Delft University of Technology

Delft Center for Systems and Control

Technical report 06-043

Towards a practical application of model predictive control to suppress shock waves on freeways*

A. Hegyi, M. Burger, B. De Schutter, J. Hellendoorn, and T.J.J. van den Boom

If you want to cite this report, please use the following reference instead:

A. Hegyi, M. Burger, B. De Schutter, J. Hellendoorn, and T.J.J. van den Boom, "Towards a practical application of model predictive control to suppress shock waves on freeways," *Proceedings of the European Control Conference 2007 (ECC'07)*, Kos, Greece, pp. 1764–1771, July 2007.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: `https://www.dcsc.tudelft.nl`

*This report can also be downloaded via `https://pub.deschutter.info/abs/06_043.html`

Towards a practical application of model predictive control to suppress shock waves on freeways

A. Hegyi, M. Burger, B. De Schutter, J. Hellendoorn, T.J.J. van den Boom

***Abstract*— We present the results of the application of model predictive control (MPC) to a micro-simulation model with a scenario where shock waves are present, and a micro-simulation model functions as a substitute for the real-world traffic system. Shock waves emerge in most cases from traffic jams at bottlenecks, propagate upstream on the freeway, and can remain existent for a long time and distance. This increases travel time, is potentially unsafe, and increases noise and air pollution.**

Previously reported results using MPC to eliminate shock waves, showed an improvement of 20% of the total time that the vehicles spent in the network. However, they were based on the assumption that the simulation model (representing the real world) and the prediction model are the same, which may have lead to overoptimistic results.

In this paper a micro-simulation model (Paramics 5.1 by Quadstone) is used to represent the real world, which results in a model mismatch between the simulation model and the prediction model. We show by simulation that even in the case of a model mismatch the controller is able to suppress or remove shock waves.

***Index Terms*— model predictive control, dynamic speed limits, freeway traffic control, shock waves**

I. INTRODUCTION

A. Shock waves

In the field of freeway traffic control most of the attention is paid to solving the fixed jams at on-ramps, whereas upstream propagating shock waves (short jams) often emerge from on-ramps *and* other types of bottlenecks, and resolving such shock waves would greatly improve the freeway traffic flow. Shock waves propagate upstream along the freeway and can remain existent for a long time and distance. As a consequence, every vehicle that enters the freeway upstream of the jammed area will have to pass through the jammed area, which increases the travel time, creates potentially unsafe situations and increases noise and air pollution by braking and accelerating vehicles. Shock waves typically have a lower outflow than the capacity of the freeway at the given location, which motivates the idea that traffic flow can be improved by resolving shock waves.

To resolve shock waves one can use dynamic speed limits in the following way. In some sections upstream of a shock wave speed limits are imposed in order to reduce the inflow to the jammed area. When the inflow of the jammed area is reduced sufficiently, i.e., to a lower value than its outflow, the jam will eventually dissolve. In other words, the speed limits create a low-density wave that propagates downstream. This low-flow, low-density wave meets and compensates the high-density shock wave. As a result, the shock wave is reduced or eliminated.

Although the above approach to speed limit control is valid, we will formulate the controller in terms of a control *goal* instead of an explicit *behavior* of the controller. The control goal will be minimization of the total time that vehicles spend on the freeway (the Total Time Spent, or TTS), with the implicit assumption that removing shock waves reduces the TTS. We will see in the experiment results that this control goal leads to the shock wave reduction behavior as described above.

The purpose of this paper is to discuss issues and choices that are related to the application of model predictive control in a real-world environment (which is here represented by a micro-simulation model).

B. State-of-the-art

In the literature, basically two views on the use of speed limits can be found. The first emphasizes the homogenization effect [1], [2], [3], [4], whereas the second is more focused on flow limitation by speed limits to prevent traffic breakdown or to solve jams [5], [6], [7].

The basic idea of homogenization is that speed limits can reduce the speed differences, by which a stabler (and safer) flow can be achieved. This approach can in theory increase the time to breakdown [8], but cannot suppress or resolve existing shock waves. According to field tests homogenization results in a somewhat more stable and safer traffic flow, but no significant improvement of traffic volume was measured [3], [9], [10] (nor can be expected based on theory). An extended overview of existing speed limit systems that aim at reducing speed differentials is given by Wilkie [11].

The flow limitation approach focuses more on preventing or removing too high densities, and also allows speed limits that are lower than the critical speed in order to limit the inflow to the high-density area (a jam or nearly a jam). By resolving these high-density areas higher flow can be achieved compared to the homogenization approach. If the breakdown cannot be prevented, the dynamic speed limits can be used to limit the inflow of the jammed area to reduce the density until the jam is resolved.

Several control methodologies are described in literature to find a control law for speed control, such as multi-layer control [12], sliding-mode control [6], [7], and optimal

All authors are with Delft Center for Systems and Control, TU Delft, Mekelweg 2, 2628 CD Delft, The Netherlands. B. De Schutter is also with the Marine and Transport Technology department of TU Delft. Email: {a.hegyi,a.j.j.vandenboom,j.hellendoorn}@tudelft.nl, m.burger@ewi.tudelft.nl, b@deschutter.info

control [13]. In [14] optimal control is approximated by a neural network in a rolling horizon framework.

Other authors use (or simplify their control law to) a control logic where the switching between the speed limit values is based on traffic volume, speed or density measurement [1], [2], [3], [4], [7], [8]. In some cases the switching between the speed limit values is also based on special circumstances, such as light and weather conditions [1].

Some authors recognize the importance of anticipation in the speed control scheme. A pseudo-anticipative scheme is used in [7], [15] by an intelligent switching scheme based on the density of the neighboring downstream segment. Real predictions are used in [13], [14], and together with [15] these are the only approaches that results in a significant flow improvement.

The approach used in this paper is the flow limitation approach. Using this approach in a model predictive control (MPC) framework, earlier simulation studies [16] resulted in an improvement of the average travel time of around 20%. In these studies it was assumed that the prediction model is exactly the same as the simulation model, which is unrealistic and may lead to overoptimistic results.

The contribution of this paper is that the simulation model (representing the real world) is now replaced by a micro-simulation model. For practical reasons the experiments cannot be performed in a real traffic system, and the micro-simulation model is considered as a substitute of the real-world. One of the consequences of the difference between the models of the controller and the real-world substitute is that the prediction model of the controller is not perfect anymore. In this paper, the performance is evaluated under this mismatch between the prediction model and the simulation model, which is considered a step towards the real application of the MPC approach.

II. Approach

To solve the problem of finding the dynamic speed limit signals that suppress the shock waves, the MPC framework is used. Below we shortly explain this framework and the two traffic flow models used in this paper, one for prediction and one for simulating of the effects of the resulting speed limits. We also explain the prediction model calibration approach.

A. Model predictive control

We use a model predictive control (MPC) scheme to solve the problem of optimal coordination of dynamic speed limits, which is schematically represented in Fig. 1.

In MPC, at each controller time step k_c the optimal control signal is computed (by numerical optimization) over a prediction horizon N_p for the given actual system (here: traffic) state. A control horizon N_c is selected over which the control variables are allowed to vary, and after the control horizon has been passed (but before the prediction horizon is reached) the control signal is typically taken to be constant. Taking the control horizon smaller reduces the number of optimization variables of the system, and typically also improves the stability of the system[1]. From the resulting optimal control signal only the first sample (with time index $k_\mathrm{c}+1$) is applied to the process. In the next time step $k_\mathrm{c}+1$, when the new system state is available, a new optimization is performed (with a prediction period that is shifted one time step ahead) and of the resulting control signal again only the first sample is applied, and so on. This scheme, called rolling horizon, allows for updating the state from measurements, or even for updating the model in every iteration step.

Note that the controller sampling time T_c and the related time index k_c are not necessarily equal to the simulation time step T and index k of the prediction model. In general, we assume that the controller sampling time is an integer multiple of the controller model time step:

$$T_\mathrm{c} = MT, \tag{1}$$

where M is a positive integer.

An essential feature of the MPC scheme is that it uses a *prediction*, which makes it possible *to perform temporarily sub-optimally in order to gain more performance in the future.* In our case, we will see (what intuitively already can be expected) that in order to suppress a shock wave, traffic flow has to be limited, but when the shock wave has dissolved, the traffic flow will be higher than without resolving the shock wave.

In this scheme an objective function is used to describe the (predicted) performance of the system given a control signal. This objective function expresses the control goal and typically formulates the cost of a (predicted) system evolution associated with given control signal. In our case we use the total time that vehicles spend in the network (TTS) as the objective function that should be minimized.

For more information on MPC we refer the interested reader to [17], [18], [19] and the references therein.

[1]It improves the stability in the sense that with fewer optimization variables the optimization result will be less likely to fluctuate between equally good but different solutions.

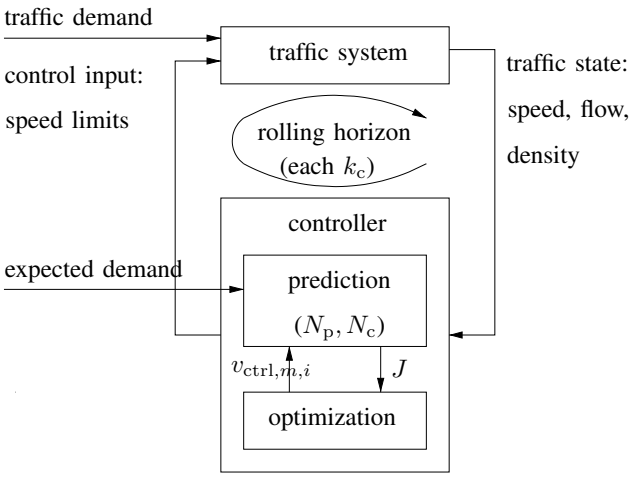

Fig. 1. Schematic view of the model predictive control (MPC) structure.

B. Prediction model

The MPC procedure includes a prediction of the network evolution as a function of the current state and a given control input. For this prediction we use an extended version of the (destination-independent version) of the macroscopic traffic flow model METANET [20], [21]. The extensions are introduced to model shock waves and freeway origins better and to include the effects of dynamic speed limits explicitly.

Note that the since MPC approach is generic, also other traffic models (that include the effect of the speed limit) could be used.

For the sake of brevity, we describe only those parts of the model that are relevant for interpreting and understanding the simulation results of our benchmark network (as described in Section III). In [20], [21] the complete METANET model is given (including lane drops, nodes, and merging and weaving processes).

1) Basic METANET model: Freeway networks are modeled by the composition of links and nodes. Links consist of several segments that have identical properties. Consider a freeway link m that is subdivided into N_m segments, each with a length L_m and λ_m lanes, and a discrete time step with length T (h). Traffic dynamics is described in terms of the aggregated variables *average speed* $v_{m,i}(k)$ (km/h), *average flow* $q_{m,i}(k)$ (veh/h), and *average density* $\rho_{m,i}(k)$ (veh/km/lane), where i is the segment index.

The METANET model equations are given by the fundamental relationship between speed, density and flow,

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

the law of conservation of vehicles,

$$\rho_{m,i}(k+1) = \rho_{m,i}(k) + \frac{T}{L_m \lambda_m} \left(q_{m,i-1}(k) - q_{m,i}(k)\right), \tag{3}$$

and a heuristic relationship for the speed dynamics,

$$\begin{aligned} v_{m,i}(k+1) = v_{m,i}(k) &+ \frac{T}{\tau}\left(V(\rho_{m,i}(k)) - v_{m,i}(k)\right) \\ &+ \frac{T}{L_m} v_{m,i}(k) \left(v_{m,i-1}(k) - v_{m,i}(k)\right) \\ &- \frac{\eta T}{\tau L_m} \frac{\rho_{m,i+1}(k) - \rho_{m,i}(k)}{\rho_{m,i}(k) + \kappa} \,, \end{aligned} \tag{4}$$

$$V\left(\rho_{m,i}(k)\right) = v_{\mathrm{free},m} \exp\left[-\frac{1}{a_m}\left(\frac{\rho_{m,i}(k)}{\rho_{\mathrm{crit},m}}\right)^{a_m}\right], \tag{5}$$

where $v_{\mathrm{free},m}$ is the free-flow speed in segment m, $\rho_{\mathrm{crit},m}$ is the critical density (the density at or above which traffic becomes unstable), and τ, η, a_m, κ, are model fitting parameters without direct physical meaning.

Origins are modeled with a simple queue model. The length of the queue $w_o(k)$ equals the previous queue length plus the demand $d_o(k)$, minus the outflow $q_o(k)$:

$$w_o(k+1) = w_o(k) + T\big(d_o(k) - q_o(k)\big) \ .$$

The outflow of on-ramp o depends on the traffic conditions on the freeway and the capacity of the origin. The flow $q_o(k)$ is the minimum of the demand and the maximal flow that can enter the freeway given the mainstream conditions,

$$q_o(k) = \min\left[d_o(k) + \frac{w_o(k)}{T}, Q_o \frac{\rho_{\max} - \rho_{\mu,1}(k)}{\rho_{\max} - \rho_{\mathrm{crit},\mu}}\right], \tag{6}$$

where Q_o is the on-ramp capacity (veh/h) under free-flow conditions, $\rho_{\max}$ is the maximum density, and μ the index of the link to which the on-ramp is connected.

2) Extensions: Since the original METANET model does not explicitly describe the effect of speed limits, we have proposed an equation for the desired speed (5) in order to incorporate speed limits:

$$\begin{aligned} V\big(\rho_{m,i}(k)\big) = \min\Bigg(&(1+\alpha) v_{\mathrm{ctrl},m,i}(k), \\ &v_{\mathrm{free},m} \exp\left[-\frac{1}{a_m}\left(\frac{\rho_{m,i}(k)}{\rho_{\mathrm{crit},m}}\right)^{a_m}\right]\Bigg) \ , \end{aligned} \tag{7}$$

where $v_{\mathrm{ctrl},m,i}(k)$ is the speed limit imposed on segment i, link m, at time k, and the factor $(1+\alpha)$ expresses the (non-)compliance, i.e., the factor that the desired speed is higher or lower than the displayed speed limit.

A second extension is introduced to express the different nature of a mainstream origin link o compared to an on-ramp (the queue at a mainstream origin is in fact an abstraction of the sections upstream of the origin of the part of the freeway network that we are considering). If o is the origin of mainstream link μ, then we have

$$q_o(k) = \min\left[d_o(k) + \frac{w_o(k)}{T},\ q_{\mathrm{lim},\mu,1}(k)\right] \ , \tag{8}$$

where $q_{\mathrm{lim},\mu,1}(k)$ is the maximal inflow determined by the limiting speed in the first segment of link μ:

$$q_{\mathrm{lim},\mu,1}(k) = \begin{cases} \gamma\, \lambda_\mu\, v_{\mathrm{lim},\mu,1}(k)\, \rho_{\mathrm{crit},\mu} \left[-a_\mu \ln\left(\dfrac{v_{\mathrm{lim},\mu,1}(k)}{v_{\mathrm{free},m}}\right)\right]^{\frac{1}{a_\mu}} \\ \qquad\qquad \text{if } v_{\mathrm{lim},\mu,1}(k) < V(\rho_{\mathrm{crit},\mu}) \\ \gamma\, q_{\mathrm{cap},\mu} \qquad \text{if } v_{\mathrm{lim},\mu,1}(k) \geq V(\rho_{\mathrm{crit},\mu}), \end{cases}$$

where $v_{\mathrm{lim},\mu,1}(k) = \min(v_{\mathrm{ctrl},\mu,1}(k), v_{\mu,1}(k))$ is the speed that limits the flow, $q_{\mathrm{cap},\mu} = \lambda_\mu V(\rho_{\mathrm{crit},\mu}) \rho_{\mathrm{crit},\mu}$ is the capacity flow, and γ is a tuning parameter.

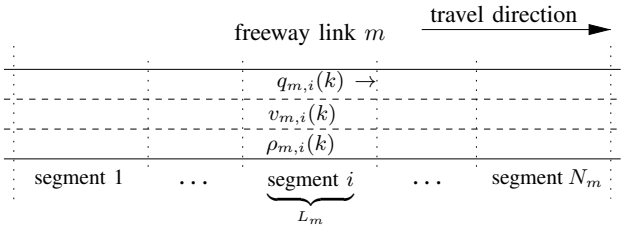

Fig. 2. In the METANET model, a freeway link is divided into segments. The main variables in the model are the average outflow of a segment $q_{m,i}(k)$, the average speed $v_{m,i}(k)$, the average density $\rho_{m,i}(k)$, and the segment length L_m.

A third extension is introduced to be able to express the different anticipation behavior of the drivers at the head and the tail of a traffic jam (i.e., a shock wave). The parameter η in (4) is replaced by the density-dependent parameter $\eta_{m,i}(k)$ according to:

$$\eta_{m,i}(k) = \begin{cases} \eta_{\text{high}} & \text{if } \rho_{m,i+1}(k) \geq \rho_{m,i}(k) \\ \eta_{\text{low}} & \text{if } \rho_{m,i+1}(k) < \rho_{m,i}(k). \end{cases}$$

A motivation of these extensions can be found in [16] and [22].

C. Simulation model

To represent the real-world, a simulation model is used, namely the microscopic traffic flow simulator Paramics v5.1, by Quadstone [23]. This model is used to generate calibration data for the prediction model, and to evaluate the speed limit control strategy.

Although the user is free to place detectors at any location in this model, in this paper all the detectors are located at the downstream end of the corresponding segment. In this paper, the detectors provide the time mean speed, the density (which is calculated based on individual vehicle speed measurements), and flow.

D. Calibration

In order to match the behavior of the prediction model with the simulation model the parameters of the prediction model have to be calibrated.

In a real-world control situation the prediction model would be calibrated by using real traffic data. Here the real world is represented by the simulation model, so Paramics is used to generate data for the calibration. Since the goal is to calibrate METANET such that it reproduces shock waves and the reaction to the speed limits, the calibration data set (or scenario) should be rich enough, i.e., should be defined such that it contains shock waves, traffic in free flow, the transitions between the two, and a variety of speed limits.

The selection of the calibration performance measure is based on the intended application of the prediction model: the prediction of the TTS for a given control signal. The TTS is in general calculated based on the number of vehicles (or densities) in the network over the predicted period. Since the TTS calculated over a given time horizon for a given network in fact only reflects the average traffic density, many different traffic state trajectories (speeds and densities) may lead to the same TTS. Therefore, the TTS is not very powerful to make a distinction between better or worse parameter sets in the sense that a well-calibrated model is expected to be able to reproduce shock wave propagation and the effects of dynamic speed limits.

For these reasons and the fact that in reality the density is usually not measured, the calibration performance measure was based on the differences of the predicted speed $\hat{v}_i(k_c)$ and flow $\hat{q}_i(k_c)$ relative to the speed $v_i(k_c)$ and flow $q_i(k_c)$ in the calibration data set.

Note that the calibration data and the controller model use different time steps, similarly to the difference in the controller sampling time and the prediction model sampling time, as described in Section II-A. For the sake of simplicity, we assume here that the time step counter in the calibration data equals the controller time step counter k_c. For the comparison of the predicted speed and flow with the speed and flow calibration data, the predictions are averaged over M time steps:

$$\hat{v}_{m,i}(k_c) = \frac{1}{M} \sum_{k=k_c M}^{(k_c+1)M-1} v_{m,i}(k), \tag{9}$$

and $\hat{q}_i(k_c)$ is defined similarly.

At step k_c the performance measure that quantifies the error of the prediction over a horizon of N_p steps is given by:

$$J_{k_c}(\theta) = \left[\frac{1}{N_p} \sum_{m=1}^{N_{\text{links}}} \sum_{i=1}^{N_m} \sum_{j=k_c}^{k_c+N_p} \left(\frac{(\hat{v}_{m,i}(j) - v_{m,i}(j))^2}{v_{\text{norm}}} + \frac{(\hat{q}_{m,i}(j) - q_{m,i}(j))^2}{q_{\text{norm}}} \right) \right]^{1/2} \tag{10}$$

where θ is the vector of variables to be calibrated, N_{links} is the number of links in the network and v_{norm} and q_{norm} normalizing factors equaling the average speed and flow in the data set.

To evaluate this measure for a data set with a longer time period $k_c = 1, \ldots, K$ ($K > N_p$) the N_p-step ahead performance of (10) is evaluated each k_c, with the initial traffic state reinitialized for each prediction according to the state in the calibration data. The overall performance is expressed as

$$J(\theta) = \sum_{k_c=1}^{K} J_{k_c}(\theta), \tag{11}$$

which is the objective function minimized in the calibration procedure.

III. Benchmark set-up

A. Lay-out and scenario

The above approach was tested with a benchmark network and scenario. The network consists of a 2-lane freeway of 14 km length, with two on-ramps, one located 1 km from the upstream end and one 1 km from the downstream end. Fig. 3 shows the lay-out of the network. The upstream on-ramp was included to be able to load the link with a flow close to capacity and the downstream on-ramp was included to create shock waves emerging from the on-ramp. The stretch

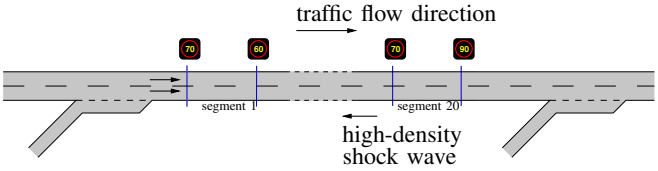

Fig. 3. The layout of the benchmark network. A 2-lane freeway with two on-ramps.

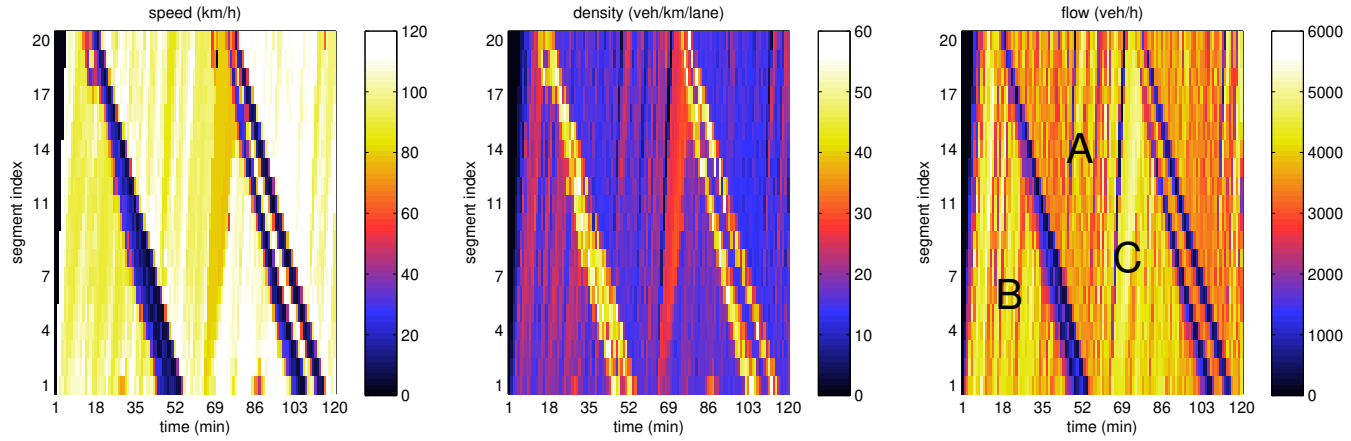

Fig. 4. The benchmark scenario contains two shock waves in a period of two hours, indicated by the dark lines in the speed diagram. The traffic flow direction is from segment 1 toward segment 20. Both shock waves originate from the downstream on-ramp.

considered for the dynamic speed limit control application is the stretch of 10 km length in the middle between the two on-ramps, and consists of 20 segments of 500 m. Since the simulation network consists of only one link, we drop the link index m from now on, for the sake of simplicity. The demand scenario was chosen such that two shock waves occur in a period of two hours. The evolution of the traffic in Paramics for this scenario is shown in Fig. 4. In each of these plots one of the main traffic variables is shown (speed, density and flow). The vertical axis shows the segment numbers and the travel direction is from segment 1 to 20. It is clear that the outflow of the shock wave (region A in the flow diagram on the right) is lower (darker) than the inflow to it (region B). So, there is a capacity drop associated with the shock wave. After the first shock wave has exited the link and passed the upstream on-ramp (not shown here) the flow from the on-ramp increased the inflow to the link and caused a forward propagating high-flow wave (through C, light color) that triggered the second shock wave at the downstream on-ramp.

B. Calibration

The calibration data set was generated based on the same scenario, but for the second hour (and the second shock wave) several speed limit patterns were applied, as shown in Fig. 5. The calibration of the METANET parameters was performed according to Section II-D, with the boundary conditions given by the speed and outflow of segment 1 and the density of segment 20 in the calibration data set. These states serve as the upstream speed $v_0(k)$ and upstream flow $q_0(k)$ and the downstream density $\rho_{N_m+1}(k)$ in (4) and (8). All parameters of the METANET model are calibrated, using the Matlab implementation of sequential quadratic programming (the fmincon function). The parameters resulting from the calibration were used for the MPC.

The controller sampling time $T_\mathrm{c} = 1$ minute and the prediction model time step $T = 10\,\mathrm{s}$ thus $M = 6$, and $N_\mathrm{p} = 15$ which were used in (10) and (11).

C. Boundary conditions for MPC

During each MPC iteration the prediction model needs a prediction of the boundary conditions. We assumed that the future boundary conditions are equal to the boundary conditions defined by the uncontrolled shock wave scenario. At first glance this may seem unrealistic because in practice the future upstream and downstream boundary conditions are unknown. However, if the network for which the current traffic state is observed is taken sufficiently larger than the controlled network, then the future demands at the boundaries of the controlled network can be deduced from the actual traffic state only. Furthermore, the predicted boundary conditions given by the shock wave scenario will typically differ from the simulated boundary conditions because of the effect of the speed limits (i.e., the improved flow and the eliminated shock waves create different patterns at the network boundaries). In practice one could also use historical data.

D. Performance criterion

The performance criterion for the MPC was the TTS defined by

$$J_\mathrm{TTS}(k_\mathrm{c}) = T \sum_{j=Mk_\mathrm{c}}^{M(k_\mathrm{c}+N_\mathrm{p})} \sum_{i=1}^{N_m} \rho_i(j)\,\lambda\,L.$$

E. Speed limits

Although all segments are speed controllable, the choice was made to control only the speed limits of the last 7 km (segments 6–20) and to assign the same speed limit per two adjacent segments (segments 6 and 7 get the same speed limit, 8 and 9 also, etc.), in order to reduce the computation time of the controller. Furthermore, the speed limits were constrained to the interval

$$40\,\mathrm{km/h} \le v_i(k) \le 120.$$

The controller sampling time T_c was 1 minute, the control and prediction horizons were respectively $N_\mathrm{c} = 8$ (equaling

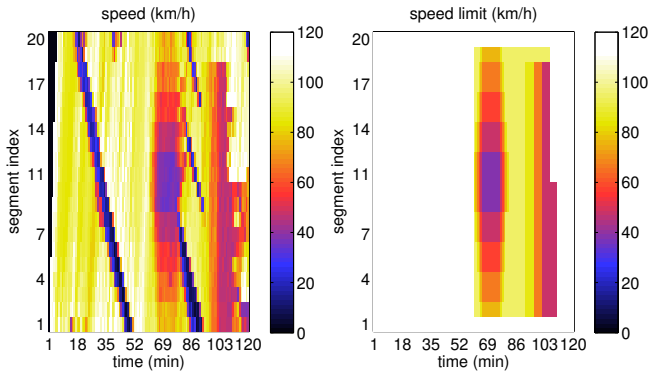

Fig. 5. The speed limits in the calibration scenario and their effects on the traffic scenario.

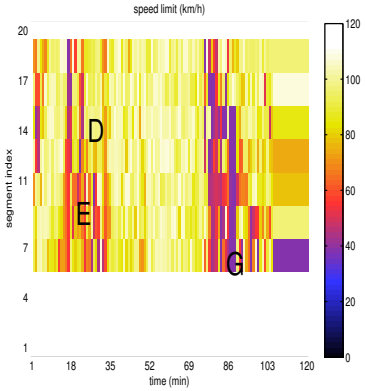

Fig. 6. The speed limit values that resulted from MPC.

8 minutes) and $N_{\mathrm{p}} = 15$ (15 minutes) which are somewhat larger than the horizons reported in [16] to be on the safe side.

IV. Results

A. Calibration

For the METANET model all parameters were optimized according to (11) by numerical optimization and this resulted in the following parameters: $\tau = 25.4$ s, $a = 3.133$, $\eta_{\mathrm{high}} = 217.3\,\mathrm{km^2/h}$, $\eta_{\mathrm{low}} = 98.9\,\mathrm{km^2/h}$, $\kappa = 69.5$ veh/km/lane, γ =2.99, $\rho_{\mathrm{crit},m} = 28.2$ veh/km/lane, $v_{\mathrm{min}} = 15.97$ km/h, $v_{\mathrm{free},m} = 114.8$ km/h, $\alpha = 0.054$.

The corresponding value of the performance function was $J(\theta) = 21.91$ and the average relative error in the TTS resulting from these parameter settings was about 10%.

In qualitative terms, this parameter setting was able to reproduce the shock waves in most cases. For some initial conditions (in the beginning, when the jam was short) the density was not high enough to trigger a shock wave in METANET, but this occurred only for three out of 105 start indices k_{c} in (11).

B. MPC simulation

Several runs were performed with the MPC controller and the result of one of the runs is shown in Fig. 7 and the corresponding speed limits are shown in Fig. 6. The results of the other runs were similar, but not exactly the same because of the random components in the simulation model and in the numerical optimization in the controller. Here we describe the interpretation of these figures in qualitative terms.

After the first shock wave just has entered the link, the speed limits switch on (the region to the left of E in Fig. 6) and create a forward-propagating low-density low-flow wave (the dark line to the left of D in the density and flow plots in Fig. 7). This low-flow wave reduces the inflow to the high-density shock wave sufficiently to eliminate it. At the same time, the speed limits cause an increase in density in the region E, which needs some time to flow out and to disappear. Eventually, all high density regions are eliminated (at approximately k_{c} =35). At the downstream end of the link (in the region above D) it can be seen that resolving the shock wave causes a higher flow (a lighter area in the flow diagram) and that this results in higher densities near the on-ramp (segment 20), but does not create new shock waves. It can also be seen that the outflow in the first hour of the controlled case is higher (lighter in the flow diagram) than in the uncontrolled case.

The controller attempts to eliminate the second shock wave in a similar way: the speed limits cause a low-flow wave that meets the high-density jam in the area to the left from F (as shown the flow diagram, i.e., the rightmost plot in Fig. 7). This eliminates the shock wave, however it also results in a high-density area (above G in the density diagram) that does not resolve. From this area another shock wave emerges that exits the link at k_{c} =120.

It can be concluded that both shock waves were resolved, but with different scenarios for the increased-density area's caused by the speed limits. There may be several reasons for these differences:

- It is unknown what the best possible result is that can be achieved within a given scenario. It may be possible that the current result cannot be improved any further. Currently there are no systematic methods to determine under which conditions the shock waves can be successfully resolved.
- The two shock waves had a different shape, the second one was somewhat longer, and it may have required a longer stretch of controlled segments to eliminate it without creating a new shock wave.
- The model mismatch may have played a role in the inability to stabilize traffic in the second case.

In quantitative terms the average outflow improved from 3483.5 veh/h to 3766.5 veh/h, which is an improvement of 8%.

Comparing the TTS based on the number of vehicles or densities in the link (as is done usually) is not a good measure here, because the controller did not only increase the outflow but also the inflow, because the shock waves did not block the inflow anymore. Therefore, we apply an equivalent formulation of the TTS, where the TTS is expressed as a

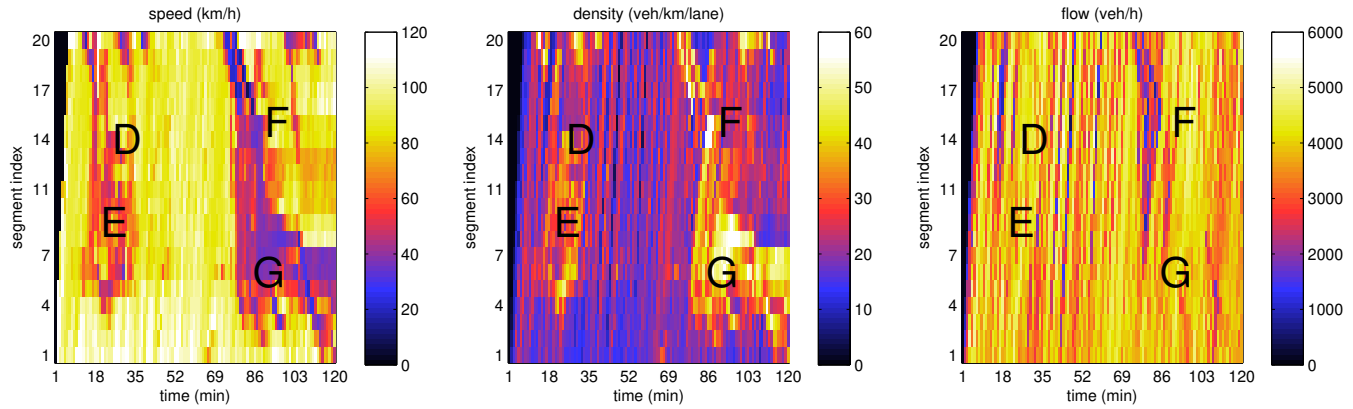

Fig. 7. The results of the MPC-based speed limits. The first shock wave is completely resolved, and the second is also resolved but at the cost of a larger high-density region from which eventually another shock wave emerges.

function of the inflows and outflows:

$$\tilde{J}_{\mathrm{TTS}} = T_{\mathrm{c}} N(0) K + T_{\mathrm{c}}^2 \sum_{k=0}^{K-1} (K-k)\big(q_{\mathrm{demand}}(k) - q_{\mathrm{out}}(k)\big)$$

where $N(0)$ is the initial number of vehicles in the network, $q_{\mathrm{demand}}(k)$ is the traffic demand (veh/h) at segment 1 at time $t = kT_{\mathrm{c}}$, $q_{\mathrm{out}}(k)$ is the outflow (veh/h) at segment 20 at time $t = kT_{\mathrm{c}}$, and K is the last sample step index of the scenario[2].

To make the uncontrolled and the controlled scenarios comparable, we assume that the demand for both scenarios is equal, and that it is also equal to the inflow in the controlled case. The interpretation of this assumption is that there were vehicles that could not enter in the uncontrolled case (by the blocking shock wave) but could enter in the controlled case. All these vehicles contribute in both cases to the TTS until they exit the link. So, this TTS calculation also accounts for the waiting time of these vehicles. This results in a TTS of 1413.2 veh.h for the uncontrolled case, 961.6 veh.h for the controlled case, which is an improvement of nearly 32%.

V. Conclusion

In this paper is we discussed issues and choices that play a role when applying model predictive control in a real-world environment (which is here represented by a micro-simulation model).

We have reported the results of applying model predictive control with METANET as the prediction model to a microscopic traffic flow model as simulation model. In the simulations a scenario of two hours was used with two shock waves propagating over a stretch of freeway of 10 km length. The controller eliminated the shock wave in both cases, but in the second case after some time another shock wave emerged from the increased-density area caused by the speed limits. Despite the imperfect resolution of the second shock wave, the TTS improvement was nearly 32% and the flow improvement 8%. Further research is necessary to investigate the reasons for the different behavior in the two cases.

[2]This expression for the TTS can be derived from $\tilde{J}_{\mathrm{TTS}} = T_{\mathrm{c}} \sum_{k=0}^{K-1} N(k)$ and $N(k) = N(k-1) + T_{\mathrm{c}}(q_{\mathrm{demand}}(k-1) - q_{\mathrm{out}}(k-1))$.

One of the reasons that may have degraded the performance of the controller is the model mismatch between the prediction model and the simulation model. Although the calibrated prediction model was able to reproduce the shock wave occurring in the calibration data in most cases, it also has failed occasionally. To correctly interpret the results of a simulation there is a need to better assess the achievable improvement, based on the observed shock wave and traffic scenario properties.

Future investigations will address these issues, together with the evaluation of the controller for a wider range of scenarios. Furthermore, the same approach will be applied for a real freeway, using the real layout and real traffic data; and the possibilities of a robust MPC approach will be investigated, to deal explicitly with the model mismatch.

VI. Acknowledgments

This research was supported by the BSIK project "Transition Sustainable Mobility (TRANSUMO)" and the Transport Research Centre Delft.

References

[1] H. Zackor, "Self-sufficient control of speed on freeways," in *Proceedings of the International Symposium on Traffic Control Systems*, vol. 2A. Berkeley, California: California University, Aug. 6–9, 1979, pp. 226–249.

[2] S. Smulders, "Control by variable speed signs — the Dutch experiment," in *Proceedings of the Sixth International Conference on Road Traffic Monitoring and Control*, ser. IEE Conference Publication. London: IEE, Apr. 28–30, 1992, pp. 99–103.

[3] E. van den Hoogen and S. Smulders, "Control by variable speed signs: results of the Dutch experiment," in *Proceedings of the 7th International Conference on Road Traffic Monitoring and Control*, ser. IEE Conference Publication No. 391, London, England, Apr. 26-28, 1994, pp. 145–149.

[4] E. J. Hardman, "Motorway speed control strategies using SISTM," in *Road Traffic Monitoring and Control*, ser. Conference Publication, no. 422. IEE, Apr.23–25, 1996, pp. 169–172.

[5] C.-C. Chien, Y. Zhang, and P. A. Ioannou, "Traffic density control for automated highway systems," *Automatica*, vol. 33, no. 7, pp. 1273–1285, 1997.

[6] H. Lenz, R. Sollacher, and M. Lang, "Nonlinear speed-control for a continuum theory of traffic flow," in *14th World Congress of IFAC*, vol. Q, Beijing, China, 1999, pp. 67–72.

[7] ——, "Standing waves and the influence of speed limits," in *Proceedings of the European Control Conference 2001*, Porto, Portugal, 2001, pp. 1228–1232.
[8] S. Smulders, "Control of freeway traffic flow by variable speed signs," *Transportation Research Part B*, vol. 24B, no. 2, pp. 111–132, 1990.
[9] Rijkswaterstaat Directie Zuid-Holland, "Evaluatie 80 km/uur-maatregel A13 Overschie, Doorstroming and Verkeersveiligheid (Evaluation of the 80 km/h measure on the A13 at Overschie, Traffic Flow and Safety)," Ministry of Transport, Public Works and Water Management, Rijkswaterstaat, Directie Zuid-Holland - VIV, Tech. Rep., 2003, in Dutch.
[10] J. Wesseling, K. Hollander, S. Teeuwisse, M. Keuken, H. Spoelstra, R. Gense, E. Burgwal, L. Hermans, J. Voerman, P. Kummu, and J. Elshout, "Onderzoek naar de effecten van de 80 km/h-maatregel voor de A13 op de luchtkwaliteit in Overschie (Investigation of the effects of the 80 km/h measure at the A13 on the air quality in Overschie)," TNO (on behalf of Rijkswaterstaat, Directie Zuid-Holland), Tech. Rep., 2003, in Dutch.
[11] J. K. Wilkie, "Using variable speed limit signs to mitigate speed differentials upstream of reduced flow locations," Department of Civil Engineering, Texas A& M University, College Station, Texas 77843, Tech. Rep., Aug. 1997, prepared for CVEN 677 Advanced Surface Transportation Systems.
[12] P. Y. Li, R. Horowitz, L. Alvarez, J. Frankel, and A. M. Robertson, "Traffic flow stabilization," in *Proceedings of the American Control Conference*, Seattle, Washington, June 1995, pp. 144–149.
[13] A. Alessandri, A. Di Febbraro, A. Ferrara, and E. Punta, "Nonlinear optimization for freeway control using variable-speed signaling," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 6, pp. 2042–2052, Nov. 1999.
[14] A. Di Febbraro, T. Parisini, S. Sacone, and R. Zoppoli, "Neural approximations for feedback optimal control of freeway systems," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 1, pp. 302–312, Jan. 2001.
[15] P. Ioannou, "Intelligent vehicles: closing the loop with the highway," in *Proceedings of the 11th IFAC Symposium on Control in Transportation Systems*, H. van Zuylen and F. Middelham, Eds., no. 265, Delft, The Netherlands, Aug. 2006, on CD.
[16] A. Hegyi, B. De Schutter, and J. Hellendoorn, "Optimal coordination of variable speed limits to suppress shock waves," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 1, pp. 102–112, Mar. 2005.
[17] E. Camacho and C. Bordons, *Model Predictive Control in the Process Industry*. Berlin, Germany: Springer-Verlag, 1995.
[18] J. Maciejowski, *Predictive Control with Constraints*. Harlow, England: Prentice Hall, 2002.
[19] F. Allgöwer, T. Badgwell, J. Qin, J. Rawlings, and S. Wright, "Nonlinear predictive control and moving horizon estimation – An introductory overview," in *Advances in Control: Highlights of ECC '99*, P. Frank, Ed. Springer, 1999, pp. 391–449.
[20] M. Papageorgiou, J.-M. Blosseville, and H. Haj-Salem, "Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris: Part II: coordinated on-ramp metering," *Transportation Research Part A*, vol. 24A, no. 5, pp. 361–370, 1990.
[21] Technical University of Crete and A. Messmer, *METANET – A simulation program for motorway networks*, Technical University of Crete, Dynamic Systems and Simulation Laboratory and A. Messmer, Nov. 2001.
[22] A. Hegyi, "Model predictive control for integrating traffic control measures," Ph.D. thesis, TRAIL Thesis Series T2004/2, Delft University of Technology, Delft, The Netherlands, Feb. 2004, ISBN 90-5584-053-X.
[23] Quadstone, "Paramics v5.1," 2006. [Online]. Available: http//:www.paramics-online.com