4318.00 ±202.56	9.39 (60)	11.05 (60)	6.89	4.23	3.00
54+490b	2214.91 ±20.22	2230.32 ±102.90	2490.00 ±142.82	0.69 (60)	12.41 (60)	7.86	6.73	4.00
56+799	4066.71 ±80.64	3947.45 ±58.53	3967.06 ±33.60	-2.93 (80)	-2.45 (80)	4.29	5.38	5.96

The discharge flow is the suitable indicator for the capacity of a bottleneck section. Although many locations present congestion episodes, only the section located immediately downstream of the bottleneck can capture the representative capacity value: i.e. section 54+490a. Then, it is expected that, under the DSL strategy, certain bottleneck discharge flow increment would be observed. This assumption is verified thanks to the 9-11% maximum flow increment reported for $V=$ 60 km/h case, when the controlled case is compared with the uncontrolled one (Figure 5.7(b)-(c)). Capacity increments of around 3% are also observed for V=50 and V=80. Remarkably, this increment occurs for speed limit values lower than 80 km/h ($\approx$ free-flow speed value, [Soriguera, et al., 2013]), which may be considered nonsensical if assuming the intuitive principle which states the lower the speed limit value is, lower maximum flows would be achievable (e.g. Figure 2.1(d)). This evidence may evoke to non-triangular FD shapes, but it falls out of the present research scope. Anyway, if considering other sections, e.g. 53+570 or 51+689, significant capacity increments of around 7-10% are reported for $V \leq$70 km/h cases.
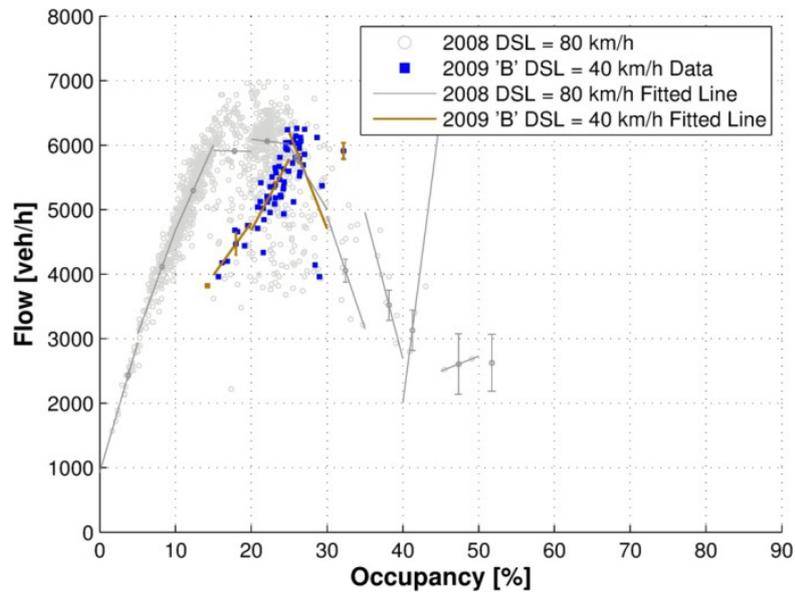
(a)
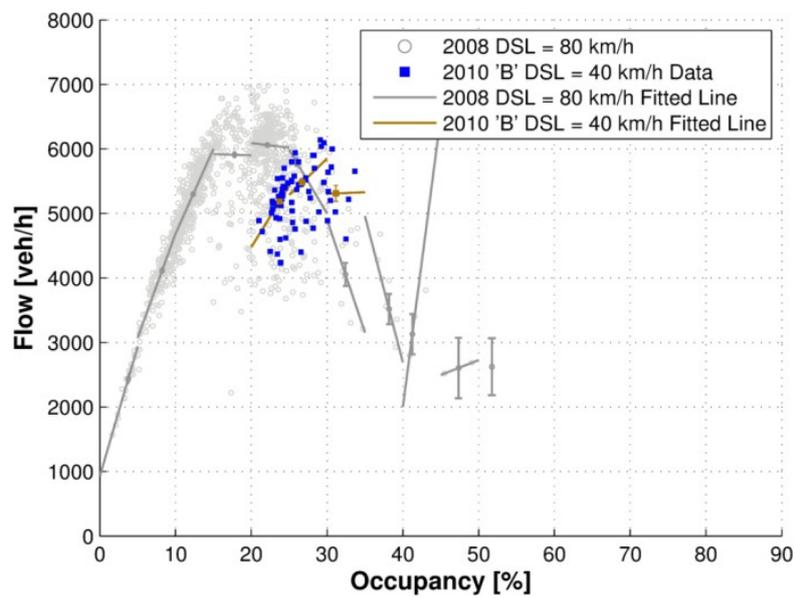
(b)
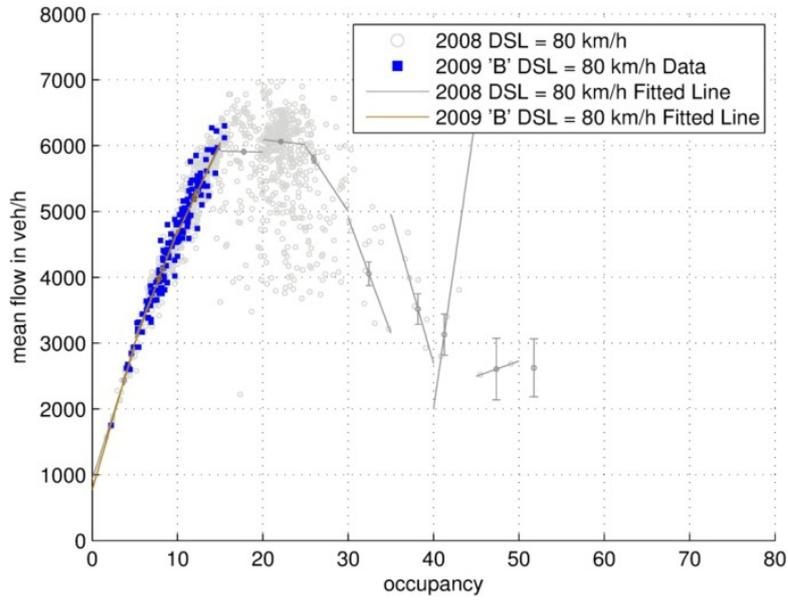
(c)
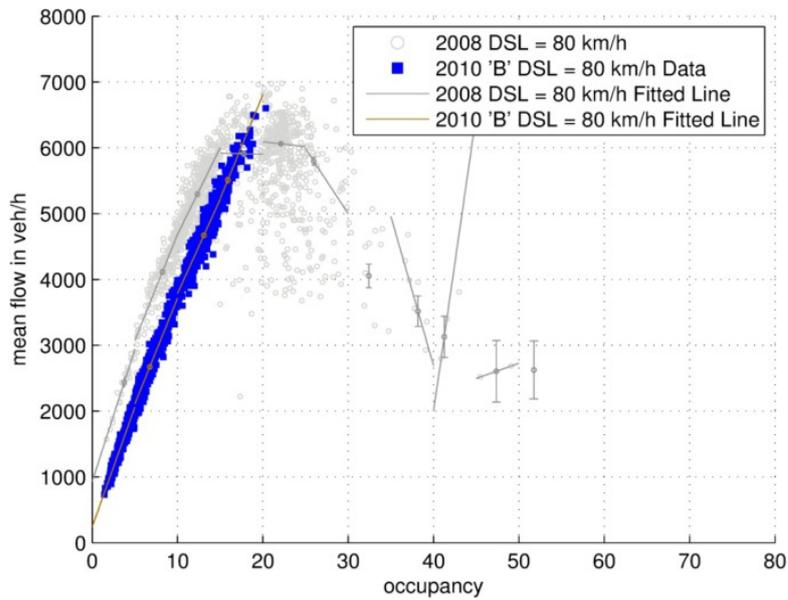
(d)

Figure 5.8 *Fundamental diagram calibration for 40 km/h speed limit (a) on year 2009, (b) on year 2010; and 80 km/h speed limit (c) on year 2009, (d) on year 2010, at section 53+570. All cases are compared with respect to year 2008 fixed 80 km/h speed limit. Error estimations, 'e(q)', are depicted for every aggregated point with a vertical line.*

Finally, the slight decrease observed at section 56+799 capacity may be justified in different ways. First, if considering the maximum flow is computed as the mean value of a normal distribution, the observed descend may be considered negligible, or at least unremarkable. Second, as section 56+799 is located more than 2 km downstream of the bottleneck, it is reasonable to assume it is unaffected for the DSL strategy controlled bottleneck. Third, the capacity estimation for year 2009 and 2010 may be biased, as the section operates most of the time in free-flow conditions, in contrast with year 2008 case (see Figure 5.7(d)). Finally, it can also be affected by the demand descent reported in the whole C32 freeway [Soriguera, et al., 2013; Torné, et al., 2011].

Another interesting conclusion may be drawn referred to section 54+440b. Remember this section is devoted to an exit ramp, i.e. S2. Focusing on year 2009 results it can be observed that almost no maximum flow increment is observed in the section due to the DSL policy. This result is coherent with the fact that DSL cannot influence in the off-ramps capacities, which are only constrained by structural factors or downstream traffic conditions (e.g. roundabouts, urban network…). However, year 2010 results seem to contradict the latter assertion, because an increase of around 12% is reported for the referred section. Two arguments can shed some light to the issue. First, the fact that, for years 2008 and 2009, the conservation of the flow principle is verified among sections 53+570 and 54+490a/b capacities. That is, the addition of the two most downstream sections capacities exhibit close values with respect the 53+570 section capacity value (only deviating -0.71% for 2008 and +1.77% for 2009). Second, the capacity value for section 53+570 remains almost constant among the two DSL-controlled scenarios, exhibiting a constant maximum flow value of around 6600 veh/h. If considering both facts together, it is reasonable to conclude that an increment in the off-ramp of around 12% is produced at S2. It may be motivated by a variation on the drivers' routes preferences, a structural improvement on the off-ramp or any other downstream element of the network.

5.5.2   Effects on critical occupancy

To better capture the critical occupancy variation within the closed congested stretch preceding the bottleneck location (i.e. between E1 location –around kilometer post 53+200– and S2 –around kilometer post 54+490–), the most representative section must be suitably selected. In that case, and in contrast with the capacity estimation, for the critical occupancy calibration it is recommended to analyze data inside the congested stretch, not downstream it. This is justified by the fact that at section 54+490a, almost no data belonging to the congested branch is observed. So, sections 53+200 and 53+570 may be a suitable election. However, these sections presents drift counting errors which produce an overestimation of the 'o_{crit}' values, except for year 2009 data at section 53+570 (Figure 5.8(a) and (c)). The obtained results show a clear tendency to increase the 'o_{crit}' value when lowering the speed limit (see Table 5.6). The slope of FD free-flow branch is lowered proportionally to the speed limit value, shifting the 'o_{crit}' to higher values. In this way, DSL-induced critical occupancies are higher than their non-DSL counterparts. Note that for the scenarios containing errors, the absolute values must be considered overestimated,

while the relative ones may still be valid for descriptive purposes. In particular, increments up to 45% are found for V =40 km/h cases at section 53+570 on year 2009, Figure 5.8(a) and (b), presents the bivariate diagram for flow and occupancy variables for the latter section on years 2009 and 2010, respectively. With this scatter representation, the 'o_{crit}' value can only be roughly determined, thus the usefulness of the data fitting process which helps to improve the FD characterization. Unfortunately, data only covering a thick range of occupancy values is observed for high speed limit values and may lead to a biased 'o_{crit}' estimation. Examples can be found for $V \geq 70$ km/h cases at section 53+570 on year 2009, where unrealistic negative increments are captured. Figure 5.8(c) and (d) show the V =80 km/h case for year 2009 and 2010, respectively. It can be clearly observed how the points do not reach 'o_{crit}' values corresponding to year 2008 case, thus non representative estimations of 'o_{crit}' are obtained.

Table 5.6 *Variation of 'o_{crit}' with respect year 2008 case (80 km/h max. speed limit in force), for all the speed limits values, two different sections and year 2009-2010 data; only considering stationary states belonging to category 'B'.*

Section	Year	'o_{crit}' for year 2008 [%]	'o_{crit}' variation with respect year 2008 [%]				
			40	50	60	70	80
53+200	2009	27,15	26.03	36.75	-18.68	3.11	-22.75
	2010		21.39	17.98	13.34	-22.82	-41.27
53+570	2009	22,11	45.49	16.64	14.58	-22.99	-45.72
	2010		20.78	19.29	16.91	-4.45	-28.15

5.5.3 Effects on lane utilization

Once the effects of the DSL strategy have been sketched, it is turn to speculate about the causes. Chapter 2 has summarized different authors' contributions stating that under active DSL scenarios the capacity in a section may increment if certain stream instabilities are mitigated (e.g. lane changing maneuvers, sudden driver brakes caused by significant differences among the speed of the vehicles, etc.). For the study purposes, two mesoscopic variables were previously defined in order to indirectly capture these operations, 'r_o' and 'r_v'. The assertion being verified is: does certain correlation exist between a capacity increment and a reduction of 'r_o' and 'r_v' values? The possible locations suitable for this purpose are the same that for the critical occupancy issue, i.e. section 53+200 and 53+570. However, section 53+200 contains the referred counting errors which affect the accuracy of the results. So, section 53+570 is the chosen one.

The selected section presents capacity increments for cases exhibiting V =50÷70 km/h values, no matter the year. The V =40 km/h case will also be included in the next analysis as it groups extremely congested traffic scenarios which are suitable for testing the homogenization effect of DSL. In its turn, Figure 5.9 presents the different bivariate plots

(i.e. 'r_v' vs 'o' and 'r_o' vs 'o') with its corresponding vertical error bar (i.e. '$e(q)$') obtained when considering all the 'V' values together with the year 2008 reference case. To better understand its behavior, each family data set is now fitted to a 2nd order function.

(a)

(b)

(c)

(d)
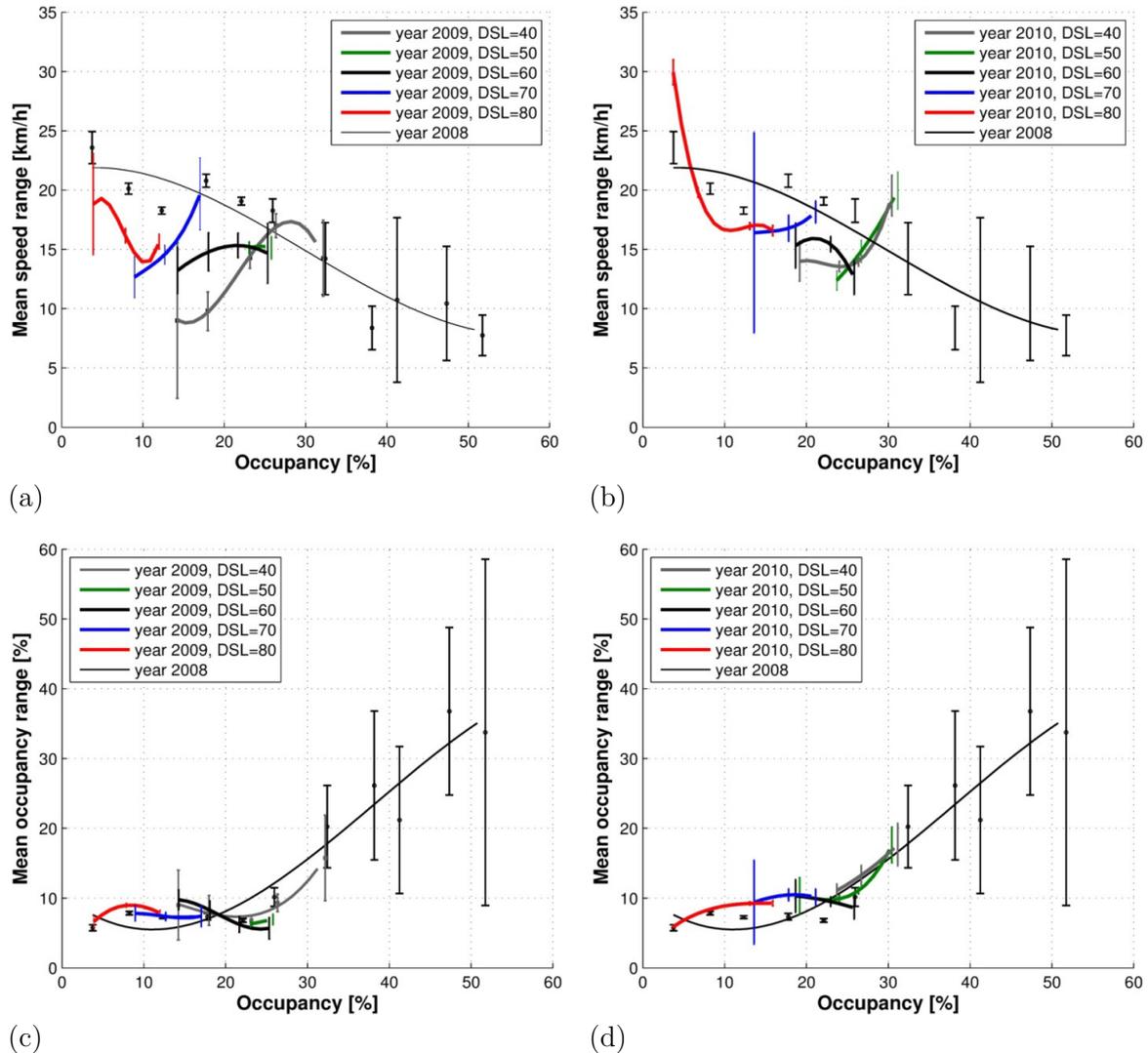
Figure 5.9 *Plots showing the regressions for (a)-(b) 'Mean speed range' vs 'occupancy' and (c)-(d) 'mean occupancy range' vs 'occupancy', with the corresponding error indicators, considering different speed limits and years, at section 53+570.*

Examining Figure 5.9 reveals two opposed trends in 'r_o' and 'r_v' reference curves. The first exhibits an increasing trend with respect to the occupancy, while the latter behaves in the opposite way. These results are coherent with the fact that in congested situations, vehicles exhibit more homogeneous speeds among the lanes but higher dispersion in terms of occupancy distribution, caused by lane changing maneuvers or, sudden brakes, among other possible causes. When analyzing the rest of the curves, it can be observed how the DSL strategy is not capable of systematically (i.e. for any speed limit value) reducing the traffic heterogeneity among the lanes. However, significant diminutions are observed up to 25% occupancy values in 'r_v' curves and within 20÷30% occupancy values. In fact, for

V=40 and V=50 km/h cases, 'r_v' curves cross the reference one. Remarkably, the obtained cross point is observed for occupancy values around 25%, roughly matching with the 'o_{crit}' value estimated for $V < 70$ km/h cases. Respect the 'r_o' curves, the cross point occurs for occupancy values around 20% when V =60 or V =70 km/h. It means the obtained cross-points are close to the 'o_{crit}' value corresponding to V =70km/h (i.e. 17% for year 2009 and 21% for year 2010). Focusing on occupancy values higher than the cross point value, 'r_v' curve tends to be positioned above the reference curve whereas 'r_o' curve tends to be positioned below. It suggests that DSL strategy helps to homogenize the occupancy distribution among lanes in congested conditions and the speed distribution in uncongested conditions. All together contributes to improve the capacity of the DSL-controlled sections.

5.6 Conclusions and further research

Despite the considerable interest and the intense debate among the scientific community in the DSL strategies, the effectiveness of this policy is still a controversial issue. The empirical approach which has been presented in the chapter aims to have illuminated the issue.

The first speed based data stratification has provided an enriching insightful about the driver's behavior when facing this kind of active traffic management (ATM) strategies. At the same time, this step has allowed to identify traffic periods affected by the DSL strategy susceptible of being classified as stationary periods. Next, a useful approach for searching stationary traffic periods has been developed in order to avoid any visual inspection tool, differing from [Cassidy, 1998]. The whole process has permitted to obtain a clear FD characterization under DSL strategies.

The critical occupancy calibration is one of the most important outputs derived from the latter issue. In fact, when set speed limits are lower than free-flow speed, the slope of the free-flow branch is lowered proportionally. This implies that critical occupancies are higher than in their non-DSL analogues. An inherent difficulty arises when measurements under certain speed limit values do not cover the whole range of possible occupancies, conclusions related with capacity increase and critical occupancy shift are difficult to be drawn in such conditions. This is the case for high speed limit values (e.g. 80 km/h), which are only in force for low occupancy values. Such inconvenience would only be solved in scenarios where a particular DSL value is fixed from the onset to the offset of the rush hour episode (e.g. [Soriguera and Sala, 2014]).

The DSL effects on freeways capacity have also been addressed. In that way, one of the most important pending questions related with such strategy was attempted to be answered. Section 54+490a was selected for the capacity assessment. Stationary periods (of 3 to 5 min of duration) exhibiting maximum flow values, were obtained for the DSL and non-DSL scenarios. DSL capacity increase of around 10% was observed. This means that the DSL strategy is capable of improving the freeway performance, at least for certain time periods when suitable speed limit values are set. In fact, as important deficiencies have

been reported in the DSL governing algorithm, they are likely to have reduced the number and length of stationary states reporting close to capacity flow values. So, the DSL algorithm enhancement is a crucial issue to be addressed in order to maximize the DSL benefits and the drivers' perceptions against this particular ATM strategy.

In order to explore the reasons for this capacity gain, per lane mesoscopic variables were defined. The chapter has showed how in DSL scenarios, the lane utilization variability (i.e. the occupancy distribution) becomes more uniform in congested conditions. The same happens with speed variability (i.e. the 'mean speed range') under non congested conditions. Both patterns outline a DSL feature which induces traffic flow to achieve higher discharge rates in the vicinity of a bottleneck.

Certainly, many issues remain for further research, but one stands out: the need of individual vehicle data for computing certain explanatory indicators. The availability of such database would allow obtaining a more comprehensive knowledge about the DSL impact on freeway traffic flow and its causes. The selected indicators may include (but not limited to): lane changes maneuvers counting, accurate speed and occupancy vehicles distributions along the section or microscopic variables distributions, for instance, vehicle headway or spacing. Anyway, it would be helpful to develop a similar analysis in a different DSL installation in order to avoid site and algorithm specific results. A good way of carrying out such studies is via cooperative research, as multiple DSL installations have already been analyzed.

Assessment of dynamic speed limit management on metropolitan freeways

6.1 Introduction

Air pollution is a major problem in most metropolitan areas, where pollutants reach recursively concentrations far above the acceptable limits (for example those established by the European Union air quality directives [European Parliament, 2008]). In developed countries, traffic is the main source of air pollution and greenhouse gas emissions [Keuken, et al., 2005]. The global problem is related to the transportation unsustainable level of energy consumption (97% coming from oil), which, with its increasing price, adds a direct perceived cost to the problem.

The development of a sustainable metropolitan transportation system must be the long term goal. The solution will be clearly multimodal. However, as far as traffic is concerned, and besides demand management strategies and the development of efficient mobility alternatives, traffic administrations must now seek to reduce the environmental impacts of traffic, while improving its safety and maintaining competitive travel times. This is a difficult task, where new surveillance and communication technologies (in short, the ITS) will play an important role.

Several initiatives have recently focused on solving problems related to freeway traffic sustainability (i.e. saving time and energy, decreasing accident risk and diminishing pollutant emissions). The common approach involves control strategies over the traffic stream by using variable traffic signs (e.g. ramp control, dynamic route guidance or dynamic speed limits - DSL). Among them, DSL is one of the most attractive policies, because of its apparent simplicity and reduced implementation costs.

Take as an example the case in the metropolitan region of Barcelona (Spain). In July 2007, a 73-measure plan to improve the air quality was passed. The most controversial policy among those included in the plan was the immediate reduction of the speed limits on major freeways around the city to 80 km/h (from the preexistent limits of 120 km/h). This was planned as the first step towards adopting a DSL system, maintaining the maximum 80 km/h limit. The DSL system became operational on a test corridor on the south access to Barcelona in January 2009. Since the very beginnings, the profitability of the policy became a topic of heated debate. All the stakeholders (e.g. the Government and the opposition, automobile associations, chamber of commerce, media, users…) have disseminated information on the consequences of the policy from their own, generally biased, point of view. None of them have given any scientific arguments to support their claims. The present paper provides an objective and quantitative assessment, grounded on empirical and scientific facts, of this type of policies.

DSL policies have been acknowledged since the early 1970s in Germany [Zackor, 1972] and around the 1980s in the Netherlands [Remeijn, 1982]. Nowadays, many European and American cities have introduced them [Smulders, 1990; Sumner and Andrew, 1990]. Whether isolated or coordinated with other control measures, the effectiveness of the policy is still a controversial issue. The usual claimed benefits imply reductions in pollutant emissions and accident rate, as well as congestion relief. It is reported [Lin, et al., 2004; Smulders, 1990] that these benefits are due to the homogenization of traffic flow, which allows for increased capacity. However, recent data in European highways are inconclusive on this point [Papageorgiou, et al., 2008]. Anyway, few evaluations of DSL management strategies that use real data can be found in the literature, which mainly focuses on control algorithms tested by means of simulation [Carlson, et al., 2010; Hegyi, et al., 2008; Hegyi, et al., 2005], being [Hegyi and Hoogendoorn, 2010] an exception to this last statement.

The present paper fills this gap in two aspects. First, a method for quantitatively evaluating the benefits and costs of active traffic management strategies is presented. The method involves the definition of an objective function that quantifies the effects of any policy in terms of travel times, energy consumption, pollutant emissions and accident risk. These components result in a trade-off, which when adequately weighted in the function allows obtaining the social profitability of the policy.

Second, the method is applied to evaluate the performance of the DSL policy introduced since 2009 on the metropolitan freeways around Barcelona, Spain. The availability of measured before and after data makes it possible to assess driver compliance with the speed limits and to monitor the expected increase in freeway capacity. Unfortunately, no evidence has been found in this direction. This means that standard DSL strategies do not

significantly help in alleviating congestion, and their benefits (in reducing emissions, energy consumption and improving safety) are restricted to free flowing episodes. These benefits must be contrasted with the increase in travel times due to lower travel speeds. This analysis is confronted in the present paper.

In the next section, the objective function is presented. After that, in Section 6.3, some key points to be considered in the cost-benefit analysis of traffic management strategies are outlined. Section 6.4 is devoted to presenting and analyzing the Barcelona DSL case study. Finally, some conclusions and issues for further research are presented in Section 6.5.

6.2 Objective function

The objective function is defined in order to quantitatively assess the efficiency of a freeway stretch in different scenarios. The function accounts for the main costs resulting from inefficiencies in a traffic stream: delays, pollutant emissions, fuel consumption and safety (in terms of accident risk). These are subject to reduction as a result of adequate active traffic management strategies.

The objective function (that is to say, the cost function to be minimized) is composed of a weighted addition of these costs. The weighting factors are, simply, the monetary valuation (direct or indirect) of each term. Therefore, the weighted addition not only adjusts the relative importance of each element, but also normalizes all of them to the same units: the monetary value.

Then, the objective function can be expressed as:

$$F = D \cdot \varepsilon_D + E \cdot \varepsilon_E + C \cdot \varepsilon_C + S \cdot \varepsilon_S \tag{6.1}$$

Where "F" is the total aggregated monetary cost of inefficiencies, and "D", "E", "C" and "S" are the terms accounting for delays [h], emissions [kg of pollutants], fuel consumption [l] and safety [expected # of injured people], respectively. The weighting factors are "ε_D", the monetary cost of delay per vehicle [€/h], "ε_E", the emission costs considered for the pollutants [€/kg], "ε_C" the fuel price [€/l] and "ε_S", the monetary quantification of the injuries resulting from an accident [€/injured person].

In fact, the terms "E", "C" and "S" in Equation (6.1) are row vectors. The emission term "E" is defined by three components, one for each pollutant considered: Carbon dioxide (CO_2), Nitrogen oxides (NO_x) and fine particulate matter (diameter below 10μm) (PM_{10}). The fuel consumption term "C" contains two components, accounting for the total gasoline and diesel consumption. And finally, the safety term "S" includes three components, related to the severity of the injuries: number of deaths, number of major injuries and number of minor injuries resulting from freeway accidents. Consequently, the monetization weighting factors are column vectors accounting for the costs of each one of these components. Then, the objective function can be reformulated as:

$$F = D \cdot \varepsilon_D + [E_{CO_2} \quad E_{NO_x} \quad E_{PM_{10}}] \cdot \begin{bmatrix} \varepsilon_{CO_2} \\ \varepsilon_{NO_x} \\ \varepsilon_{PM_{10}} \end{bmatrix} + [C_{Gas} \quad C_{Diesel}] \cdot \begin{bmatrix} \varepsilon_{Gas} \\ \varepsilon_{Diesel} \end{bmatrix}$$
$$+ [S_{Killed} \quad S_{Major} \quad S_{Minor}] \begin{bmatrix} \varepsilon_{Killed} \\ \varepsilon_{Major} \\ \varepsilon_{Minor} \end{bmatrix} \tag{6.2}$$

As a reference, the weighting factors considered in the application presented in Section 6.4 are summarized in Table 6.1. It is clear that the numeric value resulting from the objective function would be dependent on the costs assigned to each term. Because some of these costs are subjective, interpretable and variable, it is always advisable to perform a sensitivity analysis of the results. For instance it could be claimed that the value of delays depends on the magnitude of the delay (i.e. "ε_D" is some function of "D"). A sensitivity analysis in respect to this is presented in the results section (Section 4). In spite of this, one has to bear in mind that the effects of any control strategy on the freeway are only reflected in the variations of the "D", "E", "C" and "S" terms. Therefore it is in these terms where one should look for the insights.

Table 6.1 *Monetary quantification of the elements considered in the objective function*

Weighting Factor	Monetary Quantification
ε_D	14.2 €/h · veh [1]
ε_{CO_2}	0.00723 €/kg [2]
ε_{NO_x}	4.7 €/kg [3]
$\varepsilon_{PM_{10}}$	7.9 €/kg [3]
ε_{Gas}	1.28 €/liter [4]
ε_{Diesel}	1.22 €/liter [4]
ε_{Killed}	1 265 000 € [5]
ε_{Major}	125 000 € [5]
ε_{NO_x}	2 720 € [5]

[1] In accordance with [Asensio and Matas, 2008], [2] In accordance with [SENDECO2., 2011], [3] In accordance with [European Commission DG Environment, 2000], [4] Average fuel price in Barcelona (January 2011) [Spanish Ministry of Industry, Tourism and Commerce, 2001], [5] European Union official costs [European Transport Safety Council, 2003]

By simply calculating the difference in the objective function (i.e. costs) between the after "F_A" and before "F_B" scenarios, the global benefits of the policy are obtained. Note that negative "P" is desired as it implies a reduction in the total inefficiency costs.

$$P = F_A - F_B \tag{6.3}$$

To completely define the objective function it is only necessary to describe the "D", "E", "C" and "S" terms. In the formulation of each term, described next, special care has been taken in order to obtain simple functions that can be directly evaluated with the commonly available traffic data. It would make no sense to define a precise and complex function in which the data inputs are not available in most of the situations.

As magnetic loop detectors are the most commonly installed surveillance equipment worldwide, and they provide average speed and flow over short time intervals (e.g. one minute or so), these will be the only variable inputs of the objective function. The time – space domain under evaluation, of a total duration "T" and a total length "L", must be divided into cells, whose dimensions are (s,t), where "t" are the loop detector data updating intervals and "s" are the different sections that form the freeway stretch. Sections should be defined taking into account that one loop detector must be available within every section.

6.2.1   Delay term 'D'

Free flow traffic is considered to be the level of service of the freeway that allows the vehicles to reach their destination in the shortest time possible with appropriate levels of safety and comfort. Inefficiencies appear in the form of delays, defined as any travel time increase in relation to this free flow travel time. Delays represent an undesired cost for the user. The objective function accounts for this delay term as:

$$D = \sum_{s \in L} \sum_{t \in T} d_{(s,t)} \cdot n_{(s,t)} \tag{6.4}$$

Where "$n_{(s,t)}$" is the vehicle count and "$\eta_{(s,t)}$" is the average delay per vehicle, both considered on section "s" at time interval "t". In turn, "$\eta_{(s,t)}$" is obtained as:

$$\eta_{(s,t)} = d_{(s)} \left(\frac{1}{v_{(s,t)}} - \frac{1}{v_{f(s)}} \right) \tag{6.5}$$

Where "$v_{(s,t)}$" is the average speed during time period "t", "$v_{f(s)}$" is the free-flow speed and "$d_{(s)}$" is the section length, all of them related to section "s". Plugging Equation (6.5) into Equation (6.4) we obtain:

$$D = \sum_{s \in L} \sum_{t \in T} \left(\frac{1}{v_{(s,t)}} - \frac{1}{v_{f(s)}} \right) \cdot VKT_{(s,t)} \tag{6.6}$$

Where "$VKT_{(s,t)}$" stands for the production of the highway (i.e. vehicles · kilometer travelled) for the section "s" during the time period "t", computed as:

$$VKT_{(s,t)} = n_{(s,t)} \cdot d_{(s)} \tag{6.7}$$

6.2.2   Over-optimum emissions term 'E'

Traffic emissions are the amount of pollutants and greenhouse gases discharged into the atmosphere by an internal combustion engine. The objective function does not compute the vehicles' total emissions, but the excess of emissions over the minimum for a given travel production (i.e. veh · km). These over-optimum emissions are due to inefficiencies and should tend to zero. This responds to the objective of minimizing emissions while satisfying

a given traffic demand (i.e. satisfying the mobility needs at a minimum environmental cost).

The empirical measurement of the amount of emissions produced by a traffic stream is a difficult task. It would be necessary to isolate the traffic contribution from all the other sources of pollution. In addition, changes in the atmospheric conditions in the measurement site may bias the results. This leads to an indirect estimation of emissions, using directly measurable traffic variables, as an appropriated alternative.

The amount of over-optimum emissions of the pollutant "X" produced by a traffic stream, summed across all (s,t) that belong to the space-time interval under evaluation, (L,T), can be expressed as:

$$\tilde{E}_X = \sum_{s \in L} \sum_{t \in T} \sum_{y \in Y} \Big(e_{X,y}(v_{(s,t)}) - min(e_{X,y})\Big) \cdot \gamma_{(y)} \cdot VKT_{(s,t)} \tag{6.8}$$

Where "$e_{X,y}(v_{(s,t)})$" is the average hot exhaust emission factor for a type "y" vehicle and for the pollutant "X" as a function of the travelling speed "$v_{(s,t)}$" (i.e. the unitary emission per $veh \cdot km$ of a type "y" vehicle. "$\gamma_{(y)}$" is the fraction of the vehicle $\cdot$ kilometer production travelled by type "y" vehicles, which can be considered equivalent to the share of type "y" vehicles in the whole travelling fleet "Y". A better estimation could be achieved if details of the particular trips (e.g. lengths) of each type of vehicles were available.

The interpretation of Equation (6.8) is simple. If considering two scenarios with the same demand (i.e. the same production," $VKT_{(s,t)}$"), the total amount of emissions will only depend on "$e_{X,y}(v_{(s,t)})$" and "$\gamma_{(y)}$", which in turn, depend on the characteristics and relative amount of each type of travelling vehicles (e.g. fuel type, weight, age, size of the engine), and on the characteristics of the vehicles' trajectories (i.e. the instantaneous speeds). Therefore, changes in the speed behavior of vehicles will have some effect on the total amount of emissions. However, it has to be borne in mind that, reducing the travel production or updating the vehicle fleet (e.g. promoting "green" cars) are basic policies in order to reduce traffic emissions (i.e. $\tilde{E}_X$),

Then, Equation (6.8) can be expressed as:

$$E_X = \sum_{s \in L} \sum_{t \in T} \Big(e_X(v_{(s,t)}) - min(e_X)\Big) \cdot VKT_{(s,t)} \tag{6.9}$$

Where "$e_X(v_{(s,t)})$" is the average hot exhaust emission factor as a function of the travelling speed "$v_{(s,t)}$" of a virtual vehicle whose emission characteristics are those equivalent to a weighted average of the vehicle types composing the vehicle fleet. This is:

$$e_X(v_{(s,t)}) = \sum_{y \in Y} e_{X,y}(v_{(s,t)}) \cdot \gamma_{(y)} \tag{6.10}$$

Note that given these definitions, Equations (6.8) and (6.9) are not completely equivalent. While Equation (6.8) considers the minimum emissions of each vehicle type (each one achieved at a particular, and different, travelling speed), Equation (6.9) considers the

minimum of the aggregate emission factors, achieved at a particular speed which is optimal given the composition of the vehicle fleet, but which does not need to coincide with any optimal speed for a particular vehicle type. It is clear in this case that for the same inputs "$\tilde{E}_X > E_X$". However, the over-optimum emissions "$\tilde{E}_X$" can only be eliminated by posting a specific speed limit for each vehicle type. As this is not a feasible policy, the lower minimum emissions considered in "$\tilde{E}_X$" are not realistic. Instead, the larger minimum threshold in "E_X" is accepted. This is the reason why Equation (6.9) (and not Equation (6.8)) describes the "E_X" term in the objective function.

Characterization of the hot exhaust emission factors, as a function of speed, for 30 different vehicle types (divided by fuel type, vehicle weight, age of the vehicle and engine size) and for each of the pollutants (CO_2, NO_x and PM_{10}) can be found in [Ntziachristos, Samaras, Eggleston, Gorissen, Hassel and Hickman. ,2000]. As a general rule, emission factors increase with the size and with the age of the vehicle. Gasoline vehicles used to be cleaner than diesel ones, but this difference is smaller and smaller with each new generation of vehicles. Newer vehicles also tend to have larger optimal speeds in relation to their older brothers. These "$e_{X,y}(v_{(s,t)})$" functions are the inputs to construct the average function, "$e_X(v_{(s,t)})$". As an example, see Figure 6.1, valid for the specific application presented in Section 6.4. The main characteristics of the vehicle fleet in this case are 95.2% light vehicles, 63% of them using gasoline. The average age of the fleet is almost 10 years.
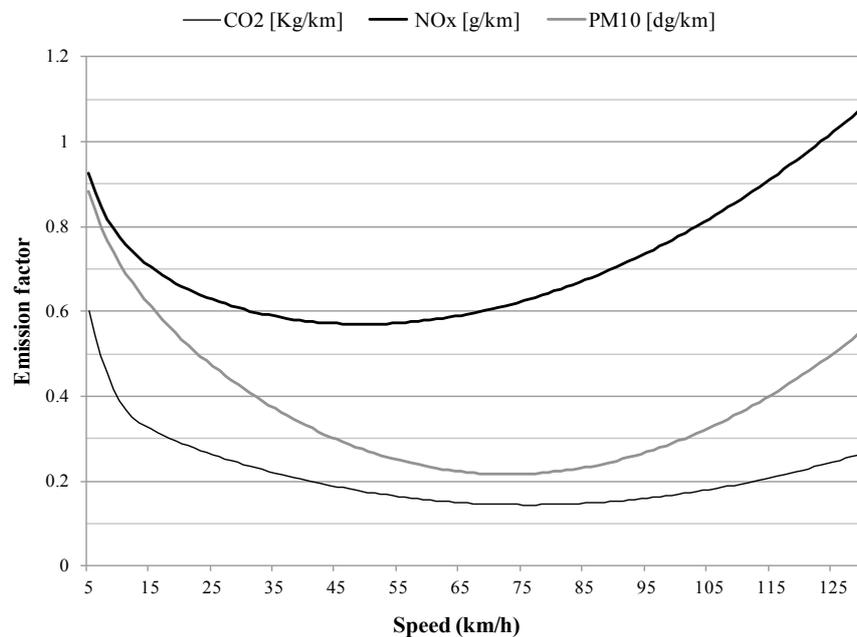
Figure 6.1 *Pollutant emission factors for the average vehicle type distribution in the C-32 freeway near Barcelona, Spain.*
Note: Different units for each pollutant.

Plugging the measured average speed "$v_{(s,t)}$" into Equation (6.9) is not enough to obtain a sufficiently accurate estimation of the total over-optimum emissions [Smit, et al., 2008]. Avoiding to consider the speed variability among vehicles would generally imply an underestimation of the total emissions, given the shape of "$e_X(v_{(s,t)})$" and typical freeway speeds. To account for this variability it is necessary to consider the speed frequency distribution, although alternatives may exist [Baldasano, et al., 2010; Barth, et al., 2000; Samaras and Ntziachristos, 2010].

The speed frequency distribution (e.g. as seen in Figure 6.2) provides the fraction of vehicles traveling at each particular speed "v" (e.g. with a 5 km/h discretization) given the average traffic speed "$v_{(s,t)}$". This can be expressed as "$P(v|v_{(s,t)})$". Given this knowledge, Equation (6.9) can be reformulated to include the speed variability as:

$$E_X = \sum_{s \in L} \sum_{t \in T} \sum_{\forall v} (e_X(v) - min(e_X)) \cdot P(v|v_{(s,t)}) \cdot VKT_{(s,t)} \qquad (6.11)$$

where "$e_X(v)$" is the average emission factor for the "v" speed bin.

However, only average speeds over "t" are usually available from loop detector measurements [Soriguera and Robusté, 2011]. In order to transform them into speed distributions, it is assumed that vehicle speed on a freeway follows a normal distribution (i.e. Normal$(v_{(s,t)}, \sigma_v^2)$, truncated between $\pm 2\sigma_v$ and always positive. This avoids the theoretically possible extreme values of the normal distribution. This assumption is supported by [Soriguera and Soriguera, 2009]. The standard deviation "σ_v" of the speed distribution must be estimated taking into account the traffic stream characteristics in each particular application, or computed using the methodology presented in [Soriguera and Robusté, 2011]. As an order of magnitude, the data reported in [Soriguera and Soriguera, 2009] suggests a standard deviation between 15-20 km/h for free flowing episodes ($v_{(s,t)} \geq 70$km/h) and around 10 km/h for congested situations with stop & go episodes ($v_{(s,t)} < 70$km/h). These approximations can be considered valid for typical loop detector aggregation intervals, on the order of one minute.

In addition, and considering the results in [Soriguera and Soriguera, 2009], a bias is introduced into the distribution to take into account the mandatory speed limits (see Figure 6.2) and the aversion to speeding. In case the speed limit falls within the range of possible speeds, an "extra-frequency" is added to the speed bins bordering (± 10 km/h) the limit. This "extra-frequency" for the bins just above the speed limit is taken from $^1/_4$ of all the vehicles that [Soriguera and Robusté, 2011]would theoretically have traveled above the speed limit, if only considering the baseline speed distribution. This 10 km/h threshold may respond to the known tolerance of the speed enforcement equipment. The "extra-frequency" for the speed bins just below the limit is taken from all the frequencies of bins of smaller speeds and its incremental amount is computed so that the original mean speed, "$v_{(s,t)}$", remains constant. The convergence of this process is granted for speed limits above the mean speed. If this is not the case, the mean speed is considered to play the role of the speed limit in relation to the balancing of the speed distribution. This assumption is supported by the data shown in Figure 6.7, where it can be qualitatively seen how, when

the mean speed is greater than the speed limit (a situation that arises frequently due to the generalized violation of DSL), the speed distribution is concentrated around the mean.
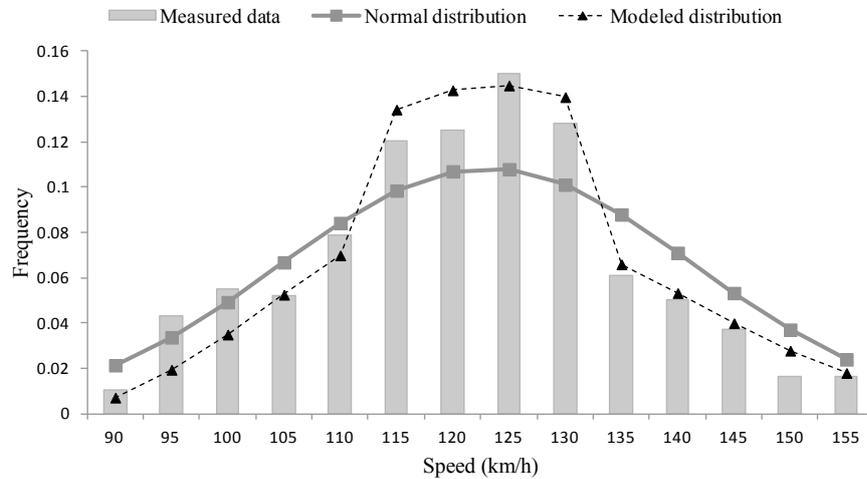
Figure 6.2 *Model for the speed frequency distribution.*
Note: Measured data from AP-7 highway, Spain [Soriguera, 2009].

Knowledge of the speed distribution allows a more accurate estimation of the emissions produced by a traffic stream by using Equation (6.11). Despite this, the proposed method does not account for all the sources of underestimation. Consider the dependence of the results on the "$e_{X,y}(v)$" functions, and the assumptions made on its measurement. The unitary emission curves are generally obtained empirically by car manufacturers in test laboratories where the ideal conditions differ much from real traffic. In particular, emission curves are constructed by interpolation between measured points at constant speeds. This means that accelerations, a huge source of increased emissions, are obviated. This could be a fatal flaw in urban traffic environments, steered by accelerations and decelerations. Although less problematic in freeway traffic, a significant underestimation should be expected in unstable congested traffic (stop & go episodes) [De Haan and Keller, 2004].

Alternatives exist. "$e_{X,y}(v)$" could be obtained for each type of vehicle on a test bench, by measuring the vehicle emissions during some characteristic "driving cycles" [Barth, An, Younglove, et al. ,2000; Samaras and Ntziachristos. ,2010]. In this case "$e_{X,y}$" would not be expressed as a continuous function in terms of speed, but as a discrete function in terms of the "driving mode". The driving cycle tries to mimic the acceleration/deceleration transitions and the speeds that characterize each driving mode. The uncertainty of this approach resides on how the driving mode approaches the actual traffic dynamics under evaluation [André, 2004]. Again, freeway traffic is privileged in this case, due to its controlled context.

Microscopic models (i.e. "$e_{X,y}$" emissions as a function of instantaneous speed, acceleration rates, engine performance, etc...) could also be proposed as an alternative to the semi-

aggregate speed distribution method or to the aggregate driving mode method. However any attempt to increase the detail in the emissions model is myopic in the sense that does not account for the actual available traffic data. If another model is necessary to obtain the inputs of the microscopic emissions model from aggregate traffic data, the benefits of the increased detail are doubtful.

6.2.3   Over-optimum fuel consumption term

Similarly to before, the objective is to reduce the over-optimum fuel consumption. This allows not penalizing the travel production (veh · km) but only the inefficiencies.

Equivalently to the emissions term, the over-optimum fuel consumption is computed as:

$$C_F = \sum_{s \in L} \sum_{t \in T} \sum_{\forall v} \big(c_F(v) - min(c_F)\big) \cdot P\big(v|v_{(s,t)}\big) \cdot VKT_{(s,t)} \qquad (6.12)$$

Where subscript "F" stands for the type of fuel (i.e. gasoline or diesel) and "$c_F(v)$" is the average consumption factor of the travelling vehicle fleet. These factors are proportional to a weighted average of the "$e_{CO_2,y}(v)$" functions, considering the actual vehicle type distribution in terms of type of fuel used, because the amount of CO_2 emitted is a direct consequence of the fuel burnt. Figure 6.3 shows such functions in the present application.

Over-optimum fuel consumption are also evaluated taking into account the vehicles' speed variability, and suffers from the same drawbacks as the emissions term in relation to not considering the acceleration over consumption.

6.2.4   Safety term

This term of the objective function includes the evaluation of the accident risk level associated with the freeway. The number of casualties is a random variable with a very low presentation rate. It is generally accepted that the expected number of casualties can be obtained by multiplying these risk rates by the total risk exposure (i.e. the travel demand in terms of vehicles · kilometers travelled). For each severity level "Z", this is:

$$S_Z = \sum_{s \in L} \sum_{t \in T} s_z(?) \cdot VKT_{(s,t)} \qquad (6.13)$$

For decades, researchers from all around the world have invested huge efforts in modeling these accident rates "$s_z(?)$" [Tarko and Songchitruksa, 2005]. Different types of functions (linear, power, logistical, extreme value...) have been proposed in order to model the relationship between the accident rates and some explicative variables and layout characteristics [Abdel-Aty, et al., 2007]. The uncertainty about which variables should be considered is represented in Equation (6.13) by the "?" symbol. Among those, the variables related to the speed (e.g. the average travelling speed, the speed limit or the speed variance) are predominant. To some extent, it is accepted that "speed kills" meaning that

at the individual vehicle level and for a given freeway layout, the higher the speed the higher the probability of having a fatal accident, and also that "speed variance kills" meaning that this probability is also higher as higher are the speed differences between vehicles. As an apparent paradox, these two postulates do not imply that average speed on the freeway have a significant impact on accident rates [Rodriguez, 1990]. Despite this intensive work, there is still no consensus about the validity of any of these models [Elvik, et al, 2005].

Figure 6.3 *Consumption factors for the average vehicle type distribution in the C-32 freeway near Barcelona, Spain, a) Gasoline, b) Diesel*

For decades, researchers from all around the world have invested huge efforts in modeling these accident rates "$s_z(?)$" [Tarko and Songchitruksa, 2005]. Different types of functions (linear, power, logistical, extreme value…) have been proposed in order to model the relationship between the accident rates and some explicative variables and layout characteristics [Abdel-Aty, et al., 2007]. The uncertainty about which variables should be considered is represented in Equation (6.13) by the "?" symbol. Among those, the variables related to the speed (e.g. the average travelling speed, the speed limit or the speed variance) are predominant. To some extent, it is accepted that "speed kills" meaning that at the individual vehicle level and for a given freeway layout, the higher the speed the higher the probability of having a fatal accident, and also that "speed variance kills" meaning that this probability is also higher as higher are the speed differences between vehicles. As an apparent paradox, these two postulates do not imply that average speed on the freeway have a significant impact on accident rates [Rodriguez, 1990]. Despite this intensive work, there is still no consensus about the validity of any of these models [Elvik, et al, 2005].

In spite of this, it seems reasonable to assume that given the extremely random behavior of accident occurrence and the extremely low rates in relation to VKT, these rates are invariant to small variations of VKT, if everything else is held constant. If this was true, it is possible to directly compute the safety impacts of the policy with a completely empirical approach. This is by computing the average before and after accident rates and directly comparing the before – after expected number of casualties using Equation (6.13). No safety model is needed in this case, which is appealing given the limitations of the "state of the art" in relation to the freeway safety modeling. Crash data is needed to compute the accident rates (see Table 6.2). This type of data is commonly available.

A rough analysis of data presented in Table 6.2 indicates that DSL strategies have implied approximately a 50% reduction in the rate of #Killed/Veh · Km and a 25% reduction for the #Major inj./Veh · Km. In contrast the rate accounting for the minor injuries (#Minor inj./ Veh · Km) has increased by a 14%. These results are "significant" at a 5% level, taking into account a Poisson distribution of the number of casualties [Tarko and Songchitruksa, 2005]. A loose interpretation of statistical significance is required here, given the small amount of data.

6.3 Setting the before/after scenarios: the need for simulation

The benefits of a traffic management strategy are obtained by comparing the results of the objective function before and after the application of the policy. Two requirements must be considered when setting these before/after scenarios:

1. Traffic demand must be the same in both scenarios.

2. Input data to the objective function must be accurate in both scenarios.

A recurrent problem in this type of assessment is the difficulty in fulfilling both requirements simultaneously. For instance, in order to obtain accurate before/after data, these should be measured on two real days. But it is almost impossible to assure that two different days are exactly the same in terms of traffic demand and absence of incidents. The alternative could be to select a typical traffic demand pattern and simulate the before / after traffic behavior. In this case, the equality of demand is assured, but the accuracy of the simulation results may be dubious. In both approaches, the assumptions made (i.e. constant activity pattern or known freeway performance functions) are not true, and should be considered only as approximations.

On the one hand, the possibility of knowing the real behavior of drivers when facing the policy favors the first approach. However, it cannot be assured that relatively small demand variations will result in equally limited freeway performance variations. Therefore, it would be difficult to discriminate between the effects of the policy and the effects of the demand variation in the after scenario. On the other hand the possibility of setting a

specific demand favors the simulation approach. But in this case the ability of traffic simulators to replicate driver behavior when facing new management strategies is uncertain. This suggests using the simulation approach when the available before / after data is scarce, different in nature or a long time apart. Besides, if the active management strategy is complex and the resulting behavior of the drivers is to some extent unknown, an empirical comparison is desired (recall that this is the case for the safety term, where it is assumed that small demand variations do not affect significantly the accident rates).

Table 6.2 *Accident data in the metropolitan freeways around Barcelona, Spain*

| Year | Before Data (120 km/h speed limit in force) | | After Data (80 km/h max. speed limit in force) | | | |
| | | | No DSL | DSL | | |
	2006	2007	2008	2009	2010	2011
# accidents	593	658	512	535	594	615
# Killed	14	12	7	4	4	6
# Major injured	58	55	33	33	40	31
# Minor Injured	783	861	646	726	837	895
Annual travel production (Billion veh · km)	8.201	8.570	7.925	7.481	7.400	7.252
Long term rate reduction without speed limit modification	# Killed & # Major injured: 8.7% # Accidents & # Minor Injured: 0.7%					
"s_{killed}" [fatalities/billion veh · km]	1.42 *		0.70			
"s_{major}" [major inj./billion veh · km]	6.15 *		4.56			
"s_{minor}" [minor inj./billion veh · km]	97.34 *		103.27			

* In order to only account for the DSL strategy effects, the long term reduction in accident rates has been discounted from the before data (i.e. expected rates if speed limits strategies were maintained).

These arguments suggest defining the following sets of before and after scenarios:

- **BS scenario:** *Before Simulated scenario.* Simulated freeway performance under typical demand pattern. Calibration parameters of the freeway performance function measured before the application of the active management policy.

- **AS actual scenario:** *After Simulated scenario with measured calibration parameters.* Same demand pattern as the previous scenario but the calibration parameters of the performance function are measured after the application of the policy.

- **AS ideal scenario:** *After simulated scenario with idealized calibration parameters.* In this case the after behavior of drivers is not measured but idealized. This simulates how drivers would behave in a perfect environment and strict compliance of the policies. This is the only after scenario available for planning purposes, as it does not need the policy to be put into practice.

The two remaining scenarios are defined by using only real measured data in order to avoid the simulation uncertainty:

- **BM scenario:** *Before measured scenario.* Measured freeway performance before the application of the active management policy, for a particular existing traffic demand.

- **AM scenario:** *After measured scenario.* Measured freeway performance after the application of the active management policy. Although similar demand periods must be selected for comparing BM and AM scenarios, exactly the same demand cannot be assured.

6.4 Case study: evaluation of dynamic speed limit management in Barcelona, Spain

Since January 1ˢᵗ 2009 a DSL management system has been operative on a test freeway on the south access to Barcelona (Spain), with the main objective of improving the air quality in its metropolitan area (see Figure 6.4). The implementation of the policy was progressive from the pre-existent 120 km/h speed limit. Firstly, it was changed into a fixed 80 km/h limit, on January 1ˢᵗ 2008. After this transitional period the DSL system became fully operative in 2009. The DSL system dynamically reduces the speed limit on a particular freeway section in accordance with the measured downstream drop of the average traffic speed due to congested conditions. A minimum speed limit of 40 km/h is established. The 80 km/h maximum speed limit is maintained in free flowing situations.

6.4.1 The freeway site and available data

The test site consists of a 14.5 km stretch corresponding to the C-32 freeway between the cities of Castelldefels and Sant Boi de Llobregat, on the south access towards Barcelona. This freeway also links the city with its airport (Barcelona – El Prat Airport) (see Figure 1.1). This is mainly a 3-lane freeway, with short sections with 2 or 4 lanes. There are 8 off-ramps and 7 on-ramps. The full stretch is equipped with loop detectors every 500-1 000 m, measuring the average speed, flow and occupancy on a one minute basis. In addition, 10 Variable Message Sign gantries (VMS) displaying the dynamic speed limits in force at every instant are also installed. Considering this layout, 14 sections are defined, taking into account that each section contains one detector. The section limits are defined either by a

junction, a VMS gantry, or a change in the number of lanes (see Figure 6.4). This allows considering each section as homogeneous.
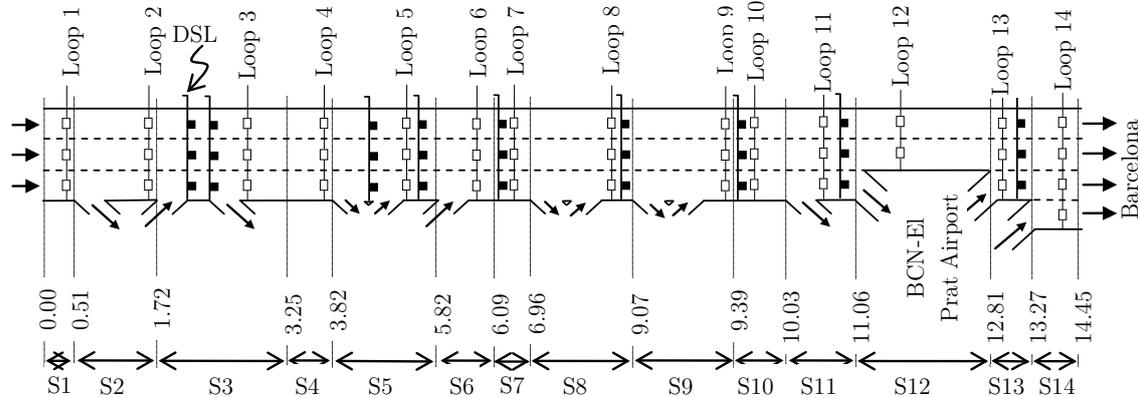
Figure 6.4 *Test site section layout.*
Note: Lengths are displayed in km.

6.4.2 Defining the before / after scenarios

Both, simulation based and empirical assessments are considered in the present application. The before / after scenarios are defined as follows:

- *Simulated scenarios*: The DSL policy is evaluated by means of simulation. Traffic demand is obtained from data corresponding to Thursday May 20th, 2010. This was an incident free day with a high percentage of valid data. The parameters of the simulation model were calibrated using the whole available database: December 2007 for the before scenario (BS) and May 2010 for the after scenario (AS actual). Regarding the AS ideal scenario, strict compliance to speed limits were assumed for all vehicles.

- *Measured scenarios*: The DSL management system is empirically evaluated. This avoids the uncertainties in modeling this strategy. The BM Scenario is defined by data measured on Tuesday December 18th 2007. Similarly, data measured on Thursday May 20th 2010, with the DSL management system active, defines the AM Scenario. Note that the demand in this last case is exactly the same as in the simulated scenarios. This allows evaluating the accuracy of the simulation model.

6.4.3 Construction of the simulation model

Simulated scenarios require a model capable of reproducing traffic dynamics. The asymmetric version of the cell transmission model [Daganzo, 1994, 1995; Gomes and Horowitz, 2006] was selected; specifically, the implementation of the model in the

AURORA software package [Chow, et al., 2010; Kurzhanskiy and Kwon., 2008]. This model has been selected due to its robustness, simplicity and acceptable accuracy relying only on three parameters for each section: the capacity (i.e. maximum possible flow), the free flow speed (i.e. average speed in the low density range) and jam density (i.e. maximum density in completely jammed traffic). These three parameters are required in order to calibrate the flow (q) – density (k) diagram according to the Newell's triangular simplification [Newell, 1993] (see Figure 6.5). Additional parameters are required at each junction. Firstly, the flow blending coefficient is required at each on-ramp. This has been set proportional to the on-ramp relative capacity in relation to the mainline capacity. Secondly and last, the capacity of each off-ramp is also a required parameter. Obviously the traffic demand is an input to the model.

Some drawbacks to the model also exist. Firstly, the lack of discrimination between different types of vehicles in free flowing conditions, with their specific performances. Secondly, the inability to replicate unstable conditions in congested episodes. Both limitations may contribute seriously to underestimating traffic emissions and fuel consumption. Both issues are addressed here by the construction of speed distributions from average simulated results, as described in the previous section. It is only needed to be taken into account an additional variance term for congested periods in simulated scenarios, to account for traffic instabilities (in addition to the inherent stochastic drivers' behavior). According to the available data, this yields a global standard deviation of approximately 15 km/h to be applied to the simulated average speeds in congested episodes. More specific approaches [Tuerprasert and Aswakul, 2010] could be used as an alternative.

In the next sections, the three basic steps in the construction of the simulation model (i.e. data validation, fundamental diagram calibration and demand characterization) are described in detail.

6.4.3.1   Data validation

The first step in any process involving traffic data is the data validation. The data validation process presented here consists of the application of 'absolute' and 'relative' filtering criteria to the raw "one minute average" loop detector data. Absolute criteria are only able to remove very abnormal observations (such as negative or very extreme values of speed, occupancy or flow). Relative ones take advantage of the existing bivariate relations between fundamental traffic variables [Cassidy, 1998]. This allows detecting unfeasible combinations of flow, speed and occupancy values, although each observation could be considered acceptable by itself.

Table **6.3** summarizes the considered filtering criteria. Take into account that absolute thresholds allowed eliminating a 12% of erroneous data (including voids), while the relative filters detected an additional 2% of invalid data. Invalid data are erased from the database. This yields a non-exhaustive database (i.e. some holes may exist).

Table 6.3 *Filtering criteria applied in the data validation process*

	Description	Criterion Definition
Absolute Criteria	Flow range $(q_{(s,t)})$	0 to 2 600 veh/h · lane
	Average speed range $(v_{(s,t)})$	0 to 180 km/h
	Occupancy range $(occ_{(s,t)})$	0 to 100% for each lane
Relative Criteria	Trapezoidal zone within the flow-speed plane	Contained in the trapezoid defined by the following (q,v) points: (0, 0), (0, 180), (2600, 50), (2600, 90)
	Range of feasible speed values in terms of extreme values of "g" the effective vehicle length	$\dfrac{q_{(s,t)}}{occ_{(s,t)}} \cdot g_{min} \leq v_{(s,t)} \leq \dfrac{q_{(s,t)}}{occ_{(s,t)}} \cdot g_{max}$ Where $g_{min} = 5$ m and $g_{max} = 17$ m

6.4.3.2   Fundamental Diagram Calibration

A critical issue for the representativeness of the simulation model is the calibration of the fundamental diagrams, an empirical curve relating, for a particular point on the freeway, the observed densities with respect to the observed flows. Densities are indirectly computed using average flows and speeds, and assuming that the available time mean speeds are almost equal to space mean speeds (i.e. avg. density = avg. flow / avg. speed). Per lane data is spatially aggregated in order to obtain a sectional measure. These diagrams completely define the behavior of drivers on the freeway section, and their characterization distinguishes one scenario from another.

Prior to the adjustment of the curve to the data, it is of particular importance to temporally aggregate as many data as possible to increase the accuracy of the computed mean, while still assuring that all data points belong to the same stationary traffic state. In [Cassidy, 1998] a methodology to deal with this trade-off is proposed.

Finally, it is necessary to adjust the realistic triangular shape [Newell, 1993] of the fundamental diagram to the accepted aggregated data. Two linear regressions are used, one for each of the linear branches of the diagram: congested and free flowing. The free flowing branch is forced to contain the (0,0) point. Similarly, the jam density is fixed for all scenarios to 150 veh/km · lane. This means that the congested branch needs to contain the pair (150 · #lanes, 0). For the AS ideal scenario, where strict compliance of the speed limits is supposed, despite the data, the slope of the free flowing branch of the diagram cannot be higher than the posted speed limit. See Figure 6.5 for some details.
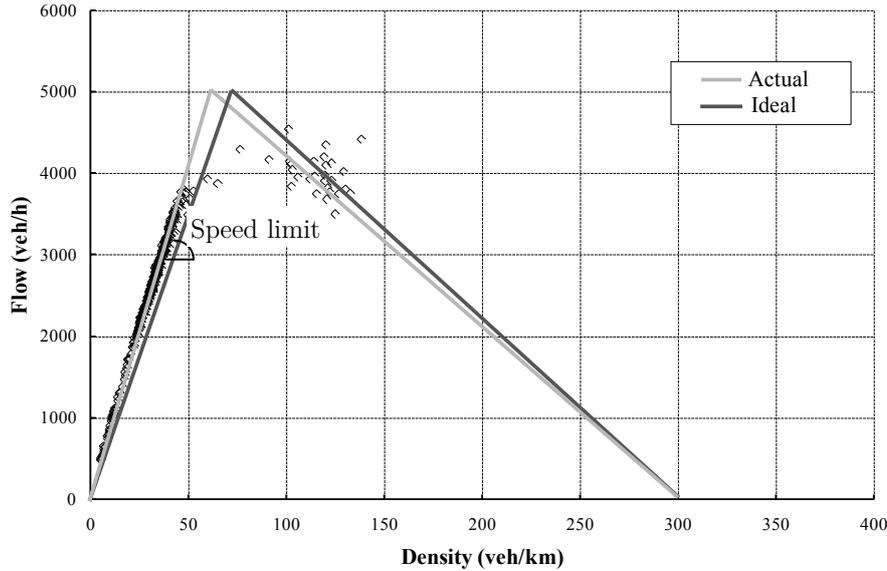
Figure 6.5 *Example of a fundamental diagram adjustment for the after scenario.*

The parameters that define the triangular diagrams (i.e. maximum flow, jam density and free flow speed) are the input variables needed for the simulation model.

6.4.3.3   Traffic demand characterization

The demand profiles at every on-ramp (and at the origin on the main trunk) must be introduced to the model. In addition, the split-ratios (i.e. the percentage of vehicles exiting in relation to the mainline flow) must be set for every off-ramp. These define the demand pattern for the freeway stretch.

The first problem to address in the demand characterization is that profiles must be defined for all periods of time. This means that the holes created during the data validation process must be filled in. The overcomplicated (and usually highly inaccurate) data reconstruction methods are avoided here. It is considered that confidence in the reconstructed data is only acceptable if some measured observation exists within the space – time region under consideration. This means that a single lane one minute hole (i.e. a single observation missing) cannot be reconstructed by itself. It is necessary to extend the observation region in space (e.g. considering all the lanes as a whole, or aggregating several consecutive detectors defining a new extended section) or in time (e.g. aggregating several minutes) until valid observations are found. When this happens, the missing data are reconstructed under the assumption of uniform behavior of traffic in the extended region. This could be seen as a simplification of the method proposed in [Chen, et al., 2003]. In the present application, the whole database could be reconstructed if the considered space – time granularity was defined as whole freeway section measurements at detector sites and 15-minute time aggregations. This time granularity represents an adequate trade-off

between registering the demand variations during short periods of time and avoiding excessive fluctuations. In relation to the split-ratios, it is known that the flow share between each OD pair remains fairly constant for longer periods of time, in relation to absolute demand. Thus, it is justified for the split-ratio to be averaged considering one-hour intervals. This allows obtaining a more stable estimation.

The second problem in this demand characterization process is that loop detectors are usually (and also in this case) solely installed on the main freeway trunk, and consequently no on/off-ramp measurements are directly available. Obtaining on-ramp demand and off-ramp split ratios only from mainline measurements is a difficult task. One possible approximation is to compute the count differences between consecutive detectors isolating a single on/off-ramp. Demand patterns could be obtained if detector count drift and variations of vehicle accumulation within the section (in case of growing or dissolving congestion) are neglected. This simplistic approach results in biased demands in those episodes. However, the short length of the sections (generally <1 000 m) implies that these situations only arise during very short periods of time, and do not imply big differences in the aggregated overall results. Despite this, the improvement of this methodology is an issue of active research [Muralidharan and Horowitz, 2009].

In addition, in some situations it is not possible to isolate a single ramp in between two consecutive detectors. Then, only the net in/out flow through these several ramps could be obtained (this only happens in junctions where the on/off-ramps are very close together). In these cases the whole net in/out flow was assigned to the on/off ramp respectively. This assumption does not imply big drawbacks, as only the segment in between on and off ramps is affected. And this segment is by definition very short.

6.4.4   Results and discussion

6.4.4.1   Analysis of the Before – After data

From the fundamental diagram adjustment to data in the different scenarios, significant preliminary results arise. By observing the obtained free flow speeds (see Figure 6.6) it can be seen that before applying any speed limitation (Before scenarios), the measured free speeds were substantially below the 120 km/h speed limit in force. This fact is accentuated between the 8th and the 14th sections, where the freeway enters a more urban environment. Figure 6.6 also shows the generalized lack of compliance of the reduced 80 km/h speed limit in the after scenario. Actually, a 40 km/h reduction of the speed limit only implied an average reduction of 6.5 km/h in the free flow speeds.

Generalized speeding is also notorious when the speed limit drops below the maximum of 80 km/h as a result of the DSL strategy (see Figure 6.7). In these situations, the compliance to the speed limit is more related to the impossibility of driving faster due to

high traffic density than to the mandatory speed limits. Figure 6.7 also shows that stop&go episodes are not eliminated by using DSL strategies.
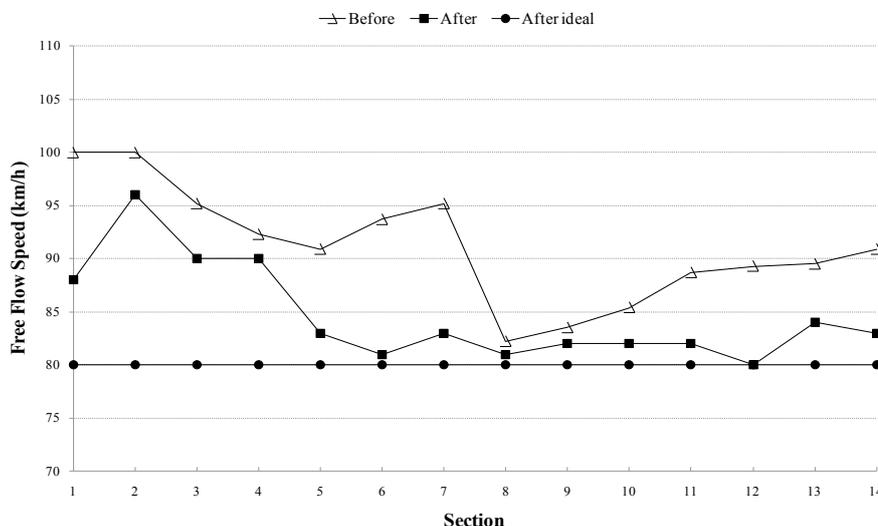
Figure 6.6 *Adjusted free flow speeds for the before – after scenarios*

Being aware of these facts, one should not expect huge variations in the freeway performance as a result of the speed limit reduction. However, and although only small differences in the free flow speeds were achieved, more significant changes appear when considering speed distributions. This will have some effects, mainly in the pollutant emission term.

Another point to focus on is the potential increase in the freeway capacity with the implementation of DSL, which is usually claimed to be a result of the reduction in the vehicle speed variance and due to a possible reduction in the amount of lane changes. This has not been found in the present application. The measured maximum flows, before and after the implementation of the policy, remain unchanged. Capacities between 2150 and 2400 veh/h · lane, depending on the section, are obtained. This result should be not taken as a postulate, but a particular result in this case study. Take into account the significant amount of speeding vehicles that may affect this result. Further research is needed in order to reach a deeper understanding of this issue.

6.4.4.2   Freeway performance in the proposed scenarios

The analysis is performed for a typical working day on the test freeway, in the northbound direction towards Barcelona, where typically the traffic presents a morning rush (approximately between 7:00 am and 10:30 am) and a smaller evening rush (between 6:00 pm and 8:00 pm). Recurrent congestion can be expected in the morning peak hour. Two locations are generally problematic. First, the off-ramp located at the very beginning of

Section 11 (see Figure 6.4(b)), which acts as a bottleneck and causes congestion to grow upstream. Queues may reach Loop 7 at the peak hour, which means a total length of approximately 4 km of congested traffic. Second, queues also spill back into the test site from a further downstream bottleneck. At peak hour these queues may have 2 km in length, reaching Loop 13.
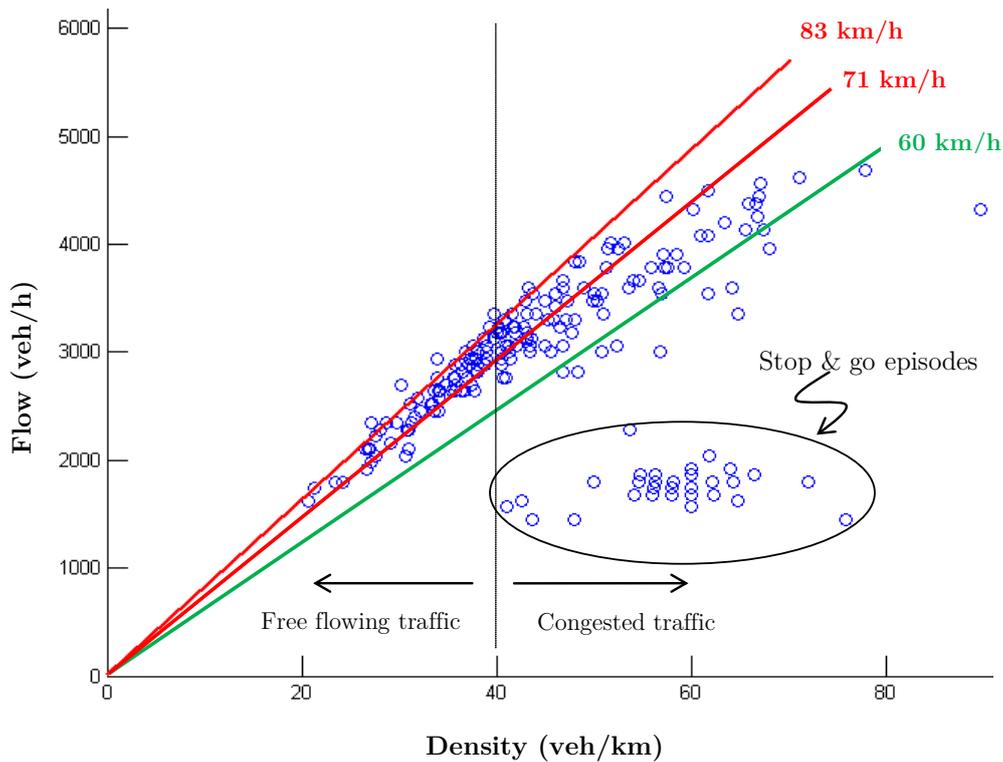
Figure 6.7 *Flow – density plot of data measured in Section 6.3 with a posted speed limit of 60 km/h*

In order to validate the adequacy (or at least know the deficiencies) of the proposed scenarios, two comparisons are relevant. First, check that the differences in the traffic demand between measured before – after scenarios (i.e. BM and AM) are minimal. This is achieved by comparing the total production of the freeway in these scenarios (in terms of veh · km). It is found that the production in the AM scenario is 1.83% less than in the BM scenario (Table 6.5, and see Figure 6.8 for a time series plot). This can be seen as a direct consequence of the economic recession. While this reduction in demand can be considered minimal, it could also have significant implications for freeway performance, as it actually has (see Figure 6.9). The significant reduction in peak travel times in the AM scenario can be mainly attributed to the differences in traffic demand, more than to the DSL strategy. If DSL strategy was to be the cause, increased discharge flows from the bottlenecks should have been measured in this scenario. And they were not, as stated previously. This means that the slight reduction in demand between BM and AM scenarios implies significant effects on traffic dynamics. This makes difficult to clearly isolate the DSL policy effects in this empirical comparison. This is the reason to consider measured results as biased.

Second, it is necessary to validate the simulation model. This is achieved by comparing the performance of the freeway (e.g. in terms of travel times) between AM and AS_actual scenarios, which both respond to the same reality, as the simulated demand was extracted from the after measured scenario. It is found that the simulation model overestimates by 1.7% the average travel time in congested conditions (see Figure 6.9 for a time series plot). This accuracy is considered more than acceptable given the proposed approach to traffic dynamics (i.e. only 3 parameters per section). It is also worth noting from Figure 6.8 (by comparing the simulated scenarios with the AM plot) the remarkable accuracy of the demand characterization process.
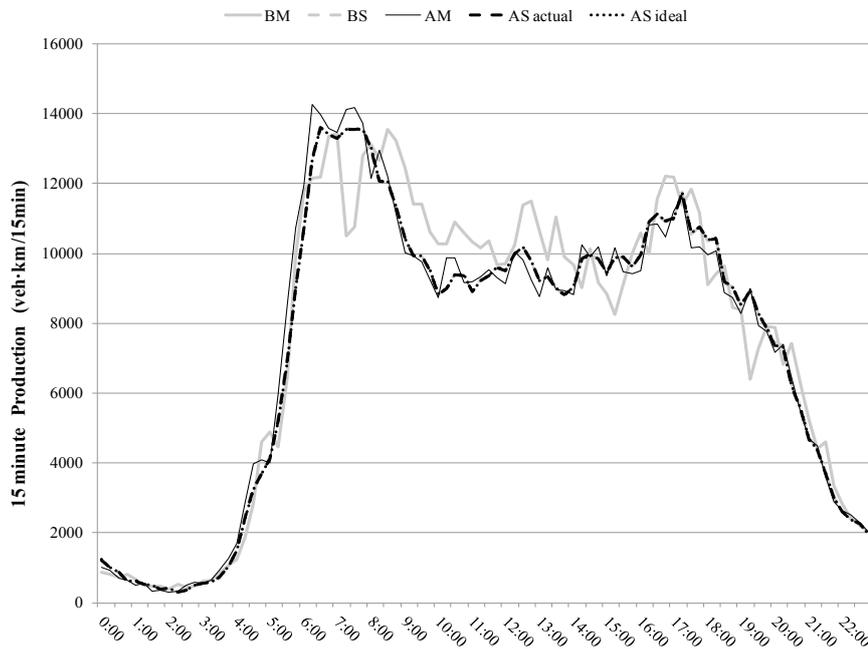
Figure 6.8 *Time series plot of the freeway stretch production.*
Note: Small variations in production for simulated scenarios are not appreciated in the plot.

6.4.4.3   Objective function evaluation

Results obtained from the application of the objective function to simulated and measured scenarios, are presented in Tables 4 and 5 respectively. Table **6.6** is only devoted to verify the accuracy of the simulation in terms of the objective function metric. Using this metric, the speed variability plays a role, which was not captured by the previous travel time validation. As the demand in the AS and in the AM scenarios is the same, differences in the objective function between these scenarios can be seen as the lack of accuracy of the traffic simulation. As shown in Table **6.6**, the accuracy with this metric is also pretty good with differences below 10% (except for delay). This means that the inability of traffic models based on the continuous traffic flow theory to replicate traffic instabilities in

congested scenarios and the effects of the variability among different types of vehicles in free flowing is attenuated by the proposed modeling of speed variability, and does not have a remarkable impact on the results.
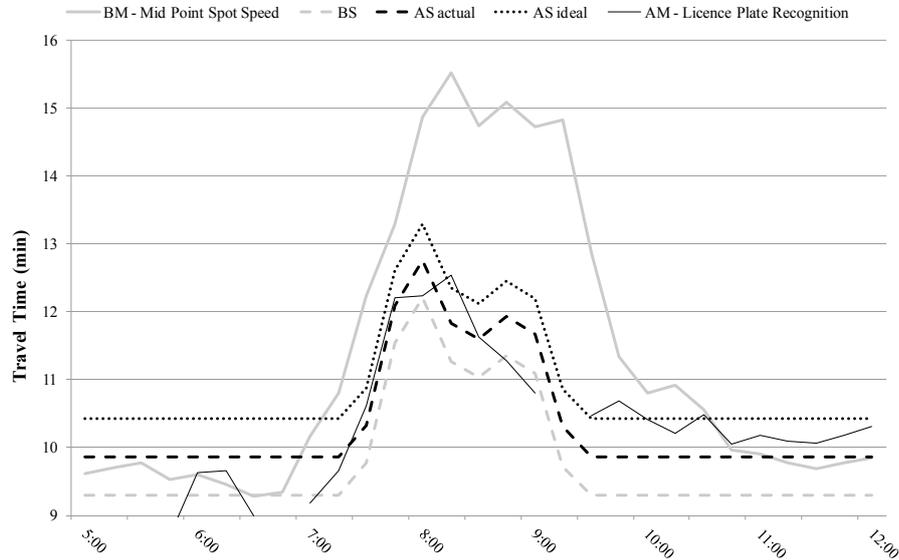
Figure 6.9 *Travel times on the test stretch during the morning rush for different scenarios.*
Note: Licence plate recognition data only available for the after scenario. To obtain measured travel times in the before scenario Mid Point Spot Speed Algorithm is used [Soriguera and Robusté, 2011].

The key points to be highlighted from the obtained results are:

- Simulated results (Table 6.4) must be considered more consistent than measured ones (Table 6.5) due to the consistency in the before – after demands. This assertion is supported by the discussion in the previous section.

- Overall profitability of DSL strategy is negative, as inefficiency costs increase by 45%. This is due to the cost of increased travel times in free flowing conditions (as a result of the reduced maximum speed limit). These costs are not compensated by the achieved reductions in pollutant emissions, fuel consumption and accident risk (of the order of 25-30% reduction of the over-optimum costs, which translates to 3-6% of total costs).

- The installed DSL strategy does not contribute to easing congestion. This means that the strategy has no significant effect in congested periods.

- The negative profits achieved in the after "actual" scenario are, in fact, smoothed down due to traffic limit violations. An ideal scenario implies more drastic negative profits, due to increased free-flowing delays.

- Measured results (Table 6.5) are diametrically opposite to simulated results. However, in the interpretation of these results one has to bear in mind that the comparison between BM and AM is not consistent in the terms of demand (as it is in the simulated scenarios). As proved previously, the slightly reduced demand in the AM scenario in relation to the BM scenario implies significant reduction of the delays in the AM scenario which is the main cause of the different sign of the results shown in Table 6.5.

6.4.4.4   Discussion and sensitivity analysis

The total results of the objective function, presented in Table 6.4 and Table 6.5, provide an overall assessment of the DSL strategy in terms of the after-before differences in total over-optimum costs. This is achieved by weighting the several terms included in the objective function by their particular costs. Despite the insights of the consequences of the strategy must be seek only in the "D", "E", "C" and "S" terms, the overall results will obviously depend on these weighting factors. Two issues must be discussed here: First, the assumption that these weighting costs are constants, which represent the rates of variation in the linear model proposed in Equation (6.1). Second, the value of these constants consider that there is neither a market nor a universally accepted cost function for most of the concepts included in the objective function. Exceptions to this statement are the fuel costs and the "CO_2" emissions, for which a market exists and a linear cost function with a known price (see Table 6.1) applies to drivers and Governments respectively (although these markets are strongly regulated at different levels which implies huge geographical variation). For the remaining terms, the assumption of a linear cost function is simple and provides the benefits of not rescaling the effects of the DSL strategy inside each composing term. This eases the interpretation of the objective function results, and makes it easy to figure out how the results would vary by considering different costs. This is appealing, as the present application considers "official costs" in the European Union (except for delays because they do not exist, see Table 6.1), but this does neither mean that these costs must be accepted everywhere, nor that they are the actually true or perceived costs.

The delay costs deserve a special treatment. Note that the overall results of the objective function (and their sign) are mainly steered by the delay costs. This means that the global assessment of the strategy strongly depends on the cost of delays "ε_D". Figure 6.10 shows how the global results would change by considering different values for "ε_D". Considering the AS Actual scenario (Figure 6.10(a)), it is interesting to highlight the breakpoint between the positive and negative profitability of the strategy for "ε_D=4.50 €/h". For values of delays lower than this threshold, the strategy would be beneficial for the society. This value is significantly lower than the accepted average value for delays of 14.2 €/h [Asensio and Matas, 2008]. However, take into account that this is only an average. There is evidence in the literature (going back to [AASHTO, 1977]) that drivers value differently the delays depending on the total time delayed (i.e. non linear delay cost function). Drivers might be indifferent to very short delays, while increasingly penalized for longer ones. As it has been shown the effects of DSL strategies only imply short additional delays in free

flowing conditions and considering the typical controlled lengths (i.e. <5minutes). This means that the marginal cost of these delays should be much below the average. Considering the values consistently reported since first published in [AASHTO, 1977], these short delays are valued by the drivers only as 5.4% of the general delay cost. This means "ε_D=0.76 €/h" in the present case. Under this assumption the strategy would result profitable, as seen in Figure 6.10(a).

Table 6.4 *Objective function evaluation in Simulated scenarios*

	Concept (in €/day, unless stated differently)	Baseline Scenario, BS Scenario	Comparison Scenarios (% difference respect baseline)	
			AS actual	AS ideal
Whole Day	**Production** (veh · km/day)	**682798**	**-0.19**	**-0.19**
	Total Objective Function	**12179**	**44.49**	**71.88**
	Delay Costs	4695	169.06	257.39
	Over-Optimum Emissions Cost	523 (3150)	-25.81 (-4.28)	-39.34 (-6.53)
	Over-Optimum Fuel Consumption	5033 (55241)	-32.54 (-2.96)	-47.24 (-4.30)
	Accident Risk Penalty	1928	-38.70	Idem as AS actual
Stop & Go Periods	**Production** (veh · km/day)	**11299**	**-0.01**	**-0.01**
	Total Objective Function	**3438**	**2.75**	**3.76**
	Delay Costs	2841	3.10	4.13
	Over-Optimum Emissions Cost	12 (55)	3.66 (0.78)	4.76 (1.01)
	PM_{10} Contribution	2 (4)	3.15 (1.48)	4.10 (1.92)
	CO_2 Contribution	8 (19)	3.30 (1.30)	4.26 (1.68)
	NO_x Contribution	2 (32)	5.13 (0.39)	6.77 (0.52)
	Over-Optimum Fuel Consumption	554 (1384)	3.30 (1.32)	4.26 (1.71)
	Accident Risk Penalty	32	-38.59	Idem as AS actual
	Killed	20	-50.64	Idem as AS actual
	Major Injuries	9	-25.91	Idem as AS actual
	Minor Injuries	3	6.09	Idem as AS actual
Free – Flowing Periods	**Production** (veh · km/day)	**671500**	**-0.19**	**-0.19**
	Total Objective Function	**8741**	**60.90**	**98.67**
	Delay Costs	1854	423.27	645.35
	Over-Optimum Emissions Cost	511 (3095)	-26.49 (-4.37)	-40.35 (-6.66)
	PM_{10} Contribution	23 (136)	-38.03 (-6.33)	-56.37 (-9.39)
	CO_2 Contribution	65 (756)	-37.30 (-3.21)	-54.27 (-4.67)
	NO_x Contribution	423 (2203)	-24.21 (-4.65)	-37.35 (-7.18)
	Over-Optimum Fuel Consumption	4480 (53857)	-36.97 (-3.07)	-53.61 (-4.46)
	Accident Risk Penalty	1896	-38.59	Idem as AS actual
	Killed	1202	-50.64	Idem as AS actual
	Major Injuries	516	-25.91	Idem as AS actual
	Minor Injuries	178	6.09	Idem as AS actual

Note: [1] Potential reduction for over-optimum emissions is 100%.

[2] Values in brackets represent total values (not over-optimum as usual)

[3] Stop & Go periods are defined in relation to the baseline scenario when $v_{(s,t)} \leq 45$ km/h.

[4] Safety term comes from data on actual casualties. No model has been developed to obtain AS ideal values.

Table 6.5 *Objective function evaluation in Measured scenarios*

	Concept (in €/day, unless stated differently)	Baseline Scenario, BM Scenario	Comparison Scenario (% difference) AM Scenario
Whole Day	**Production** (veh · km/day)	**701565**	**-1.83**
	Total Objective Function	**26883**	**-39.25**
	Delay Costs	19531	-42.27
	Over-Optimum Emissions Cost	406 (3102)	-1.32 (-0.17)
	Over-Optimum Fuel Consumption	4965 (56485)	-30.30 (-2.66)
	Accident Risk Penalty	1981	-39.71
Congested Periods	**Production** (veh · km/day)	**31537**	**34.71**
	Total Objective Function	**8197**	**-68.03**
	Delay Costs	6793	-68.11
	Over-Optimum Emissions Cost	27 (148)	-33.02 (-5.98)
	PM_{10} Contribution	4 (9)	-69.24 (-29.45)
	CO_2 Contribution	18 (50)	-71.66 (-25.29)
	NO_x Contribution	5 (89)	127.20 (7.40)
	Over-Optimum Fuel Consumption	1288 (3605)	-71.80 (-25.65)
	Accident Risk Penalty	89	-17.27
	Killed	56	-33.51
	Major Injuries	24	-0.19
	Minor Injuries	8	42.91
Free – Flowing Periods	**Production** (veh · km/day)	**670028**	**-3.55**
	Total Objective Function	**18686**	**-26.63**
	Delay Costs	12737	-28.48
	Over-Optimum Emissions Cost	380 (2954)	0.91 (0.12)
	PM_{10} Contribution	17 (129)	-8.15 (-1.04)
	CO_2 Contribution	52 (741)	-14.84 (-1.05)
	NO_x Contribution	311 (2083)	4.06 (0.60)
	Over-Optimum Fuel Consumption	3677 (52880)	-15.77 (-1.10)
	Accident Risk Penalty	1892	-40.77
	Killed	1200	-52.39
	Major Injuries	515	-28.54
	Minor Injuries	177	2.33

Note: [1] Potential reduction for over-optimum emissions is 100%.

[2] Values in brackets represent total values (not over-optimum as usual)

[3] Stop & Go periods are defined in relation to the baseline scenario when $v_{(s,t)} \leq 45$ km/h.

Table 6.6 *Comparison of objective function results between After Measured and Simulated scenarios*

Concept (in €/day, unless stated differently)		AM	AS actual	Difference (%)
Whole Day	**Production** (veh · km/day)	**688753**	**681553**	**-1.05**
	Total Objective Function	**16332**	**17598**	**7.75**
	Delay Costs	11276	12632	12.03
	Over-Optimum Emissions Cost	401 (3097)	388 (3015)	-3.28 (-2.64)
	Over-Optimum Fuel Consumption	3 461(54980)	3 396 (53604)	-1.88 (-2.50)
	Accident Risk Penalty	1 195	1 182	-1.05

The monetization of the reduction in pollutant emissions could be another candidate to apply non-linear cost functions. It can be claimed that these benefits are a function of the existing pollutant concentration. This would imply the DSL strategies not only steered by traffic conditions but also by existing pollution. Further research in the quantification of these cost functions is necessary in this case.

6.5 Conclusions and further research

This paper presents a general methodology to assess the overall profitability of advanced traffic management strategies. Efforts have been made to keep the method simple and use only inputs that are usually available in most traffic management centers. This should allow a wider implementation of before – after quantitative assessments in the current practice.
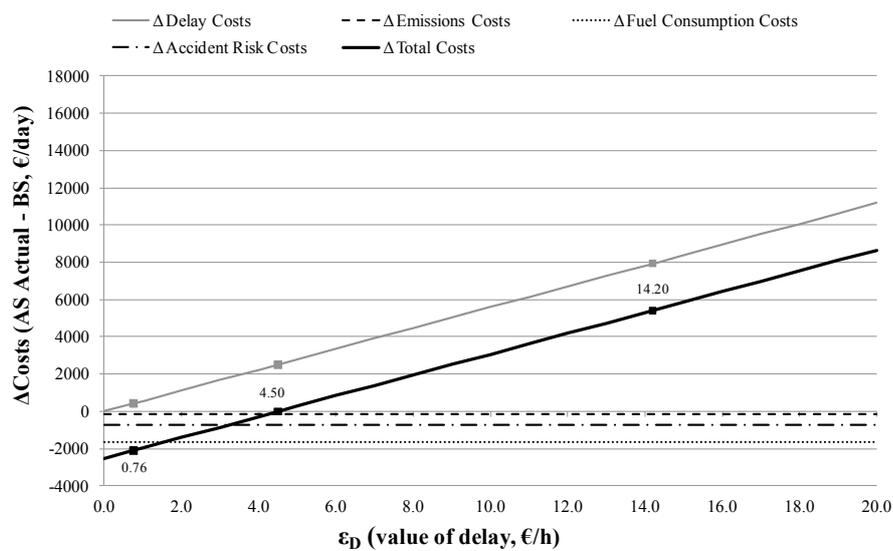
a)
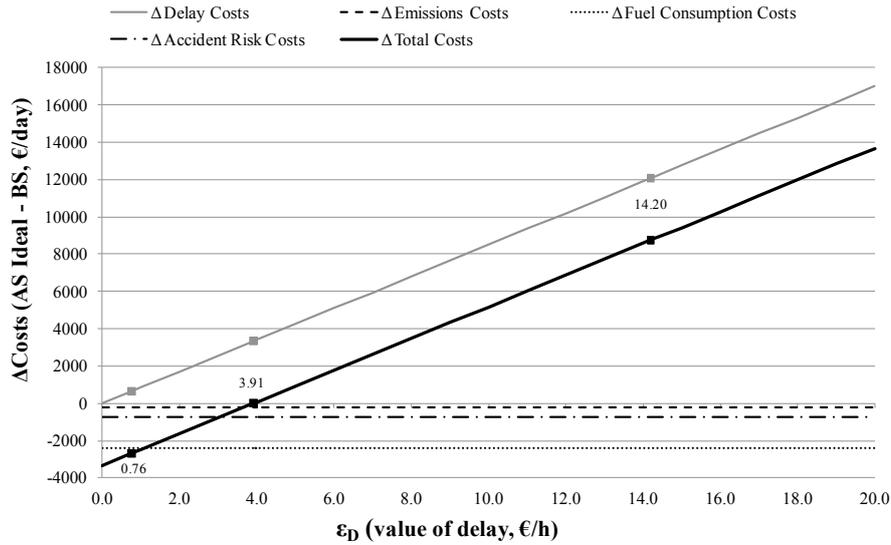
b)

Figure 6.10 *Sensitivity analysis of the objective function results in relation to the value of delay. a) Actual scenario, b) Ideal Scenario*

The proposed objective function focuses on the gross overall result. Detail is only addressed when the impact on the final result is significant (e.g. modeling of speed variability). It is stated in the text that the several terms of the objective function (which account for delays, pollutant emissions, fuel consumption and safety) may be somehow inaccurate. In addition, more "advanced" models may exist. Despite this, the added level of complexity by using these models does not generally translate into more accuracy in the overall result, the main cause being the scarcity of detailed data to feed the models.

The proposed methodology is used to assess the Dynamic Speed Limit (DSL) system installed on a freeway accessing Barcelona (Spain). The DSL algorithm implemented can be considered as "standard", limiting the speed in accordance to the downstream traffic conditions. Readers must be warned that the case study and its results are site specific. This includes a specific composition of the vehicle fleet, a particular calibration of the safety function, specific changes of the speed limits, and a specific freeway with its traffic. All these specificities (that can be considered fairly general and consistent for all metropolitan freeways) may affect the details of the numeric results. However they do not affect the conceptual behavior of the DSL strategies. For instance, it is usually claimed that DSL implies lower but more uniform speeds, and less desirability for lane changing. This may lead to increased freeway capacity and easing of congestion episodes. This theoretically reasonable and practically appealing assumption has not been proven here. No empirical effect is seen in congested periods. The possible inability of DSL strategies in congestion mitigation clearly cuts down its potential, as many of the undesired effects of traffic are related to congested traffic. The reason could be attributed to the difference between the idealized "infinite and uniform" freeway traffic and the real environment with

bottlenecks, weaving sections, on/off ramps, lane drops, drivers prone to traffic violation, etc.

This means that DSL effects are reduced to free flowing episodes (at least in the case the system uses the current standard algorithms). Clearly a trade-off exists in this case for freeways whose design speed is around 120 km/h. Reduction of emissions, fuel consumption, and accident risk may be achieved by reducing allowed speeds. However, this will imply increased free flow travel times. Therefore, the application of such policies should not be indiscriminate but respond to a detailed analysis, as it may result in negative social profits. This is the case in the presented analysis. Despite these discouraging results, one has to bear in mind their strong dependence on the different monetary values given to each term of the objective function. If the cost of traffic externalities increases, the sign of the profits could change. Or alternatively, if marginal costs are applied to the short additional delays caused by the DSL control (instead of the higher average costs) the strategy is shown to be profitable.

More research efforts should be dedicated to how to modify the real effects of DSL strategies. If DSL could be used to obtain more stable traffic speeds in congested periods (avoiding stop & go oscillations), significant emission and fuel consumption reductions would be achieved (note that the reduction potential per veh·km in congested periods is at least 40% higher than in free flowing conditions). Furthermore, the capacity drop effect would be avoided, leading also to improved travel times. All effects turn to benefits in this case.

Smart DSL strategies would be needed. And probably they will not be powerful enough to achieve these results by themselves either. A combination with other traffic management strategies (e.g. lane and hard shoulder management, demand management…) appears to be a way to go.

Finally, it is also concluded from the paper that DSL policies are prone to violation. This will compromise the DSL efficiency. It is claimed here that speeding, while DSL strategy is active, is more related to the limits being counter-intuitive than to the lack of enforcement. The key issue in a so significant amount of speeding is that the drivers are forced to drive at low speeds along an infrastructure designed for much greater speeds. This is a source of anger and an incentive for speeding. A well designed policy implies a great level of compliance even without limitation signals. Infrastructures subject to advanced traffic management strategies should be redesigned for being auto-explicative. Common sense should guide drivers to the correct decisions. Dynamic lane widths, illumination, signals, protective elements, pavement colors and textures … are some possibilities to explore.

PART IV

Conclusions and future research

CHAPTER 7

Conclusions and future research

"Dynamic speed limits adequately coordinated with ramp metering strategies improve the performance of freeway merging bottlenecks, promoting social and environmental sustainability."

This conclusion of the thesis is relevant and encouraging. The topic has raised much controversy among drivers, researchers, practitioners, together with media and politicians. The outcomes of the thesis shed some light on the debate. This and other research contributions should promote the willingness of traffic agencies to improve their decision making processes, making use of solidly grounded analyses. A better infrastructure management is absolutely essential in the present context.

Another issue to be addressed is driver's compliance with Active Traffic Management (ATM) policies. For instance drivers are forced to drive slowly along infrastructures designed for much greater speeds. This is an incentive for speeding, as has been observed with the field data, and is one of the causes for drivers' misunderstanding of the measure. Infrastructures subject to ATM strategies should be physically redesigned to become auto-explicative, guiding drivers to the correct decisions. All these measures together would certainly lead to an improved public opinion towards ATM strategies.

Definitely ATM strategies will play a crucial role in managing the increasing traffic congestion in metropolitan freeways where the "add a lane" solutions are exhausted.

7.1 Main findings and conclusions

In this section the main findings and conclusions are summarized. Detailed conclusions of each part can be found in the corresponding chapters.

Chapter 2 describes the main ATM strategies. A summary of the most remarkable techniques, addressing the most controversial topics is provided. Three ATM strategies are explored in detail: Dynamic Speed Limits (DSL) and managed lanes (on the supply side) and ramp metering (RM) on the demand side. The claimed benefits of DSL strategies are verified in Chapter 5, whereas the main ramp metering principles are applied in the coordinated strategy presented in Chapter 4.

In Chapter 3, the physical feasibility of the Assymetric Cell Transmission Model (ACTM) at uncontrolled merges is recovered. With the proposed calibration procedure the queue length estimation at on-ramps is reliable. The standard simplistic calibration of the ACTM parameters fails on this attempt, basically because uncontrolled merges were implicitly considered out of scope of the model. In fact, when on-ramps are controlled, the merging dynamics are affected, the uncontrolled merging priorities invalidated and the relationship between on-ramp and mainline flow modified. The proposed methodology is simple, condensed in a unique explicit expression and fully consistent with the Newell-Daganzo merge model. In this way, the ACTM suitability for operational planning and management on controlled and uncontrolled freeway corridors is preserved.

The main goal of Chapter 4 is to present an extension of the Cell Transmission Model, i.e. the Capacity Lagged CTM. The model is capable of simulating DSL strategies by introducing a modified inverse-lambda shaped fundamental diagram. It also reproduces the capacity drop phenomenon in a first order model. The model is fully consistent with CTM. It only differs in the addition of some capacity restriction rules at merges. In order to improve accuracy, lagged flow values are considered in order to compute traffic conditions.

A new coordinated RM and DSL strategy is also proposed in Chapter 4. The strategy mainly takes advantage of all the endogenous merge capacity. The coordinated strategy succeeds in increasing the total discharge flow. This implies significant equity and mobility benefits for all the freeway users.

With Chapter 5 starts the empirical part of the thesis. The first objective is to evaluate the effectiveness of DSL policies via an empirical approach. Remarkably, this topic has almost never been addressed in the literature, due to the scarcity of empirical DSL data. The analysis consists in a stratification of loop detector data, considering the records of dynamic speed limits. A novel approach for searching stationary periods in order to characterize traffic fundamental diagrams is also developed. The approach is simple, effective and easy to code. Results show a significant capacity increase for 60 km/h speed limit. Capacity flows are sustained over 3-4 min. This means that DSL strategies are capable of improving freeway performance, at least for certain time periods when suitable speed limit values are set. It is also found that the critical occupancy under DSL is higher than their non-DSL counterpart. The causes of these effects may be attributed to the more

uniform lane utilization observed within the traffic stream in congested conditions and to the higher speed homogeneity in uncongested conditions. These patterns confirm the intuitions referring the DSL features. However, all these assertions cannot be considered conclusive, because data collected during specific speed limit values may not cover the whole range of possible occupancies. This is the case for high speed limit values. Hence, conclusions related to capacity increase or critical occupancy shift should be taken with care.

The case study in Chapter 6 combines all the methods developed in the previous chapters and validates the traffic flow model. In essence, a general methodology to assess the social profitability of any ATM strategy is proposed. The method is simple and uses only inputs that are usually available in most traffic management centers. More advanced models may be appealing. However, the scarcity of detailed data to feed these models makes them difficult to implement widely. The proposed methodology is used to assess the DSL system installed on a freeway accessing Barcelona. Note that the case study and its results are site specific, as it includes a specific composition of the vehicle fleet, a particular calibration of the safety function, specific changes of the speed limits, and a specific freeway with its traffic. All these specificities may affect the details of the numeric results, although they do not affect the conceptual behavior of the DSL strategies. The assessment methodology has been applied both for real and simulated scenarios. The simulation experiment neglects the capacity drop. This means that the obtained results underestimate the true benefits. Recall that one of the benefits of DSL strategies is the capacity drop avoidance. In any case, results show that DSL effects are only significant in free flowing episodes (at least when the system uses the current standard algorithms). A trade-off arises: the reduction of emissions, fuel consumption, and accident risk may be achieved by reducing allowed speeds and increasing travel times. Therefore, special attention should be paid when implementing DSL policies because they may result in negative social profits, as is the case in the presented analysis. However, the strong dependence existing on the different monetary values given to each term of the objective function play a crucial role. Actually, increased costs of traffic externalities could make the sign of the profits change.

After all, it is concluded that DSL strategies probably will not be powerful enough to achieve all the expected benefits by themselves. Despite the fact that smarter DSL algorithms are required, a combination with other ATM strategies (e.g. managed lanes, demand management, route guidance…) seems a promising direction in order to exploit the existing synergies.

7.2 Recommendations for future research

The final section of this dissertation is devoted to outline directions for further research.

Topics related to **modeling** ATM strategies**:**

- *Validation of modeling assumptions.* In Chapter 4 a modeling assumption for the CL-CTM is made: the inverse lambda fundamental diagram shape. This model assumption needs to be validated with real data from freeways with DSL installations.

- *Adapt the coordinated strategy in order to consider the critical value of density as an input.* A main drawback of the coordinated strategy is that the capacity is assumed known in the formulation. The critical value of occupancy is less sensitive and more stable than capacity. Using the critical occupancy as an input would allow a more effective control.

- *Modeling of other control measures.* Other ATM strategies may also show potential synergies capable of further improving the traffic flow throughput. For instance dynamic lane assignment (e.g. hard-shoulder utilization during peak hours) or non-segregated HOV lanes appear as interesting candidates.

- *Modeling and control of capacity drop at other bottleneck types.* Different freeway elements can cause a capacity drop (e.g. merges, diverges, bridges, tunnels, curves or grades). The detailed modeling of these bottlenecks may open up the possibility of better control performance. Empirical validation is also required.

Topics related to **empirical** research:

- *Measure the full range of the flow-occupancy relationship under different speed limits.* It is necessary to obtain fundamental diagrams where the whole range of occupancy values is covered. This needs scenarios where the speed limit is fixed from the onset to the offset of the rush hour episode.

- *Microscopic DSL effects.* Individual vehicle data, without any type of aggregation, makes it possible to compute interesting traffic features, e.g. the homogeneity of speed and occupancy values within the traffic stream or to count the number of lane changes. The macroscopic research on this thesis needs to be completed with the microscopic approach.

- *Empirical characterization of the capacity drop phenomenon.* There is a need to estimate the capacity drop values for a particular freeway in order to improve the ATM control schemes.

References

AASHTO, A. (1977) *Manual on User Benefit Analysis of Highway and Bus-Transit Improvements.* American Association of State Highway and Transportation Officials, Standing on Highways, Washington, D.C., USA.

Abdel-Aty, M., A. Pande, C. Lee, V. Gayah and C.D. Santos. (2007) Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. *Journal of Intelligent Transportation Systems,* 11, (3), 107-120.

Alecsandru, C., A. Quddus, K.C. Huang, B. Rouhieh, A.R. Khan, et al. (2011) An Assessment of the Cell-Transmission Traffic Flow Paradigm: Development and Applications. *Proceedings of the 90th Transportation Research Board Annual Meeting.* Washington, D.C., USA

André, M. (2004) The ARTEMIS European driving cycles for measuring car pollutant emissions. *Science of the total Environment,* 334, 73-84.

Aron, M., R. Seidowsky and S. Cohen. (2013) Safety impact of using the hard shoulder during congested traffic. The case of a managed lane operation on a French urban motorway. *Transportation Research Part C,* 28, 168-180.

Asensio, J., and A. Matas. (2008) Commuters' valuation of travel time variability. *Transportation Research Part E: Logistics and Transportation,* 44, (6), 1074-1085.

Aw, A. and M. Rascle. (2000) Resurrection of "Second Order" Models of Traffic Flow. *SIAM Journal on Applied Mathematics* Vol. 60 (3), pp. 916-938.

Baldasano, J.M., M. Gonçalves, A. Soret and P. Jiménez-Guerrero. (2010) Air pollution impacts of speed limitation measures in large cities: The need for improving traffic data in a metropolitan area. *Atmospheric Environment,* 44(25), 2997-3006.

Banks, J.H. (2000) Are minimization of delay and minimization of freeway congestion compatible ramp metering objectives? *Transportation Research Record: Journal of the Transportation Research Board,* 1727, 112-119.

Banks, J.H. (1990) Flow processes at a freeway bottleneck. *Transportation Research Record: Journal of the Transportation Research Board,* 1287, 20-28.

Bar-Gera, H. and S. Ahn. (2010) Empirical macroscopic evaluation of freeway merge-ratios. *Transportation Research Part C,* 18, (4), 457-470.

Barth, M., F. An, T. Younglove, C. Levine, G. Scora, M. Ross and T. Wenzel. (2000). *The development of a comprehensive modal emissions model.* Final Report Number: 25-21. National Cooperative Highway Research Program, Washington, DC.

Bogenberger, K. and A. D. May. (1999) *Advanced coordinated traffic responsive ramp metering strategies.* UC Berkeley: California Partners for Advanced Transit and Highways (PATH).

Brinckerhoff, P. (2010) *Synthesis of active traffic management experiences in europe and the united states.* Final Report Number: FHWA-HOP-10-031. Federal Highway Administration.

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

Cassidy, M. J., and Ahn, S. (2005) Driver turn-taking behavior in congested freeway merges. *Transportation Research Record: Journal of the Transportation Research Board,* 1934, 140-147.

Cassidy, M. J. and R. L. Bertini. (1999) Some traffic features at freeway bottlenecks. *Transportation Research Part B,* 33, (1), 25-42.

Cassidy, M. J. (1998) Bivariate relations in nearly stationary highway traffic. *Transportation Research Part B,* 32(1), 49-59.

Cassidy, M. J. and J. R. Windover. (1995) Methodology for assessing dynamics of freeway traffic flow. *Transportation Research Record: Journal of the Transportation Research Board,* 1484, 73-79.

Cassidy, M.J., K. Jang and C.F. Daganzo. (2010) The smoothing effect of carpool lanes on freeway bottlenecks. *Transportation Research Part A,* 44(2), 65-75.

Chen, C., J. Kwon, J. Rice, A. Skabardonis and P. Varaiya. (2003) Detecting errors and imputing missing data for single-loop surveillance systems. *Transportation Research Record: Journal of the Transportation Research Board,* 1855, 160-167.

Chen, C., P. Varaiya and J. Kwon. (2005) An empirical assessment of traffic operations. *Proceedings of the 16th International Symposium on Transportation and Traffic Theory.* Maryland, USA, 105-123.

Chow, A.H., G. Gomes, A.A. Kurzhanskiy, P. Varaiya, G. Gomes, et al. (2010) AURORA RNM - A macroscopic simulation tool for arterial traffic modeling and control. *Proceedings of the 89th Transportation Research Board Annual Meeting.* Washington, D.C., USA.

Chung, C. L. and W. Recker. (2011) State-of-the-Art Assessment of Toll Rates for High-Occupancy and Toll Lanes. *Proceedings of the 90th Transportation Research Board Annual Meeting.* Washington, D.C., USA.

Cremer, M. (1979) *Der verkehrsfluss auf schnellstrassen: modelle, Überwachung, regelung.* Springer-Verlag. In German.

Daganzo, C. F. (2005) A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B,* 39, (2), 187-196.

Daganzo, C. F. (1999) The lagged cell-transmission model. *14th International Symposium on Transportation and Traffic Theory.* 81-104. Jerusalem, Israel.

Daganzo, C. F. (1997) A continuum theory of traffic dynamics for freeways with special lanes. *Transportation Research Part B,* 31(2), 83-102.

Daganzo, C. F. (1995) The cell transmission model, part II: network traffic. *Transportation Research Part B,* 29(2), 79-93.

Daganzo, C. F. (1994) The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B,* 28(4), 269-287.

Daganzo, C.F., J. Laval and J.C. Muñoz. (2002) Ten strategies for freeway congestion mitigation with advanced technologies. *Traffic Engineering and Control,* 43, 397-403.

De Haan, P. and M. Keller. (2004) Modelling fuel consumption and pollutant emissions based on real-world driving patterns: the HBEFA approach. *International Journal Environmental Pollution,* 22(3), 240-258.

Dervisoglu, G., G. Gomes, J. Kwon, R. Horowitz and P.P. Varaiya. (2009) Automatic calibration of the fundamental diagram and empirical observations on capacity. *Proceedings of the 88th Transportation Research Board Annual Meeting.* Washington, D. C.

Desnouailles, C. and S. Cohen. (2007) Managed Lanes: A French project to reduce congestion on motorways. *Proceedings of the i2tern Conference.* Aalborg, Denmark.

Edie, L. C. (1961) Car-Following and Steady-State Theory for Noncongested Traffic. *Operations Research,* 9(1), 66-76.

Elvik, R., P. Christensen and A. Amundsen. (2005) Speed and road accidents: an evaluation of the power model. *Nordic Road & Transport Research,* 17(1), 9-11.

European Commission DG Environment. (2000). *BeTa Version E1.02a.* Benefits Table database: Estimates of the marginal external costs of air pollution in Europe. Available online: http://ec.europa.eu/environment/enveco/air/.

European Parliament, C. (2008) *Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe.*

European Transport Safety Council. (2003) *Cost effective EU transport safety measures.* Available online: http://www.etsc.eu/documents/costeff.pdf

Federal Highway Administration. (2008) *Federal-Aid Highway Program Guidance on High Occupancy Vehicle (HOV) Lanes.* U.S. Department of Transportation. Federal Highway Administration,Washington, DC.

Geroliminis, N., A. Srivastava and P. Michalopoulos. (2011) A Dynamic-Zone-Based Coordinated Ramp-Metering Algorithm With Queue Constraints for Minnesota's Freeways. *IEEE Transactions on Intelligent Transportation Systems,* 12, (4), 1576-1586.

Godunov, S. (1959) A difference method for numerical calculation of discontinuous solutions of hydrodynamic equations. Matematicheskii Sbornik. *Matematicheskii Sbornik,* 47, 271-290.

Gomes, G. C. (2004) Optimization and Microsimulation of On-ramp Metering for Congested Freeways. University of California at Berkeley. PhD thesis,

Gomes, G. and R. Horowitz. (2006) Optimal freeway ramp metering using the asymmetric cell transmission model. *Transportation Research Part C,* 14(4), 244-262.

Gomes, G., R. Horowitz, A.A. Kurzhanskiy, P. Varaiya and J. Kwon. (2008) Behavior of the cell transmission model and effectiveness of ramp metering. *Transportation Research Part C,* 16(4), 485-513.

Greenshields, B.D., J. Thompson, H. Dickinson and R. Swinton. (1934) The photographic method of studying traffic behavior. *Proceedings of the 13th Annual Meeting of the Highway Research Board.* 382-399.

Hadi, M. A. (2005) Coordinated Traffic Responsive Ramp Metering Strategies-An Assessment Based on Previous Studies. *Proceedings of the 12th World Congress on Intelligent Transport Systems.* San Francisco, USA,

Hadiuzzaman, M., T. Qiu and X. Lu. (2013) Variable Speed Limit Control Design for Relieving Congestion Caused by Active Bottlenecks. *Journal of Transportation Engineering* 139(4), 358-370.

Halbert, M., and Orme, L. (2008) Safety on the M42 Active Traffic Management Project. *Traffic Engineering & Control,* 49, (10), 397-401.

Hall, R. W. and C. Caliskan. (1999) Design and evaluation of an automated highway system with optimized lane assignment. *Transportation Research Part C,* 7, (1), 1-15.

Hegyi, A. (2004) Model Predictive Control for Integrating Traffic Control Measures. PhD thesis, Delft University of Technology (Netherlands).

Hegyi, A. and S. Hoogendoorn. (2010) Dynamic speed limit control to resolve shock waves on freeways-Field test results of the SPECIALIST algorithm. *Proceedings of the 13th*

International IEEE Conference on Intelligent Transportation Systems. 519-524. Madeira Island, Portugal

Hegyi, A., S.P. Hoogendoorn, M. Schreuder, H. Stoelhorst and F. Viti. (2008) SPECIALIST: A dynamic speed limit control algorithm based on shock wave theory. *Proceedings of the 11th International IEEE Conference on Intelligent Transportation Systems.* 827-832. Beijing, China.

Hegyi, A., B. De Schutter and H. Hellendoorn. (2005) Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C,* 13, (3), 185-209.

Heydecker, B. G. and J. D. Addison. (2011) Analysis and modelling of traffic flow under variable speed limits. *Transportation Research Part C,* 19(2), 206-217.

Highways Agency. (2004). *M25 Controlled Motorways. Summary report, issue 1.* Highways Agency Publications. Department of Transport. Birmingham, United Kingdom.

Ishak, S., C. Alecsandru and D. Seedah. (2006) Improvement and evaluation of cell-transmission model for operational analysis of traffic networks: freeway case study. *Transportation Research Record: Journal of the Transportation Research Board,* 1965(1), 171-182.

Jacobson, L., K. Henry and O. Mehyar. (1989) Real-time metering algorithm for centralized control. *Transportation Research Record: Journal of the Transportation Research Board,* 1732, 20-32.

Jella, R.K., S.R. Sunkari, W.L. Gisler, N.J. Rowan and C.J. Messer. (1993) *Space Management: An Application of Dynamic Lane Assignment.* Research Report 1232-18. University of Texas at Austin, Center for Transportation Research, Austin, Texas, USA.

Jin, W. L., and H. M. Zhang. (2003) On the distribution schemes for determining flows through a merge. *Transportation Research Part B,* 37(6), 521-540.

Keller, J., S. Andreani-Aksoyoglu, M. Tinguely, J. Flemming, J. Heldstab, et al. (2008) The impact of reducing the maximum speed limit on motorways in Switzerland to 80 km h−1 on emissions and peak ozone. *Environmental Modelling & Software,* 23(3), 322-332.

Kerner, B. S. and H. Rehborn. (1996) Experimental features and characteristics of traffic jams. *Physical Review E,* 53, (2), R1297-R1300.

Kerner, B. S. (2013) Criticism of generally accepted fundamentals and methodologies of traffic and transportation theory: A brief review. *Physica A: Statistical Mechanics and its Applications* Vol. 392 (21), pp. 5261-5282.

Keuken, M., E. Sanderson, R. Aalst, J. Borken and J. Schneider. (2005) *Contribution of traffic to levels of ambient air pollution in Europe.* In: M. Krzyzanowski, B. Kuna-Dibbert and J. Schneider, Editors, Health effects of transport-related air pollution, World Health Organization, Geneva.

Kotsialos, A., M. Papageorgiou and A. Messmer. (1999) Optimal coordinated and integrated motorway network traffic control. *Proceedings of the 14th International Symposium on Transportation and Traffic Theory.* 621-644.

Kotsialos, A., M. Papageorgiou, C. Diakaki, Y. Pavlis and F. Middelham. (2002) Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET. *IEEE Transactions on Intelligent Transportation Systems,* 3(4), 282-292.

Kuhne, R. D. and T. Beckschule. (1993) Non-linearity stochastics of unstable traffic flow. *Proceedings of the International Symposium on Transportation and Traffic Theory.* 367-386. New York, USA.

Kurzhanskiy, A. A. and J. Kwon. (2008). *AURORA Road Network Modeler, User guide.* TOPL Group. UC Berkeley.

Kurzhanskiy, A.A., J. Kwon and P.P. Varaiya. (2009) Aurora Road Network Modeler. *Control in Transportation Systems.* 204-210.

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

Kurzhanskiy, A. A. and P. Varaiya. (2010b) Using Aurora road network modeler for active traffic management. *Proceedings of the American Control Conference.* 2260-2265.

Kwon, J., and P. Varaiya. (2008) Effectiveness of California's High Occupancy Vehicle (HOV) system. *Transportation Research Part C,* 16(1), 98-115.

Laval, J. A., and C.F. Daganzo. (2006) Lane-changing in traffic streams. *Transportation Research Part B,* 40(3), 251-264.

Laval, J. A. and L. Leclercq. (2008) Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. *Transportation Research Part B,* 42(6), 511-522.

Lebacque, J. (2003) Two-phase bounded-acceleration traffic flow model: analytical solutions and applications. *Transportation Research Record: Journal of the Transportation Research Board,* 1852(1), 220-230.

Lebacque, J. (1996) The Godunov scheme and what it means for first order traffic flow models. *Proceedings of the Internaional symposium on transportation and traffic theory.* 647-677. Lyon, France.

Leclercq, L., J.A. Laval and N. Chiabaut. (2011) Capacity drops at merges: An endogenous model. *Transportation Research Part B,* 45(9), 1302-1313.

Lee, C., B. Hellinga and F. Saccomanno. (2006) Evaluation of variable speed limits to improve traffic safety. *Transportation Research Part C,* 14(3), 213-228.

Lennie, S., Han, C., Luk, J., Pyta, V. and P. Cairney. (2009). *Best Practice for Variable Speed Limits: Literature Review.* Austroads Incorporated. Sydney, Australia.

Lighthill, M. and G. Whitham. (1955) On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences,* 317-345.

Lin, P., K. Kang and G. Chang. (2004) Exploring the Effectiveness of Variable Speed Limit Controls on Highway Work-Zone Operations. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations,* 8(3), 155-168.

Long, S., L. Gentry and G.H. Bham. (2012) Driver perceptions and sources of user dissatisfaction in the implementation of variable speed limit systems. *Transport Policy,* 23, 1-7.

Lu, X.Y., P. Varaiya, R. Horowitz, D. Su and S.E. Shladover. (2010) A new approach for combined freeway Variable Speed Limits and Coordinated Ramp Metering. *Proceedings of the 13th International IEEE Conference on Intelligent Transportation System.* 491-498. Madeira Island, Portugal.

Masher, D.P., D.W. Ross, P.J. Wong, P.L. Tuan, H.M. Zeidler, et al. (1975) *Guidelines for design and operating of ramp control systems.* SRI Project 3340. SRI, Menid Park, CA, Stanford Res. Inst. Rep. NCHRP 3-22,

May, A. D. (1975) *A proposed dynamic freeway control system hierarchy.* Institute of Transportation and Traffic Engineering, University of California, Berkeley.

Menendez, M. and C. F. Daganzo. (2007) Effects of HOV lanes on freeway bottlenecks. *Transportation Research Part B: Methodological,* 41(8), 809-822.

Messmer, A. and M. Papageorgiou. (1990) METANET: a macroscopic simulation program for motorway networks. *Traffic Engineering & Control,* 31(9), 466-470.

Mirshahi, M., J.T. Obenberger, C.A. Fuhs, C.E. Howard, R.A. Krammes, et al. (2007) *Active Traffic Management: The Next Step in Congestion Management.* Federal Highway Administration. Washington, DC.

Monamy, T., H. Haj-Salem and J. Lebacque. (2012) A Macroscopic Node Model Related to Capacity Drop. *Procedia - Social and Behavioral Sciences,* 54, 1388-1396.

Muralidharan, A. and R. Horowitz. (2009) Imputation of ramp flow data for freeway traffic simulation. *Transportation Research Record: Journal of the Transportation Research Board,* 2099(1), 58-64.

Newell, G. F. (1993) A simplified theory of kinematic waves in highway traffic, Part II: Queueing at freeway bottlenecks. *Transportation Research Part B,* 27(4), 289-303.

Newell, G. F. (1982) *Applications of queueing theory.* Chapman & Hall,New York.

Ni, D. and J. D. Leonard II. (2005) A simplified kinematic wave model at a merge bottleneck. *Applied Mathematical Modelling,* 29, (11), 1054-1072.

Ntziachristos, L., Samaras, Z., Eggleston, S., Gorissen, N., Hassel, D. and A. Hickman. (2000). *COPERT III Computer Programme to calculate emissions from road*

transport, methodology and emission factors (version 2.1). Rep. No. 74, European Environment Agency (EEA), Copenhagen, Denmark.

Olde, M., P. van Beek, M. Stemerding and P. Havermans. (2005) Reducing speed limits on highways: Dutch experiences and impact on air pollution, noise-level, traffic safety and traffic flow. *Proceedings of ETC.* Strasbourg, France.

Papageorgiou, M. (1983) A hierarchical control system for freeway traffic. *Transportation Research Part B,* 17(3), 251-261.

Papageorgiou, M., H. Hadj-Salem and J.M. Blosseville. (1991) ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record: Journal of the Transportation Research Board,* 1320, 58-64.

Papageorgiou, M. and I. Papamichail. (2007). *Handbook of Ramp Metering.* Report for EUropean RAmp Metering Project (EURAMP).

Papageorgiou, M., E. Kosmatopoulos and I. Papamichail. (2008) Effects of Variable Speed Limits on Motorway Traffic Flow. *Transportation Research Record: Journal of the Transportation Research Board,* 2047, 37-48.

Papageorgiou, M. and A. Kotsialos. (2002) Freeway ramp metering: an overview. *IEEE Transactions on Intelligent Transportation Systems,* 3(4), 271-281.

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

Payne, H. J. (1971) Models of freeway traffic control. *Mathematical Models of Public System of Simulation Council Proceedings Series,* (28), 51-61.

Piotrowicz, G. and J. Robinson. (1995) *Ramp Metering Status in North America* Report Number: DOT-T-95-17, Department of Transportation, Washington, D.C., USA.

Ramoneda, D. (2013) *Anàlisi del flux del trànsit a la autopista C32 amb sistemes actius de velocitat variable.* ETSEIB-Barcelona Tech. Master thesis. Directed by: JM Torné and F. Soriguera. In Catalan.

Remeijn, H. (1982) *The Dutch Motorway Control and Signalling System.* Rijkswaterstaat, Traffic Engineering Division.

Richards, P. I. (1956) Shock waves on the highway. *Operations Research,* 4(1), 42-51.

Richardson, H. W. and C. H. C. Bae. (2008) *Road congestion pricing in Europe: implications for the United States.* Edward Elgar Publishing.

Rodriguez, R. J. (1990) Speed, speed dispersion, and the highway fatality rate. *Southern Economic Journal,* 349-356.

Samaras, Z. and L. Ntziachristos. (2010). *EMEP/EEA air pollutant emission inventory guidebook.1.A.3.b Road transport update June 2010.* Report Number: 9/2009, European Environment Agency.

Schijns, S. and P. Eng. (2006) High Occupancy Vehicle Lanes–Worldwide Lessons for European Practitioners. *Urban Transport XII: Urban Transport and the Environment in the 21st Century,* 89, 181.

SENDECO2. (2011). *European Bourse for European Unit Allowances and Carbon Credits.* www.sendeco2.com. Accessed April, 2011.

Smit, R., M. Poelman and J. Schrijver. (2008) Improved road traffic emission inventories by adding mean speed distributions. *Atmospheric Environment,* 42(5), 916-926.

Smulders, S. (1992) Control By Variable Speed Signs — The Dutch Experiment. *Proceedings of the 6th International Conference on Road Traffic Monitoring and Control.* 99-103. London, United Kingdom.

Smulders, S. (1990) Control of freeway traffic flow by variable speed signs. *Transportation Research Part B,* 24(2), 111-132.

Soriguera, F. (2010) Gestión variable del límite de velocidad en autopistas metropolitanas: un análisis conceptual. *Proceedings of the IX Congreso de Ingeniería del Transporte.* Madrid, Spain. In Spanish.

Soriguera, J. and F. Soriguera. (2009) *Traffic stream analysis in AP7 highway.* ETSECCPB-UPC. Master Thesis. Directed by F. Soriguera. Available online: http://hdl. handle. net/2099.1/8473, .

Soriguera, F. and F. Robusté. (2011) Estimation of traffic stream space mean speed from time aggregations of double loop detector data. *Transportation Research Part C,* 19(1), 115-129.

Soriguera, F., J.M. Torné and D. Rosas. (2013) Assessment of Dynamic Speed Limit Management on Metropolitan Freeways. *Journal of Intelligent Transportation Systems,* 17(1), 78-90.

Soriguera, F. and M. Sala. (2014) Experimenting with Dynamic Speed Limits on Freeways. *Submitted for publication in the 93th Annual Meeting Transportation Research Board.* Washington, D.C., USA.

Spanish Ministry of Industry, Tourism and Commerce. (2001). *Precios de carburantes y combustibles. Datos de Enero de 2011. www.mityc.es/energia.* Accessed April 2011. In Spanish.

Srivastava, A. and N. Geroliminis. (2013) Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C,* 30, 161-177.

Stoelhorst, H. (2008) Reduced speed limits for local air quality and traffic efficiency. *Procedings of the 7th European Congress and Exhibition on Intelligent Transport Systems and Services.* Genova, Italy.

Sultan, B., R. Meekums and M. Brown. (2008) The impact of Active Traffic Management on motorway operation. *Road Transport Information and Control-RTIC and ITS United Kingdom Members' Conference, IET.* 1-8. Manchester, United Kingdom.

Sumner, R. L. and C. M. Andrew. (1990) Variable Speed Limit System. *US Department of Transportation, Federal Highway Administration,* .

Szeto, W. Y. (2008) Enhanced lagged cell-transmission model for dynamic traffic assignment. *Transportation Research Record: Journal of the Transportation Research Board,* 2085(1), 76-85.

Tampère, C.M.J., R. Corthout, D. Cattrysse and L.H. Immers. (2011) A generic class of first order node models for dynamic macroscopic simulation of traffic flows. *Transportation Research Part B,* 45(1), 289-309.

Tarko, A. and P. Songchitruksa. (2005) Measuring Roadway Safety. *Conference proceedings: Road Safety on Four Continents,* organized by Swedish National Road and Transport Research Institute. Warsaw, Poland.

Thomas, R. (2006) Assessing the benefits of options to improve the UK's air quality. *Local Transport Today,* (441), 1-10.

Torné, J.M., D. Rosas and F. Soriguera. (2011) Evaluation of speed limit management on C-32 highway access to Barcelona. *Proceedings of the 90th Transportation Research Board Annual Meeting.* Washington, D.C., USA.

Torné, J.M., F. Soriguera and N. Geroliminis. (2012) Modifications of the asymmetric cell transmission model for modeling variable speed limit strategies. *Proceedings of the 1st European Symposium on Quantitative Methods in Transportation Systems.* Lausanne, Switzerland.

Torné, J.M., F. Soriguera and D. Ramoneda. (2013) Empirical evidences of dynamic speed limits impact on a metropolitan motorway. *Submitted in the 18th Pan-American Conference of Traffic and Transportation Engineering and Logistics.*

Torné, J.M., F. Soriguera and D. Rosas. (2010) Evaluación de la gestión variable del límite de velocidad en la autopista metropolitana C-32 en el acceso a Barcelona. *Proceedings of the IX Congreso de Ingeniería del Transporte.* Madrid, Spain. In Spanish.

Treiterer, J. and J. A. Myers. (1974) The hysteresis phenomena in traffic flow. *Proceedings of the 6th Symposium on Transportation and Traffic Theory.* 13-38.

Tuerprasert, K., and Aswakul, C. (2010) Multiclass Cell Transmission Model for Heterogeneous Mobility in General Topology of Road Network. *Journal of Intelligent Transportation Systems,* 14(2), 68-82.

Turnbull, K.F., H.S. Levinson, R.H. Pratt, I. Evans and K.U. Bhatt. (2006) *HOV Facilities-Traveler Response to Transportation System Changes.* Transportation Research Board TCRP Report 95. Washington, D.C., USA

van den Hoogen, E. and S. Smulders. (1994) Control by variable speed signs: Results of the Dutch experiment. *Proceedings of the 7ᵗʰ International Conference on Road Traffic Monitoring and Control,* 145-149.

van Wageningen-Kessels, F. (2013) *Multi-class continuum traffic flow models: Analysis and simulation methods.* PhD thesis, Delft University of Technology (Netherlands).

Wang, Z. and C. M. Walton. (2006). *An Investigation on the Environmental Benefits of a Variable Speed Control Strategy* Rep. No. SWUTC/06/473700-00072-1*,* University of Texas at Austin, Center for Transportation Research, Austin, Texas, USA.

Zackor, H. (1979) Self-sufficient control of speed on freeways. *Proceeding of the International Symposium on Traffic Control Systems,* 226-249.

Zackor, H. (1972) Beurteilung verkehrsabhängiger Geschwindigkeitsbeschränkungen auf Autobahnen. *Strassenbau Und Strassenverkehrstech,* (128), 1-61. In German.

Zackor, H. and M. Papageorgiou. (1991) Speed limitation on freeways: Traffic-responsive strategies. *Concise Encyclopedia of Traffic and Transportation Systems,* 507–511.

Zarean, M., M.D. Robinson and D. Warren. (2000) Applications of Variable Speed Limit Systems to Enhance Safety. *Proceedings of the 7th World Congress on ITS.* Turin, Italy.