Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits

Han, Yu; Hegyi, Andreas; Yuan, Yufei; Hoogendoorn, Serge; Papageorgiou, Markos; Roncoli, Claudio

DOI
10.1016/j.trc.2017.02.009

Publication date
2017

Document Version
Accepted author manuscript

Published in
Transportation Research. Part C: Emerging Technologies

Resolving freeway jam waves by discrete first-order model-based predictive control of variable speed limits

Yu Han[a], Andreas Hegyi[a], Yufei Yuan[a], Serge Hoogendoorn[a], Markos Papageorgiou[b], Claudio Roncoli[c]

[a]*Department of Transport and Planning, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands.*
[b]*Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania 73100, Greece.*
[c]*Department of Built Environment, School of Engineering, Aalto University, 02150 Espoo, Finland*

Abstract

In this paper we develop a fast model predictive control (MPC) approach for variable speed limit coordination to resolve freeway jam waves. Existing MPC approaches that are based on the second-order traffic flow models suffer from high computation load due to the non-linear and non-convex optimization formulation. In recent years, simplified MPC approaches which are based on discrete first-order traffic flow models have attracted more and more attention because they are beneficial for real-time applications. In literature, the type of traffic jam resolved by these approaches is limited to the standing queue in which the jam head is fixed at the bottleneck. Another type of traffic jam known as the jam wave, has been neglected by the discrete first-order model-based MPC approaches. To fill this gap, we develop a fast MPC approach based on a more accurate discrete first-order model. The model keeps the linear property of the classical discrete first-order model, meanwhile takes traffic flow features of jam waves propagation into consideration. A classical non-linear MPC and a recently proposed linear MPC are compared with the proposed MPC in terms of computation speed and jam wave resolution by a benchmark problem. Simulation results show that the proposed MPC resolves the jam wave with a real-time feasible computation speed.

Keywords: Model predictive control, variable speed limit, capacity drop, linear model, jam wave

1. Introduction

Due to increasing demand, traffic jams happen frequently at current freeway networks. Traffic jams generate many negative effects such as reducing freeway capacity, inducing travel delays and potentially unsafe traffic conditions. Infrastructure expansion is an effective way to combat traffic jams. However, adding additional lanes and new freeways are not always viable due to economical or environmental concerns. Instead, traffic control measures are playing more and more important roles in improving traffic operation efficiency.

To implement effective traffic control measures to resolve traffic jams, the features of traffic jams have to be investigated. In general, two types of traffic jams on freeways can be identified. As has been presented in (Hegyi et al., 2008), traffic jams with the head fixed at the bottleneck are known as standing queues, and jams that have an upstream moving head and tail are known as jam waves (also known as wide moving jams in some studies, e.g., (Kerner and Rehborn, 1996)). A standing queue can be formed at a busy on-ramp, a lane drop bottleneck, or any kind of infrastructural bottleneck. According to empirical studies in literature (Cassidy and Bertini, 1999; Chung et al., 2007; Srivastava and Geroliminis, 2013), the outflow of a standing queue is typically 10-15 percent lower than the free flow capacity, which is known as the capacity drop. Jam waves usually originate from breakdowns, and in most cases from standing queues. Possible triggers for the formation of a jam wave could be a vehicle suddenly braking or a lane-changing vehicle in a high-demand traffic flow situation. Therefore, since so many (partially unobservable) factors influence the formation process, the formation of a jam wave is considered stochastic and difficult to be reproduced. However, from different empirical studies, some common patterns of its propagation can be distilled. For example,

Email address: y.han-1@tudelft.nl (Yu Han)

the propagation speed of jam waves is relatively constant, typically between 15-20 km/h (Kerner and Rehborn, 1996). The queue discharge rate of a jam wave is typically around 30 percent lower than the free flow capacity (Schönhof and Helbing, 2007). It can propagate for a long time period and distance, and resolves only with the traffic demand decreasing (Kerner and Rehborn, 1996). Jam waves are considered more difficult to reproduce than standing queues, because the bottleneck of jam waves moves dynamically rather than fixed at specific locations.

The common feature of standing queues and jam waves is that both of them are associated with a capacity drop. Thus, the aim of many freeway traffic control methods are often to prevent the activation of bottlenecks, so as to avoid the occurrence of the capacity drop and maximize the outflow of the bottleneck. Variable speed limit (VSL) and ramp metering are commonly used control measures on freeway networks to regulate the flow and delay the onset of congestion. Ramp metering regulates the inflow from the on-ramp to the mainstream of the freeway to delay or prevent the activation of the bottleneck. The performance of ramp metering control is limited by available storage space at on-ramps (Carlson et al., 2011). Normally VSLs do not suffer from the storage capacity restriction since the mainstream of freeways are wider and longer than on-ramps. VSLs may be used to reduce the mainstream flow, so as to resolve traffic jams or delay or prevent the activation of bottlenecks. There are two mechanisms of VSLs in reducing mainstream flows that are explained as follows. First, when lower VSLs are applied on a freeway stretch at under-critical densities, there is a transition to a new traffic state, which serves the same flow at lower speed and high density. During this transition, the density increase leads to a temporary reduction of the mainstream flow (Carlson et al., 2010). Second, sub-critical VSLs lead to a fundamental diagram that has lower capacity with lower VSLs.

There are various theories and algorithms to determine the approprate values of the VSLs. SPECIALIST is an analytical approach for VSL control to resolve jam waves based on the shock wave theory (Hegyi et al., 2008). While the approach was successfully tested in practice (Hegyi and Hoogendoorn, 2010), its disadvantage is that due to its feed-forward structure, disturbances during activation of the VSLs are not handled. Later on, the algorithm has been extended to a feedback structure, which is known as COSCAL v2 (Mahajan et al., 2015), to overcome the deficiency of the SPECIALIST algorithm. Analytical approaches are usually efficient in computation and easy to be implemented. However, they are not easy to be adapted to new situations (e.g., changes in infrastructure). Traffic model predictive control (MPC) is another control approach which has attracted the attention of many researchers. It predicts the evolution of traffic dynamics and calculates the optimal control scheme for the time period in which the relevant traffic dynamics occurs. This feature enables the controller to take advantage of potentially larger future gains at a current (smaller) cost, so as to avoid a myopic control actions. The coordination of VSL by MPC to suppress jam waves has been investigated by (Hegyi et al., 2005b), where the design is based on a non-linear second-order model METANET (Kotsialos et al., 2002). Another advantage of MPC is that various of control measures can be easily integrated into one control system. Consequently, VSL control is often integrated with other control measures such as ramp metering by MPC (Hegyi et al., 2005a; Carlson et al., 2010; Roncoli et al., 2015b).

Optimal control formulations based on the discrete macroscopic second-order model, METANET, have become a popular choice since the model has been calibrated and validated with a reasonable accuracy. However, the non-linear and non-convex optimization formulation of METANET-based MPC might result in high computation load, especially if the optimization is solved by the standard SQP algorithm (Hegyi et al., 2005b). Researchers have managed to solve this problem by developing more efficient computational algorithms, such as the feasible-direction algorithm (Kotsialos et al., 1999a), the parameterization algorithm (Zegeye et al., 2012; van de Weg et al., 2015). A noticeable problem for METANET-based MPC of VSL is that the solution of the optimization may depend on the selection of the initial guess trajectory, in the sense that if the initial guess is not appropriately chosen, the solution of the optimization might get stuck at a local minimum. In recent years, linear or quadratic optimizations for traffic control have attracted more and more attention. The origin of this type of approach dates back to (Gazis and Potts, 1963), who suggested a store-and-forward modeling approach to traffic control problems. Following this line, Papageorgiou (1995) proposed a linear optimization approach for integrated traffic control based on a store-and-forward model. Later on, Lo (1999); Ziliaskopoulos (2000) embedded the cell-transmission model into a linear optimization problem for dynamic traffic assignment purpose. Since the capacity drop was not considered, the only gain of the controller was coming from increasing the off-ramp flows. Recent research has incorporated the capacity drop into linear MPC for VSL and ramp metering control (Muralidharan and Horowitz, 2015; Roncoli et al., 2015b). The controllers have been demonstrated to be efficient in terms of computation speed and reducing total travel delay by the designed case studies. However, the type of traffic jam considered in their case studies is only the standing queue.

As far as we are concerned, there is no specific research in literature which aimed to resolve freeway jam waves

through a linear MPC formulation. To fill this gap, in this paper we develop a linear-quadratic MPC based on an extended discrete first-order traffic flow model. The extended model keeps the linear property of the classical discrete first-order model, meanwhile takes traffic flow features of jam waves propagation into consideration. The performance of the proposed controller is compared with a classical non-linear MPC and one recently proposed linear MPC in terms of resolving jam waves and the computation speed for a benchmark problem.

The rest of this paper is set up as follows. Section 2 describes the mechanism of the extended model and the models for comparison. Section 3 elaborates on the formulation of the MPC approach. In Section 4 the presented MPC approach is compared with other MPC approaches for a benchmark problem, in terms of the capability of resolving jam waves, the resulting total travel time, and the required computation time. Section 5 presents the conclusions and some topics for future research.

2. Model description

In this section, the mechanism of several discrete first-order models in literature are briefly described. The classical cell-transmission model is described in Section 2.1. Then some extended cell-transmission models which have incorporated the capacity drop are summarized in Section 2.2. The new extended CTM and the VSL assumption are described in Section 2.3.

2.1. The cell-transmission model

The origin of discrete first-order models is the well-known cell-transmission model (CTM) developed by (Daganzo, 1994), which is consistent with the hydrodynamic theory of LWR model (Lighthill and Whitham, 1956; Richards, 1956). The roadway stretch is divided into cells, which are increasingly numbered from upstream to downstream, as shown in Fig. 1. In the following, cell i is referred to as an upstream cell, and cell $i + 1$ is interpreted as a downstream cell. Traffic states of the roadway stretch are represented by the densities of cells at discrete time steps indexed by k. The density of cell i at time step $k + 1$, which is represented as $\rho_i(k + 1)$ [veh/km], is updated as:

$$\rho_i\,(k+1) = \rho_i\,(k) + \frac{T}{L_i}[f_{i-1}(k) - f_i(k)], \tag{1}$$

where, T is the duration of the discrete time steps. L_i is the length of cell i. f_{i-1} and f_i [veh/h] are the downstream boundary flows of cell $i - 1$ and cell i.

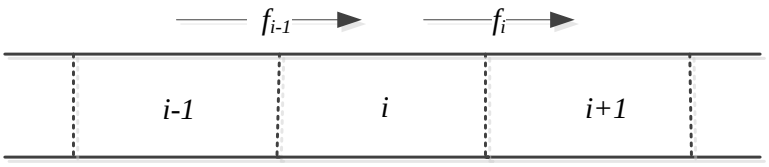

Figure 1: A freeway stretch divided into discrete cells.

Lebacque (1996) showed that CTM is the first-order discrete Godunov approximation of the LWR model. According to the Godunov scheme, the flow $f_i(k)$ between cell i and cell $i + 1$ is determined by the minimum of the sending flow (represented by $S_i(k)$) and the receiving flow (represented by $R_i(k)$), also known as the demand of sending cell and the supply of receiving cell. $f_i(k)$ is represented as,

$$f_i(k) = \min\left(S_i(k), R_{i+1}(k)\right), \tag{2}$$

The demand of sending cell, $S_i(k)$, is represented as:

$$S_i(k) = \min\left(\rho_i(k)v, c\right), \tag{3}$$

which states that the outflow of cell i is restricted by the vehicles present in the sending cell and the capacity of the sending cell. The supply of receiving cell, $R_{i+1}(k)$, is represented as:

$$R_{i+1}(k) = \min\left(c, \beta(\rho^J - \rho_{i+1}(k))\right) \tag{4}$$

which states that the inflow to cell $i + 1$ is restricted by capacity of the receiving cell and the amount of space in the receiving cell. The CTM assumes a triangular-shaped fundamental diagram (FD) for the above equations, which is characterized by the free flow speed v, the capacity c, the congestion wave speed β, and the jam density ρ^J. The FD and the demand and supply functions are depicted in Fig. 2. ρ^{cr} denotes the critical density, which separates the FD into free-flow branch and congestion branch.

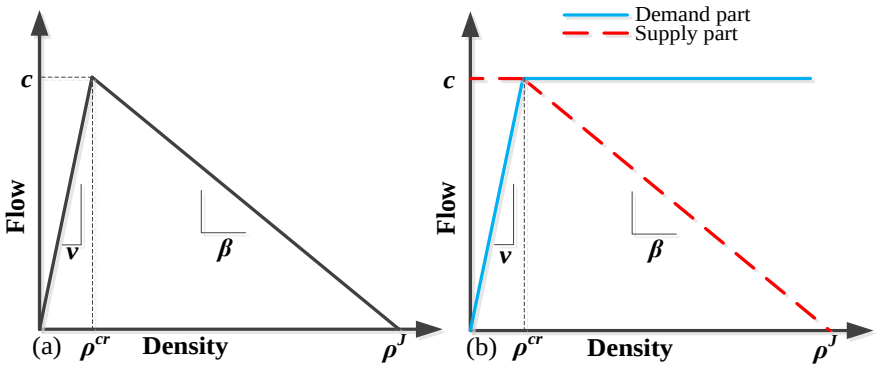

Figure 2: The triangular-shaped FD of the CTM (a), and the demand and the supply parts (b).

2.2. *Extended versions of the CTM*

It is commonly known that the classical CTM does not take the capacity drop phenomenon into consideration. Thus, many researchers have presented extensions of the CTM to incorporate the capacity drop. Most of those approaches try to incorporate the capacity drop by changing the demand function, which is first proposed by (Lebacque, 2003). Schreiter et al. (2010) assumed two constant capacity values for the cell-transmission model, one for free-flow condition and the other for congested condition. Srivastava and Geroliminis (2013) extended the LWR model by defining a FD with two capacity values. A memory-based approach was proposed to choose the appropriate capacity value, so as to distinguish congested and uncongested states. However due to non-linearities, these models have not been used for model-based traffic control. Muralidharan (2012) incorporated the capacity drop into a link-node cell-transmission model, which used a discontinuous fundamental diagram. The capacity drop was assumed to be triggered only at the bottleneck cell, and with a constant value. The model has been developed into a mixed-integer linear programming for network traffic optimization. However, since bottlenecks have to be predefined at fixed locations, this approach is not suitable for the jam wave type of congestion under which bottlenecks move dynamically. Roncoli et al. (2015a) proposed an optimization-oriented first-order model, which also represented the capacity drop by a reduced demand function. They assumed that the extent of the demand reduction of a cell has a linear relation with the vehicle density of the same cell. When the density of a cell is beyond the critical value, the demand decreases linearly with increasing density. Srivastava et al. (2014) used a similar assumption for urban traffic to capture the headway behavior. The empirical study by (Yuan et al., 2015) has shown that the magnitude of capacity drop has a negative relation with the speed in congestion. Thus, the state-dependent capacity drop assumption by (Roncoli et al., 2015a) is considered to be more reasonable than the assumptions with a constant extent of capacity drop. Han et al. (2015) used a similar capacity drop assumption for lane drop bottlenecks. In that study, the demand reduction which represented the capacity drop, was assumed to take place at the downstream cell of the cell with over-critical density, rather than in the cell with over-critical itself.

The model from (Roncoli et al., 2015a) has considered both lateral flows and longitudinal flows. Stronger capacity drop maybe reproduced when there are lateral flows due to lane changes or on-ramp traffic. In this paper we do not take lane changing into consideration since the focus is on jam waves for which the lateral flows do not play an important role. In the following we only introduce their assumptions on longitudinal flows. In their model, the capacity drop was characterized by variable c^{jam}, which represented the flow discharge rate of a cell in jam state. As shown in Fig. 3 (a), the demand reduces gradually as the density increases at the over-critical part. The supply remains the same as the original cell-transmission model, which is shown in Fig. 3 (b).

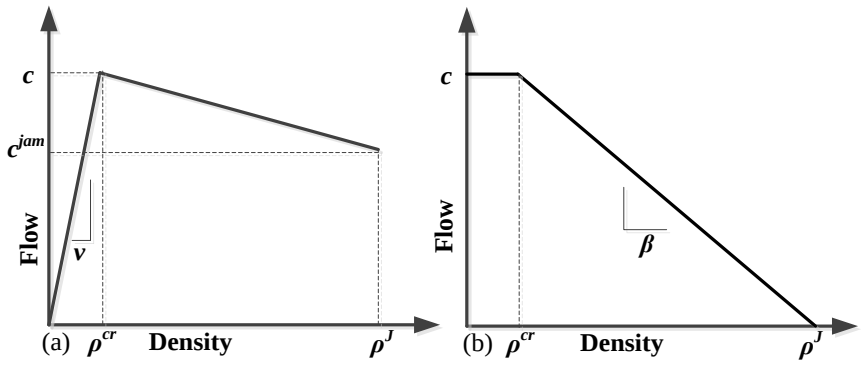

Figure 3: The demand (a) and supply (b) parts of the FD of the model from (Roncoli et al., 2015a).

It is found that the type of traffic jam considered by most of the above mentioned models is only the standing queue. In (Han et al., 2016), it has been tested that most of the above mentioned models can not reproduce the propagation of jam waves accurately. Thus, we have proposed a new extension of the discrete first-order model to reproduce the propagation of jam waves more accurately, which will be briefly introduced in the next section.

2.3. The new extended CTM

As has been discussed in (Han et al., 2016), the cell-transmission model and many of its variants have difficulties in reproducing the propagation of jam waves accurately. One reason is that capacity drop has not been adequately incorporated. The capacity drop assumption that we propose is a linear reduction of the outflow of cell $i+1$ as a function of the density of cell i in case the densities are over-critical. For simplicity, in this paper we assume that all of the cells have same properties. However, note that it is possible to make the model parameters cell-dependent. For a general presentation of the model, readers are referred to (Han et al., 2016). Without loss of generality, the index of cell number remains in the following equations. According on the capacity drop assumption, the maximum outflow of cell $i+1$ can be represented as:

$$Q_{i+1}(k) = \min\left\{c_{i+1}, c_{i+1}(1-\alpha\frac{\rho_i(k)-\rho_i^{cr}}{\rho_i^J-\rho_i^{cr}})\right\}, \tag{5}$$

where, $Q_{i+1}(k)$ is the maximum outflow of cell $i+1$ at time step k. α denotes the maximum extent of the capacity drop. The capacity drop model is depicted by the sub-figures (a,c,e,g) of Fig. 5. In this way, the capacity drop can not only be reproduced at on-ramp bottlenecks, but also other infrastructural bottlenecks such as lane drops (Han et al., 2015). When cell i is in an over-critical state and cell $i+1$ is in an under-critical state, $Q_{i+1}(k)$ represents the flow discharge rate of cell $i+1$, which is lower than the free-flow capacity c_{i+1}. Accordingly, the critical density of cell $i+1$ under such a condition, represented by $\rho_{i+1}^{cr-dch}(k)$, is lower than ρ_i^{cr} because the free-flow speed does not change. $\rho_{i+1}^{cr-dch}(k)$ is calculated as:

$$\rho_{i+1}^{cr-dch}(k) = \frac{Q_{i+1}(k)}{v_{i+1}}. \tag{6}$$

Thus, the demand of the proposed model is represented as:

$$S_{i+1}(k) = \begin{cases} \rho_{i+1}(k)v_{i+1} & \text{if } \rho_{i+1}(k) < \rho_{i+1}^{cr-dch}(k) \\ Q_{i+1}(k) & \text{if } \rho_{i+1}(k) \geq \rho_{i+1}^{cr-dch}(k). \end{cases} \tag{7}$$

The other aspect that plays a role in the proper reproduction of jam wave propagation is the supply function when a cell is in a discharging state. According to the cell-transmission model, the supply function (4) is dominating in

determining the flow if both the sending cell and the receiving cell have over-critical densities. In other words, the boundary flow between two over-critical cells only depends on the average density of the receiving cell. However in practice, traffic states at the over-critical part can be inhomogeneous if a cell has a long length (typically several hundred meters when applied to traffic control). For example, if the jam head is in the middle of cell $i+1$, then cell $i+1$ is in such a state that the back part is congested while the front part is not, which is shown in Fig. 4. We define such a state as the discharging state. Note that the discharging state may be avoided by making the cell size small. However, the proposed model is applied to a linear optimization problem for traffic control purpose. If a smaller cell length is used, the linear optimization framework results in higher number of cells, control variables, and constraints. This increases the computation time and may render the control problem too complex to be solved in real-time. From this point of view, our way of modelling gives good trade-off between modelling accuracy and controller performance (in terms of computation speed).

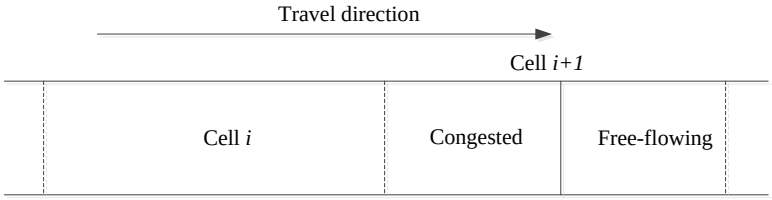

Figure 4: An example to explain the discharging state. Cell $i+1$ is in the discharging state.

Suppose traffic states at the downstream of cell $i+1$ are free-flow states. The discharging state of cell $i+1$ is captured by the condition $\rho_i(k) > \rho_{i+1}(k) > \rho_{i+1}^{cr-dch}$ in the proposed model. The inequality is explained as follows. According to (5) and (6), if $\rho_i(k) > \rho_{i+1}^{cr-dch}$, then $\rho_i(k) > \rho_i^{cr}$. Thus, cell i is in an over-critical state. If the density of cell $i+1$ is lower than the density of cell i, but higher than the critical discharging density, ρ_{i+1}^{cr-dch}, then cell $i+1$ should be in a mixed state with a congested upstream part and free-flow downstream part. According to the definition in Fig. 4, cell $i+1$ is in the discharging state. Under the discharging state, it is assumed that the supply of cell $i+1$ is overestimated by (4) because the average density $\rho_{i+1}(k)$ is lower than the density upstream part of cell $i+1$. Empirical analysis in (Han et al., 2016) has confirmed that this assumption is reasonable. Thus, we assume two supply functions $R_{i+1}^{(1)}(k)$ and $R_{i+1}^{(2)}(k)$, which apply for the discharging state and the non-discharging state respectively when cell $i+1$ is over-critical. The non-discharging supply function is represented as,

$$R_{i+1}^{(1)}(k) = \beta_{i+1}^{(1)}(\rho_{i+1}^{J} - \rho_{i+1}(k)), \quad \text{if } \rho_i(k) \le \rho_{i+1}(k) < \rho_{i+1}^{J}, \tag{8}$$

where, $\beta_{i+1}^{(1)}$ represents the slope of the line which connects the critical state and the jam state. It expresses the congested states of traffic in a jam or getting in a jam. When cell $i+1$ is in the discharging state, the supply from (8) would overestimate the real value. It was also shown in (Han et al., 2016) that the extent of the supply overestimation has a positive relation with the density difference between the sending cell and the receiving cell. To keep the linear property of the model, we assume a linear relation between the density difference $\rho_i(k) - \rho_{i+1}(k)$ and the supply reduction $R_{i+1}^{(1)}(k) - R_{i+1}^{(2)}(k)$. The discharging supply function is represented as,

$$R_{i+1}^{(2)}(k) = R_{i+1}^{(1)}(k) - (\beta_{i+1}^{(1)} - \beta_{i+1}^{(2)}) \cdot (\rho_i(k) - \rho_{i+1}(k)), \quad \text{if } \rho_{i+1}^{cr-dch}(k) \le \rho_{i+1}(k) < \rho_i(k). \tag{9}$$

where, $\beta_{i+1}^{(2)}$ represents the slope of the line which connects the state of a discharging cells and the state of its upstream cell. It expresses the congested states of traffic leaving a jam wave, which in free-flow will lead to the queue discharge rate that is lower than the free-flow capacity. When traffic of cell $i+1$ is getting in a jam, the traffic state lies on the line with the slope $\beta_{i+1}^{(1)}$. When traffic in cell $i+1$ is leaving a jam, the traffic state lies on the line with the slope $\beta_{i+1}^{(2)}$. Thus, the line of $\beta_{i+1}^{(1)}$ represents the deceleration of traffic whereas $\beta_{i+1}^{(2)}$ represents the acceleration of traffic. In real traffic, deceleration is usually faster than acceleration, thus $\beta_{i+1}^{(1)}$ is higher than $\beta_{i+1}^{(2)}$. The overall supply function

of cell $i+1$, $R_{i+1}(k)$ is formulated as:

$$R_{i+1}(k) = \begin{cases} Q_{i+1}(k) & \text{if } \rho_{i+1}(k) < \rho_{i+1}^{cr-dch}(k) \\ R_{i+1}^{(2)}(k) & \text{if } \rho_{i+1}^{cr-dch}(k) \le \rho_{i+1}(k) < \rho_i(k) \\ R_{i+1}^{(1)}(k) & \text{if } \rho_i(k) \le \rho_{i+1}(k) \le \rho_{i+1}^{J}. \end{cases} \tag{10}$$

The above equations can be further explained by considering the following three conditions:

1. $\rho_i(k) \le \rho_i^{cr}$. Under this condition, neither capacity drop nor supply reduction happens for cell $i+1$. Thus $Q_{i+1}(k) = c_{i+1}(k)$, and the demand and supply of cell $i+1$ are the same as which from the CTM, which are shown as sub-figure (b) of Fig. 5.
2. $\rho_i(k) = \rho_i^J$. Cell i is in the jam state. Thus, the demand of cell $i+1$ is represented by the blue line of sub-figure (d) in Fig. 5, which has a maximum flow $Q_{i+1}(k) = c_{i+1}(k)(1-\alpha)$ and a critical density $\rho_{i+1}^{cr-dch}(k) = c_{i+1}(k)(1-\alpha)/v_{i+1}$. Under this condition, ρ_i can not be smaller than ρ_{i+1}, so the supply of cell $i+1$ at the over-critical part is represented by $R_{i+1}^{(2)}$, which equals to $\beta_{i+1}^{(2)}\left(\rho_i^J - \rho_{i+1}(k)\right)$ according to the third equation of (10). The supply of cell $i+1$ at the under-critical part equals to F_{i+1}, and the overall supply of cell $i+1$ is represented by the red line of sub-figure (d) in Fig. 5.
3. $\rho_i^{cr} < \rho_i(k) < \rho_i^J$. The demand of cell $i+1$ is calculated by (6) and (7), and represented by the blue line in Fig. 5 (f). Under this condition, the supply at the over-critical part is represented by $R_{i+1}^{(1)}(k)$ when $\rho_i(k) \le \rho_{i+1}(k)$, and is represented by $R_{i+1}^{(2)}(k)$ when $\rho_i(k) > \rho_{i+1}(k)$. The line of $R_{i+1}^{(1)}(k)$ and $R_{i+1}^{(2)}(k)$ intersect at $\rho_i(k) = \rho_{i+1}(k)$. The supply of cell $i+1$ at the under-critical part equals to $Q_{i+1}(k)$, and the overall supply of cell $i+1$ is represented by the red line in Fig. 5 (f).

Variable speed limits (VSLs) are used as the control measure in this paper. Thus, the effect of VSLs have to be taken into account by the model. The influence of VSL on the shape of FD was investigated by (Papageorgiou et al., 2008). Some findings of the reported investigations can be summarized as: (i) Speed limits decrease the slope of the flow-occupancy diagram at under-critical occupancies. (ii) The VSL-affected flow-occupancy curve crosses the no-VSL curve, shifting the critical occupancy to higher values in the flow-occupancy diagram. (iii) There is no conclusive result whether VSL will increase the flow capacity or not, since a slight increase is only visible at some locations. Since the effect of VSL on aggregated traffic flow features is complex, simplified VSL model assumptions were often made. Hegyi et al. (2005a) assumed the VSL impact was to replace the left part of the flow-occupancy curve by a straight line with slope corresponding to the displayed VSL. Carlson et al. (2010) incorporated a VSL model into the METANET model based on the empirical findings of (Papageorgiou et al., 2008).

For VSL models incorporated in discrete first-order models, the approach is often that the demand function is modified while the supply function is left unchanged (Muralidharan, 2012; Hadiuzzaman and Qiu, 2013). The VSL value in this paper is, however, derived from the optimization outcome. Specifically, if the optimizer determines a flow reduction and the flow is determined by demand constraints, then the sending cell is under VSL control and the corresponding VSL value is calculated based on the vehicle density.

A VSL model can be derived from the optimization. As shown in Fig. 5 (g-h), if a VSL signal is applied to cell $i+1$ with the value $V_{i+1}(k)$, the demand of the cell is the minimum of $V_{i+1}(k)\rho_{i+1}(k)$ and $Q_{i+1}(k)$, which is represented by the blue line in Fig. 5 (h). The new extended CTM incorporated with VSLs can be used for model calibration and validation.

3. The optimal control formulation

This section shows the MPC formulation based on the model proposed in the previous section. MPC-based controllers predict traffic flow evolutions by traffic flow models, and determine the optimal control signals based on current system states. The length of prediction horizon is represented by N_p, and the length of control horizon is represented as N_c. The duration of a control time step is represented as T_c. After optimization the first sample of the optimal control signals is applied to the process. The remaining part of the control signal is recalculated in the rolling horizon scheme. For a detailed description of MPC approach, readers are referred to (Camacho and Bordons, 2012). MPC-based controllers are formulated into linear or non-linear optimizations, depending on the model properties.

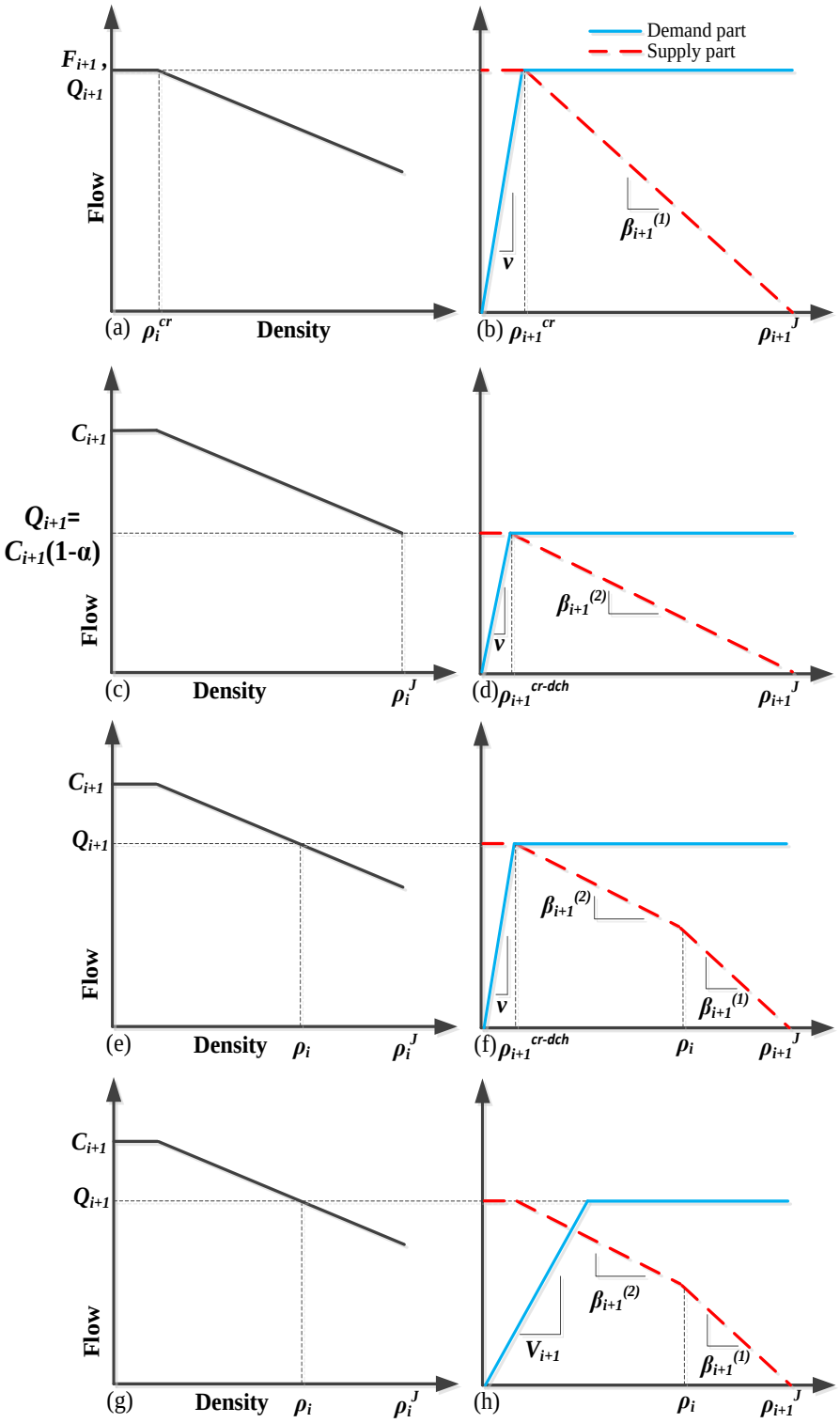

Figure 5: The left column, figures (a,c,e,g) show the relation between $\rho_i(k)$ and $Q_{i+1}(k)$, which represents the capacity drop assumption. The right column shows the demand and supply of cell $i+1$, (b) for condition $\rho_i(k) \leq \rho_i^{cr}$, (d) for condition $\rho_i(k) = \rho_i^J$, (f) for condition $\rho_i^{cr} < \rho_i(k) < \rho_i^J$, (h) for condition $\rho_i^{cr} < \rho_i(k) < \rho_i^J$ and cell $i+1$ is under VSL control with the value $V_{i+1}(k)$.

The MPC approaches that are formulated into linear optimization problems can be solved very efficiently (Roncoli et al., 2015b). The proposed model has a linear property, which will be described in Section 3.1. The MPC controller in this paper is formulated as a linear-quadratic optimization problem. The objective function of the proposed MPC is presented in Section 3.2.

3.1. The linear property of the model

In literature, extended cell-transmission models have been reformulated as linear optimization problems for traffic control (Gomes and Horowitz, 2006; Muralidharan and Horowitz, 2015; Roncoli et al., 2015b). The most important feature of this model type is the linear property. Traffic states can be predicted by the following linear equation,

$$\rho_i(k+n) = \rho_i(k) + \sum_{j=k}^{k+n-1} \frac{T}{L_i}[f_{i-1}(j) - f_i(j)], \tag{11}$$

where, flows between cells, $f_i(j)$, $j = k : k+n-1$, are decision variables of the model. The calculation of $f_i(j)$ are delegated to the controller, and only upper bounds are specified to $f_i(j)$ as follows:

$$f_i(j) \leq \rho_i(j) v_i(j) \tag{12}$$

$$f_i(j) \leq c_i \left(1 - \alpha \frac{\rho_{i-1}(j) - \rho_{i-1}^{cr}}{\rho_{i-1}^{J} - \rho_{i-1}^{cr}}\right) \tag{13}$$

$$f_i(j) \leq c_{i+1} \left(1 - \alpha \frac{\rho_i(j) - \rho_i^{cr}}{\rho_i^{J} - \rho_i^{cr}}\right) \tag{14}$$

$$f_i(j) \leq \beta_{i+1}^{(1)} \left(\rho_{i+1}^{J} - \rho_{i+1}(j)\right) \tag{15}$$

$$f_i(j) \leq \beta_{i+1}^{(1)} \left(\rho_i^{J} - \rho_i(j)\right) + \beta_{i+1}^{(2)} \left(\rho_i(j) - \rho_{i+1}(j)\right). \tag{16}$$

The modeling approach for flows is based on the demand and supply functions in Equation (7) and (10). More specifically, equation (12) and (13) correspond to two terms of the demand function (7) and equation (14), (15), (16) correspond to three terms of the supply function (10). Thus, the lines of the demand and the supply function in Fig. 5 (b), (d), (f) represent the upper bounds of $f_i(j)$, corresponding to different values of $\rho_i(j)$. In case no control actions are applied, the actual flow equals the minimum between the demand and supply flows and the solution of $f_i(j)$ lies in the lines of the upper bounds. If the solution of $f_i(j)$ is lower than the minimum of the demand and supply flows, then cell i at time step j is under VSL control. The displayed VSL value, $V_i(j)$, is calculated by:

$$V_i(j) = \frac{f_i(j)}{\rho_i(j)}. \tag{17}$$

The above equations show the linear property of the model, which enables the formulation of a linear optimization problem. Thus the state variables are calculated in one step for the whole prediction horizon. In the formulation of non-linear optimizations, the state variables have to be calculated sequentially and iteratively, which might result in a higher computation load.

3.2. Objective function

The objective function of a controller needs to reflect the level of congestion in the freeway. Total travel time (TTT) captures the aggregate performance of traffic conditions on all users in the freeway. Minimizing TTT is equivalent to minimize the total number of vehicles in the entire network. The number of vehicles in the network is represented by a vector $X(k)$, where $X(k) = [L_1\rho_1(k), L_2\rho_2(k), \cdots L_I\rho_I(k)]'$. The evolution of $X(k)$ is represented by a linear matrix formulation:

$$X(k+n) = X(k) + \sum_{j=k}^{k+n-1} B * U(j) + \sum_{j=k}^{k+n-1} d(j), \tag{18}$$

where, $U(j) = [f_1(j), f_2(j), \cdots f_I(j)]'$ is a decision vector. B is a matrix that contains the freeway structure information. The non-zero elements in matrix B are the duration of the discrete time step. $d(j)$ is the traffic demand vector. To approximate the minimization of the TTT, the objective of the proposed controller is to minimize a quadratic function of the number of vehicles in the network, which is shown as:

$$\min \sum_{j=k}^{k+N_p-1} [X(j)'PX(j) + GU(j)], \tag{19}$$

where, P is a matrix with all 1 elements, so the term $X(j)'PX(j)$ aims to minimize the quadratic function of the number of vehicles in the network in the prediction horizon. G is a vector of the cell lengths multiplied by a small negative value. The term $GU(j)$ aims to maximize the flows, so as to address the so-called holding back problem. Holding back may occur for a congested state, where the VSLs cannot have an influence. The second term in (19) (which was initially suggested in (Papageorgiou, 1995)) mitigates this undesirable occurrence by rewarding the solutions where the flows take higher values. Note that this second term in (19) has a physical meaning, as it is proportional to the total travel distance (TTD). Thus, (19) may be perceived as a weighted combination of TTT minimization and TTD maximization. A more detailed explanation of this objective function is in (Le et al., 2013). In that paper, the linear quadratic objective function is applied to an urban signal control problem. The objective function (19) under the constraints in Section 3.1 can be solved by an appropriate solver (e.g., CPLEX). The outputs of the controller are optimal flows between cells.

3.3. The minimum VSL constraint

Typical real-world speed limit systems have some pre-defined minimum speed that can be displayed. The optimization variables of the proposed controller are flows, thus the minimum VSL constraints cannot be applied directly, instead minimum flow constraints are applied to achieve the same purpose. However, minimum flow constraints might cause infeasibility due to conflicting constraints. For example, the flows prediction made at time step k is denoted by $f_i(j|k)$ (the same meaning as $f_i(j)$ in (12-16)). If the minimum flow constraints restrict $f_i(j|k) \geq A_i(j|k)$, but the demand and supply constraints restrict $f_i(j|k) \leq B_i(j|k)$, then the optimization problem will not have any solution if $A_i(j|k) > B_i(j|k)$.

To overcome this problem, here we introduce a method to apply the minimum flow constraints based on a forward prediction. The basic idea is to find possibly conflicting flows by predicting future traffic states, then exclude them from the minimum flow constraints. Therefore, before running the optimization at each control step, the traffic states are predicted by the prediction model of the controller. The flows from the forward prediction are denoted as $\widehat{f_i}(j|k)$. The following equation is used as the minimum flow constraints for the optimization:

$$f_i(j|k) \geq A_i(j|k), \qquad \text{if } f_i(j|k) \geq A_i(j|k) \tag{20}$$

$$f_i(j|k) \geq 0, \qquad \text{if } f_i(j|k) < A_i(j|k). \tag{21}$$

The values of $A_i(j|k)$ are set based on the predicted density of every cell at every time step. For any cell at which VSLs can be applied, $A_i(j|k)$ are set to the value that $A_i(j|k) = \text{VSL}_{min} \times \rho_i(j|k)$. It can be guaranteed that the optimization problem with the above constraint has a solution, because $\widehat{f_i}(j|k)$ is a feasible solution of the optimization problem with the above constraints. Note, however, that in this way, it cannot be ensured that every $f_i(j|k)$ under VSL control is restricted by the minimum VSL constraints. For flows in the prediction that are lower than $A_i(j|k)$, the minimum flow constraint cannot be guaranteed. This can be perceived as a shortcoming of linear optimizations. But still, the proposed method provides a way to realize the minimum flow constraints to some extent.

4. Benchmark problem

A benchmark problem is set up to test the performance of the proposed controller. A classical non-linear MPC and one recently proposed linear MPC, are compared with the proposed MPC in terms of resolving jam waves and the computation speed. The benchmark network we choose is a simple homogeneous freeway stretch, which is sufficient to test the relevant contributions we present in this paper. The 6km freeway stretch contains three lanes, and it is divided into 20 cells, which is depicted in Fig. 6. Note that in practice, deceleration areas are needed for VSLs for safety reasons. Deceleration areas for VSLs are not considered in this paper because the transition period is much shorter than the normal control period. For example, the SPECIALIST theory also did not consider deceleration areas (Hegyi et al., 2008). However in practice, the SPECIALIST field test imposed a very short deceleration period for safety. The field test results show that the deceleration did not have much effect (Hegyi and Hoogendoorn, 2010). To reproduce a jam wave, the second-order model METANET is used as the process model to represent the real world. The process of reproducing the jam wave is presented in Section 4.1. The prediction models of the proposed controller and the linear MPC for comparison are calibrated with the simulation results of the METANET. The calibration of the models' parameters and the setup of the controllers are introduced in Section 4.2. The performance of the controllers are shown in Section 4.3.

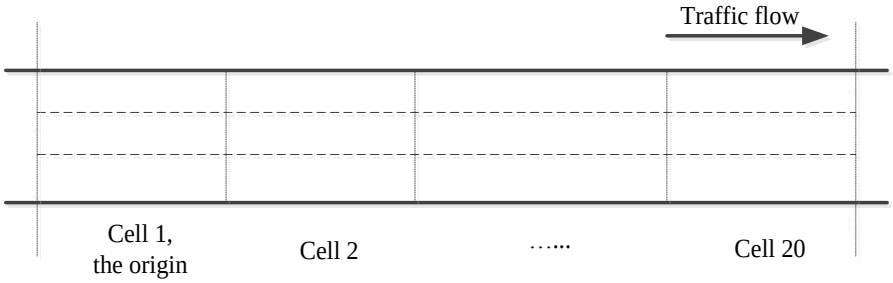

Figure 6: A graphical representation of the freeway stretch for the benchmark problem.

4.1. Reproducing of a jam wave

The second-order macroscopic traffic flow model METANET, which can reproduce capacity drop and the propagation of jam waves with a reasonable accuracy, is used as the process model to represent the real world. The METANET model is briefly introduced in the following.

In the METANET model, the freeway network is divided into segments which have a uniform geometry structure. Each segment i is characterized by the traffic density $\rho_i(k)$ [veh/lane/km], the mean speed $v_i(k)$ [km/h], and the outflow $q_i(k)$ [veh/h], where k is the index of the discrete time step. The following equations describe the evolution of the freeway stretch over time. The outflow of each segment equals to the density times the mean speed and the number of lanes of that segment (represented by λ_i):

$$q_i(k) = \rho_i(k) v_i(k) \lambda_i, \tag{22}$$

The density of a segment follows the vehicle conservation law, which is represented as:

$$\rho_i(k+1) = \rho_i(k) + \frac{T}{L_i \lambda_i} \left(q_{i-1}(k) - q_i(k) \right), \tag{23}$$

The mean speed of segment i at time step $k+1$ depends on the mean speed at time step k, the speed of the inflow of vehicles, and the density downstream, which is represented as:

$$v_i(k+1) = v_i(k) + \frac{T}{\tau}(V(\rho_i(k)) - v_i(k)) + \frac{T}{L_i} v_i(k)(v_{i-1}(k) - v_i(k)) - \frac{\vartheta T}{\tau L_i} \frac{\rho_i(k+1) - \rho_i(k)}{\rho_i(k) + \kappa}, \tag{24}$$

where, τ, ϑ, κ are model parameters. $V(\rho_i(k))$ denotes the desired speed that drivers try to achieve, which is:

$$V(\rho_i(k)) = v_{free,i} \cdot \exp[-\frac{1}{a_m}\left(\frac{\rho_i(k)}{\rho_{cr,i}}\right)^{a_m}], \tag{25}$$

where, a_m is a model parameter. $v_{free,i}$ represents the free flow speed of segment i, and $\rho_{cr,i}$ denotes the critical density at which the flow is maximal. The effect of VSL follows the assumption of (Hegyi et al., 2005a), that the desired speed is the minimum of (25) and the speed limit displayed on variable message sign:

$$V(\rho_i(k)) = \min\left(VSL_i(k), v_{free,i} \cdot \exp[-\frac{1}{a_m}(\frac{\rho_i(k)}{\rho_{cr,i}})^{a_m}]\right). \tag{26}$$

The first segment of the freeway stretch is perceived as the origin, and modeled by simple queue model. The number of vehicles in the first segment is calculated as:

$$\omega_1(k+1) = \omega_1(k) + T(d(k) - q_1(k)), \tag{27}$$

where, $d(k)$ is the demand of the origin. $q_1(k)$ is determined by:

$$q_1(k) = \min[d(k) + \frac{\omega_1(k)}{T}, \lambda_1 V(\rho_{cr,1})\rho_{cr,1}], \tag{28}$$

where, the first term is the available traffic in time period k. The second term is the capacity flow. For the upstream boundary condition, the speed of the origin segment is assumed to be the same as the second segment, so $v_1(k) = v_2(k)$. For the downstream boundary condition, the density downstream of the freeway stretch is defined according to the simulation scenario.

The parameters of METANET model are set up as follows. In (23), L_i=0.3 km, λ_i=3. T=5 s. In (24), τ is set to 18 s, κ is set to 40 veh/km/lane, and ϑ is set to 30 km^2/h. These parameters are referred to (Kotsialos et al., 1999b). In (25), ρ_{cr}=27.6 veh/km/lane, a_m= 2.5, v_{free}=108 km/h. These parameters are selected based on the experience of a reasonable shape of the fundamental diagram. The input of the system are the traffic demand at the origin and the density downstream of the freeway stretch. Traffic demand profile is shown in Fig. 7 (a). The downstream density equals to the critical value constantly, except for a pulse from time step 380 to 400 which is used to trigger the jam wave. The pulse is chosen large enough to cause an upstream-propagating jam wave. The downstream density profile is shown in Fig. 7 (b).

The entire simulation process contains 1440 time steps, which correspond to 2 hours. Simulation results by the METANET model are shown in Fig. 8. The jam wave propagates from the downstream end to the upstream end in about 200 time steps. The flow contour plot shows a conspicuous capacity drop, that the outflow of the jam (the yellow area) is much lower than the flow after the jam wave is resolved (the white area).

As presented, three model predictive controllers are compared in this benchmark problem. To tune the parameters of N_p and N_c, we follow the rule that N_p should be larger than the typical travel time from the controlled cells to the exit of the network, and N_c should has a good trade-off between the computational effort and the performance. The prediction horizon and control horizon are set to 10 minutes, and the duration of control step T_c is set to 10 seconds. For a fair comparison the parameters are set to same values for three controllers. The controllers are assumed to be activated at time step 420 when the jam wave has formed and deactivated if the density of every cell in the freeway stretch is lower than the critical value so there will be no risk to form a new jam. In the following we briefly describe the mechanisms of three controllers.

1. Controller 1, the METANET-based MPC. The optimal control formulation we refer to (Hegyi et al., 2005b), thus both the process model and the prediction model in controller 1 are the METANET model. The optimization problem is solved by MATLAB implementation of the SQP algorithm (fmincon). The lowest VSL value is set to 35 km/h.
2. Controller 2, a model predictive controller, which uses (19) as the objective function and the model from (Roncoli et al., 2015a) (referred to as model R in the following) as the prediction model. Model R has taken lane-changing into consideration. Since lateral flows are not considered in this paper, we only use their longitudinal flow model which is introduced in Section 2.2. $A_i(j)$ are set to values based on the predicted density of every cell at every time step.

3. Controller 3, a model predictive controller, which uses (19) as the objective function and the new extended CTM as the prediction model. $A_i(j)$ are set to values based on the predicted density of every cell at every time step. The linear quadratic formulation of controllers 2 and 3 are solved by CPLEX from MATLAB toolbox.

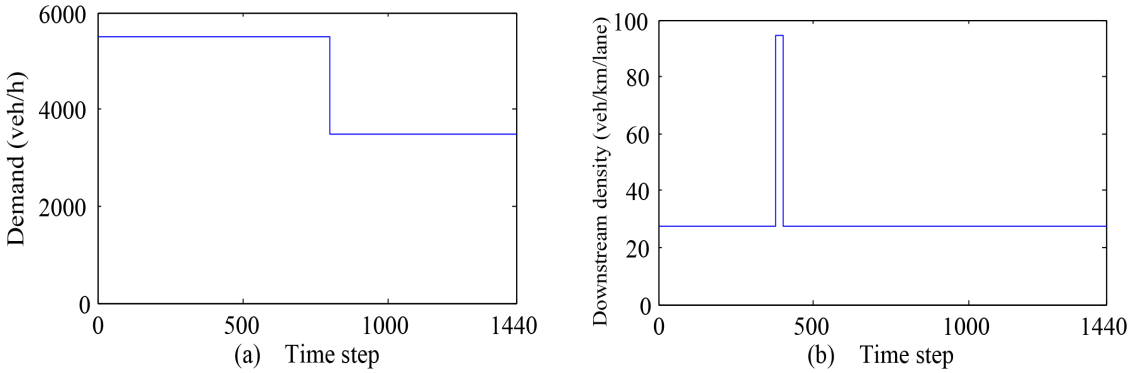

Figure 7: The traffic demand at the origin (a), and the density downstream of the freeway stretch (b).

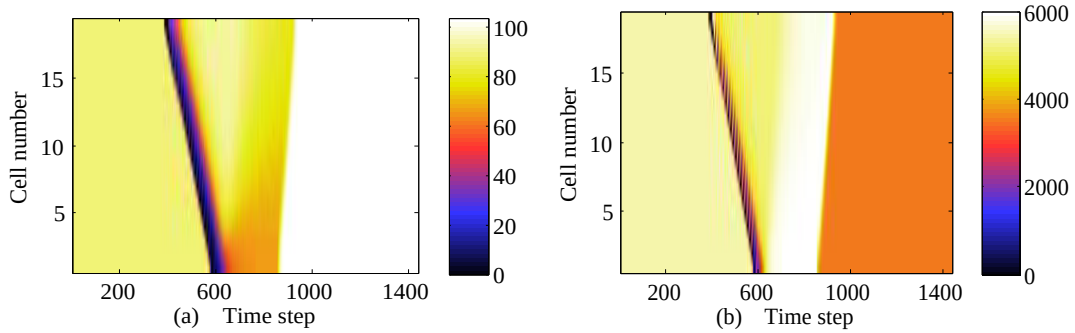

Figure 8: The speed (km/h) contour plot (a), and the flow (veh/h) contour plot (b) of the simulation without VSL control.

4.2. The set up of controllers

The model parameters of controllers 2 and 3 are calibrated with the simulation results from the METANET model. The calibration is performed based on an optimization-based method to best match the model parameters with the simulation results of the METANET model. The root mean square error (RMSE) of density is chosen as the performance indicator of the calibration, which is shown as the following equation:

$$\min \frac{1}{(I-1)\times(N-k+1)} \sqrt{\sum_{i=2}^{I}\sum_{j=k}^{N}(\rho_i(j)-\widehat{\rho_i}(j))^2}, \tag{29}$$

where, $\widehat{\rho_i}(j)$ is the density of cell i at time step j from the METANET simulation, and $\rho_i(j)$ is the simulated density from the prediction models of controllers 2 and 3. Data are extracted from time step 401 to 700, thus $N = 700$, $k = 401$. The optimization problem is solved by MATLAB implementation of the SQP algorithm (fmincon). Multi-starting points have been tried to avoid the optimization results getting stuck in a local optimum. The calibration results of controllers 2 and 3 are shown in Table 1. The parameter which characterizes the extent of capacity drop in the model of model R is c^{jam}, while in the new extended CTM it is α. Since they represent the same physical meaning,

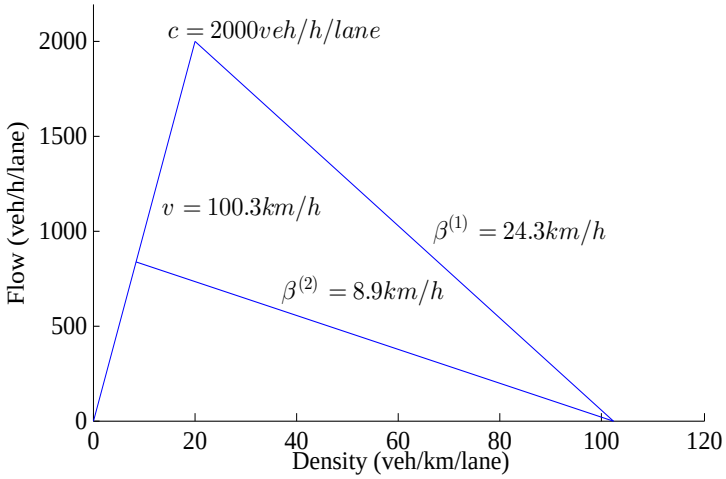

Figure 9: The calibated fundamental diagram of controller 3

Table 1: The calibration results of controllers 2 and 3

	v (km/h)	α	c (veh/h)	$\beta^{(1)}$ (km/h)	RMSE
Controller 2	100.32	0.58	6000	24.33	10.30
Controller 3	100.75	0.79	6000	23.9	8.41

for consistency both of them are represented by α in the table. If α is known, $\beta^{(2)}$ can be calculated accordingly. The values of parameters of controller 3 are also shown in Fig. 9

The RMSE shows that the new extended CTM has a more accurate simulation result than model R. Fig. 10 shows the density contour plots which are simulated by the predictions models. Both the new extended CTM and model R show dispersions of the jam wave, but qualitatively speaking, the dispersion of the new extended CTM is less.

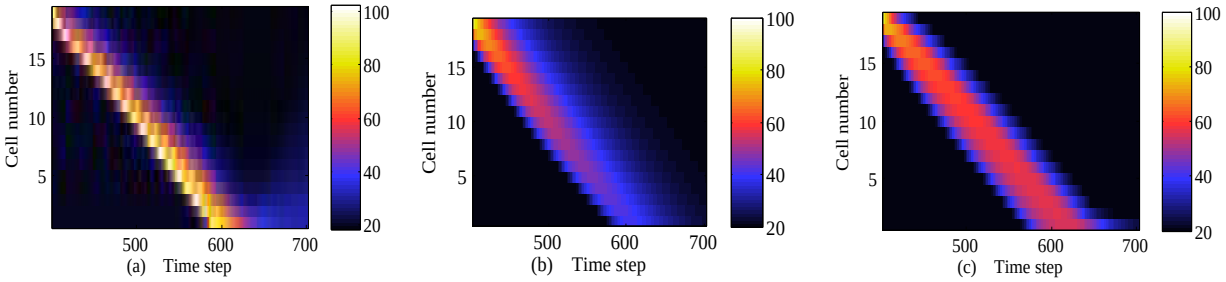

Figure 10: The density (veh/km/lane) contour plots (without VSL control) from (a) the METANET model, (b) Roncoli's model, and (c) the new extended CTM.

Fig. 11 shows the outflow of the freeway stretch simulated by those three models. The METANET model reproduces a substantial capacity drop, where the outflow of the network stays around 5200 veh/h till time step 650. The outflow reproduced by model R gradually increases, and recovers to the free-flow capacity value at time step around 550. The outflow reproduced by the new extended CTM stays in the range of 4700 to 5500 veh/h, from time step 450 to 650. The oscillation of flow happen due to the numerical issue, which can be briefly explained as follows: Suppose cell i is in a discharging state, and cell $i+1$ is in free-flow state. Under this circumstance, the outflow of the network is dominated by the outflow of cell $i+1$, f_{i+1}, because cell $i+1$ is the most upstream cell of all downstream free-flow cells. Since cell i is in the discharging state, the vehicle density ρ_i is decreasing and according to (5), f_{i+1} is also decreasing. That is the reason for the increasing of the outflow. When ρ_i decreases to an under-critical value,

cell $i-1$ becomes the discharging cell and the outflow of the network is dominated by f_i. That's the reason for the sudden decreasing of the outflow. This process repeats as the discharging cell moves upstream, causing the oscillation of flow. The numerical issue can be perceived as a drawback of the model, nevertheless, the new extended CTM still reproduces capacity drop during jam waves propagation, which is the main target traffic control aims to avoid.

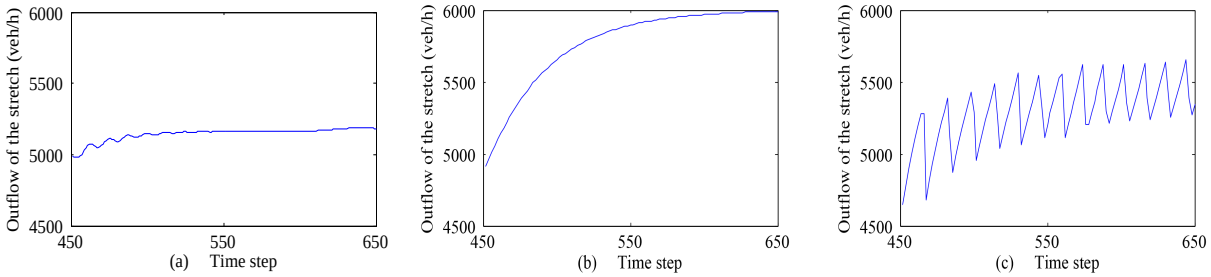

Figure 11: The outflow of the freeway stretch from the simulation of the METANET model (a), Roncoli's model (b), and the new extended CTM (c).

4.3. Control performance

The simulation results of three controllers are shown in Fig. 12. It can be seen that the jam wave is resolved by controllers 1 and 3. VSL control signals are applied to the upstream of the jam area to limit the inflow to the jam, creating such a condition that the outflow of the jam is higher than the inflow, so that the jam wave will be resolved. Fig. 13 shows the VSL signals generated by the controllers. The white color indicates areas where no VSL signal was applied. At the beginning of traffic control, the VSL control area generated by controller 1 is longer than that by controller 3, so the jam wave is resolved faster by controller 1. Controller 2 narrows the width of the jam, but it is unable to resolve it. In order to exclude the possibility that the length of the simulated stretch was causing the inability of controller 2 to resolve the jam, we have tried a scenario which the freeway stretch was extended to 40 cells. The result was that the jam wave was still not resolved by the VSL control even when the length of the freeway stretch was doubled. The reason is that the VSL signal generated by controller 2 is too short in length and time to resolve the jam wave. This could be attributed to the fact that the prediction model of controller 2 has underestimated the extent of the capacity drop, thus required VSLs (in length and time) to resolve the jam is underestimated. For controllers 2 and 3, most of the control signals are higher than the minimum VSL value, which are shown in Fig. 13 (b-c). A few control signals which are lower than the minimum VSL value is because of that the flow bound cannot be set to cells which have lower values of flows, as explained in Section 3.3. In this case, the proportion of the VSL signals which are lower than the desired minimum VSL value (35 km/h) is less than 2%, and the minimum value of contradicted signals is higher than 30 km/h.

More results are summarized in Table 2. It can be seen that all of the three controllers reduce the total travel delay compared to the no-control scenario. Total travel delay is calculated from the difference between the total travel time in the simulations and the free flow total travel time. Controller 1 achieves less total travel delay, but at a cost of a higher computation time which is not real-time feasible. Controller 2 has a less computation time, but it does not reduce the total travel delay as much as controllers 1 and 3. In addition, since the jam wave is not resolved, the traffic upstream of the network still suffers bad influences. Controller 3 has a more balanced performance, because it reduces total travel delay substantially compared to controller 2 while it can maintain a low computation time. The computation time of controllers 2 and 3 are both possible for a real time application. The computer with an E5-1620 processor and and 16 GB RAM is used for the optimizations. Note that one of the reason which causes the longer computation time of controller 1 is that the non-linear optimization problem is solved by a standard SQP algorithm. As mentioned earlier, there are some algorithms which could increase the computation speed of the non-linear optimization. More comparisons with different kinds of algorithms remain a subject for future research.

Table 2: The total travel delay and computation time of three controllers

	No control	The non-linear MPC (Controller 1)	The linear MPC (Controller 2)	The proposed MPC (Controller 3)
Total travel delay (h)	242	112.7	136.1	116.2
Computation time	-	6-8 min	6 s	9 s

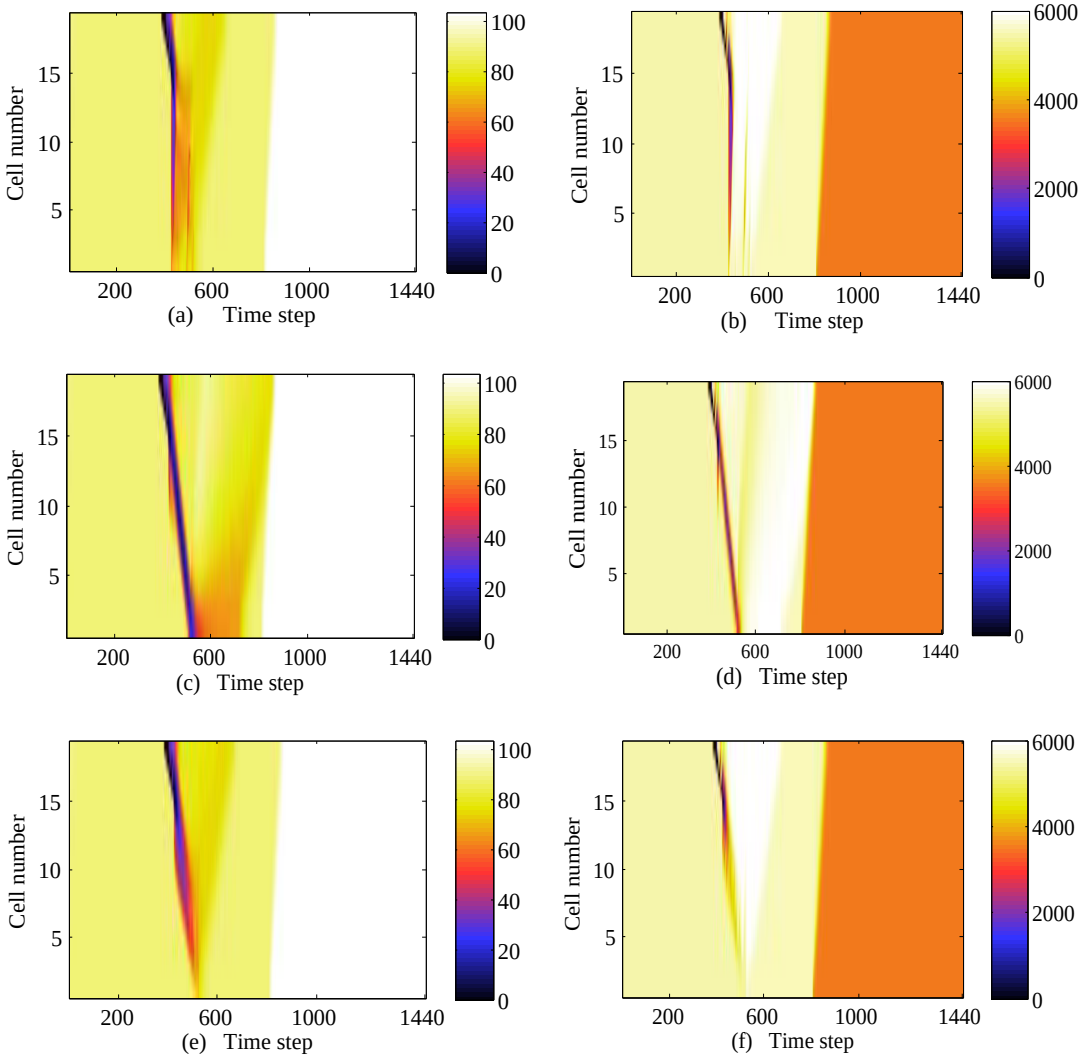

Figure 12: The speed contour plots (with VSL control) of controllers 1, 2, 3 (a,c,e), and the flow contour plots of controllers 1, 2, 3 (b,d,f).

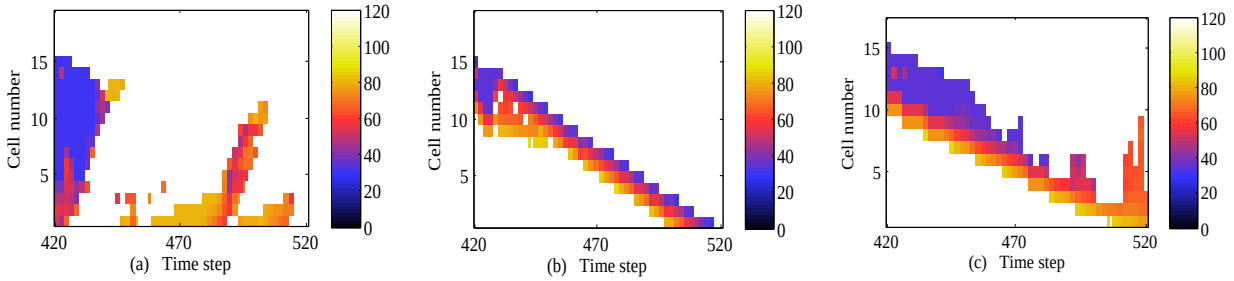

Figure 13: The VSL control signals (km/h) generated by controller 1 (a), controller 2 (b), and controller 3 (c).

5. Conclusion and future research

In this paper we apply a linear MPC for optimal VSL control to resolve jam waves on freeways. The proposed MPC is developed based on a new extended discrete first-order model which takes the capacity drop into account and is able to reproduce jam wave propagation accurately. The proposed MPC controller is applied to a benchmark freeway stretch to test the performance in terms of computation speed and jam wave resolution. The second-order model METANET that can reproduce the capacity drop phenomenon, is used as the process model to represent the real world. Two other controllers, one non-linear MPC based on the METANET model and the other linear MPC based on a first-order model, were also tested in the benchmark problem for comparison.

In the benchmark problem, the proposed controller shows a better performance in resolving the jam wave and reducing the total travel delay than the compared linear MPC. Meanwhile, the proposed controller keeps a real time feasible computation time, which has a significant improvement comparing to the METANET-based MPC. This feature enables the proposed controller to be implemented in practice.

This paper also proposes a way to implement the minimum flow constraints to linear MPC formulations. A forward prediction is performed before the optimization process, so as to ensure that the minimum flow constraints do not conflict with the demand and supply constraints.

Future research will consider: implementing safety constraints into the controller. For safety, normally the VSL variation between the neighboring cells should be lower than a certain value. At this moment it is not trivial to implement such constraints into the controller. The proposed controller also needs to be tested for integrated control measures (such as ramp metering, route guidance, etc.) in a large-scale network.

Acknowledgments

The research leading to these results are sponsored by Chinese Scholarship Council (No. 201307090008).

References

Camacho, E. F. and Bordons, C. *Model predictive control in the process industry*. Springer Science & Business Media, 2012.

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

Carlson, R. C., Papamichail, I., and Papageorgiou, M. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1261–1276, 2011.

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

Chung, K., Rudjanakanoknad, J., and Cassidy, M. J. Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological*, 41(1):82–95, 2007.

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

Gazis, D. and Potts, R. The oversaturated intersection. *Proceedings of the 2nd International Symposium of Traffic Theory*, 1963.

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

Hadiuzzaman, M. and Qiu, T. Z. Cell transmission model based variable speed limit control for freeways. *Canadian Journal of Civil Engineering*, 40(1):46–56, 2013.

Han, Y., Yuan, Y., Hegyi, A., and Hoogendoorn, S. P. Linear quadratic MPC for integrated route guidance and ramp metering. In *Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems (ITSC 2015)*, pages 1150–1155. IEEE, 2015.

Han, Y., Yuan, Y., Hegyi, A., and Hoogendoorn, S. A new extension of discrete first-order model to reproduce the propagation of jam wave. In *Transportation Research Board 95th Annual Meeting*, number 16-3482, 2016.

Hegyi, A. and Hoogendoorn, S. Dynamic speed limit control to resolve shock waves on freeways-field test results of the SPECIALIST algorithm. In *Proceedings of the IEEE 13th International Conference on Intelligent Transportation Systems (ITSC 2010)*, pages 519–524. IEEE, 2010.

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

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

Hegyi, A., De Schutter, B., and Hellendoorn, J. Optimal coordination of variable speed limits to suppress shock waves. *IEEE Transactions on Intelligent Transportation Systems*, 6(1):102–112, 2005b.

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

Kotsialos, A., Papageorgiou, M., and Messmer, A. Optimal coordinated and integrated motorway network traffic control. In *14th International Symposium on Transportation and Traffic Theory*, 1999a.

Kotsialos, A., Papageorgiou, M., and Messmer, A. Optimal coordinated and integrated motorway network traffic control. In *14th International Symposium on Transportation and Traffic Theory*, 1999b.

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

Le, T., Vu, H. L., Nazarathy, Y., Vo, Q. B., and Hoogendoorn, S. Linear-quadratic model predictive control for urban traffic networks. *Transportation Research Part C: Emerging Technologies*, 36:498–512, 2013.

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

Lebacque, J.-P. The godunov scheme and what it means for first order traffic flow models. In *Internaional symposium on transportation and traffic theory*, pages 647–677, 1996.

Lighthill, M. J. and Whitham, G. B. On kinematic waves. II. a theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 229, pages 317–345. The Royal Society, 1956.

Lo, H. A dynamic traffic assignment formulation that encapsulates the cell-transmission model. In *14th International Symposium on Transportation and Traffic Theory*, 1999.

Mahajan, N., Hegyi, A., De Weg, V., Sterk, G., and Hoogendoorn, S. P. Integrated variable speed limit and ramp metering control against jam waves–a COSCAL V2 based approach. In *Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems (ITSC 2015)*, pages 1156–1162. IEEE, 2015.

Muralidharan, A. *Tools for modeling and control of freeway networks*. PhD thesis, 2012.

Muralidharan, A. and Horowitz, R. Computationally efficient model predictive control of freeway networks. *Transportation Research Part C: Emerging Technologies*, 2015.

Papageorgiou, M. An integrated control approach for traffic corridors. *Transportation Research Part C: Emerging Technologies*, 3(1):19–30, 1995.

Papageorgiou, M., Kosmatopoulos, E., and Papamichail, I. Effects of variable speed limits on motorway traffic flow. *Transportation Research Record: Journal of the Transportation Research Board*, (2047):37–48, 2008.

Richards, P. I. Shock waves on the highway. *Operations research*, 4(1):42–51, 1956.

Roncoli, C., Papageorgiou, M., and Papamichail, I. Traffic flow optimisation in presence of vehicle automation and communication systems–part i: A first-order multi-lane model for motorway traffic. *Transportation Research Part C: Emerging Technologies*, 57:241–259, 2015a.

Roncoli, C., Papageorgiou, M., and Papamichail, I. Traffic flow optimisation in presence of vehicle automation and communication systems–part ii: Optimal control for multi-lane motorways. *Transportation Research Part C: Emerging Technologies*, 57:260–275, 2015b.

Schönhof, M. and Helbing, D. Empirical features of congested traffic states and their implications for traffic modeling. *Transportation Science*, 41 (2):135–166, 2007.

Schreiter, T., Smits, E.-S., Van Lint, H., and Hoogendoorn, S. The cell-transmission model with capacity drop. In *Proceedings of the 11th International Congress of the Research School Transportation, Infrastructure and Logistics(TRAIL)*, 2010.

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

Srivastava, A., Jin, W.-L., and Lebacque, J. P. A modified cell transmission model for signalized intersections. In *Transportation Research Board 93rd Annual Meeting*, number 14-4713, 2014.

van de Weg, G. S., Hegyi, A., Hoogendoorn, S. P., and De Schutter, B. Efficient model predictive control for variable speed limits by optimizing parameterized control schemes. In *Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems (ITSC 2015)*, pages 1137–1142. IEEE, 2015.

Yuan, K., Knoop, V. L., and Hoogendoorn, S. P. Capacity drop: A relation between the speed in congestion and the queue discharge rate. In *Transportation Research Board 94th Annual Meeting*, 2015.

Zegeye, S. K., De Schutter, B., Hellendoorn, J., Breunesse, E. A., and Hegyi, A. A predictive traffic controller for sustainable mobility using parameterized control policies. *IEEE Transactions on Intelligent Transportation Systems*, 13(3):1420–1429, 2012.

Ziliaskopoulos, A. K. A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation science*, 34(1):37–49, 2000.