Delft University of Technology

Delft Center for Systems and Control

Technical report 03-001

Suppressing shock waves on the A1 in The Netherlands – Model calibration and model-based predictive control*

A. Hegyi, D. Ngo Duy, B. De Schutter, J. Hellendoorn, S.P. Hoogendoorn, and S. Stramigioli

If you want to cite this report, please use the following reference instead:
A. Hegyi, D. Ngo Duy, B. De Schutter, J. Hellendoorn, S.P. Hoogendoorn, and S. Stramigioli, "Suppressing shock waves on the A1 in The Netherlands – Model calibration and model-based predictive control," *Proceedings of the IEEE 6th International Conference on Intelligent Transportation Systems (ITSC'03)*, Shanghai, China, pp. 672–677, Oct. 2003.

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.bartdeschutter.org/abs/03_001.html

Suppressing Shock Waves on the A1 in The Netherlands – Model Calibration and Model-Based Predictive Control

A. Hegyi*, D. Ngo Duy†, B. De Schutter*, J. Hellendoorn*, S.P. Hoogendoorn† and S. Stramigioli‡

*Delft Center for Systems and Control, Delft Univ. of Technology, Mekelweg 2
2628 CD Delft, The Netherlands. Email: {A.Hegyi,B.DeSchutter,J.Hellendoorn}@dcsc.tudelft.nl

†Transportation Planning and Traffic Engineering Section, Fac. Civil Eng. & Geosciences, Delft Univ. of Technology
P.O. Box 5048, 2600 GA Delft, The Netherlands. Email: {D.Ngoduy,S.P.Hoogendoorn}@ct.tudelft.nl

‡Control Laboratory, EL/CE, Fac. of Electrical Eng., Mathematics & Computer Science, Univ. of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands. Email: S.Stramigioli@utwente.nl

***Abstract*—We now apply the model predictive control (MPC) of speed limits that we have presented in previous publications [1], [2] to a calibrated METANET model of a 19 km stretch of the real-world freeway A1 in The Netherlands. This freeway regularly suffers from shock waves originating mainly from on-ramps, and speed limits are now used to suppress these shock waves. First, we calibrate and validate the extended METANET model with data from the A1 freeway, and we use the Delft OD method to estimate the origin-destination patterns that are needed for the simulation of the destination oriented traffic. Next, we verify from data whether the necessary conditions for applying speed limits against shock waves are satisfied. We show that the MPC controller performs well even under the assumption that the traffic demand is not known on the on-ramps and is known for only a few kilometers upstream and downstream of the controlled stretch. This approach results in an improvement of the total time spent in the network with about 15 %.**

I. Introduction

A. Shock waves

Shock waves on freeways are often created at bottlenecks, and can remain existent for a long time and propagate upstream many kilometers from the location of their creation. Besides the fact that shock waves can trigger a new shock wave and/or cause traffic jams, they are potentially unsafe and lead to increased travel times.

In previous publications [1], [2] we have presented a synthetic case study, where we applied a methodology called model predictive control (MPC) to control speed limits in order to suppress or to eliminate shock waves on the freeway. The results of the synthetic study are very promising: the total time that vehicles spend in the link (TTS) was reduced by 15-20 %, but there were some assumptions made about modeling and traffic scenario (traffic demand) that were not yet validated with real data. In this paper we apply the same methodology to control speed limits, but we also verify the traffic conditions necessary for speed limit control against shock waves and calibrate/validate the model based on real data from the Dutch A1 freeway. Furthermore, we demonstrate the approach with a traffic scenario taken from real data. The use of real data results in a better assessment of the improvement that can be achieved in practice.

Dynamic speed limits can be used for several purposes, such as: homogenization, prevention of breakdown, reduction of pollution and noise, increased safety, reduction of the number of incidents, etc. In this paper we consider the traffic breakdown prevention approach, which focuses on preventing too high densities so as to improve traffic flow.

B. Coordination and prediction

In practice, dynamic traffic management often operates based on local data only. However, considering the effect of the measures on the network level has in general many advantages compared to local control. Hence, a network-wide *coordination* of the control measures based on global data is certainly useful. The coordination of the control signals is obtained by reformulating the control design problem over a given time horizon as an optimization problem that yields the optimal speed limit settings (see Section II-B1).

Since we want to determine control signal settings that are optimal for a given freeway network and since the effect of a given control measure on more distant parts of the network might only be visible or measurable after some time, an accurate *prediction* of the future evolution of the traffic flows in the network is necessary. In particular, prediction is needed for two reasons: first, if the formation or the arrival of a shock wave in the controlled area can be predicted, then preventive measures can be taken. Second, the positive effect of speed limits on the traffic flow can not be observed instantaneously, so the prediction should be made at least up to the point where the improvement can be observed.

Besides prediction and coordination the speed limit control problem has some other characteristics that impose specific requirements to the control strategy:

- There is a direct relation between the outflow of a

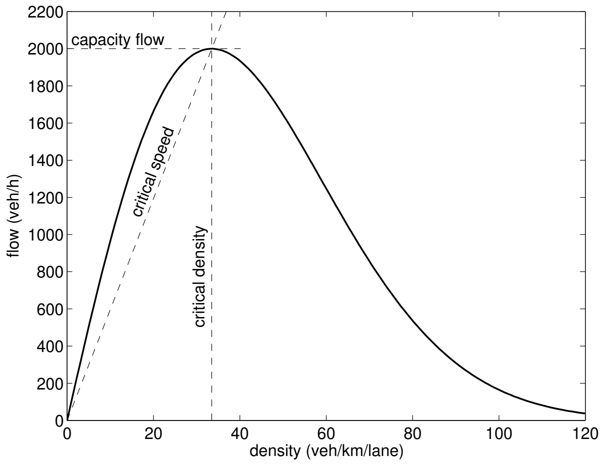

Figure 1. A typical example of the fundamental diagram, which represents the relation between density and flow for a given freeway section. The slope of the line connecting the origin and a point on the fundamental diagram represents the speed corresponding to that point. The critical speed is the speed that corresponds to capacity flow.

network and the total time spent (TTS) in the network, assuming that the traffic demand is fixed. Papageorgiou [3] showed that in a traffic network an increase of outflow of 5 % may result in an decrease of the TTS of 20 %. This effect can be explained by the fact that the number of vehicles in the network is equal to the accumulated net inflow of the network. But the congestion after a breakdown usually has an outflow that is about 5–10 % lower than the available capacity (this is the so-called *capacity drop phenomenon*, so the queue grows faster, and consequently congestion will last longer, and the outflow will be low for a longer time (the time that the queue needs to dissolve). So we can conclude that any control method that resolves (or reduces) congestion will at best achieve a flow improvement of approximately 5–10 %, but this improvement can decrease the TTS significantly. Furthermore, since this flow improvement is relatively small, and since there are always disturbances present in the traffic flow, *feedback control* is required, i.e., the result of any control action is monitored and fed back into the controller. In this way imprecisions of control and traffic disturbances can be observed and appropriate control actions can be taken.

- The speed limit signs used in practice display speed limits in increments of, e.g., 10 or 20 km/h. Therefore, the controller should produce discrete-valued control signals.
- For safety it is often required that the driver should not encounter a decrease in the speed limit larger than a prespecified amount. The controller should be able to take this kind of constraints into account.

The MPC control strategy we use addresses these issues.

C. Speed limits against shock waves

It is well known [4] that some types of traffic jams move upstream with approximately 15 km/h. These moving jams are called waves. As they can remain existent for a long time, every vehicle that enters the freeway upstream of the congested area will have to pass through the jammed area, which increases the travel time. Furthermore, these moving jams may cause unsafe situations. Lighthill and Whitham [5] introduced the term shock wave for waves that are formed by several waves running together. At the shock wave fairly large reductions in velocity occur very quickly. In this paper we will use the term "shock wave" for any wave (the moving congested areas) and not distinguish between waves and shock waves, because in practice any wave is undesired.

Based on previous experiments [1], [2] the following qualitative description can be given of how the MPC control of speed limits resolves shock waves. On some sections upstream of a shock wave speed limits are imposed with a value that is low enough to ensure that the inflow of the jammed area is smaller than its outflow. Consequently, the density in the jammed area will decrease and the jam will eventually dissolve. The speed limits create a low-density wave (with a density that is lower than it would be in the uncontrolled situation) that propagates downstream. This low-density wave meets the shock wave and compensates its high density, which reduces or eliminates the shock wave. A point of criticism could be that this approach reduces the shock wave, but at the cost of creating new shock waves upstream of the sections controlled by speed limits. However, as the speed limits are optimized properly, they will never create a shock wave that gives rise to higher delays than in the uncontrolled case.

II. A MODEL-BASED PREDICTIVE CONTROL APPROACH

A. Prediction model

The model-based predictive control procedure considered in this paper requires a prediction of the network evolution, for which we use a slightly modified version of the (destination-oriented) METANET model [6], [7]. Below we will briefly described the (extended) METANET model, limiting ourselves to only those parts of the model that are relevant for the case study of Section III below.

1) Original METANET model: The METANET model represents a network as a directed graph with the links corresponding to freeway stretches. Each freeway link has uniform characteristics, i.e., no on-ramps or off-ramps and no major changes in geometry. Where major changes occur in the characteristics of the link or in the road geometry (e.g., an on-ramp or an off-ramp), a node is placed. Each link m is divided into N_m segments of length L_m (see Figure 2). Each segment i of link m is characterized by the *partial traffic density* $\rho_{m,i,j}(k)$ (veh/lane/km), expressing the portion of traffic heading to destination j, by the *mean speed* $v_{m,i}(k)$ (km/h), and the *flow* $q_{m,i}(k)$ (veh/h), where k indicates the time instant $t = kT$, and T is the simulation time step (typically $T = 10\,\mathrm{s}$). Furthermore, the *total density* $\rho_{m,i}$ is defined as:

$$\rho_{m,i}(k) = \sum_{j \in J_m} \rho_{m,i,j}(k) \ ,$$

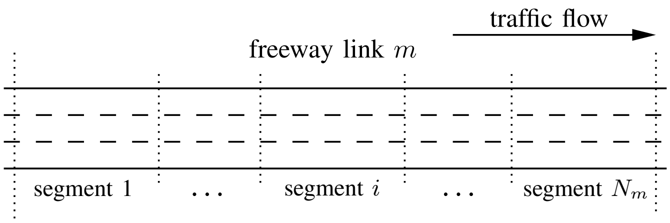

Figure 2. In the METANET model, a freeway link is divided into segments.

where J_m is the set of destinations reachable from link m. We have

$$q_{m,i}(k) = \rho_{m,i}(k)\, v_{m,i}(k)\, \lambda_m \quad , \tag{1}$$

where λ_m denotes the number of lanes of segment m. The principle of conservation of vehicles yields:

$$\rho_{m,i,j}(k+1) = \rho_{m,i,j}(k) + \frac{T}{L_m \lambda_m}\big(\gamma_{m,i-1,j}(k)\, q_{m,i-1}(k) - \gamma_{m,i,j}(k)\, q_{m,i}(k)\big) \quad , \tag{2}$$

where $\gamma_{m,i,j}(k) = \rho_{m,i,j}(k)/\rho_{m,i}(k)$. The mean speed depends on the previous mean speed plus a relaxation term, a convection term, and an anticipation term:

$$v_{m,i}(k+1) = v_{m,i}(k) + \frac{T}{\tau}\Big(V\big(\rho_{m,i}(k)\big) - v_{m,i}(k)\Big) + \frac{T}{L_m} v_{m,i}(k)\big(v_{m,i-1}(k) - v_{m,i}(k)\big) - \frac{\eta T}{\tau L_m} \frac{\rho_{m,i+1}(k) - \rho_{m,i}(k)}{\rho_{m,i}(k) + \kappa} \quad , \tag{3}$$

where τ, η and κ are model parameters, and

$$V\big(\rho_{m,i}(k)\big) = v_{\text{free},m} \exp\left[-\frac{1}{a_m}\left(\frac{\rho_{m,i}(k)}{\rho_{\text{crit},m}}\right)^{a_m}\right] \quad , \tag{4}$$

with a_m a model parameter, $v_{\text{free},m}$ the free-flow speed, and $\rho_{\text{crit},m}$ the critical density.

Origins are modeled with a destination dependent queue model. The evolution of queue length $w_o(k)$ at origin o is described by

$$w_{o,j}(k+1) = w_{o,j}(k) + T\big(\gamma_{o,j}(k)\, d_o(k) - \gamma_{o,j}(k)\, q_o(k)\big),$$

where $d_o(k)$ is the traffic demand at origin o, $\gamma_{o,j}(k)$ the fraction of the demand traveling to destination j, and $q_o(k)$ the outflow of origin o, which is given by

$$q_o(k) = \min\left[d_o(k) + \frac{w_o(k)}{T}, Q_o \frac{\rho_{\max} - \rho_{\mu,1}(k)}{\rho_{\max} - \rho_{\text{crit},\mu}}\right], \tag{5}$$

where Q_o is the on-ramp capacity (veh/h) under free-flow conditions, $\rho_{\max}$ is the maximum density, μ the index of the link to which the on-ramp is connected, and $w_o(k) = \sum_{j \in J_{\text{o},o}} w_{o,j}(k)$, where $J_{\text{o},o}$ is the set of destinations reachable from origin o.

The coupling equations that connect links are as follows. A node provides the incoming links with a downstream density, and the leaving links with an inflow and a upstream speed. More specifically, the flow that enters node n is distributed among the leaving links according to

$$Q_{n,j}(k) = \sum_{\mu \in I_n} q_{\mu,N_\mu}(k)\, \gamma_{\mu,N_\mu,j}(k)$$

$$q_{m,0}(k) = \sum_{j \in J_m} Q_{n,j}(k)\, \beta_{n,j,m}(k)$$

$$\gamma_{m,0,j}(k) = \beta_{n,j,m}(k)\, Q_{n,j}(k)/q_{m,0}(k)$$

for all $m \in O_n$, $j \in J_m$ where $Q_{n,j}(k)$ is the total flow with destination j flow that enters node n at time k, I_n is the set of links that enter node n, O_n is the set of links leaving node n, $\beta_{n,j,m}(k)$ is the split rate (the fraction of the total flow through node n with destination j that leaves via link m), and $q_{m,0}(k)$ is the flow that leaves node n via link m. When node n has more than one leaving link, the virtual downstream density $\rho_{m,N_m+1}(k)$ of entering link m is given by

$$\rho_{m,N_m+1}(k) = \frac{\sum_{\mu \in O_n} \rho^2_{\mu,1}(k)}{\sum_{\mu \in O_n} \rho_{\mu,1}(k)} \quad .$$

2) Extensions: Since the original METANET model does not describe the effect of speed limits, we modify the equation for the desired speed (4) to incorporate speed limits. The second extension regards the modeling of a mainstream origin, which has a different nature than an on-ramp origin. The third extension describes the different effects of a positive or negative downstream density gradient on the speed (cf. the anticipation term in (3)). Finally, the fourth extension is made to be able to use the available speed and flow data to express incoming shock waves. For a more extensive motivation of extensions 1 to 3 we refer to [1], [2].

1) To model the effects of the speed limits, we assume that

$$V\big(\rho_{m,i}(k)\big) = \min\Bigg((1+\alpha) v_{\text{ctrl},m,i}(k), v_{\text{free},m} \exp\left[-\frac{1}{a_m}\left(\frac{\rho_{m,i}(k)}{\rho_{\text{crit},m}}\right)^{a_m}\right]\Bigg) \quad , \tag{6}$$

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

2) A mainstream origin link is modeled by

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

where $q_{\text{lim},\mu,1}(k)$ is the maximal inflow determined by the speed limit or the traffic situation in the first segment of link μ (see [2]).

3) In (3), η is a global parameter and has the same value for all segments. However, here we take different values for $\eta_{m,i}(k)$, depending on whether the downstream density is higher or lower than the density in the actual segment:

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

4) For the (virtual) downstream density of the last segment before the destination we assume that the destination is in free-flow:

$$\rho_{m,N_m+1}(k+1) = \begin{cases} \rho_{m,N_m}(k) & \text{if } \rho_{m,N_m}(k) \leq \rho_{\mathrm{crit},m} \\ \rho_{\mathrm{crit},m} & \text{if } \rho_{m,N_m}(k) > \rho_{\mathrm{crit},m} \end{cases}$$

When the free-flow assumption does not hold (i.e., $v_d(k+1) < v_{\mathrm{crit},m}$) , and the speed $v_d(k)$ and flow $q_d(k)$ of destination d is available, the following equations are used in addition to (1), (3):

$$q_{\mathrm{new},m,N_m}(k+1) = \begin{cases} q_d(k+1) & \text{if } v_d(k+1) < v_{\mathrm{crit},m} \\ & \text{and } q_d(k+1) < q_{m,N_m}(k+1) \\ q_{m,N_m}(k+1) & \text{otherwise,} \end{cases}$$

$$v_{\mathrm{new},m,N_m}(k+1) = \begin{cases} \frac{q_d(k+1)}{q_{m,N_m}} v_{m,N_m}(k+1) & \text{if } v_d(k+1) < v_{\mathrm{crit},m} \text{ and} \\ & q_d(k+1) < q_{m,N_m}(k+1) \\ v_{m,N_m}(k+1) & \text{otherwise,} \end{cases}$$

where $q_{\mathrm{new},m,N_m}(k)$ and $v_{\mathrm{new},m,N_m}(k)$ are respectively the outflow and the speed of the last segment of link m that are used in (1), (2).

B. Model Predictive Control

1) Approach: We use a model predictive control (MPC) scheme to solve the problem of optimal coordination of speed limits. In MPC, at each time step k the optimal control signal is computed (by numerical optimization) over a prediction horizon N_p. A control horizon N_c ($< N_\mathrm{p}$) is selected to reduce the number of variables and to improve the stability of the system: after the control horizon has been passed, the control signal is usually taken to be constant. From the resulting optimal control signal only the first sample $k+1$ is applied to the process. In the next time step $k+1$, a new optimization is performed (with a prediction horizon that is shifted one time step further) 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. Updating the state results in a controller that has a low sensitivity to prediction errors, and updating the model results in an adaptive control system, which could be useful in situations where the model significantly changes, such as in case of incidents or changing weather conditions. For more information on MPC we refer the interested reader to [8], [9] and the references therein.

2) Objective function: We consider the objective function

$$J(k) = T \sum_{l=k}^{k+N_\mathrm{p}-1} \left\{ \sum_{(m,i)\in I_{\mathrm{all}}} \rho_{m,i}(l) L_m \lambda_m + \sum_{o\in O_{\mathrm{all}}} w_o(l) \right\} + a_{\mathrm{speed}} \sum_{l=k}^{k+N_\mathrm{c}-1} \sum_{(m,i)\in I_{\mathrm{speed}}} \left(\frac{v_{\mathrm{ctrl},m,i}(l) - v_{\mathrm{ctrl},m,i}(l-1)}{v_{\mathrm{free},m}} \right)^2 ,$$

where O_{all} is the set of indices of all origins, and I_{speed} is the set of pairs of indices (m,i) of the links and segments where speed control is applied. This objective function contains a term for the TTS (on the freeway and in the origin queues), and a term that penalizes abrupt variations in the speed limit control signal. The nonnegative parameter a_{speed} expresses the relative importance of each term.

3) Constraints: In general, for the safe operation of a speed control system, it is required that the maximum decrease in speed limits that a driver can encounter (v_{maxdiff}) is limited. There are three situations where a driver can encounter a different speed limit value: (1) when the speed limit changes in a given segment (and there are more speed limit signs on the same segment), (2) when a driver enters a new segment, (3) when the driver enters a new segment and the speed limit changes. The resulting maximum speed difference constraints are:

$$v_{\mathrm{ctrl},m,i}(l-1) - v_{\mathrm{ctrl},m,i}(l) \leq v_{\mathrm{maxdiff}}$$
$$v_{\mathrm{ctrl},m,i}(l) - v_{\mathrm{ctrl},m,i+1}(l) \leq v_{\mathrm{maxdiff}}$$
$$v_{\mathrm{ctrl},m,i}(l-1) - v_{\mathrm{ctrl},m,i+1}(l) \leq v_{\mathrm{maxdiff}}$$

for the appropriate indices m, i, and for $l \in [k, \ldots, k+N_\mathrm{c}-1]$. In addition to the above safety constraints the speed limits are often subject to a minimum value v_{ctrlmin}:

$$v_{\mathrm{ctrl},m,i}(l) \geqslant v_{\mathrm{ctrlmin}}$$

for $l \in [k, \ldots, k+N_\mathrm{c}-1]$. In practice, the variable speed limit signs display speed limits in increments of, e.g., 10 or 20 km/h. Therefore, the controller should produce discrete-valued control signals. This is expressed by the constraint

$$v_{\mathrm{ctrl},m,i}(l) \in \mathcal{V}_{m,i}$$

for $l \in [k, \ldots, k+N_\mathrm{c}-1]$, where $\mathcal{V}_{m,i}$ is the set of possible discrete speed limit values.

III. Case study

A. Network description

The freeway stretch under consideration is part of the A1 freeway in The Netherlands. It has a length of 19 km (see Figure 3), one mainstream origin and one mainstream destination, six on-ramps, six off-ramps, and two lanes for the whole stretch. The inductive loop-detectors measure average speed and flow (over 1 min), and are located at approximately every 500 m, except on one location where their distance is approximately at 200 m. Near the downstream end of the stretch there is a bridge of 1 km. The on-ramps and off-ramps are not equipped with loop detectors, except for O1, D1, O2 and D2, which connect the A1 stretch to another freeway.

B. Traffic scenario

Figure 4 shows a typical traffic scenario. Low-speed regions can be seen at and upstream of the main on-ramps at 96.2 km and 90.3 km. The shock waves are created at the on-ramps and propagate backwards with a speed of approximately 20 km/h.

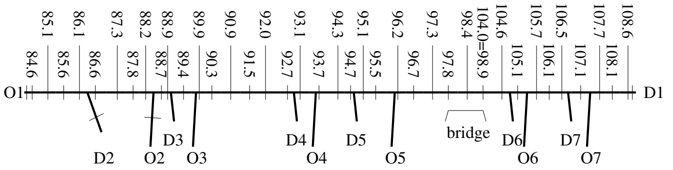

Figure 3. A schematic view of the considered stretch of the A1. The detector locations are in units of km's, and the on/off-ramps are indicated by respectively O1–O7 and D1–D7.

C. Calibration

As mentioned in Section III-A six of the seven on-ramps and off-ramps complexes are not equipped with detectors. As a consequence, the origin-destination (OD) flows need to be estimated before the traffic flow model can be calibrated. The OD relations are estimated with the Delft OD method (see [10]) that uses statistical techniques to extract time dependent OD relationships based on detector data.

For the calibration (parameter identification) we construct a nonlinear least-squares minimization problem in which we (numerically) minimize

$$\sum_{l=1}^{N_{\text{samp}}} \sum_{(m,i)\in I_{\text{all}}} \big((\hat{q}_{m,i}(l) - \tilde{q}_{m,i}(l)\big)^2 + \xi\big((\hat{v}_{m,i}(l) - \tilde{v}_{m,i}(l)\big)^2$$

with N_{samp} is the number of data samples, I_{all} is the set of indices of all pairs of links and segments, $\tilde{q}_{m,i}(l)$ and $\tilde{v}_{m,i}(l)$ denote the measured flow and speed data, ξ is a tuning weight, and l corresponds to the time instant $t = lT_{\text{samp}}$ where T_{samp} is the sampling time. We choose T_{samp} and T such that $T_{\text{samp}}/T \in \mathbb{N}$, and we compute the simulated values $\hat{q}_{m,i}(l)$ and $\hat{v}_{m,i}(l)$ as

$$\hat{q}_{m,i}(l) = \frac{T}{T_{\text{samp}}} \sum_{k=\frac{T_{\text{samp}}}{T}l}^{\frac{T_{\text{samp}}}{T}(l+1)-1} q_{m,i}(k)$$

$$\hat{v}_{m,i}(l) = \frac{T}{T_{\text{samp}}} \sum_{k=\frac{T_{\text{samp}}}{T}l}^{\frac{T_{\text{samp}}}{T}(l+1)-1} v_{m,i}(k) \quad .$$

D. Verification of the conditions for successful speed limit control

In the synthetic study [1], [2] several assumptions were made that need to be satisfied when applying speed limit control to a real-world scenario:

- The traffic situation downstream the controlled area should be generally congestion-free, since otherwise any improvement achieved by efficient control will cause a more severe congestion downstream. This condition is verified by visual inspection of traffic data (speed and flow) in figures similar to Figure 4.
- Capacity drop (cf. Section I-B) should be observed in the traffic data, and the traffic model used in the MPC controller should be able to reproduce the capacity drop. Otherwise, if there is no capacity drop, the outflow of the shock wave will equal the maximal flow and there will be no possibility to improve the flow any further.
The capacity drop is estimated by comparing outflow of a shock wave with the maximum flow of freely flowing traffic. The time and location for the outflow of the shock wave have to be such that there are no on-ramps or off-ramps between the shock wave and the measurement point, otherwise the entering or exiting traffic could bias the estimation. Another condition is that the traffic should be in free flow, to be sure that the flow drop is not caused by a downstream bottleneck.
We explain the capacity drop estimation by an example. In Figure 5 (which is a zoom-in of Figure 4) a shock wave is shown, which starts at point A and is caused by excessive on-ramp traffic just before 96.2 km (see Figures 4 and 3). There is an off-ramp just after 92.7 km, and on-ramps before respectively 89.9 km and 88.7 km, which can be recognized by the sudden increase of the flow at the detector locations after the on-ramps. Between 89.9 km and 92.7 km there are no on-ramps or off-ramps and this is the area that we consider in this example for capacity drop estimation[1]. The area where the traffic is in free flow again is around point B: the speed is approximately 100 km/h and the flow around 3000 veh/h. An example of the capacity flow is around point C, where the flow is approximately 4200 veh/h. We can conclude that there is a capacity drop of roughly 30 %.
- The minimum value of the speed limit should result in a flow that is lower than the outflow of a congested area (a shock), otherwise the density will not decrease even when the speed limit is set to its minimum value. This condition is satisfied as the lowest speed limit is 50 km/h,

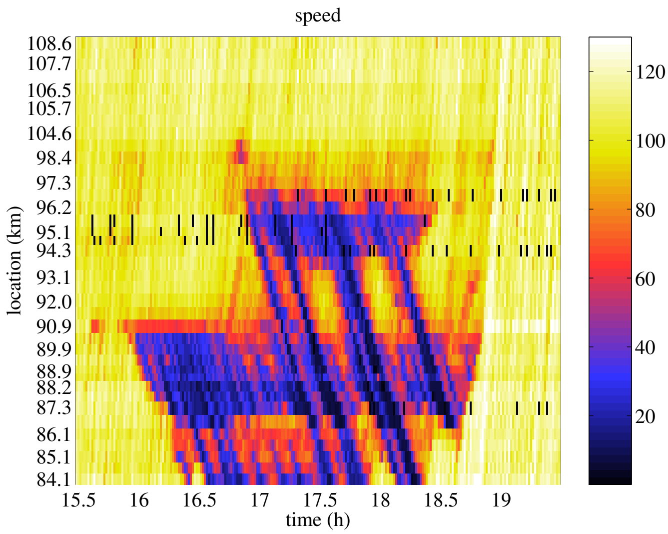

Figure 4. A typical traffic scenario on the A1. The shock waves usually start at on-ramps and propagate backwards through the link.

[1]It is interesting to note in Figure 5 that after the shock wave has passed the on-ramp upstream of 88.7 km (which connects another freeway) the flow suddenly increases because of the additional vehicles from the on-ramp.

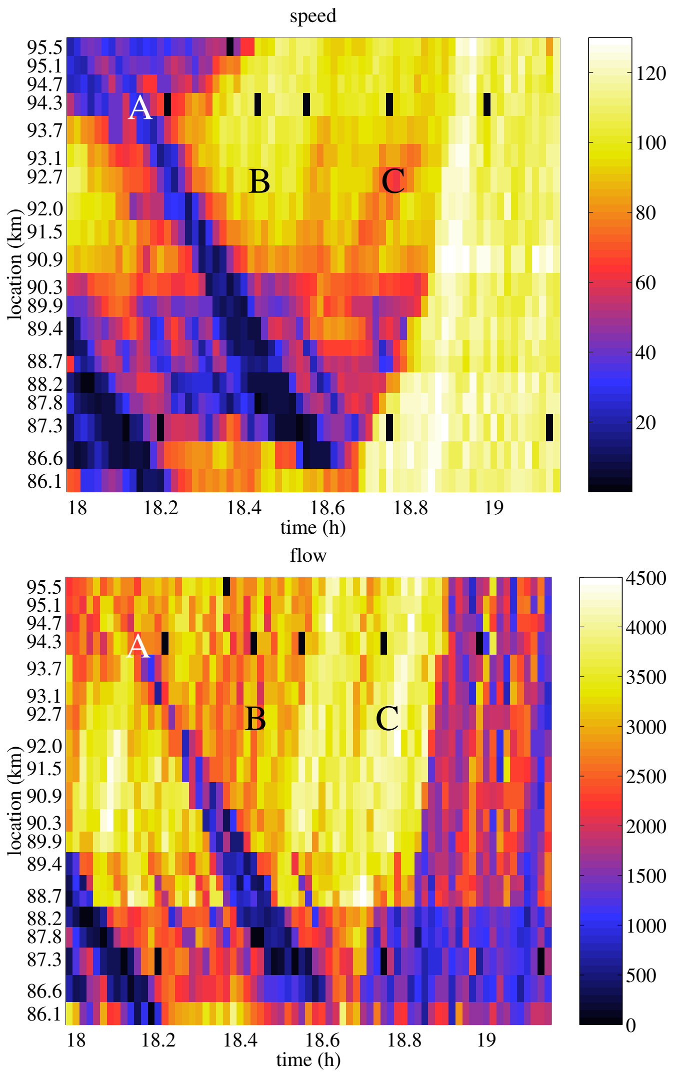

Figure 5. Example of the capacity drop.

and the average flow at this speed is 2900 veh/h.

- Traffic flow should be in metastable state[2], and the traffic model should be able to represent this. In other words, the speed limit controller should have the possibility to 'convert' the (unstable) shock wave into a wider, less intense and stable wave. It is unknown how to verify this precondition from the data. However, is seems plausible to say that this precondition is satisfied if there is a capacity drop. For this, we assume that the (reduced) outflow of a congested area is stable (in the sense that a reasonable disturbance will not create a new congestion) and flow at capacity is marginally stable which means that even a very small disturbance can cause a breakdown. It seems reasonable to expect the closer the flow is to capacity, the smaller the disturbance needed cause a breakdown, which means metastability.

[2]Metastability means that small disturbances in the traffic flow disappear, but large disturbances cause instability, shock waves, or congestion.

E. MPC results

Based on the theoretical results of Section III-D a reduction of shock waves and an improvement of the total time spent is expected, as the conditions for successful control with speed limits are satisfied. This is confirmed by simulation results, which show that the MPC approach for the chosen scenario results in significantly less shock waves and gives an improvement of the TTS of up to 15 %.

IV. Conclusions

We have applied model predictive control to optimally coordinate variable speed limits to suppress/eliminate shock waves. The objective of the controller was to find the control signals that minimize the total time that vehicles spend in the network. The control was applied to a calibrated model of a 19 km stretch of the A1 freeway in The Netherlands. Constraints were incorporated that prevent sudden drops of speed limits and thus improve safety for the drivers. Furthermore, before applying speed limit control against shock waves, the necessary conditions for successful control were verified.

Acknowledgments

Research supported by the Traffic Research Centre (AVV) of the Dutch Ministry of Transport, Public Works and Water Management, and by the NWO-CONNEKT project "Advanced multi-agent control and information for integrated multi-class traffic networks (AMICI)".

References

[1] P. Breton, A. Hegyi, B. De Schutter, and H. Hellendoorn, "Shock wave elimination/reduction by optimal coordination of variable speed limits," in *Proceedings of the IEEE 5th International Conference on Intelligent Transportation Systems (ITSC'02)*, Singapore, Sep. 2002, pp. 225–230.

[2] A. Hegyi, B. De Schutter, and J. Hellendoorn, "MPC-based optimal coordination of variable speed limits to suppress shock waves in freeway traffic," in *Proceedings of the 2003 American Control Conference*, Denver, Colorado, Jun. 2003, pp. 4083–4088.

[3] M. Papageorgiou, J. Blosseville, and H. Hadj-Salem, "La fluidification des rocades de l'Ile de France: Un projet d'importance," Dynamic Systems and Simulation Laboratory, Technical University of Crete, Chania, Greece, Tech. Rep. No. 1998-17, 1998.

[4] B. Kerner and H. Rehborn, "Experimental features and characteristics of traffic jams," *Physical Review E*, vol. 53, no. 2, pp. R1297–R1300, Feb. 1996.

[5] M. Lighthill and G. Whitham, "On kinematic waves: II. A theory of traffic flow on long crowded roads," *Proceedings of the Royal Society of London*, vol. 299A, pp. 317–345, May 1955.

[6] A. Kotsialos, M. Papageorgiou, and A. Messmer, "Optimal coordinated and integrated motorway network traffic control," in *Proceedings of the 14th International Symposium of Transportation and Traffic Theory (ISTTT)*, Jerusalem, Israel, Jul. 1999, pp. 621–644.

[7] M. Papageorgiou, J. Blosseville, and H. Hadj-Salem, "Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris: Part I: Modelling," *Transportation Research Part A*, vol. 24, no. 5, pp. 345–359, Sep. 1990.

[8] E. Camacho and C. Bordons, *Model Predictive Control in the Process Industry*. Berlin, Germany: Springer-Verlag, 1995.

[9] J. Maciejowski, *Predictive Control with Constraints*. Harlow, UK: Prentice Hall, 2002.

[10] N. van der Zijpp, "Dynamic origin-destination matrix estimation on motorway networks," Ph.D. dissertation, Delft University of Technology, Delft, The Netherlands, 1996.