

AN OPTIMISATION-ORIENTED FIRST-ORDER MULTI-LANE MODEL FOR MOTORWAY TRAFFIC

Claudio Roncoli, Corresponding Author

Dynamic Systems and Simulation Laboratory

Technical University of Crete

Chania, 73100, Greece

Tel: +30 28210 37289; Fax: +30 28210 37584; Email: croncoli@dssl.tuc.gr

Markos Papageorgiou

Dynamic Systems and Simulation Laboratory

Technical University of Crete

Chania, 73100, Greece

Email: markos@dssl.tuc.gr

Ioannis Papamichail

Dynamic Systems and Simulation Laboratory

Technical University of Crete

Chania, 73100, Greece

Email: ipapa@dssl.tuc.gr

Word count: 5518 words + 10 tables/figures x 250 words (each) = 8018 words

November 11, 2014

ABSTRACT

Emerging vehicle automation and communication systems (VACS) may contribute to the mitigation of motorway traffic congestion on the basis of appropriate traffic control strategies. Based on appropriate VACS-based actuators, future traffic control may incorporate vehicle speed control and lane-assignment or lane-changing recommendations. To this end, an appropriate traffic flow model is needed, both for control strategy design and as a no-control base case for comparative evaluation studies. In this context, this paper presents a novel first-order multi-lane macroscopic traffic flow model for motorways which is mainly intended for use within a related optimal control problem formulation. The starting point is close to the well-known cell-transmission model (CTM), which is modified and extended to consider additional aspects of the traffic dynamics, such as lane changing and the capacity drop. The model has been derived with a view to combine realistic traffic flow description with a simple (piecewise linear) mathematical form, which can be exploited for efficient optimal control problem formulations. Although the model has been primarily derived for use in future traffic conditions including VACS, it may also be used for conventional traffic flow representation. In fact, the accuracy of the proposed modelling approach is demonstrated through calibration and validation procedures using conventional real data from an urban motorway located in Melbourne, Australia.

Keywords: macroscopic traffic flow modelling, multi-lane motorway traffic, capacity drop

INTRODUCTION

Traffic congestion is a major problem of modern motorway systems, causing serious infrastructure degradation in and around metropolitan areas. The European Commission estimates that the yearly cost of road traffic congestion in Europe exceeds 120 billion €, and similar figures apply to USA as well. Despite a multitude of practised traffic control measures, improvements in combating traffic congestion and its detrimental consequences have been relatively moderate.

Macroscopic traffic flow modelling has been the subject of many research works, pursuing the re-production of different traffic phenomena and aiming at achieving an increased accuracy. At the same time, numerous control approaches have been proposed in the last decades, which employ model-based optimisation techniques and algorithms of various types, depending on the features of the employed model; clearly, the required computation time is a major concern in these developments.

On the other hand, there has been a significant and increasing interdisciplinary effort by the automotive industry as well as by numerous research institutions around the world to plan, develop, test and start deploying a variety of Vehicle Automation and Communication Systems (VACS) that are expected to revolutionise the features and capabilities of individual vehicles within the next decades. A wide description of VACS may be found in Bishop (1). Although most VACS will have no direct impact on traffic flow, as they are aiming at improving safety or driver convenience, some of them will change actively the vehicle behaviour, offering new opportunities for innovative traffic control actions, such as individual vehicle speed and lane-change advices, among others. The uncertain future evolution path of VACS deployment calls for the development of modelling and control approaches that are robust, to the extent possible, to evolving VACS and corresponding penetration rates.

The main purpose of this paper is to describe a novel macroscopic traffic flow model to be used within a model-based optimal control approach, which enables the introduction of different, partly novel, control measures. Some related control actions can be possibly implemented using conventional road-side actuators (traffic lights, variable message signs), whereas some others may need emerging devices installed on-board of VACS-equipped vehicles. Due to this intended utilisation, the model should be able to reproduce relevant controllable phenomena, such as lane changing, speed control, and different levels of capacity drop, which may depend on the kind and penetration rates of the considered VACS, and should be described even under the no-control scenario. Because of its intended usage within an optimal control scheme (see Roncoli et al. (13) for the corresponding developments)

The rest of the paper is structured as follows. Some available models are reviewed and the proposed multi-lane macroscopic traffic flow model for motorways is described, highlighting its novel aspects. Then, the proposed model is calibrated and validated using conventional real data from a motorway network. The paper ends with a summary of the main conclusions.

A MULTIPLE-LANE TRAFFIC FLOW MODEL FOR MOTORWAYS

Literature review of multiple-lane traffic flow models

The vast majority of the proposed macroscopic motorway traffic flow models consider all the variables aggregated across all the lanes (for a classification of existing models, see Treiber and Kesting (2)); only a limited number of studies consider the different lanes in a motorway as independent entities. One of the first important works that addressed this topic was by Gazis et al. (3), where lane densities on a multi-lane highway were assumed to oscillate around an equilibrium density.

This work inspired Michalopoulos et al. (4), who proposed different models characterised by different complexity, for capturing the lane-changing behaviour; however, they were formulated in a continuous space-time domain, without applying any discretisation scheme. In a more recent work, Laval and Daganzo (5) exploited the kinematic wave (KW) theory, proposing a multi-lane KW-based model as a first module of a more complex model that considers also moving block-ages, treated as particles and characterised by bounded acceleration rates. Yet another approach is presented by Jin (6), in which the author developed a multi-commodity model based on the LWR theory; in this case the lane-changing flow is defined based on the concept of lane-changing intensity (that is introduced as a new variable which affects the speed-density relationship), defining a new fundamental diagram (FD) and introducing a so-called entropy condition.

The generic modelling framework

A multiple-lane motorway is considered, which is subdivided into segments, while each segment comprises a number of lanes. The index $i = 1, \dots, I$ is introduced for segments and the index $j = 1, \dots, J$ for lanes. The model is formulated in discrete time, considering the discrete time step T for a simulation horizon K indexed by $k = 1, \dots, K$, where the simulation time is $t = kT$. The motorway is discretised in space by defining the segment-lane entities, which are characterised by the following variables (a graphical representation of the segment-lane variables is presented in Figure 1):

- the density $\rho_{i,j}(k)$ [veh/km], i.e. the number of vehicles in the segment i , lane j , at time step k , divided by the segment length L_i ;
- the longitudinal flow $q_{i,j}(k)$ [veh/h], i.e. the traffic volume leaving segment i and entering segment $i + 1$ during time interval $(k, k + 1]$, remaining in lane j ;
- the lateral flow $f_{i,j,\bar{j}}(k)$ [veh/h] ($\bar{j} = j \pm 1$), i.e. the traffic volume moving from lane j to lane \bar{j} (vehicles changing lane remain in the same segment during the current time interval); and
- the on-ramp flow $r_{i,j}(k)$ [veh/h], i.e. the traffic volume entering from the on-ramp located at segment i , lane j , during the time interval $(k, k + 1]$.

The off-ramp flow is determined according to given turning rates $\gamma_{i,j}(k)$, as a percentage of the total flow passing through all the lanes of the segment. According to the aforementioned notation, the following conservation equation is introduced:

$$\rho_{i,j}(k+1) = \rho_{i,j}(k) + \frac{T}{L_i} \left[q_{i-1,j}(k) + r_{i,j}(k) - q_{i,j}(k) - \gamma_{i,j}(k) \sum_{j=1}^J q_{i,j}(k) + f_{i,j+1,j}(k) + f_{i,j-1,j}(k) - f_{i,j,j-1}(k) - f_{i,j,j+1}(k) \right]. \quad (1)$$

The methodology to compute the lateral and longitudinal flows is described in the next subsections.

Computation of lateral flows

When appropriate communications between equipped vehicles and the control strategy are available, efficient lane-changing advice or control may be effectuated for selected vehicles. However,

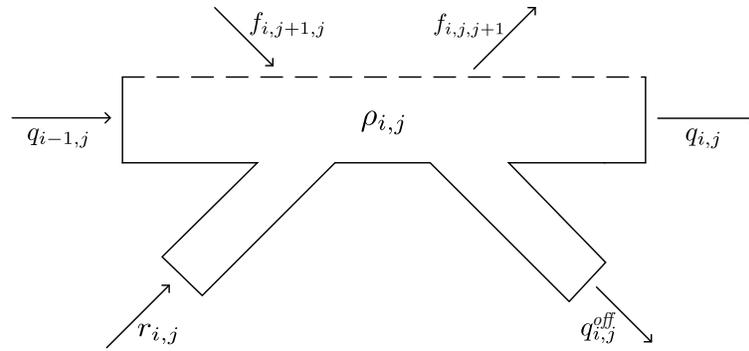


FIGURE 1 The segment-lane variables used in the model formulation.

if the corresponding VACS penetration rates are small or if a no-control base-case scenario is required for comparative evaluation, the model should be enabled to reproduce also the “natural” aggregate lane-changing behaviour of traffic. Lateral flows due to “natural” lane-changes are considered among adjacent lanes of the same segment, and corresponding rules must be defined in order to properly assign and bound the lateral flows. To start with, the maximum available flow for lateral movements is calculated based on the current amount of vehicles in the segment-lane:

$$F_{i,j}(k) = \frac{L_i}{T} \rho_{i,j}(k). \tag{2}$$

The pursued approach is to compute the lateral flow from segment-lane (i, j) to (i, \bar{j}) as a rate $A_{i,j,\bar{j}} \in [0, 1]$ of the value $F_{i,j}(k)$. We call this rate the attractiveness rate.

As a matter of fact, “natural” (uncontrolled) lane-changing behaviour and flows are extremely hard to model accurately (both macroscopically and microscopically) because they depend on a very high number of partly interdependent factors. To start with, different vehicle types (e.g. cars and trucks) and related restrictions (e.g. regarding lane usage) may give rise to a variety of different lane-changing characteristics. Human driver behaviour is another source of uncertainty and variation. Furthermore, lane-changing behaviour is different in case of road curvature or grade, lane drop, as well as in tunnels, on bridges etc.; it is also dependent on the number of lanes, environmental conditions (weather, lighting), traffic conditions (free-flow, dense, congested) and traffic signs. Last not least, lane-changing activity is quite particular in the vicinity of on-ramps and off-ramps, or at weaving sections. Considering this diversity and complexity, it appears appropriate, in the present context, to include a simple basic lane-changing flow model, capable of capturing with some accuracy many “ordinary” situations; a space-time dependent parameter is also included, allowing to influence the model behaviour appropriately whenever needed (e.g. near on- and off-ramps).

For the basic lane-changing flow model, it is assumed that drivers may consider a lane change when one of the adjacent lanes offers a higher speed or a lower traffic density. The latter option is preferred here because traffic densities are state variables of the proposed model. Thus, under “ordinary” conditions, the current attractiveness rate $A_{i,j,\bar{j}}$ may be deemed to increase proportionally to the current density difference $\rho_{i,j}(k) - \rho_{i,\bar{j}}(k)$ of adjacent lanes j and \bar{j} . This basic assumption, however, may be subject to variations due to various local effects, as mentioned earlier. For example, vehicles driving on the slow lane may consider a lane change upstream of on- or off-ramps,

to avoid interference with entering or exiting vehicles, respectively; similarly, vehicle lane assignment may deviate from the basic rule upstream of lane-drop locations; also, exiting vehicles have to move towards the exit lane(s) irrespective of the prevailing traffic density there. Some of these or other variations from the basic rule may additionally depend on time. To capture this variety of potential situations, the attractiveness rate is modelled to depend on the weighted density difference $P_{i,j,\bar{j}}(k)\rho_{i,j}(k) - \rho_{i,\bar{j}}(k)$, where the introduced factor $P_{i,j,\bar{j}}(k)$ is mostly equal to 1, but should be tuned to reflect particular location- or time-dependent effects where needed. This factor takes into account the same difference in density between the adjacent lanes, irrespectively of the considered lane-changing direction, thus it respects the relation $P_{i,j,\bar{j}}(k) = \frac{1}{P_{i,\bar{j},j}(k)}$ for all the adjacent segment-lane couples (i, j) , (i, \bar{j}) .

Finally, the attractiveness rate $A_{i,j,\bar{j}}$ is computed as:

$$A_{i,j,\bar{j}}(k) = \mu \max \left[0, \frac{P_{i,j,\bar{j}}(k)\rho_{i,j}(k) - \rho_{i,\bar{j}}(k)}{P_{i,j,\bar{j}}(k)\rho_{i,j}(k) + \rho_{i,\bar{j}}(k)} \right] \quad (3)$$

where the coefficient μ is a unique parameter in the range $[0, 1]$. Note that in Equation 3, the fraction within the bracket is a number in the range $[-1, 1]$, however only positive values are considered (which means that for all adjacent couples of segments-lanes, only one has a positive value).

The modelling function in Equation 3 is non-linear; nevertheless this is no impediment for the intended optimal control problem formulation because, as mentioned earlier, the lateral flows will be considered controllable in the control problem. Thus, Equation 3 is intended to reflect the “natural” lane-changing flow behaviour. Since the proposed methodology is not strongly dependent on the function used here for computing $A_{i,j,\bar{j}}(k)$, other functions could be used for this computation if deemed more appropriate.

As mentioned earlier, vehicles bound for specific off-ramps should change lanes in a due manner. In this model it is considered that the off-ramp flow is computed as a percentage of the total flow passing through all lanes of the segment and not only of the shoulder lane. If the values $P_{i,j,\bar{j}}(k)$ are properly tuned and the turning rates $\gamma_{i,j}(k)$ do not feature strong oscillations, this should not create any noticeable problem. In the general case, however, it could happen that the number of vehicles in the shoulder lane is insufficient to guarantee the necessary exiting flow in accordance with the corresponding pre-fixed exiting rate. In order to cope with this possibility, an approach is introduced, which incorporates a forecast of the entering and leaving flows in the off-ramp location. It is based on the assumption that $T \approx \frac{L_i}{v_{i,j}^{free}}$; this means that, in free-flow conditions, we have the flows $q_{i,j}(k) \approx q_{i-1,j}(k-1)$ and consequently $q_{i-1,j}(k) \approx q_{i-2,j}(k-1)$. Therefore, considering the conservation Equation 1, the required amount of lateral flow needed to satisfy the off-ramp flow is given by:

$$F_{i,j,\bar{j}}^{off}(k) = \gamma_{i,j}(k) \sum_{j=1}^J q_{i-1,j}(k-1) - q_{i-2,j-1}(k-1). \quad (4)$$

This procedure is applied only in the direction towards the lane that includes the ramp and it could be iteratively applied to further upstream segments in case the flow is not sufficient to match the exit rate.

In conclusion, the lateral demand flow, i.e. the flow that will actually materialise if there is enough space in the target segment-lane, is assigned according to the following formula:

$$D_{i,j,\bar{j}}(k) = \max \left[A_{i,j,\bar{j}}(k) F_{i,j}(k), F_{i,j,\bar{j}}^{off}(k) \right] \quad (5)$$

where $\bar{j} = j \pm 1$.

To complete the lateral flow modelling development, we need to account for the space available in the segment-lane that is receiving the vehicles moving laterally, considering the function:

$$S_{i,\bar{j}}(k) = \left[\rho_{i,\bar{j}}^{jam} - \rho_{i,\bar{j}}(k) \right] \frac{L_i}{T}. \quad (6)$$

Since the available space may not be sufficient for accepting the lateral flow entering from both sides, the assigned quantity is proportionally distributed according to the following relation:

$$f_{i,j-1,j}(k) = \min \left[1, \frac{S_{i,j}(k)}{D_{i,j-1,j}(k) + D_{i,j+1,j}(k)} \right] D_{i,j-1,j}(k) \quad (7)$$

$$f_{i,j+1,j}(k) = \min \left[1, \frac{S_{i,j}(k)}{D_{i,j-1,j}(k) + D_{i,j+1,j}(k)} \right] D_{i,j+1,j}(k).$$

These values may be directly used for updating the densities according to Equation 1. Note that, in case of limited available space in the target segment-lane, the described approach assigns higher priority to the lateral flows via Equation 7. Naturally, these flows will be considered in the computation of the longitudinal flows in the next subsection.

Computation of longitudinal flows

As previously mentioned, the longitudinal flows are defined as the flows going from a segment to the next downstream one, while remaining in the same lane. These flows may be considered controllable via appropriate speed control of equipped vehicles; but they need to be appropriately modelled for the no-control case. The main logic for their computation is based on the conventional cell transmission model (CTM), as proposed by Daganzo (7), where the longitudinal flow is computed as the minimum between an upstream demand flow and a downstream supply flow. However, one important phenomenon that regularly appears in real traffic, but is not reproduced by the classical CTM, is the capacity drop phenomenon, i.e. the reduction of discharge flow once queues start forming at a bottleneck location. The reasons for this phenomenon are not exactly clear, however it seems to be caused by the limited and varying acceleration of vehicles leaving a congested area while trying to reach their desired speed. In second-order models, such as METANET by Papageorgiou and Messmer (8), the capacity drop appears as a consequence of the spatiotemporal evolution of the speed dynamics. Since the speed is not dynamically calculated, this option is not available for first-order LWR models; hence, several attempts have been made in order to introduce it in some other ways. Hall and Hall (9) suggested the introduction of an inverse lambda FD, whereas Leclercq et al. (10) proposed an analytical model that incorporates endogenously the capacity drop. Lebacque (11) addressed the problem by imposing an upper bound to the acceleration depending on the traffic phase, distinguishing between traffic equilibrium and maximum acceleration conditions. Laval and Daganzo (5) proposed the introduction of discrete particles in the traffic flow, treating them as moving temporary blockages. Srivastava and Geroliminis (12)

extended the LWR model by defining a FD with two values of capacity and providing a memory-based methodology to choose the appropriate value in the numerical computation, distinguishing between congested and uncongested states.

A common feature of most of the aforementioned approaches is the non-linear formulation that makes them unattractive for the purpose of this paper. For this reason, in this work, the capacity drop phenomenon is represented by a modified demand part of the FD in the following way, as proposed by Lebacque (11) and illustrated in Figure 2. In case of congestion ($\rho_{i,j}(k) > \rho_{i,j}^{cr}$, where $\rho_{i,j}^{cr}$ is the critical density for the segment-lane (i, j)), the flow is linearly decreased according to a fixed slope $-w^D$, instead of being equal to capacity flow as in the classical CTM. This leads to the definition of an additional point in the FD: $q_{i,j}^{jam}$, i.e. the flow that is allowed to leave a segment in a completely congested state (when $\rho_{i,j}(k) = \rho_{i,j}^{jam}$).

In conclusion, the demand part for the longitudinal flow calculation is computed as:

$$Q_{i,j}^D(k) = \min \left[v_{i,j}^{free} \rho_{i,j}(k), -w^D \rho_{i,j}(k) + g^D \right] \quad (8)$$

where:

$$w^D = \frac{v_{i,j}^{free} \rho_{i,j}^{cr} - q_{i,j}^{jam}}{\rho_{i,j}^{jam} - \rho_{i,j}^{cr}} \quad (9)$$

$$g^D = \frac{\rho_{i,j}^{cr} (v_{i,j}^{free} \rho_{i,j}^{jam} - q_{i,j}^{jam})}{\rho_{i,j}^{jam} - \rho_{i,j}^{cr}}.$$

The supply function is computed based on the density of the downstream segment-lane as in the classical CTM:

$$Q_{i,j}^S(k) = \min \left[v_{i+1,j}^{free} \rho_{i+1,j}^{cr}, -w^S \rho_{i+1,j}(k) + g^S \right] \quad (10)$$

where:

$$w^S = \frac{v_{i+1,j}^{free} \rho_{i+1,j}^{cr}}{\rho_{i+1,j}^{jam} - \rho_{i+1,j}^{cr}} \quad (11)$$

$$g^S = \frac{v_{i+1,j}^{free} \rho_{i+1,j}^{cr} \rho_{i+1,j}^{jam}}{\rho_{i+1,j}^{jam} - \rho_{i+1,j}^{cr}}.$$

In presence of vehicle-speed controlling VACS (such as Adaptive Cruise Control – ACC – systems), the slope $-w^D$ is likely to depend on the employed acceleration of the automated system; while for conventional vehicles, it depends on the “natural” acceleration by human drivers. Thus, the parameter $-w^D$ may need to assume different values in presence of such VACS than in conventional traffic. In any case, the accurate calibration of the slope $-w^D$ is a non-trivial task because of the complexity in measuring the effects of a capacity drop; therefore a first attempt may be to simply define it in relation with the slope of the supply part of the FD $-w^S$ (according to some observations of traffic data, a reasonable value is $w^D \approx \frac{w^S}{3}$). Model calibration exercises, based

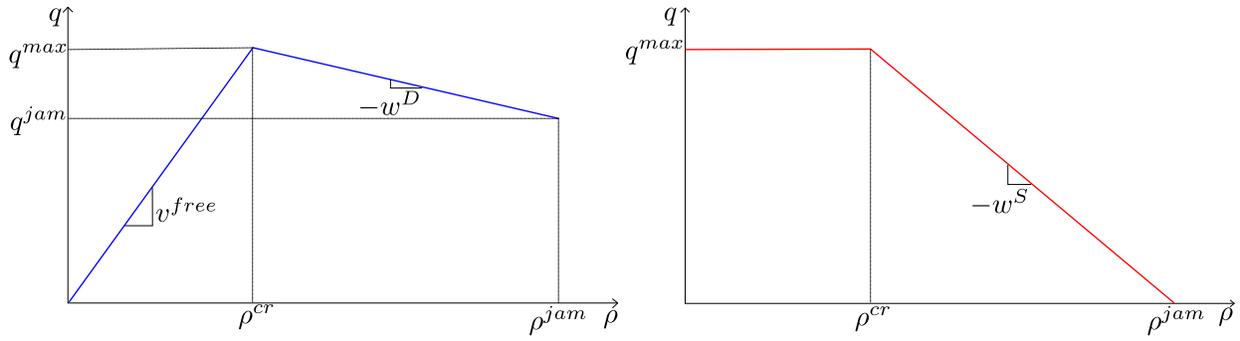


FIGURE 2 The proposed demand (left) and supply (right) parts of the Fundamental Diagram; the demand part includes the linear term for capacity drop.

on real traffic data (as in this paper), may lead to a factual quantification of the capacity drop impact. In this work, the value q^{jam} is calibrated as a model parameter using conventional (no VACS) traffic flow data.

On another issue, note that CTM considers the speed to be equal to the free speed for all under-critical densities. In the present context, this is not a strict requirement; different polygons could be used for the FD of Figure 2, the strict requirements being continuity and piecewise linearity of the employed functions. In particular, flattening the free speed line while approaching the critical density may lead to more realistic results.

Since the proposed model considers a higher priority associated to lateral flows, their values must be subtracted (or respectively added) from (to) the values previously computed:

$$\hat{Q}_{i,j}^D(k) = \min \left\{ Q_{i,j}^D(k), \frac{L_i}{T} \rho_{i,j}(k) + [f_{i,j-1,j}(k) + f_{i,j+1,j}(k)] - [f_{i,j,j-1}(k) + f_{i,j,j+1}(k)] \right\} \quad (12)$$

$$\hat{Q}_{i,j}^S(k) = Q_{i,j}^S(k) + [f_{i+1,j,j-1}(k) + f_{i+1,j,j+1}(k)] - [f_{i+1,j-1,j}(k) + f_{i+1,j+1,j}(k)]$$

Thus, the final longitudinal flows $q_{i,j}(k)$ are calculated according to the following equation:

$$q_{i,j}(k) = \min \left[\hat{Q}_{i,j}^D(k), \hat{Q}_{i,j}^S(k) \right]. \quad (13)$$

The model admits the possibility of introducing upper bounds $q_{i,j}^{off,max}$ for exit flows at off-ramps, which leads to the modelling of possible congestion due to limited-capacity off-ramps flows. In this case, the upper-bounds $q_{i,j}^{off,max}$ will affect the outflow in all the mainstream lanes at the off-ramp location; to reproduce this phenomenon, the longitudinal flow must be updated as follows:

$$\delta_{i,j}^{off}(k) = \min \left[1, \frac{q_{i,j}^{off,max}}{\gamma_{i,j}(k) \sum_{j=1}^J q_{i,j}(k)} \right] \quad (14)$$

$$\hat{q}_{i,j}(k) = \delta_{i,j}^{off}(k) q_{i,j}(k).$$

Also lateral flow into a segment-lane may affect its capacity. To reflect this possibility within our approach, we may readily render the flow capacity of a segment-lane linearly dependent on the

entering lateral flows. Again, a linear relation can be considered, with the purpose of decreasing the capacity proportionally to the entering flow (and similarly, the concept can be extended also for the flow entering from on-ramps); the following general formula can be applied, that updates the computation of the demand part for the longitudinal flow used in Equation 8:

$$\hat{g}^D = g^D - \alpha_f [f_{i,j-1,j}(k) + f_{i,j+1,j}(k)] - \alpha_r r_{i,j}(k) \quad (15)$$

where α_f and α_r are parameters to be opportunely tuned for entering lateral flow and on-ramp flow respectively.

The presented model has been used as a basis for the development of a linearly constrained optimal control problem for the traffic flow in presence of novel control actions enabled by specific VACS, see Roncoli et al. (13) for details.

MODEL CALIBRATION AND VALIDATION

Network description

The motorway traffic flow model introduced in the previous section is now applied to a particular motorway stretch in order to calibrate its parameters and validate the model equations under conventional traffic conditions. The chosen network is a stretch of the Monash Freeway (M1) located in the area of Melbourne, Victoria, Australia. Specifically, it is an urban motorway characterised by a traffic pattern that is strongly dependent on the demand due to commuters driving to and from the city centre.

The considered stretch, sketched in Figure 3, is 5.26 km long and is composed by four lanes. It starts immediately upstream of the on-ramp connected to Gully Road, it includes the on- and off-ramps connecting M1 to Blackburn Road and Forster Road, and it terminates about 1 Km downstream of the off-ramp connecting Huntingdale Road. Data is collected through 11 traffic sensors that deliver measurements of flow and speed (per lane) every minute; sensors are also located on the ramps, measuring the corresponding ongoing and outgoing flows. According to its topology, the network is subdivided in 20 segments, of lengths between 259 m to 314 m, as seen in Figure 3, that shows also the locations of the sensors (black dots), on-ramps, and off-ramps.

It was decided to take into account the morning peak, specifically from 5 AM to 9 AM of the 12th of August 2013 for model calibration. During this day, major congestion is created from about 6:15 AM. The reason why a major congestion is created is quite complex and it is due to a combination of the increased demand that interacts with the high number of trucks and the slope of the motorway. In fact, in the last part of the network, the motorway starts rising, causing the slowing down of trucks, and, because of that, a high amount of cars leave the shoulder lane to avoid the reduction of speed. Thus, a reduction of the capacity in the shoulder lane takes place due to the slope and the high percentage of trucks, and an increase of flow in the other lanes due to escaping (lane-changing) cars. This creates a congestion that spills back and spreads all over the considered network, reaching the merging areas of on-ramps. The congestion lasts for a couple of hours, until it is resolved thanks to the reduced overall demand at the end of the peak period.

The data retrieved from detector D1 is used for providing entering flows for the 4 lanes of the first segment. The downstream boundary conditions at the network exit are treated considering the measurements taken from detector D11. Since only flow and speed measurement are available, we can assume that a congestion is back-spilling from downstream whenever a reduction of speed appears; thus, in case the measured speed is lower than the free speed (in order to account for the measurement noise, a value slightly lower than the free speed is chosen), the outflow of the

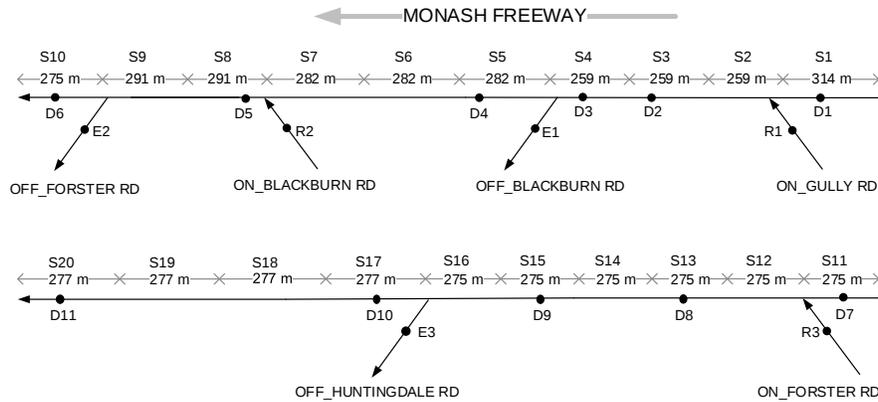


FIGURE 3 A graphical representation of the motorway stretch considered.

last segment is set equal to the minimum between the computed demand flow for the last segment and the measured flow. Moreover, in order to obtain similar outflows at the off-ramps, the measurements E1, E2, and E3, together with the respective measurements on the mainstream D3, D5, and D9, are used to compute the turning rates $\gamma_{i,j}(k)$. No off-ramp blocking occurs in this stretch, hence Equation 14 is not needed. Also, no capacity reduction in dependence of the lateral flows is considered in the model validation study.

Calibration results

The calibration of the model parameters to best match the measurement data has been performed with the help of an optimisation-based methodology for the calibration and validation of macroscopic traffic flow models. This methodology attempts to minimise a cost function that measures the difference between the results obtained applying the model versus the measurements taken from the detectors, via appropriate specification of the model parameters. In order to cope with the possibility of falling into local minima, the Nelder-Mead method, in the version proposed by Lagarias et al. (14), has been chosen; a detailed explanation of the applied methodology is presented by Spiliopoulou et al. (15) and its origin goes back to the pioneering work by Cremer and Papageorgiou (16).

The proposed model allows for different parameter values to be introduced for each segment-lane entity, which, however, would probably lead to over-parametrisation. Thus, in order to reduce the number of parameters, segment-lanes are grouped as shown in Table 1. The displayed grouping is firstly and foremost based on the lane number; further sub-grouping was based on some observed patterns which were mostly dependent on the network topology; the subdivision of lane 1 into many smaller groups is mainly due to the changing of traffic pattern implied by the proximity of on- and off-ramps.

The time step used for updating the dynamic equation is $T = 5$ s. The value of the jam density is set to $\rho_{i,j}^{jam} = 120$ veh/km for all the segments-lanes except for segments 17–20 lane 1, where a value $\rho_{i,j}^{jam} = 80$ veh/km is chosen in order to reflect their aforementioned peculiarity. For the computation of lateral flows, the coefficients $P_{i,j,\bar{j}}(k)$ are tuned according to some observations on the measured densities; more specifically, these values are set to $P_{i,j,\bar{j}}(k) = 1$ for lateral movements between lanes 2, 3, and 4, whereas in some specific locations between lanes 1 and 2, the factors

TABLE 1 The group of segments-lanes that share the same parameters for calibration and validation

Lane	Segments	Detectors	ρ^{cr} [veh/km]	v^{free} [km/h]	ρ^{jam} [veh/km] (pre-fixed)	q^{jam} [veh/h]
1	1 – 3	D1 – D2	24.5	92.3	120	1849
	4 – 7	D3 – D4	21.4	94.1	120	1597
	8 – 9	D5	19	97.9	120	1462
	10 – 11	D6	15.3	97.9	120	1185
	12 – 16	D7 – D9	21.9	95.1	120	1863
	17 – 20	D10 – D11	19	95.1	80	997
2	1 – 4	D1 – D3	22.2	94	120	1492
	5 – 20	D4 – D11	23.1	96	120	1594
3	1 – 4	D1 – D3	20	96.5	120	1718
	5 – 20	D4 – D11	22.9	99.5	120	1830
4	1 – 4	D1 – D3	24.1	103	120	1733
	5 – 20	D4 – D11	25.3	105.5	120	2092

assume a different value in order to cope with the increased density because of frequent on- and off-ramps (a value of $P_{i,j,\bar{j}}(k) = 1.5$ is set in the proximity of ramps). The real demand profiles, fed as corresponding on-ramp flows $r_{i,j}(k)$, are shown in Figure 4.

The cost function takes into account the difference in terms of both speed and longitudinal flow, allowing indirectly to capture also the lateral flow movements. More specifically, the cost function is calculated according to the following formula:

$$\min_{\bar{q}, \bar{v}} \left\{ \alpha^q \sqrt{\frac{\sum_{k=1}^K \sum_{j=1}^J \sum_{d=2}^D [\bar{q}_j^d(k) - \hat{q}_j^d(k)]^2}{KJ(D-1)}} + \alpha^v \sqrt{\frac{\sum_{k=1}^K \sum_{j=1}^J \sum_{d=2}^D [\bar{v}_j^d(k) - \hat{v}_j^d(k)]^2}{KJ(D-1)}} \right\} \quad (16)$$

where $\bar{q}_j^d(k)$ and $\bar{v}_j^d(k)$ are the flow and the speed computed by the model (the flow is calculated directly, whereas an approximation of the speed is calculated as $v_{i,j}(k) = \frac{q_{i,j}(k)}{\rho_{i,j}(k)}$) at the location of detector d ($d = 2, \dots, 11$) and $\hat{q}_j^d(k)$ and $\hat{v}_j^d(k)$ are the corresponding values measured by the detectors. Weigh coefficients α^q [h/veh] and α^v [h/km] are chosen in order to normalise the results; in this study, a good choice for the (normalised) weight of the speed term was found to be the double of the weight of the flow term; specifically $\alpha^q = 0.04$ h/veh and $\alpha^v = 2$ h/km. Equation 16 is also used as a performance index for evaluating the achieved results.

A graphical representation of the results obtained in the model calibration may be seen in Figures 5 and 6. The tuned parameters that are obtained are given in Table 1. Hereafter, some more detailed explanations on the real congestion pattern and the related reflection by the calibrated model are presented.

- Congestion is first created at about 6 AM because of the increase of the on-ramp flow at segment 2, and possibly because of some weaving phenomena due to the downstream exit (off-ramp at segment 4).

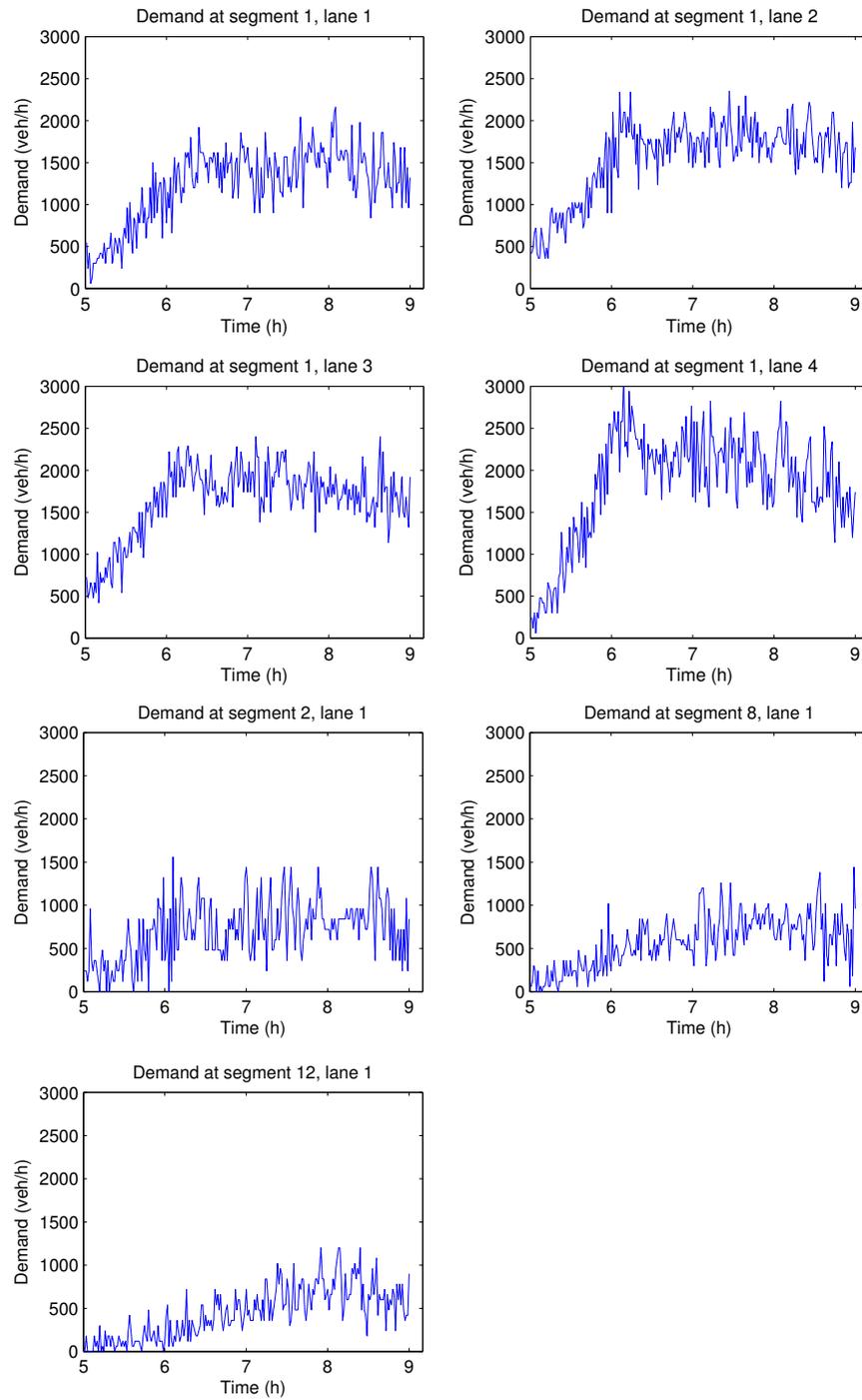


FIGURE 4 The demand profiles computed from detectors D1, R1, R2, and R3, that are used in the calibration phase.

- As previously mentioned, the main congestion starts in segment 17 (captured by Detector 10) at about 6:15 AM because of the high number of trucks reducing their velocity because of the slope of the motorway. In the calibrated model, this phenomenon is reproduced, firstly, via the relatively low calibrated $\rho_{i,j}^{cr}$ value, and then via the decreased value $\rho_{i,j}^{jam}$ set for that specific location that causes a rapid diminishing of the outflow when the critical density is exceeded.
- Since lane 1 of segments 17–20 is overcrowded with trucks, most car drivers take the decision to move to the adjacent lanes; this causes an increase of density that, quite rapidly, triggers a capacity drop due to the starting of a strong congestion. While rendering the model capable of dealing with the described complex phenomenon, the lane-changing factor $P_{i,j,\bar{j}}(k)$ plays a fundamental role. In fact, for each segment-lane located in segments 17–20, lane 1, these values are set to $P_{i,j,\bar{j}}(k) \approx 2$ (in the direction $\bar{j} > j$) during the time period in which trucks are the predominant vehicles (until about 6:30 AM), and are decreased progressively to $P_{i,j,\bar{j}}(k) = 1$ while the traffic composition is moving to normal. These parameters generate accordingly strong lateral flows directed from lane 1 to 2 (despite the density in lane 2 becoming higher than in lane 1), and, consequently, also to the other lanes in the direction of the road median. This generates a density increase in all the lanes and, once the density exceeds its critical value, the segment-lanes enter in congested state, causing the appearance of the capacity drop. This can be observed in Figure 5, by close inspection of the flow at Detector 10 for lanes 2, 3, and 4.
- Once the congestion in the last segments has started, it spills back, covering the whole stretch at about 6:35 AM. The back-spilling main congestion worsens further because of the increased flows entering from the on-ramps. More specifically, it may be seen that the increased ramp flow at segment 12 at about 7:30 AM contributes to strengthen the speed reduction (this can be seen comparing the speed diagrams at Detectors D8 and D7). Another contribution to the main congestion is again due to an increased flow entering from the on-ramp R2 that increases after 7 AM, causing a further reduction of speed (see the speed diagram for Detector D5). In this situation, the model is reproducing the congestion pattern because of the proper calibration of its parameters, above all the critical density. During this period, lateral movements are present almost exclusively in proximity of on- and off-ramps, where lane-changings are assigned from and towards lane 1; therefore, at those locations, the chosen values for $P_{i,j,j+1}(k)$ are greater than 1 (e.g. $P_{i,1,2}(k) \approx 1.5$).
- Around 8 AM, the demand starts to decrease, and this causes the gradual disappearance of the congestion, restoring the maximum flow and speed. Again, the model is able to follow properly this congestion dissolution pattern.

It should be also pointed out that, because of the first-order dynamics considered in this model, some quick changes of speed cannot be perfectly reproduced, e.g., the negative speed peak appearing around 6:30 AM is smoothed while back-spilling until the related density becomes under-critical, causing its disappearance.

The choice of having a different value of free speed for each segment-lane may represent an issue while using this model as the modelling part of an optimal control problem. Specifically, using

TABLE 2 Comparison of cost function values for different days used in the model validation. The RMSE values do not include the weights α^q and α^v , and they are expressed in [veh/h] and [km/h] respectively.

Day	Overall		Flow [veh/h]		Speed [km/h]	
	Cost Function Value	Diff.	RMSE	Diff.	RMSE	Diff.
12/08	39.7	–	285	–	14.1	–
14/08	45.1	13.6%	293	2.8%	16.7	18.4%
15/08	42.3	6.5%	298	4.6%	15.2	7.8%
16/08	36.4	-8.3%	278	-2.5%	12.6	-10.6%

different free speeds for different lanes may result in underutilisation of the slower lanes (in low-density situations) in an attempt of the optimiser to decrease the vehicle travel times. In case the difference in speed among lanes is not excessive (like in the presented example) a unique free speed for each segment (constant for all the lanes) may be a reasonable choice, without much loss in modelling accuracy, see e.g. Roncoli et al. (17).

Model validation

In order to test and demonstrate the validity of the proposed model, the parameters resulting from the calibration process are applied to the same motorway stretch but for different days. Specifically, the model is fed with boundary data stemming from different (incident-free) days, and its results, obtained with the same parameter values specified in the calibration procedure, are compared with corresponding real data from stretch-internal detectors. For all days, the (recurrent) traffic behaviour is similar to the one described in the previous section. Again, the results obtained are satisfactory, since the traffic flow model is able to reproduce correctly the traffic congestion patterns. For a quantitative comparison, the cost function presented in Equation 16 is used for the different days, obtaining the results shown in Table 2.

For the sake of brevity, only the graphical comparisons of resulting flow and speed (Figures 7 and 8) related to the 15th of August 2013 are presented here.

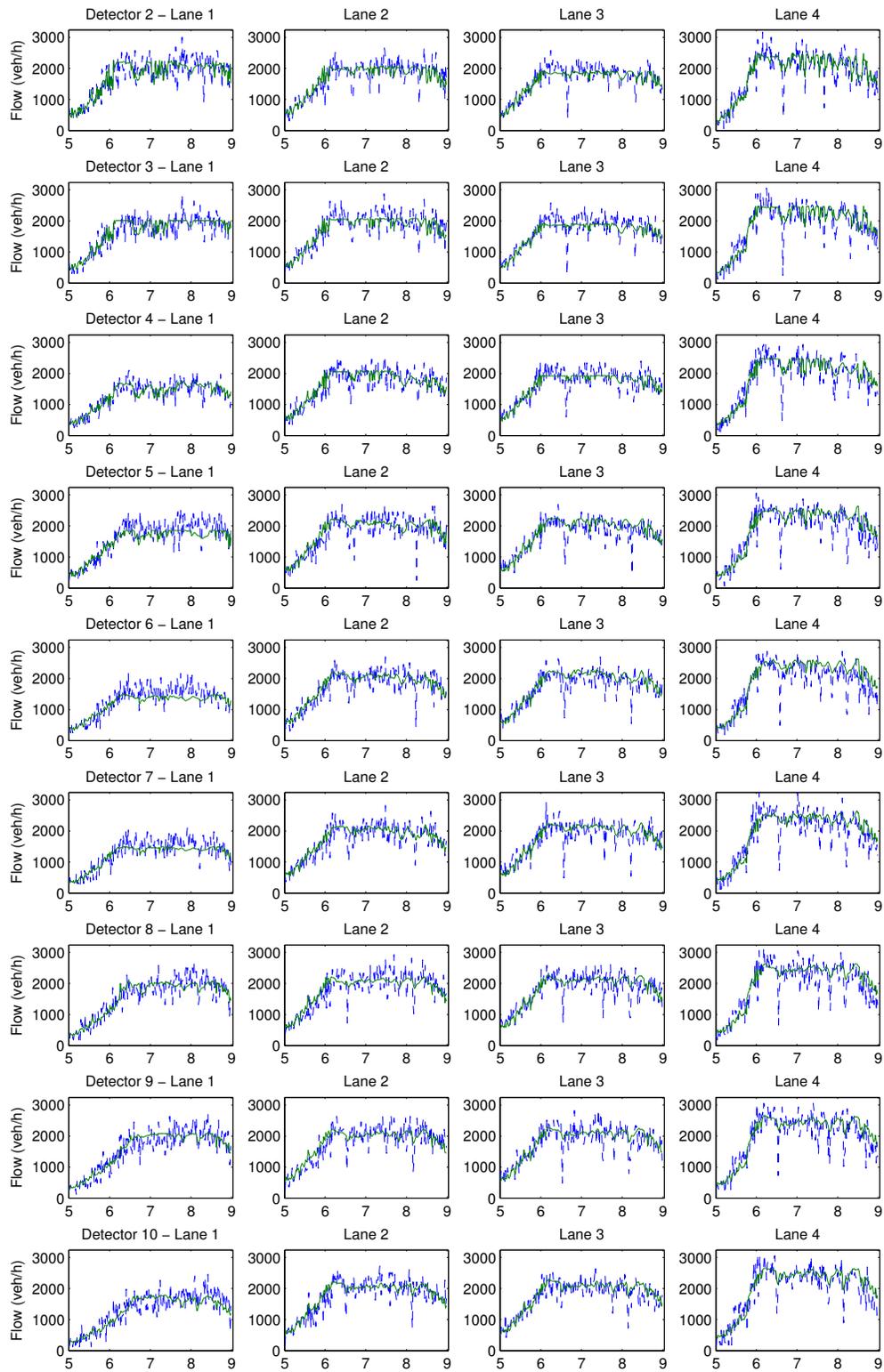


FIGURE 5 Comparison between the real flow data (dashed blue line) and the calibrated model results (solid green line).

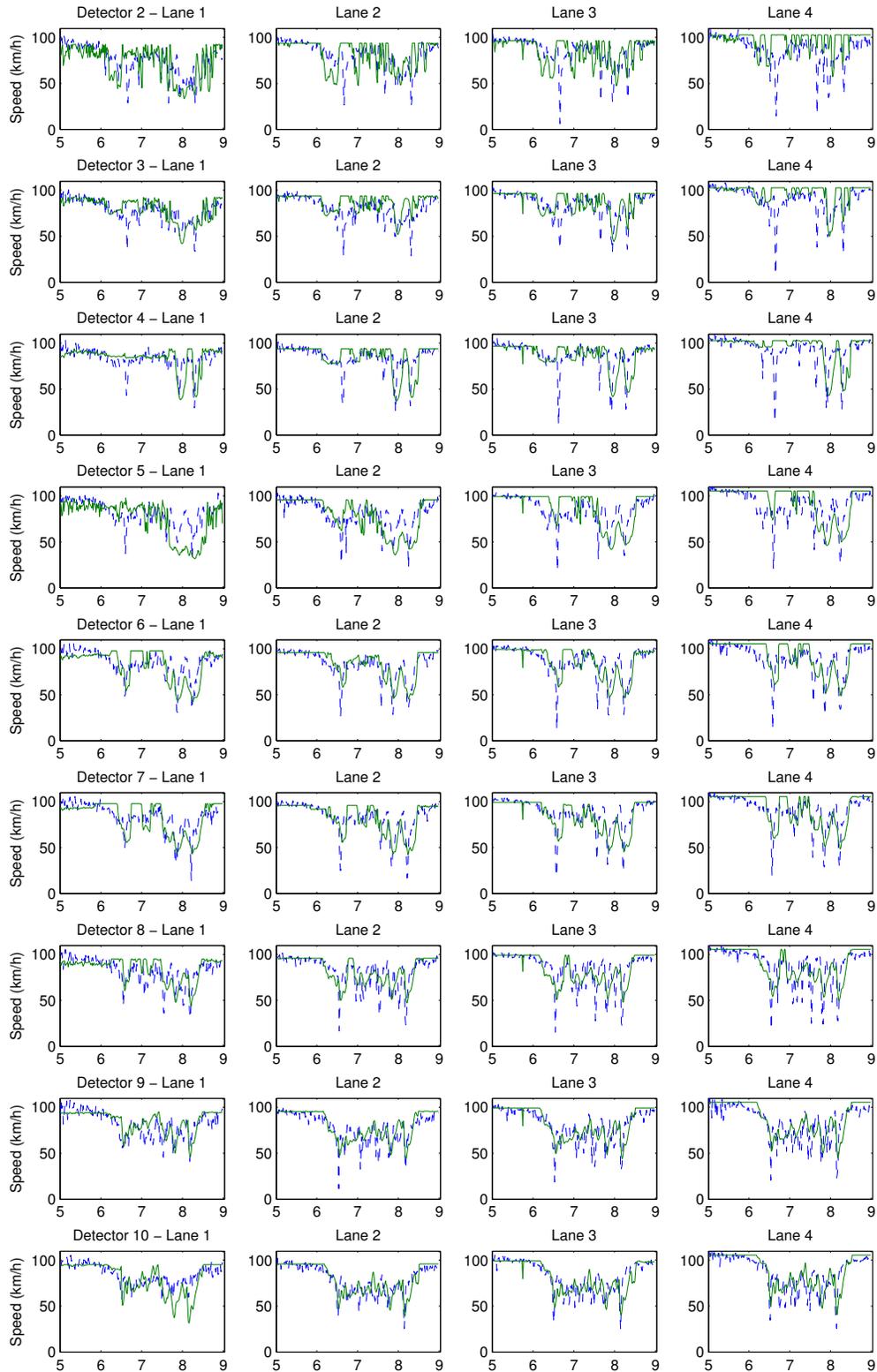


FIGURE 6 Comparison between the real speed data (dashed blue line) and the calibrated model results (solid green line).

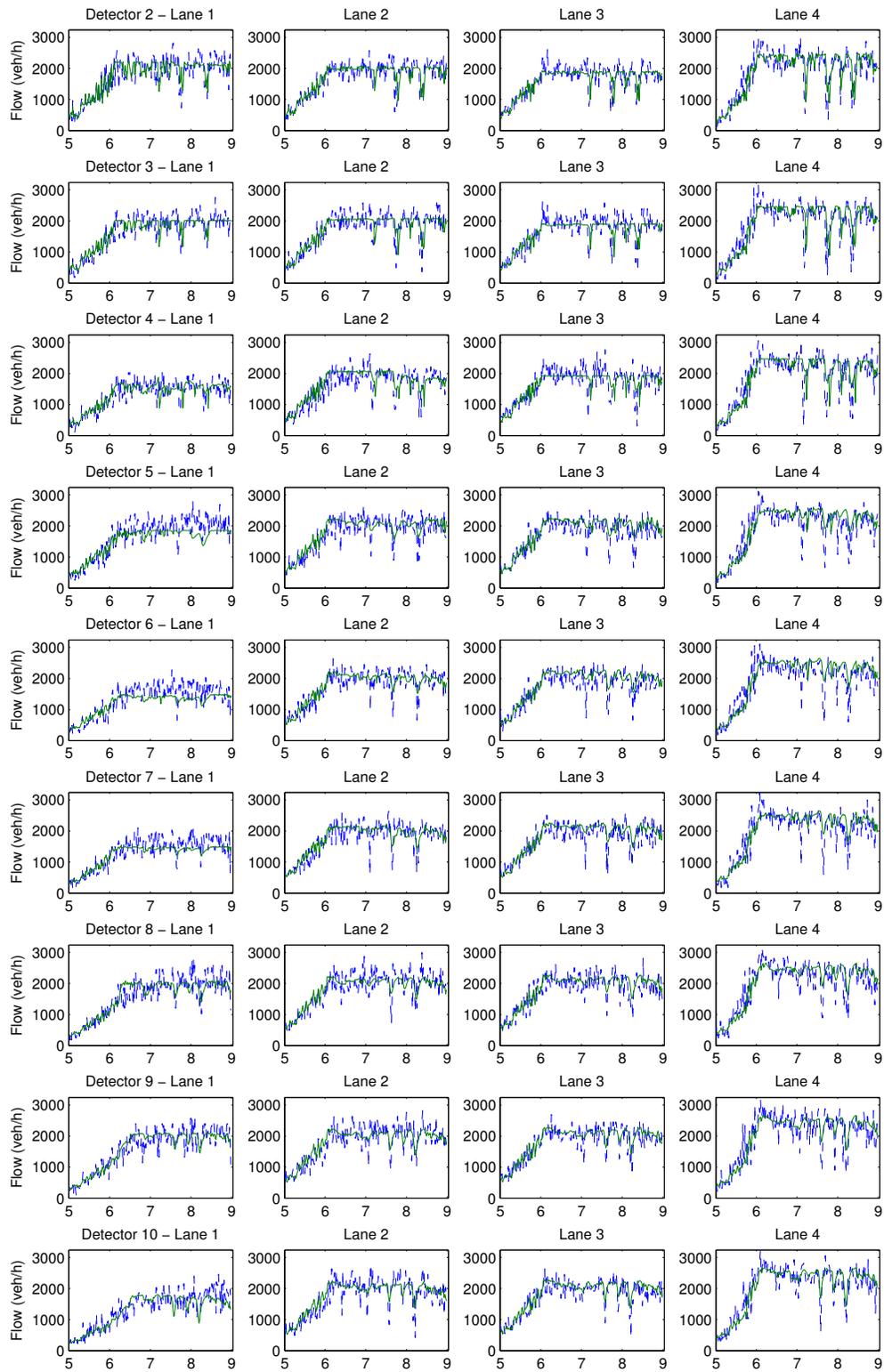


FIGURE 7 Comparison between the real flow data (dashed blue line) and the validation results (solid green line).

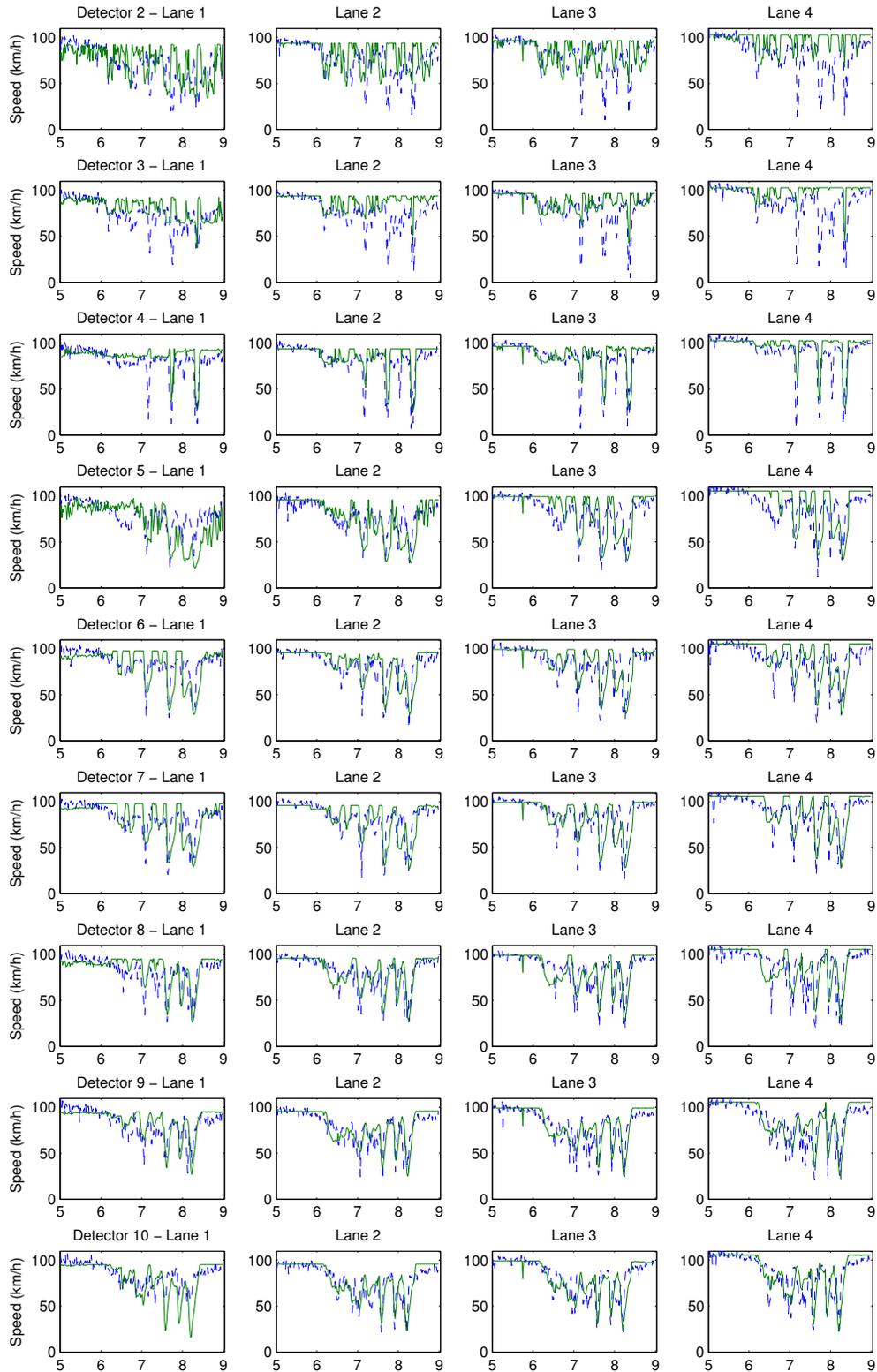


FIGURE 8 Comparison between the real speed data (dashed blue line) and the validation results (solid green line).

CONCLUSIONS

This paper introduced a novel first-order multi-lane macroscopic modelling approach for motorways, with the main purpose of applying it in a model-based optimal control scheme. The model originates from the conservation equation, and methodologies for the computation of lateral and longitudinal flows are proposed; a noteworthy ingredient is the inclusion of the capacity drop phenomenon in the utilised first-order model. In order to test and demonstrate the accuracy of the proposed model, a calibration procedure, based on conventional real data collected by a set of detectors in an urban motorway, has been conducted. Despite the complex causes behind the occurring congestion, the model is capable to replicate the flows and speeds for the analysed scenario with satisfactory accuracy. Validation results, while applying the model with data from different days at the same stretch, confirm the robustness of the modelling approach.

The formulation of a linearly-constrained optimal control problem, directly derived from this model, is proposed by Roncoli et al. (13); the purpose of mitigating traffic congestion is pursued assuming that vehicles are equipped with VACS and that they are able to materialise different control actions, i.e. ramp metering, mainstream traffic flow control, and lane-changing control.

ACKNOWLEDGEMENT

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 321132, project TRAMAN21.

REFERENCES

1. Bishop, R., *Intelligent vehicle technology and trends*. Artech House Publishers, 2005.
2. Treiber, M. and A. Kesting, *Traffic flow dynamics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
3. Gazis, D. C., R. Herman, and G. H. Weiss, Density oscillations between lanes of a multilane highway. *Operations Research*, Vol. 10, 1962, pp. 658–667.
4. Michalopoulos, P. G., D. E. Beskos, and Y. Yamauchi, Multilane traffic flow dynamics: Some macroscopic considerations. *Transportation Research Part B: Methodological*, Vol. 18, No. 4-5, 1984, pp. 377–395.
5. Laval, J. A. and C. F. Daganzo, Lane-changing in traffic streams. *Transportation Research Part B: Methodological*, Vol. 40, No. 3, 2006, pp. 251–264.
6. Jin, W.-L., A multi-commodity Lighthill-Whitham-Richards model of lane-changing traffic flow. *Procedia - Social and Behavioral Sciences*, Vol. 80, 2013, pp. 658–677.
7. Daganzo, C. F., The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, Vol. 28, No. 4, 1994, pp. 269–287.
8. Papageorgiou, M. and A. Messmer, METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering & Control*, Vol. 31, No. 9, 1990, pp. 466–470.
9. Hall, F. L. and L. M. Hall, Capacity and speed-flow analysis of the Queen Elizabeth Way in Ontario. *Transportation Research Record*, Vol. 1287, 1990, pp. 108–118.
10. Leclercq, L., J. A. Laval, and N. Chiabaut, Capacity drops at merges: an endogenous model. *Procedia - Social and Behavioral Sciences*, Vol. 17, 2011, pp. 12–26.
11. Lebacque, J., Two-phase bounded-acceleration traffic flow model: analytical solutions and applications. *Transportation Research Record*, Vol. 1852, No. 1, 2003, pp. 220–230.
12. Srivastava, A. and N. Geroliminis, Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C: Emerging Technologies*, Vol. 30, 2013, pp. 161–177.
13. Roncoli, C., M. Papageorgiou, and I. Papamichail, Optimal control for multi-lane motorways in presence of vehicle automation and communication systems. In *Proceedings of the 19th IFAC World Congress, 2014*, 2014.
14. Lagarias, J. C., J. A. Reeds, M. H. Wright, and P. E. Wright, Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on Optimization*, Vol. 9, No. 1, 1998, pp. 112–147.
15. Spiliopoulou, A., M. Kontorinaki, M. Papageorgiou, and P. Kopelias, Macroscopic traffic flow model validation at congested freeway off-ramp areas. *Transportation Research Part C: Emerging Technologies*, Vol. 41, 2014, pp. 18–29.
16. Cremer, M. and M. Papageorgiou, Parameter identification for a traffic flow model. *Automatica*, Vol. 17, No. 6, 1981, pp. 837–843.
17. Roncoli, C., M. Papageorgiou, and I. Papamichail, Traffic flow optimisation in presence of vehicle automation and communication systems - Part I: A first-order multi-lane model for motorway traffic. *Submitted to: Transportation Research Part C: Emerging Technologies*, 2014.