Transportation Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering

Rodrigo C. Carlson, Ioannis Papamichail, Markos Papageorgiou, Albert Messmer,

To cite this article:
Rodrigo C. Carlson, Ioannis Papamichail, Markos Papageorgiou, Albert Messmer, (2010) Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering. Transportation Science 44(2):238-253. https://doi.org/10.1287/trsc.1090.0314

TRANSPORTATION SCIENCE
Vol. 44, No. 2, May 2010, pp. 238–253
ISSN 0041-1655 | EISSN 1526-5447 | 10 | 4402 | 0238

informs

DOI 10.1287/trsc.1090.0314

Optimal Motorway Traffic Flow Control Involving Variable Speed Limits and Ramp Metering

Rodrigo C. Carlson
Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania 73100, Greece, and The Capes Foundation, Ministry of Education of Brazil, Brasília DF 70359-970, Brazil, rcarlson@dssl.tuc.gr

Ioannis Papamichail, Markos Papageorgiou
Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania 73100, Greece {ipapa@dssl.tuc.gr, markos@dssl.tuc.gr}

Albert Messmer
D-82402 Seeshaupt, Germany, albert.messmer@t-online.de

The impact of variable speed limits (VSL) on aggregate traffic flow behaviour on motorways is shown to bear similarities to the impact of ramp metering, in particular, when addressing potentially active bottlenecks. A quantitative model of the VSL impact is proposed that allows for VSL to be incorporated in a macroscopic second-order traffic flow model as an additional control component. The integrated motorway network traffic control problem involving ramp metering and VSL control measures is formulated as a constrained discrete-time optimal control problem and is solved efficiently even for large-scale networks by a suitable feasible direction algorithm. An illustrative example of a hypothetical motorway stretch is investigated under different control scenarios, and it is shown that traffic flow efficiency can be substantially improved when VSL control measures are used, particularly in integration with coordinated ramp metering.

Key words: variable speed limits; ramp metering; motorway traffic control; optimal control; fundamental diagram
History: Received: January 2009; revision received: July 2009; accepted: November 2009. Published online in *Articles in Advance* March 23, 2010.

1. Introduction

In the last decades, motorways have become notorious sites of extensive daily traffic congestion, particularly in and around metropolitan areas. Traffic congestion degrades the available infrastructure in the sense of reducing the motorway throughput. Thus, the expensive motorway infrastructure is underutilised ironically exactly at the times (peak hours) it is most urgently needed. The efficient, safe, and less-polluting transportation of persons and goods on motorways calls for an optimal utilisation of the available infrastructure via suitable application of a variety of traffic control measures such as ramp metering, driver information, route guidance, and variable speed limits (VSL). A number of methodological approaches including optimal control, expert systems, fuzzy systems, neural networks, and feedback control have been developed in the past for the design of related control strategies. In the present paper, the optimal control approach is employed for integrated ramp metering and VSL measures.

Ramp metering aims at improving the traffic conditions by appropriately regulating the inflow from the on-ramps to the motorway mainstream. Because ramp storage space may be limited but also because of equity considerations, ramp metering should be applied at multiple ramps (coordinated ramp metering) for maximum efficiency; coordinated ramp metering strategies make use of measurements from an entire region of the network to control all metered ramps included therein (Papageorgiou and Kotsialos 2002). Despite the ramp queue delays induced by ramp metering actions, the higher motorway throughput (and reduced mainstream delays) because of congestion reduction or avoidance may lead to shorter total travel times for most drivers. Coordinated ramp metering has been extensively studied in the past and involves sophisticated methods such as multivariable control strategies (Papageorgiou, Blosseville, and Haj-Salem 1990b; Diakaki and Papageorgiou 1994) and optimal control strategies (Papageorgiou and Mayr 1982; Chen et al. 1997; Zhang and Recker 1999; Kotsialos et al. 2002a; Kotsialos and Papageorgiou 2004; Gomes and Horowitz 2006).

Another control measure for traffic flow on motorways is the display of variable speed limits (VSL) on

appropriate variable message signs (VMS) in response to the prevailing traffic conditions. In most cases, VSL are mandatory, i.e., legally equivalent to fixed speed limits, and may even be automatically enforced to increase driver compliance and hence impact. VSL installations were first introduced in Germany more than three decades ago; today, numerous VSL installations are encountered in many European countries (e.g., a total of more than 800 km of VSL-equipped motorway stretches are currently in operation in Germany) and in North America and elsewhere, and their number is increasing at an accelerated pace.

A main targeted result of VSL is enhanced traffic safety and indeed the selection of motorway stretches for VSL installation in several countries is guided by the frequency of registered accidents. The positive impact of VSL on traffic safety is because of speed reduction and speed homogenisation that are correlated with a reduction of accident probability. Multiyear evaluations of the VSL impact on traffic safety indicate a reduction in accident numbers by as much as 20% to 30% after VSL installation. VSL are also envisaged by some authorities as a means to reduce vehicle emissions and road noise. On the other hand, to the best of the authors' knowledge, there is no evaluation of the VSL impact of available installations that would demonstrate a consistent and measurable improvement of traffic flow efficiency, e.g., in the sense of reduced travel times.

The ideal exploitation of the opportunities offered by VSL would be to preserve the safety and environmental benefits offered by the current systems along with an increase of traffic flow efficiency. The fact that an efficiency increase could not be demonstrated in the conducted field assessments does not necessarily mean that VSL per se is not an appropriate measure for the enhancement of traffic flow efficiency. As a matter of fact:

- The impact of VSL on aggregate traffic behaviour, e.g., on the fundamental diagram, has not been sufficiently investigated with real data. As a consequence, our understanding of even qualitative (let alone quantitative) impacts of VSL is limited to conjectures and assumptions; this lack of reliable understanding, by its turn, hinders the insightful development of VSL control strategies that would target an increase of traffic flow efficiency.
- Current VSL installations employ simple rule-based control strategies for VSL switching, which base their real-time decisions on preselected thresholds of traffic flow or occupancy or mean speed. The utilised thresholds are usually selected in an ad hoc way that does not necessarily exploit the (anyhow unknown) potential impact of VSL on traffic flow efficiency.

The design of pertinent control strategies that may increase traffic flow efficiency calls for a sufficiently accurate description of the VSL impact on the aggregate (macroscopic) traffic conditions. There were very few investigations in the past addressing the precise impact of VSL on aggregate traffic flow behaviour, e.g., on the fundamental diagram. Recently, the effect of VSL on the aggregate traffic flow behaviour (in form of the flow-occupancy diagram) was investigated by Papageorgiou, Kosmatopoulos, and Papamichail (2008) on the basis of traffic data from a VSL-equipped European motorway. This enhanced understanding of the VSL impact may eventually be exploited to assess and enhance the current VSL control strategies or to develop new control strategies that would target traffic flow efficiency while delivering similar benefits for traffic safety and environmental improvements as achieved by current systems.

The identified impact of VSL on the aggregate traffic flow behaviour is reflected in the quantitative model proposed in the present paper. Section 2 addresses some background issues related to VSL and highlights its resemblance to ramp metering and opportunities to avoid the detrimental capacity drop at active bottlenecks. VSL are incorporated in a general second-order traffic flow model (§3) as an additional control component based on the results of the aforementioned validation study. The augmented model leads to an accordingly extended optimal control formulation that is solved by a powerful numerical algorithm (§4). An illustrative example is discussed in §5 under different control scenarios and conclusions are summarised in §6.

2. Background Issues

2.1. The Fundamental Diagram

Under the assumption that traffic conditions do not change substantially in space (i.e., along a motorway stretch) and time (e.g., because of the arrival of shock waves from downstream), the traffic flow states may be approximated by the so-called fundamental diagram, which may be a flow-occupancy (or flow-density) diagram (inverse U shape) or a speed-flow diagram (left-turned U shape) (see Figure 1). Recall that the mean speed of a particular traffic state on the flow-occupancy diagram (Figure 1(a)) is proportional to the slope of the line that connects the particular traffic state point with the origin. A fundamental diagram may be (partially) obtained by collecting measurements of the related traffic variables (flow, occupancy, mean speed) at a specific motorway location and fitting an appropriate mathematical function. This procedure, however, may lead to flawed results if the underlying spatiotemporal traffic flow phenomena are not appropriately considered. In particular, the area around the critical occupancy (capacity flow) is properly visible in real data only at active

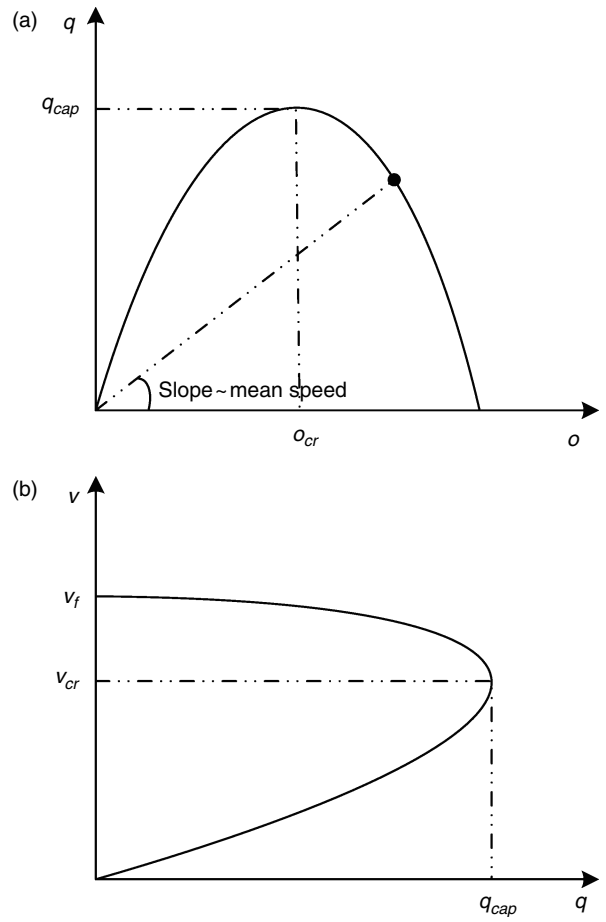

Figure 1 Flow Occupancy (a) and Speed-Flow (b) Diagrams

Notes. In the diagrams above, q is flow (veh/h); o is occupancy (%); v is mean speed (km/h); q_{cap} is capacity flow; o_{cr} is critical occupancy; v_f is free speed; v_{cr} is critical mean speed.

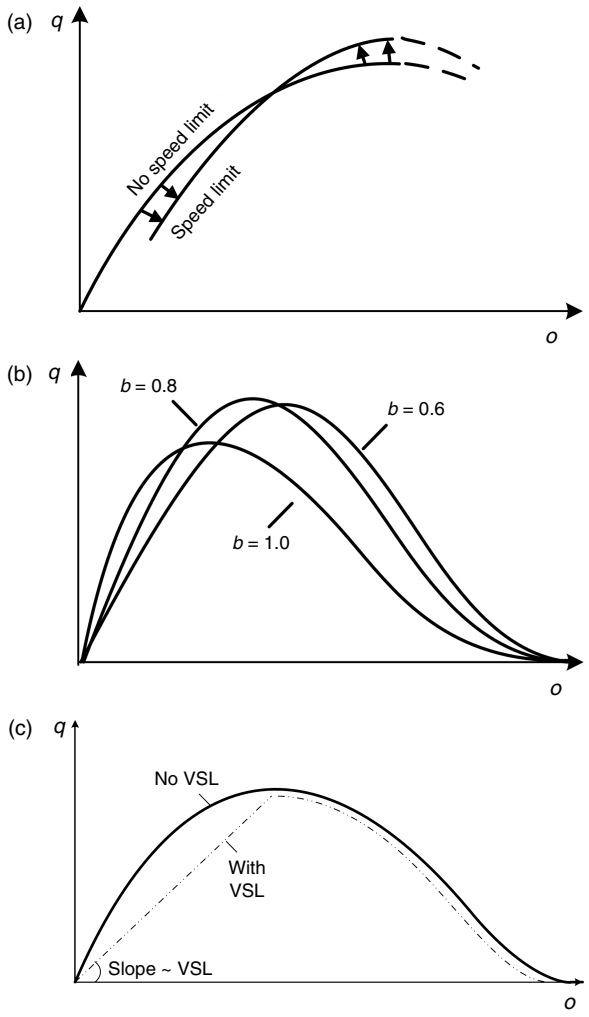

Figure 2 Early Results of the Impact of VSL on the Fundamental Diagram

Notes. In the diagrams above (a) represents a change of the fundamental diagram because of speed limits (Zackor 1991); (b) represents the Cremer (Cremer 1979) model for VSL impact, in which $b = 1$ and 0.8 and 0.6 correspond to no speed limit, VSL $= 0.8\ v_f$ and VSL $= 0.6\ v_f$, respectively; (c) Hegyi (2004) model for VSL impact is represented.

bottleneck locations (see Papageorgiou et al. 2006 for more details).

2.2. The Impact of VSL

2.2.1. Early Results. As mentioned earlier, there were very few investigations in the past addressing the precise impact of VSL on aggregate traffic flow behaviour, e.g., on the fundamental diagram (flow-density curve). Some early investigations (Zackor 1972) based on traffic data from a two-lane German motorway with and without VSL were summarised by Zackor (1991). The results indicate a speed homogenisation effect (less speed differences) for individual vehicles as well as for motorway lanes under the impact of VSL. These results are useful for a better understanding of the VSL impact on individual vehicle speed distribution, but they do not reveal the impact of VSL on aggregate traffic flow behaviour. The latter was also addressed in Zackor (1991) (see Figure 2(a)) but in a rather qualitative way. Figure 2(a) illustrates that "at lower or mean traffic volumes, the mean speed is lower due to the reduction effect whereas, at higher volumes, an increase is detected due to the stabilising effect. Thus, both capacity and speed rise by about 5% to 10% at the same time" (Zackor 1991). Zackor (1991) did not comment on the possible increase of the critical occupancy (or critical density) under the influence of VSL.

The results reported in Zackor (1972) were the basis for Cremer (1979) to propose a quantitative model for the VSL-induced fundamental diagram change as displayed in Figure 2(b) where b is the ratio of the applied VSL divided by the free speed without VSL and, by convention, $b = 1$ corresponds to the no-VSL case. It is quite likely that the displayed increase of flow capacity is rather exaggerated. In fact, later Dutch investigations could not identify any capacity increase that could be attributed to VSL (Smulders 1990), albeit

under advisory (not mandatory) VSL. This VSL model was incorporated in a general dynamic model leading to an optimal control formulation (Cremer 1979; Alessandri et al. 1998). However, a heuristically fixed control law was eventually used because of the size of the problem, and its parameters were optimised based on particular scenarios.

In more recent research work regarding VSL control, the assumed VSL impact was to merely replace the left part of the flow-occupancy curve by a straight line with slope corresponding to the displayed VSL, (see Figure 2(c)) (Hegyi 2004). Some interesting aspects and partly unexpected outcomes of VSL application are discussed by Lebacque and Haj-Salem (2001, 2004).

In conclusion, there seems to be very limited empirical evidence and indeed no factual consensus on the potential impact of VSL on aggregate traffic behaviour, let alone quantitatively reliable results that could be used for efficient control strategy development. The expectations of the VSL impact along with their implications for potentially more efficient traffic flow are examined next, followed by a summary of the main findings by Papageorgiou, Kosmatopoulos, and Papamichail (2008).

2.2.2. Reduction of Mean Speed at Undercritical Occupancies. It seems quite reasonable to assume that a VSL displayed at undercritical occupancies will reduce (with reasonable driver compliance) the (otherwise higher) mean speed (Figure 3(a)). The magnitude of this effect is likely to depend on the displayed VSL as well as on driver compliance. The new VSL-affected states serve the same flow at lower speed and higher occupancy than the original states, which implies that the travel time increases accordingly. Thus, applying VSL at undercritical traffic states is likely to increase travel times and hence deteriorate traffic flow efficiency.

The described state transition when applying VSL at undercritical occupancies could, however, be exploited in a different context. The application of VSL upstream of a bottleneck that is close to becoming active will temporarily (for the duration of the traffic state transition triggered by the VSL) decrease the mainstream flow arriving in the bottleneck area, thus retarding the bottleneck activation and the resulting congestion. Note that the temporary flow decrease during the VSL-triggered traffic state transition is because of the fact that occupancy (and density) in the VSL state is higher than in the original non-VSL state; thus, during the transition the flow is temporarily reduced to "create" the higher traffic density of the VSL state. It should be noted that this is the main VSL impact exploited by Hegyi (2004).

It is quite important to emphasize that the impact of VSL activation upstream of a potential bottleneck (e.g., an on-ramp merge area) in undercritical traffic

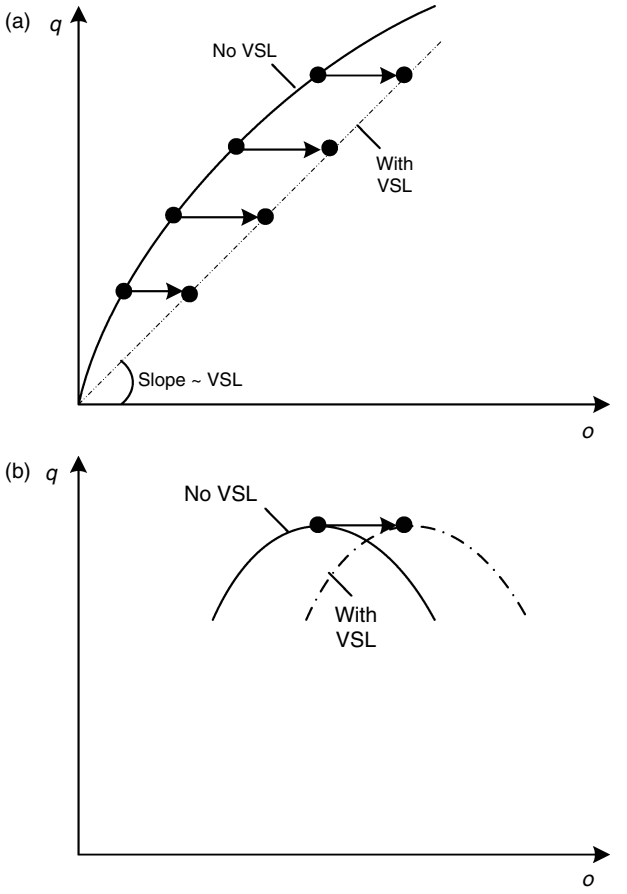

Figure 3 (a & b) Expectations of the VSL Impact on the Fundamental Diagram

Notes. (a) shows potential VSL impact on undercritical mean speeds; (b) shows the cross-point of diagrams with and without VSL.

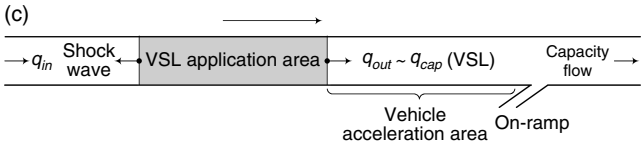

Figure 3 (c) Persistent Flow Control via VSL

conditions as illustrated in the state transition of Figure 3(a), bears great similarities to the impact of local ramp metering in the case of limited ramp storage space. More particularly:

- During the state transition of Figure 3(a), the mainstream flow toward the bottleneck is reduced similarly to ramp metering where instead it is the ramp flow that is reduced to avoid or delay the onset of a merge area congestion.
- During the state transition of Figure 3(a), the mainstream density is increased similarly to ramp metering where instead vehicles are stored in the ramp rather than in the mainstream.
- After the state transition, the mainstream flow returns to its pretransition values (essentially equal

to the upstream arriving mainstream demand) similarly to ramp metering being released when the ramp queue covers the whole ramp to avoid interference with the adjacent street traffic. More precisely, when the free ramp storage is about to be exhausted, ramp metering switches to queue control mode, attempting to maintain a maximum admissible ramp queue (see Smaragdis and Papageorgiou 2003), in which case the ramp outflow becomes essentially equal to the arriving ramp demand.

• After the state transition, the mainstream density remains at its increased value, similarly to ramp metering where the on-ramp queue remains full until the arriving demand drops to sufficiently low values. This way, the activation of VSL upstream of a mainstream bottleneck in undercritical conditions induces some delays to the concerned vehicles because of lower speed similarly to ramp metering inducing delays to the vehicles queuing on the ramp, but this may be more than counterbalanced by the avoidance or retarding of the bottleneck congestion and its associated vehicle delays.

2.2.3. Increase of Throughput and Retarding of Congestion at Overcritical Occupancies. According to the Hegyi (2004) model (Figure 2(c)), both flow-density curves (for VSL and non-VSL) meet but do not actually cross while Zackor (1991) suggests that there is actually a genuine cross-point of both curves somewhere near the critical occupancy (Figure 2(a)). The cross-points (if any) are likely to lie at increasing occupancy values for decreasing VSL because of the accordingly decreasing slope of the undercritical VSL-affected curves. In fact, there may be no cross-point for very low VSL. As we evaluate the VSL impact at occupancies near or higher than the cross point, the following (partly overlapping) questions are of interest (see also Figure 3(b)):

• Where is the cross-point (if any) located with respect to the non-VSL critical occupancy?

• Are VSL-induced critical occupancies higher than their non-VSL counterparts?

• Are VSL-induced flows higher at overcritical occupancies than their non-VSL counterparts?

• Is there a flow capacity increase for some VSL?

These issues were partly studied by Papageorgiou et al. (2006) and Papageorgiou, Kosmatopoulos, and Papamichail (2008), and the related results are summarized in the next section.

2.3. The Effect of VSL on Aggregate Traffic Flow Behaviour

The effect of VSL on aggregate traffic flow behaviour (in form of the flow-occupancy diagram) was investigated (Papageorgiou et al. 2006; Papageorgiou, Kosmatopoulos, and Papamichail 2008) on the basis of traffic data from a European motorway where a flow/speed threshold-based VSL control algorithm is currently used. A main focus of the reported work was on verifying some long-held conjectures (§2.2) regarding the VSL impact on the shape of the flow-occupancy diagram.

Some findings of the reported investigations are summarised as follows (Figure 5, which will be explained in more detail later, may be useful for illustration of the following issues):

(i) Speed limits—when applied at undercritical occupancies—have the effect of decreasing the slope of the flow-occupancy diagram. Moreover, the smaller the imposed speed limit, the larger the decrease in the slope of the flow-occupancy diagram. This impact may be exploited to hold back traffic flow to retard the onset of congestion at downstream bottlenecks, as explained in §2.2.2 and practiced, e.g., by Hegyi (2004).

(ii) The VSL-affected flow-occupancy curve crosses (at least for some VSL) the non-VSL curve, shifting the critical occupancy to higher values in the flow-occupancy diagram. The major cross-points were found to lie around or beyond the non-VSL critical occupancy. This impact may be exploited to hold more vehicles in the motorway without falling into congestion. It may sound paradoxical but these cross-points imply that the mean speed at overcritical densities is higher when a speed limit is imposed than in no-VSL cases; this may happen because of the homogenisation effects mentioned earlier.

(iii) Regarding the potential increase of flow capacity, the data analysis was rather inconclusive because as a slight increase is indeed visible at some locations for some VSL values while at other locations no capacity increase can be observed for any VSL value. In locations where VSL indeed yields a capacity increase, this may be exploited by a suitably designed control strategy for throughput increase, as practiced, e.g., by Papamichail et al. (2008).

(iv) Independently of whether flow capacity is actually increased for some VSL values or not, sufficiently low VSL lead to accordingly lower flow capacity in the fundamental diagram than in non-VSL cases.

Finding (iv) implies that if the mainstream demand q_{in} (Figure 3(c)) arriving from upstream is higher than the VSL-induced capacity q_{cap} (VSL), then the VSL application area becomes an active mainstream bottleneck that limits the area's outflow q_{out} to values corresponding to the (lower) VSL-induced capacity. Recall that the state transition discussed in §2.2.2 yields a *temporary* VSL-induced mainstream flow decrease, after which the outflow from the VSL application area returns essentially to values equal to the upstream arriving demand q_{in}. In contrast, in the case discussed

here, the arriving demand q_{in} is higher than the VSL-induced capacity q_{cap} (VSL); hence, we have the possibility of applying a more durable mainstream flow control that persists even after the transition period, with mainstream flow values q_{out} depending on the applied VSL values. Such a controllable mainstream bottleneck may be deliberately created with benefit upstream of an uncontrolled potential bottleneck (e.g., an on-ramp merge area in Figure 3(c)) to avoid its activation and the related reduction of throughput because of the capacity drop. More specifically, if the outflow of the upstream VSL-controlled bottleneck is regulated such that capacity flow can be established at the downstream bottleneck, then the final mainstream outflow is maximized, leading to a corresponding decrease of the total time spent in the system.

It should be noted that the capacity drop phenomenon, the occurrence of which has been repeatedly confirmed with traffic data (see, e.g., Cassidy and Rudjanakanoknad 2005), is deemed to occur because of the need for vehicles to accelerate from low speeds within the bottleneck congestion to higher speeds as they reach the congestion head (Papageorgiou et al. 2008). Note also that mean speeds downstream of the congestion head are around the critical speed v_{cr} (Figure 1(b)) with typical values around 70 km/h. However, the VSL-controlled vehicle speed within the VSL application area may be lower than v_{cr}. Hence, to avoid the need for vehicle accelerations and the related capacity drop at the downstream bottleneck of Figure 3(c), the head of the upstream VSL-induced bottleneck (i.e., the downstream end of the VSL application area) should be located sufficiently upstream to allow for vehicles exiting the VSL bottleneck area to accelerate and enter the downstream bottleneck area with a roughly critical speed v_{cr}. According to Figure 2.7 of Hall (2001), a distance of around 700 m should be sufficient for vehicles to accelerate from low speeds to 70 km/h.

The application of VSL is usually subject to constraints regarding the maximum admissible temporal and spatial changes of displayed VSL values to avoid driver confusion and increase compliance. For example, it may not be allowed for drivers to face consecutive VMS displaying VSL that differ by more than 20 km/h. Therefore, the core "VSL application area" of Figure 3(c) that is responsible for the creation of a controllable mainstream bottleneck may have to be complemented by two supplementary VSL application areas upstream and downstream, respectively, to comply with the mentioned VSL application constraints. This is particularly sensible upstream of the VSL bottleneck area to improve traffic safety conditions in view of the related, deliberately created shock wave (Figure 3(c)). Note that, in absence of VSL control, a more serious and more durable shock wave would result from the activation of the downstream bottleneck. Hence, the VSL-induced mainstream bottleneck is likely to improve traffic safety conditions as well.

3. Traffic Flow Modelling

3.1. Preliminaries

A macroscopic second-order traffic flow model is used in this study. The model was validated against real traffic data at several instances (Papageorgiou, Blosseville, and Haj-Salem 1990a; Kotsialos et al. 2002b) and was found to reproduce the whole range of real traffic conditions (free flow, critical, congested) with remarkable accuracy. The model is included in the METANET motorway traffic flow simulator (Messmer and Papageorgiou 1990) and is extended here to incorporate VSL control measures.

The motorway network is represented by a directed graph whereby the links of the graph represent motorway stretches. Each motorway stretch has uniform characteristics, i.e., no on- or off-ramps and no major changes in geometry. The nodes of the graph are placed at locations where major changes in road geometry occur as well as at junctions, on-ramps, and off-ramps.

The macroscopic description of traffic flow implies the definition of adequate variables expressing the aggregate behaviour of traffic at certain times and locations. The time and space arguments are discretised. The discrete-time step is denoted by T (typically, $T = 10$ s). A motorway link m is divided into N_m segments of equal length L_m (typically, $L_m = 500$ m) (Figure 4). The traffic in each segment i of link m at discrete-time $t = kT$, $k = 0, 1, \ldots, K$, where K is the time horizon, is macroscopically characterised via the following variables: the *traffic density* $\rho_{m,i}(k)$ (veh/km/lane) is the number of vehicles in segment i of link m at time $t = kT$ divided by L_m and by the number of lanes λ_m; the *mean speed* $v_{m,i}(k)$ (km/h) is the mean speed of the vehicles included in segment i of link m at time $t = kT$; and the *traffic volume* or *flow* $q_{m,i}(k)$ (veh/h) is the number of vehicles leaving segment i of link m during the time period $[kT, (k+1)T)$ divided by T.

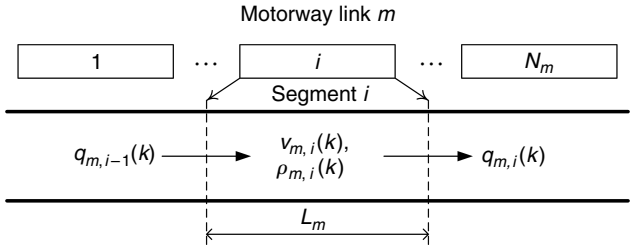

Figure 4 Discretised Motorway Link

3.2. The Motorway Link Model

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

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

$$q_{m,i}(k) = \rho_{m,i}(k) v_{m,i}(k) \lambda_m, \quad (2)$$

$$v_{m,i}(k+1) = v_{m,i}(k) + \frac{T}{\tau}\{V[\rho_{m,i}(k)] - v_{m,i}(k)\} + \frac{T}{L_m}[v_{m,i-1}(k) - v_{m,i}(k)]v_{m,i}(k) - \frac{\nu T}{\tau L_m}\frac{\rho_{m,i+1}(k) - \rho_{m,i}(k)}{\rho_{m,i}(k) + \kappa}, \quad (3)$$

$$V[\rho_{m,i}(k)] = v_{f,m} \exp\left[-\frac{1}{\alpha_m}\left(\frac{\rho_{m,i}(k)}{\rho_{cr,m}}\right)^{\alpha_m}\right], \quad (4)$$

where (1) is the conservation equation; (2) is the transport equation to be replaced in (1); (3) is an empirical dynamic mean speed equation where (4), the static speed-density relationship corresponding to the fundamental diagram, must be replaced; and τ (a time constant), ν (an anticipation constant), and κ are model parameters that are equal for all the network links. Two further terms may be added to (3) for higher accuracy under certain conditions (Papageorgiou, Blosseville, and Haj-Salem 1990a).

The original (non-VSL) model includes three link-specific constant parameters in the speed-density curve (4): the free speed $v_{f,m}$ encountered at zero density ($\rho_{m,i} = 0$), the critical density $\rho_{cr,m}$ at which traffic flow is close to capacity $q_{cap,m}$, and α_m. Combining (2)–(4) under stationary (i.e., $v_{m,i}(k+1) = v_{m,i}(k)$) and spatially homogeneous (i.e., $v_{m,i-1} = v_{m,i}$ and $\rho_{m,i+1} = \rho_{m,i}$) conditions for $\rho_{m,i} = \rho_{cr,m}$ (i.e., the critical density) yields the capacity of the fundamental diagram (flow-density curve)

$$q_{cap,m} = v_{f,m} \cdot \rho_{cr,m} \exp(-1/\alpha_m). \quad (5)$$

3.3. Incorporating the VSL Impact

The described link model may be extended to incorporate the impact of displayed VSL values on the traffic flow behaviour under the assumption that a single VSL value (if any) is displayed in each link (using, in real implementations, as many VMS gantries as necessary, depending on the link length). It should be noted that this assumption is not really restrictive because:

- If a higher spatial granularity (resolution) of VSL values is desired, links may be selected accordingly short.
- If a lower spatial granularity (resolution) of VSL values is desired, there is a possibility for the user of the related software tools METANET (simulator) and AMOC (optimal control; see §4) to create clusters of links, each cluster having a common VSL value.

To start with, particular VSL values are reflected in the link-specific *VSL rates* $b_m(k)$ that prevail, by definition, during $[kT, (k+1)T)$. The VSL rates are naturally control variables with an admissible value range $b_m(k) \in [b_{\min,m}, 1]$, where $b_{\min} \in (0, 1)$ is a lower admissible bound for VSL rates (see further below for a physical interpretation of b_m).

Using the defined VSL rates, we now proceed to the appropriate inclusion of these control variables into the link model (1)–(4). Pursuing the lines of previous works (Cremer 1979; Alessandri et al. 1998; Hegyi 2004; Papamichail et al. 2008), this is materialized by rendering the static speed-density relationship (4) b_m-dependent. Based on available real data evidence from Papageorgiou et al. (2006); Papageorgiou, Kosmatopoulos, and Papamichail (2008); and Kampitaki (2008), this is enabled by actually rendering the three parameters included in (4) b_m-dependent by use of the following linear (affine) functions:

$$v_{f,m}[b_m(k)] = v^*_{f,m} b_m(k), \quad (6)$$

$$\rho_{cr,m}[b_m(k)] = \rho^*_{cr,m}\{1 + A_m[1 - b_m(k)]\}, \quad (7)$$

$$\alpha_m[b_m(k)] = \alpha^*_m[E_m - (E_m - 1)b_m(k)], \quad (8)$$

where $v^*_{f,m}$, $\rho^*_{cr,m}$, α^*_m denote the specific non-VSL values for these parameters while A_m, E_m are constant parameters to be estimated based on real data.

As (6) reveals, b_m is equal to the VSL-induced $v_{f,m}$ divided by the non-VSL $v^*_{f,m}$ or approximately equal to the displayed VSL divided by the legal speed limit without VSL. Thus, if $b_m(k) = 1$, no VSL is applied, else $b_m(k) < 1$ and, in fact, for $b_m(k) = 1$, all parameters are seen in (6)–(8) to attain their respective non-VSL values. Equations (7) and (8) suggest that for $A_m > 0$ and $E_m > 0$, $\rho_{cr,m}$ and α_m are linear increasing functions for decreasing b_m starting with their usual non-VSL values for $b_m(k) = 1$.

The extended speed-density curve encompassing (4), and (6)–(8) was validated by use of traffic data taken from a European VSL-equipped motorway location, where the legal speed limit is 70 mph and the applied VSL values are 60 mph, 50 mph, and 40 mph. These values correspond to VSL rates $b_m \in \{1, 0.86, 0.71, 0.57\}$. The validation exercise furnished the estimated parameter values $v^*_{f,m} = 115$ km/h, $\rho^*_{cr,m} = 28.2$ veh/km/lane, and $\alpha^*_m = 2.15$ (leading via (5) to a non-VSL capacity value $q^*_{cap,m} = 2{,}036$ veh/h/lane) while the estimated VSL-related parameters were $A_m = 0.7$ and $E_m = 1.9$ for (7) and (8), respectively. Figure 5 displays the corresponding flow-density curves, generalised for a broader range $b_m \in [0.2, 1.0]$. In accordance with §2, the following may be observed in Figure 5.

(i) Free speeds are decreasing with decreasing VSL.

(ii) Cross-points of VSL affected with the non-VSL curve appear near or beyond the non-VSL critical

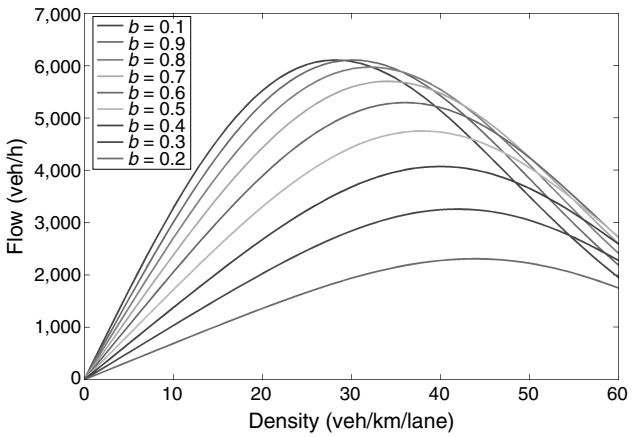

Figure 5 Fundamental Diagrams for Different VSL Rates

density; critical densities are increasing with decreasing VSL.

(iii) No capacity increase (for any VSL value) was observed at this location. Note that data from a different location of the same motorway yielded capacity increases for $b_m \in (0.6, 1)$ with a maximum increase of 8% at $b_m = 0.82$ (Kampitaki 2008; Papamichail et al. 2008).

(iv) For $b_m = 0.9$, capacity is virtually equal as in the non-VSL case but is seen to monotonically decrease with (further) decreasing b_m values.

All the other model parameter values used in this study are taken from a previous model calibration for a real motorway (Kotsialos et al. 2002), namely, $\tau = 18$ s, $\nu = 60$ km^2/h, $\kappa = 40$ veh/km/lane, and $\rho_{\max} = 180$ veh/km/lane. It should be noted that the isolated validation of the fundamental diagram before incorporating it with in the more comprehensive dynamic model (1)–(4) may have a quantitative impact on the accuracy of the VSL-extended dynamic model. Nevertheless, because the static speed-density relationship (4) is known to dominate within the dynamic speed-density relationship (3), the overall dynamic model is deemed to reflect the impact of VSL sufficiently accurately for the requirements of the present study. As a matter of fact, the validation results by Papageorgiou et al. (2006), Papageorgiou, Kosmatopoulos, and Papamichail (2008) were derived with traffic data collected at one single motorway. Hence, more validation work, using data from different motorways and countries, is necessary before arriving at a quantitatively accurate and reliable description of the VSL impact on traffic flow. However, the control results obtained in simulation in the next section exploit two particular impacts of VSL that are quite certain to occur (at least qualitatively) in reality as well (with appropriate enforcement measures) for obvious physical reasons: the decreasing of free speeds with decreasing VSL (issue (i)) and the decreasing of capacity with (sufficiently) decreasing VSL (issue (iv)).

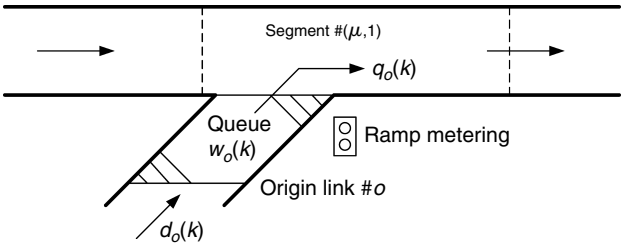

Figure 6 The Origin Link Queue Model

It is also noted that the maximum mainstream flow created by the dynamic model (e.g., at on-ramp merge areas) may be higher than the fundamental diagram's q_{cap} because of the impact of other terms in (3). Thus, the "factual capacity" of merge areas (without VSL) in the simulations of §5 is about 2,130 veh/h/lane, which is higher than $q^*_{cap} = 2{,}036$ veh/h/lane. Finally, as observed in numerous previous simulation studies with the METANET simulator, the overall dynamic model (1)–(4) (with or without VSL-related extensions) automatically create a (factual) capacity drop (of typically 8% to 10%) at active bottlenecks.

3.4. The Origin Link Model

For origin links, i.e., links that receive traffic demand d_o and forward it into the motorway network, a simple queue model is used (Figure 6). The outflow q_o of an origin link o depends on the arriving demand, on the traffic conditions of the corresponding mainstream segment $(\mu, 1)$, and on the existence of ramp metering control measures. If ramp metering is applied, then the outflow $q_o(k)$ that leaves origin o during period k is a portion $r_o(k)$ of the outflow $\hat{q}_o(k)$ that would leave in absence of ramp metering. Thus, $r_o(k) \in [r_{\min,o}, 1]$ is the metering rate for the origin link o, i.e., a control variable where $r_{\min,o}$ is a minimum admissible value. If $r_o(k) = 1$, no ramp metering is applied; else, $r_o(k) < 1$. The queuing model is described by the following conservation equation:

$$w_o(k+1) = w_o(k) + T[d_o(k) - q_o(k)], \tag{9}$$

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

$$q_o(k) = r_o(k)\hat{q}_o(k), \tag{10}$$

with

$$\hat{q}_o(k) = \min\{\hat{q}_{o,1}(k), \hat{q}_{o,2}(k)\}, \tag{11}$$

and

$$\hat{q}_{o,1}(k) = d_o(k) + w_o(k)/T, \tag{12}$$

$$\hat{q}_{o,2}(k) = Q_o \min\left\{1, \frac{\rho_{\max} - \rho_{\mu,1}(k)}{\rho_{\max} - \rho_{cr,\mu}}\right\}, \tag{13}$$

where Q_o (veh/h) is the on-ramp's flow capacity, i.e., the on-ramp's maximum possible outflow under free-flow traffic conditions in the mainstream, and $\rho_{\max}$ (veh/km/lane) is the maximum density in the network. According to (11)–(13), the uncontrolled outflow $\hat{q}_o(k)$ is determined by the current origin demand if $\hat{q}_{o,1}(k) < \hat{q}_{o,2}(k)$; else it is determined by the geometrical capacity Q_o (if the mainstream density is undercritical, i.e., $\rho_{\mu,1}(k) < \rho_{cr,\mu}$) or by the reduced capacity because of congestion of the mainstream (if $\rho_{\mu,1}(k) > \rho_{cr,\mu}$).

3.5. The Node Model

Motorway bifurcations and junctions (including on-ramps and off-ramps) and more generally link bounds are represented by nodes. Traffic enters a node n through a number of input links and is distributed to the output links according to the following equations:

$$Q_n(k) = \sum_{\mu \in I_n} q_{\mu,N_\mu}(k), \tag{14}$$

$$q_{m,0}(k) = \beta_n^m(k) Q_n(k) \quad \forall m \in O_n, \tag{15}$$

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

At a network node n, the upstream influence of the downstream link density (e.g., in case of congestion spillback) has to be taken into account in the last segment of the incoming links (see (3) for $i = N_m$). This is provided via

$$\rho_{m,N_m+1}(k) = \sum_{\mu \in O_n} \rho_{\mu,1}^2(k) \Big/ \sum_{\mu \in O_n} \rho_{\mu,1}(k), \tag{16}$$

where $\rho_{m,N_m+1}(k)$ is the virtual density downstream of any entering link m to be used in (3) for $i = N_m$ and $\rho_{\mu,1}(k)$ is the density of the first segment of the leaving link μ. The quadratic form is used to account for the fact that congestion on one leaving link may spill back into the entering link even if there is free flow in the other leaving links.

Similarly, at a network node n, the downstream influence of the upstream link speed has to be taken into account according to (3) for $i = 1$. The required upstream mean speed value is calculated from the flow-weighted average

$$v_{m,0}(k) = \sum_{\mu \in I_n} v_{\mu,N_\mu}(k) q_{\mu,N_\mu}(k) \Big/ \sum_{\mu \in I_n} q_{\mu,N_\mu}(k), \tag{17}$$

where $v_{m,0}(k)$ is the virtual speed upstream of any leaving link m that is needed in (3) for $i = 1$.

3.6. The Overall Dynamic Model

Combining the equations developed above, a nonlinear macroscopic discrete-time state-space model

$$\mathbf{x}(k+1) = \mathbf{f}[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)], \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{18}$$

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

4. The Integrated Optimal Control Problem

The integrated motorway network traffic control problem is formulated as a discrete-time dynamic optimal control problem with constrained control variables over a given optimisation horizon K_P, which is solved very efficiently even for large-scale networks by a suitable feasible direction algorithm (Papageorgiou and Marinaki 1995). This extended formulation (to incorporate the VSL impact) and the numerical solution algorithm are incorporated in an accordingly extended version of the open-loop optimal control tool AMOC (Kotsialos et al. 2002a) which is able to consider coordinated ramp metering, system optimum route guidance, and variable speed limits (using the introduced extension via (4)–(8)) as well as integrated control combining all control measures simultaneously.

If only ramp metering and VSL are considered, the general discrete-time formulation of the optimal control problem is the following:

Given disturbance predictions $\mathbf{d}(k)$, $k = 0, 1, \ldots, K_p - 1$, and the initial state $\mathbf{x}_0 = \mathbf{x}(0)$, minimise

$$J = \vartheta[\mathbf{x}(K)] + \sum_{k=0}^{K_P-1} \varphi[\mathbf{x}(k), \mathbf{u}(k), \mathbf{d}(k)], \tag{19}$$

subject to (18) and the inequality constraints imposed on the ramp metering rates $r_{\min,o} \le r_o(k) \le 1$ and the VSL rates $b_{\min,m} \le b_m(k) \le 1$.

The chosen cost criterion is the total time spent (TTS) by all vehicles in the network (including the waiting time experienced in the ramp queues), which is a natural objective for the traffic systems considered. The maximum ramp queue constraints may be taken into account via the introduction of penalty terms in the cost criterion penalising queue lengths

larger than $w_{\max,o}$, which is a predetermined maximum admissible queue for origin o. Another penalty term may be added to suppress high-frequency oscillations of the optimal control trajectories. More precisely, the cost criterion used as (19) is the following:

$$J = T\sum_{k=1}^{K_P-1}\sum_{m}\sum_{i}\rho_{m,i}(k)L_m\lambda_m + T\sum_{k=1}^{K_P-1}\sum_{o}w_o(k) + T\sum_{k=1}^{K_P-1}\sum_{o}\alpha_f[r_o(k)-r_o(k-1)]^2 + T\sum_{k=1}^{K_P-1}\sum_{m}\alpha_b[b_m(k)-b_m(k-1)]^2 + T\sum_{k=1}^{K_P-1}\sum_{o}\alpha_w[\max\{0, w_o(k)-w_{\max,o}\}]^2, \quad (20)$$

where α_f, α_b, and α_w are weighting factors for the corresponding penalty terms.

The solution determined by AMOC consists of the optimal ramp metering and VSL rate trajectories as well as the corresponding optimal state trajectory. It is interesting to note that the solution algorithm can account for control variables that change their value less frequently than the state variables. Moreover, for the VSL rates, common control variables can be considered for clusters of links as stated earlier.

It should be stressed that the extension of AMOC to also consider VSL rates necessitated the specialized and rigorous incorporation of accordingly extended generic necessary optimality conditions along with the corresponding Jacobian matrices, etc. This has led to a universal optimal control tool that is readily applicable to any, (even large-scale) motorway network to deliver optimal control results with quite low computational effort that would even allow for real-time application of the tools. This marks a clear progress compared to most previous works involving optimal VSL control. More details on the integrated traffic control problem (without VSL) and the numerical solution algorithm may be found in Kotsialos et al. (2002a).

As mentioned earlier, the field application of VSL is subject to a number of constraints regarding the maximum admissible extent of VSL changes, both temporally and spatially, that are not considered in AMOC. In addition, VSL can only attain prespecified discrete values in practice rather than the continuous (real-valued) $b_m(k)$ values delivered by AMOC. Although the optimal $b_m(k)$ values delivered by AMOC could be eventually treated appropriately to satisfy these constraints, we preferred to leave them unaltered to assess the full possibilities provided by VSL.

It must be emphasized that optimisation and optimal control methods applied adequately to specific application problems are distinguished by an inherent intelligence. More specifically, while searching for the mathematical optimum, optimal control methods (based on adequate modelling) may "discover" the particular application ways, measures, and combinations of control variables that lead to optimal performance. Thus, an engineering application-specific interpretation of the optimal results may lead to new general insights regarding the specific application. In our case, some of the insights presented in §2 were, in fact, revealed upon interpretation of the obtained optimal AMOC results. The next section presents a relatively simple application case where AMOC is seen to automatically behave as if it were aware of all the issues discussed in §2.

5. Application Results

5.1. The Test Motorway Axis

For the purposes of this study, a hypothetical three-lane motorway axis of 6.5 km, sketched in Figure 7, is considered. The mainstream is divided into five links (L0 to L4). There are two on-ramps (O1 and O2) on this motorway and one off-ramp (D1) in between. The demand profiles shown in Figure 8 are used for the motorway input (U1) and for the two on-ramps. The TTS is calculated over the 2.5 hours of simulation depicted in Figure 8, whereby the last 30 min is a cooldown period with 0 inflows in the on-ramps and only 1,000 veh/h in the mainstream. This was introduced to have equal traffic conditions on the stretch at the end of the simulation and hence comparable TTD values for all investigated scenarios. The exit rate, i.e., the percentage of the mainstream flow that leaves the motorway at the off-ramp D1 is set to 5%; the model time step used is $T = 10$ s. A number of different control scenarios are examined in the following, each for a time horizon of 2.5 hours.

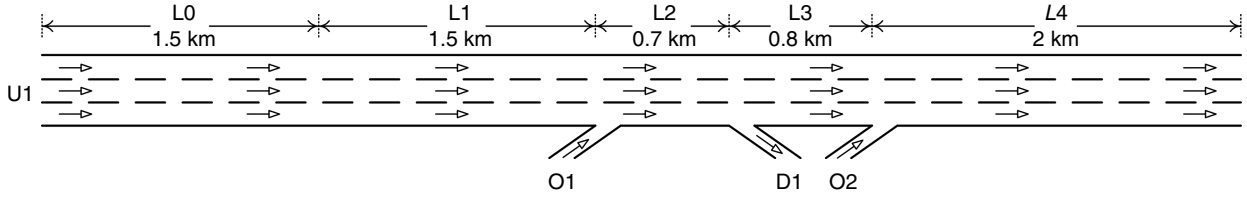

Figure 7 The Test Three-Lane Motorway Axis with Two On-Ramps

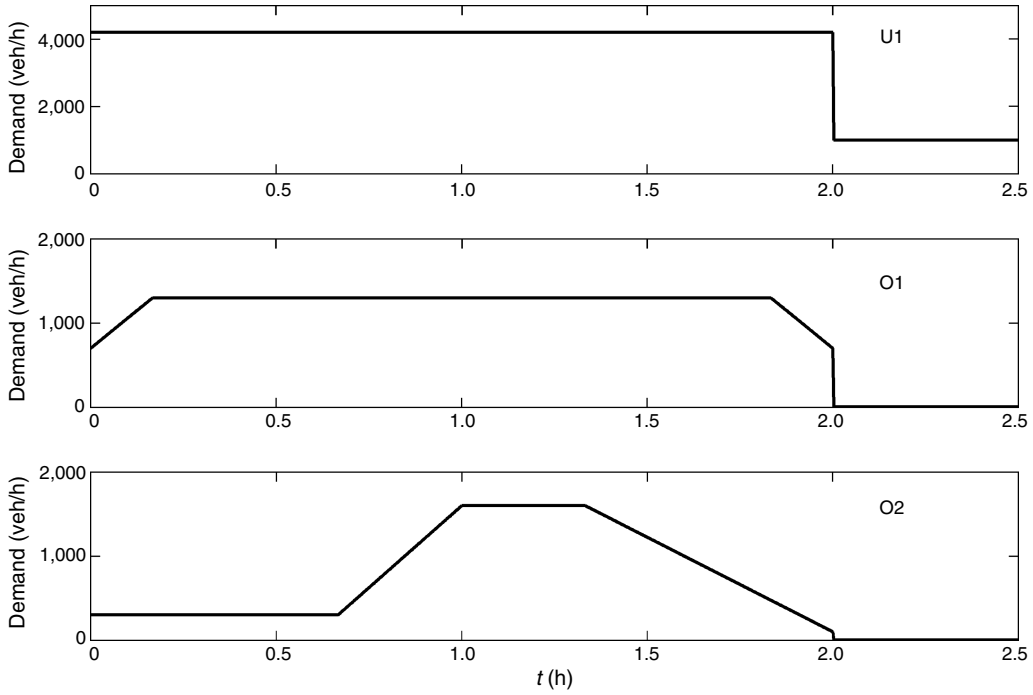

Figure 8 Demand Profiles

5.2. No-Control Case

Figure 9 shows the resulting ramp queue, density, speed, and flow profiles for both merge areas When no control measures are applied. The flow in the merge area of O2 is seen to reach the factual capacity (6,400 veh/h) at $t = 1$ h. As arriving flows continue to increase, a mainstream congestion appears after 1 h in the merge area of the O2 on-ramp; this leads to a visible gradual mainstream flow decrease (capacity drop). The created congestion (shock wave) travels upstream and reaches the merge area of the O1 on-ramp at around $t = 1.2$ h, leading to a visible speed

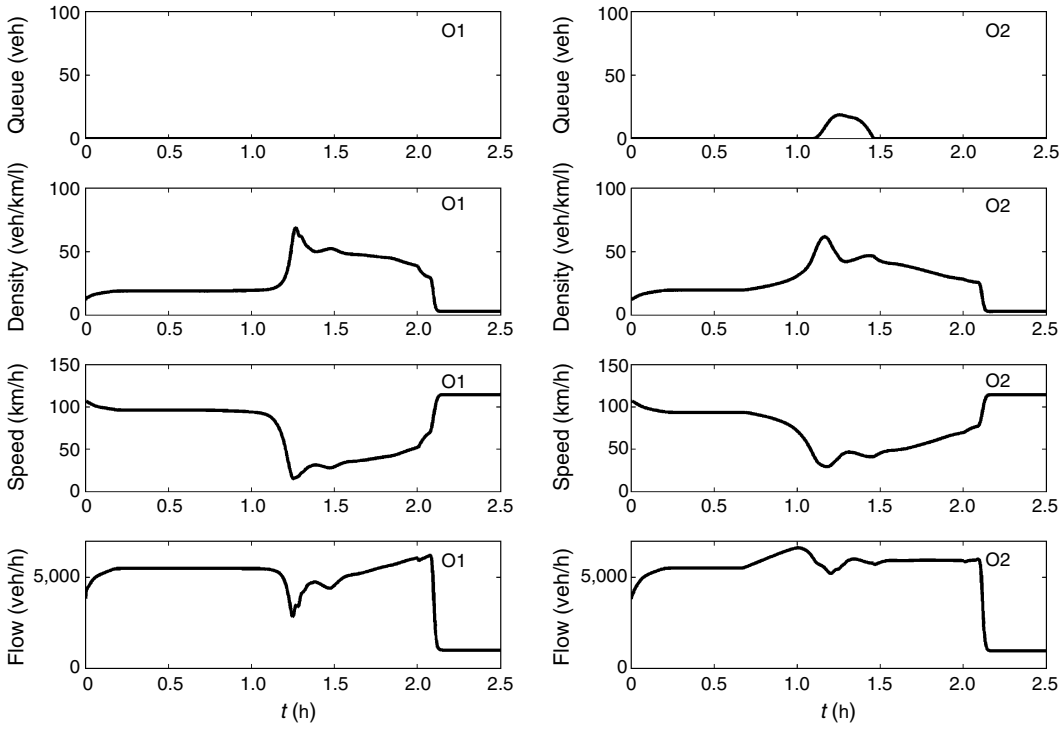

Figure 9 No-Control Case

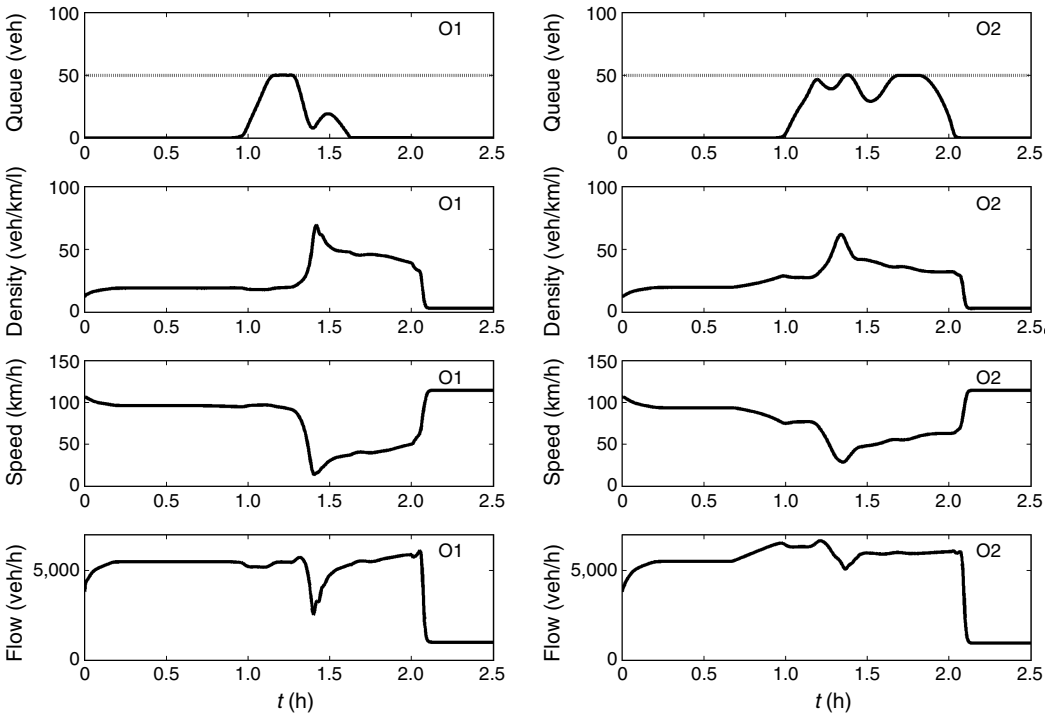

Figure 10 Coordinated Ramp Metering

drop and flow decrease there as well. In this scenario, the short queue (18 veh) that forms at the O2 on-ramp is because of the reduction of the on-ramp's flow capacity caused by the mainstream congestion (see (13)). The resulting TTS is equal to 1,167 veh · h.

5.3. Coordinated Ramp Metering

AMOC is now applied for coordinated ramp metering with maximum admissible queue equal to 50 veh for each on-ramp. The ramp metering rates are allowed to change every 30 s with a minimum admissible value equal to 0.05 (to avoid ramp closure). The resulting TTS value is equal to 1,060 veh · h, which is a 9.2% improvement compared to the no-control case. The related ramp queue, density, speed, and flow profiles for both merge areas are shown in Figure 10. The dotted curves appearing in the queue plots correspond to the utilised maximum admissible ramp queues.

The situation is identical to the no-control case until shortly before $t = 1$ h, but eventually the optimal solution maintains the density and the flow at the O2 merge area close to the critical density and factual capacity values, respectively, as long as possible to maximise the motorway exit flow (which leads to minimisation of TTS). To achieve this, ramp queues are created quasisimultaneously in both ramps. The congestion appearing at the O2 merge area at around $t = 1.2$ h is unavoidable in view of the high involved demands and the exhausted limited ramp storage.

5.4. VSL Control

For the application of VSL control, the mainstream is divided into four clusters of links, each with its own VSL rate. The first cluster comprises L0 where no VSL control is applied, i.e., $b_{L0}(k) = 1$, $k = 0, 1, \ldots, K_p - 1$; the second cluster comprises L1; the third cluster comprises L2 and L3, i.e., one single VSL rate is used for both L2 and L3; and the fourth cluster comprises L4. The VSL rates are allowed to change every 300 s and two cases are examined for the value of the minimum admissible VSL. In the first case, $b_{\min,m} = 0.5$ for all controlled link clusters, and in the second-case, $b_{\min,m} = 0.2$. Maximum ramp queue constraints are not taken into account when only VSL control is applied.

When $b_{\min,m} = 0.5$ is adopted, the resulting TTS value is equal to 1,078 veh · h, which is a 7.6% improvement compared to the no-control case. The related ramp queue, density, speed, and flow profiles for both merge areas are shown in Figure 11 and the optimal VSL rate trajectories are shown in Figure 12.

The situation is here also virtually identical to the no-control case until $t = 1$ h, because no congestion appears yet. At $t = 1$ h, the VSL rate of L4 is seen to switch gradually to values around 0.85; this allows the O2 merge area to accommodate a higher number of vehicles (because of higher critical density) without any real loss in flow capacity (Figure 5).

The next issue is the need to keep the O2 merge area density close to its (increased) critical density for

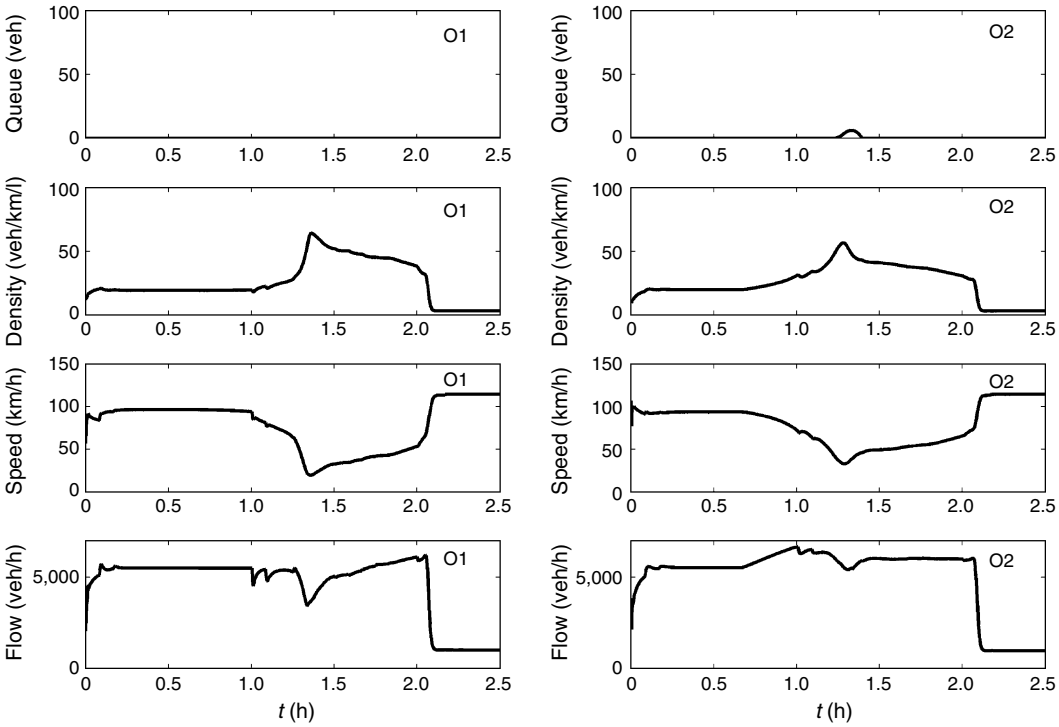

Figure 11 **VSL** ($b_{\min,m} = 0.5$) **Control**

as long as possible to avoid the congestion appearing there after $t = 1$ h in the no-control case scenario and enable maximum motorway exit flow (which leads to TTS minimisation). Indeed, the VSL rate for L1 is seen to switch from one to (the lower admissible bound) 0.5 in essentially two steps (Figure 12) shortly after $t = 1$ h. According to §2.2.2, the corresponding state transitions create temporary mainstream flow reductions that are clearly visible in the flow curve of merge area O1 (Figure 11) as short negative pulses. Because the (factual) mainstream capacity for a VSL rate of 0.5 is still higher than the upstream arriving demand, the mainstream flow is seen to fully recover after the state transition. These temporary flow reductions are effectuated to keep the O2 merge area density close to its critical value for as long as possible, but after reaching the lower bound of 0.5, VSL control has exhausted its ammunition (no further flow decrease is possible) and the O2 merge area congestion (and corresponding flow reduction) after $t = 1.2$ h becomes unavoidable similar to the ramp metering case.

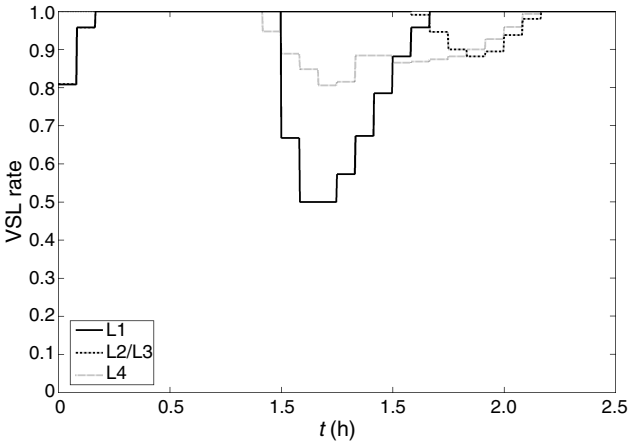

Figure 12 **Optimal VSL Rates for VSL** ($b_{\min,m} = 0.5$) **Control**

Remarkably, AMOC decides to create a mainstream bottleneck by use of the VSL of L1 rather than of L2/L3. The reason for this may be found in §2.3: The VSL application area of L2/L3 is used as an acceleration area for vehicles exiting the low-VSL link L1 to avoid capacity drop at the O2 merge area. In fact, if AMOC is rerun with uncontrolled L1 (not shown here), it refuses to create a VSL-induced mainstream bottleneck at L2/L3, because this would not bring real benefits (capacity drop is not avoided). Further interpretation details (e.g., regarding the VSL switch-off period after $t = 1.5$ h) are deemed less significant and are omitted here for the sake of brevity.

Allowing the VSL control to go to even lower values ($b_{\min,m} = 0.2$) results in a TTS value equal to 988 veh · h, which is a 15.3% improvement compared to the no-control case. The related ramp queue, density, speed and flow profiles for both merge areas are shown in Figure 13 and the optimal VSL rate trajectories are shown in Figure 14.

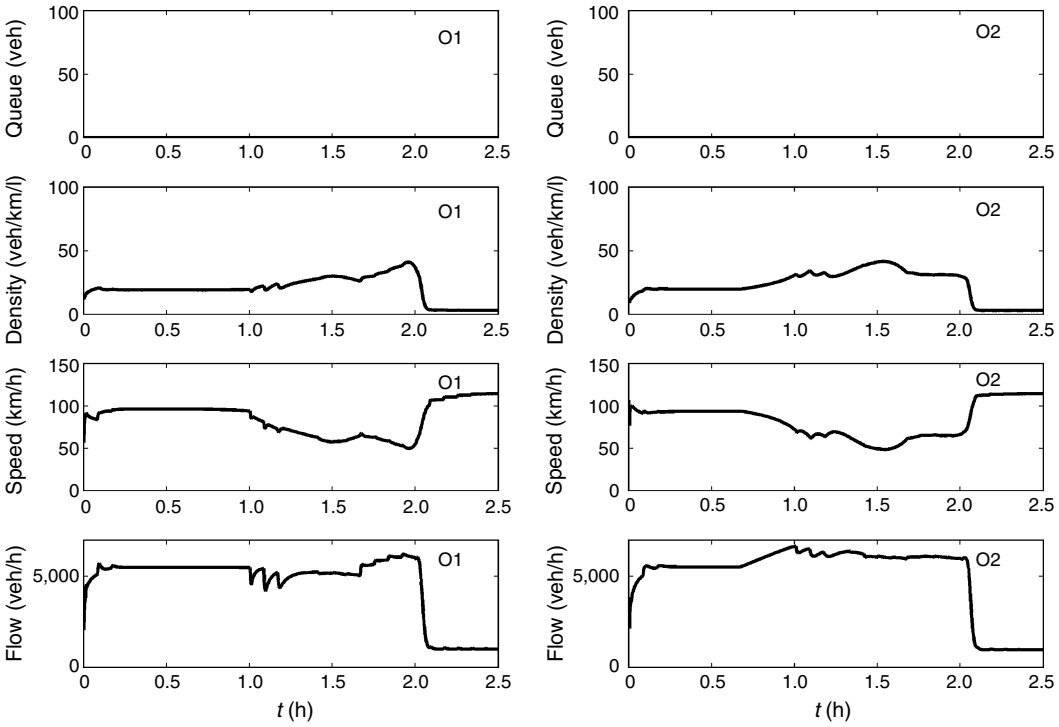

Figure 13 **VSL** ($b_{\min,m} = 0.2$) **Control**

The interpretation of these results is similar to the previous case with one notable difference. The (factual) mainstream capacity for the new (lower) admissible limit of 0.2 is now lower than the arriving mainstream demand. Hence, the flow at the O1 merge area (Figure 13) is seen to not recover fully after the third VSL switch at L1 (Figure 14). This leads to a stronger and more durable flow reduction, because of which the O2 merge area density is maintained near critical values (with maximum motorway exit flow) over a longer period and the (unavoidable) congestion appearing there around $t = 1.4$ h is seen (Figure 13) to be weaker than in previous cases. Note also the more prolonged VSL application at L1 (Figure 14) and its gradual switch off leading to increased flow in the O1 merge area at $t \in [1.6\text{ h}, 2\text{ h}]$ because opposite state transitions (from lower to higher VSL values).

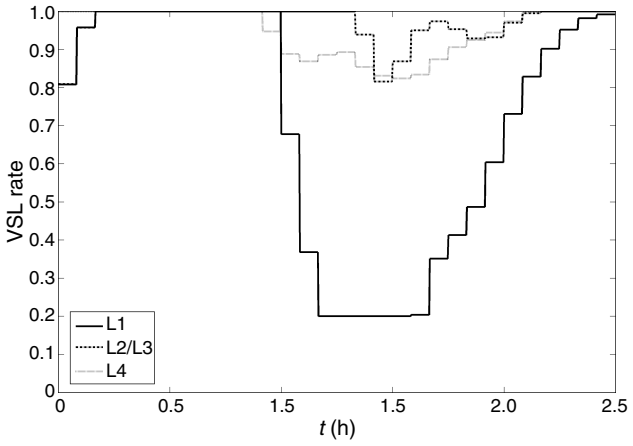

Figure 14 **Optimal VSL Rates for VSL** ($b_{\min,m} = 0.2$) **Control**

5.5. Integrated Control

When both coordinated ramp metering and VSL control with $b_{\min,m} = 0.5$ are applied, i.e., integrated traffic control, TTS is reduced to 992 veh·h, which is a 15.0% improvement compared to the no-control case. The related profiles are omitted for the sake of brevity.

The optimal results indicate a similar behaviour of each control measure (ramp metering and VSL) to previous control scenarios. Because the integration of both measures increases the possibilities to hold more traffic back from the O2 merge area, the density there can be maintained close to its (now VSL modified) critical value up to $t = 1.4$ h after which an unavoidable (but less strong) congestion appears.

Finally, when both coordinated ramp metering and VSL control with $b_{\min,m} = 0.2$ are applied, TTS is reduced further to 939 veh · h, which is a 19.5% improvement compared to the no-control case. The related ramp queue, density, speed, and flow profiles for both merge areas are shown in Figure 15 and the optimal VSL rate trajectories are shown in Figure 16.

The optimal results in Figures 15 and 16 indicate a similar control behaviour as in previous cases, with

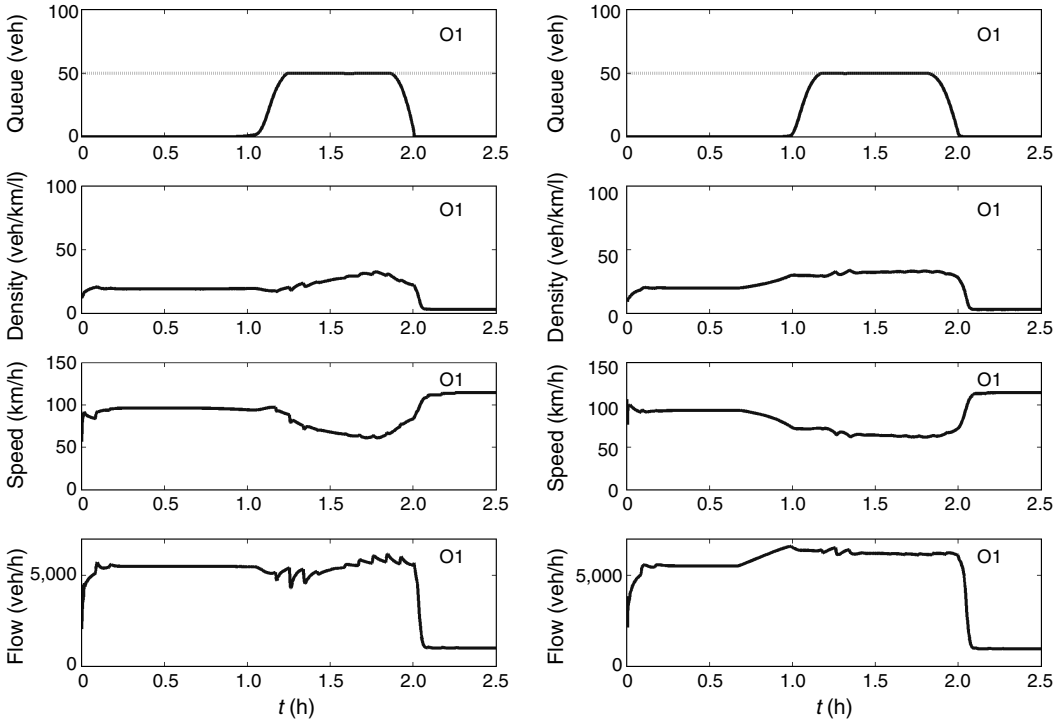

Figure 15 **Integrated Control, i.e., Coordinated Ramp Metering and VSL ($b_{\min,m} = 0.2$) Control**

the notable difference that traffic can be held back at a sufficient level to completely avoid congestion. In fact, the O2 merge area density is maintained close to its (now VSL modified) critical value for as long as is necessary and the motorway exit flow is accordingly maximised, which lead to the above mentioned further decrease in TTS. Remarkably, the available storage space for ramp metering is fully utilised while the L1 VSL rate values (which are responsible for creating the mainstream bottleneck) do not reach the admissible lower bound of 0.2 because that is not needed.

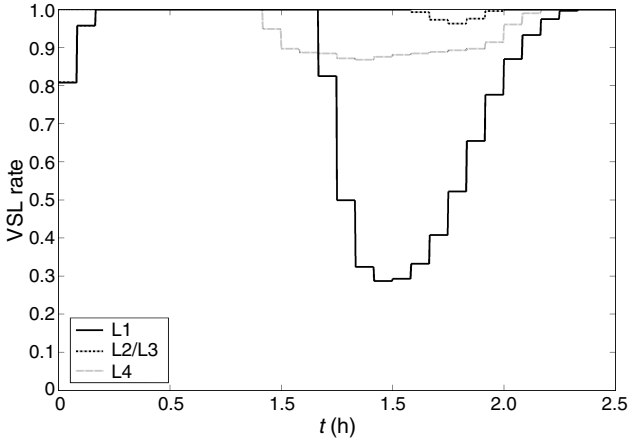

Figure 16 **Optimal VSL Rates for the Integrated Control, i.e., Coordinated Ramp Metering and VSL ($b_{\min,m} = 0.2$) Control**

6. Conclusions

Variable speed limit (VSL) control on motorways was shown resemble (if appropriately applied) ramp metering actions, albeit by holding traffic back on the mainstream rather than on the ramps. Conditions for temporary or durable mainstream flow reduction were elaborated. The mainstream flow reduction via VSL may be exploited to avoid the appearance of the detrimental capacity drop at active bottlenecks. These insights may be useful for the design of a new generation of efficient VSL control strategies.

After the qualitative elaboration of the opportunities offered by the operation of VSL to improve the traffic flow efficiency on motorways, a quantitative model for the impact of VSL on aggregate traffic flow behaviour, which resulted from on a related previous validation study with real traffic data was proposed. VSL were incorporated in a general macroscopic second-order traffic flow model as an additional control component leading to an accordingly extended general optimal control formulation and its numerical solution via the also-extended AMOC tool. An illustrative example was presented under different control scenarios. It was shown in detail how traffic flow efficiency can be substantially improved when VSL control measures are used appropriately (particularly in integration with coordinated ramp metering) mainly by retarding or avoiding the capacity drop at active bottlenecks.

Because of various inherent uncertainties, the open-loop optimal solution delivered by optimal control approaches becomes suboptimal when directly applied to the motorway traffic process (see, e.g., Papamichail et al. 2009). However, the optimal results can be used in a rolling horizon mode or can be utilised to extract useful conclusions for the development of efficient feedback control strategies. This along with applications of the developed concepts to large-scale networks that will lead to additional insights are left for future publications.

Acknowledgments

The first author was supported by The Capes Foundation, Ministry of Education of Brazil (BEX 5022-06-1).

References

Alessandri, A., A. Di Febbraro, A. Ferrara, E. Punta. 1998. Optimal control of freeways via speed signalling and ramp metering. *Control Engrg. Practice* **6** 771–780.

Cassidy, M. J., J. Rudjanakanoknad. 2005. Increasing the capacity of an isolated merge by metering its on-ramp. *Transportation Res. Part B* **39**(10) 896–913.

Cremer, M. 1979. *Der Verkehrsfluß auf Schnellstrassen*. Springer, Berlin.

Chen, O., A. Hotz, M. Ben-Akiva. 1997. Development and evaluation of a dynamic metering. Control model. *Proc. 8th IFAC/IFIP/IFORS Sympos. Transportation Systems*, Chania, Greece.

Diakaki, C., M. Papageorgiou. 1994. Design and simulation test of coordinated ramp metering control (METALINE) for A10-West in Amsterdam. Internal Report 1994-2, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.

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

Hall, F. L. 2001. Traffic stream characteristics. N. Gartner, C. J. Messer, A. K. Rathi, eds. *Traffic Flow Theory: State-of-the-Art Report*, Chapter 2. Federal Highway Administration/Transportation Research Board/Oak Ridge National Laboratory.

Hegyi, A. 2004. Model predictive control for integrating traffic control measurers. Ph.D. thesis, TRAIL Thesis Series T2004/2, Delft University of Technology, Delft, The Netherlands.

Kampitaki, K. 2008. Integrated control of traffic flow using ramp metering and variable speed limits. M.Sc. thesis, Technical University of Crete, Chania, Greece.

Kotsialos, A., M. Papageorgiou. 2004. Efficiency and equity properties of freeway network-wide ramp metering with AMOC. *Transportation Res. Part C* **14**(6) 401–420.

Kotsialos, A., M. Papageorgiou, M. Mangeas, H. Haj-Salem. 2002a. Coordinated and integrated control of motorway networks via nonlinear optimal control. *Transportation Res. Part C* **10**(1) 65–84.

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

Lebacque, J. P., H. Haj-Salem. 2001. Speed limit control: A problem formulation and theortical disscussions. *Proc. 4th Triennial Sympos. Transportation Anal.*, Vol. 2. *São Miguel, Azores*, TRISTAN, 421–426.

Lebacque, J. P., H. Haj-Salem. 2004. Ramp metering, speed management, and Braess-like paradoxes. *Proc. 5th Triennial Sympos. Transportation Anal.* CD-ROM, TRISTAN V, *Le Gosier, Guadeloupe, French West Indies*.

Messmer, A., M. Papageorgiou. 1990. METANET: A macroscopic simulation program for motorway networks. *Traffic Engrg. Control* **31**(8) 466–470.

Papageorgiou, M., A. Kotsialos. 2002. Freeway ramp metering: An overview. *IEEE Trans. Intelligent Transportation Systems* **3**(4) 271–281.

Papageorgiou, M., M. Marinaki. 1995. A feasible direction algorithm for the numerical solution of optimal control problems. Internal report 1995–4, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.

Papageorgiou, M., R. Mayr. 1982. Optimal decomposition methods applied to motorway traffic control. *Internat. J. Control* **35**(2) 269–280.

Papageorgiou, M., J. M. Blosseville, H. Haj-Salem. 1990a. Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris—Part I: Modelling. *Transportation Res. Part A* **24**(4) 345–359.

Papageorgiou, M., J. M. Blosseville, H. Haj-Salem. 1990b. Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris—Part II: Coordinated on-ramp metering. *Transportation Res. Part A* **24**(4) 361–370.

Papageorgiou, M., E. Kosmatopoulos, I. Papamichail. 2008. Effects of variable speed limits on motorway traffic flow. *Transportation Res. Record* **2047** 37–48.

Papageorgiou, M., E. Kosmatopoulos, M. Protopappas, I. Papamichail. 2006. Effects of variable speed limits (VSL) on motorway traffic. Internal report 2006-25, Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece.

Papageorgiou, M., I. Papamichail, A. D. Spiliopoulou, A. F. Lentzakis. 2008. Real-time merging traffic control with applications to toll plaza and work zone management. *Transportation Res. Part C* **16**(5) 535–553.

Papamichail, I., K. Kampitaki, M. Papageorgiou, A. Messmer. 2008. Integrated ramp metering and variable speed limit control of motorway traffic flow. *Proc. 17th IFAC World Congress, Seoul, Korea*, accessed July 24 2009, http://www.ifac-papersonline.net/detailed/38066.html.

Papamichail, I., A. Kotsialos, I. Margonis, M. Papageorgiou. 2009. Coordinated ramp metering for freeway networks—A model-predictive hierarchical control approach. *Transportation Res. Part C*. Forthcoming.

Smaragdis, E., M. Papageorgiou. 2003. A series of new local ramp metering strategies. *Transportation Res. Record* **1856** 74–86.

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

Zackor, H. 1972. Beurteilung verkehrsabhängiger Geschwindigkeits beschränkungen auf Autobahnen. *Strassenbau-und-Strassenverkehrstechnik* **128** 1–61.

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

Zhang, H. M., W. W. Recker. 1999. On optimal freeway ramp control policies for congested traffic corridors. *Transportation Res. Part B* **33**(6) 417–436.