

Model predictive control for ramp metering of motorway traffic: A case study

T. Bellemans^{a,*}, B. De Schutter^b, B. De Moor^a

^a*ESAT-SCD (Signals, Identification, System Theory and Automation), Katholieke Universiteit Leuven, Kasteelpark Arenberg10, B-3001 Leuven, Belgium*

^b*Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands*

Received 15 November 2002; accepted 15 March 2005

Available online 10 May 2005

Abstract

A real-life motorway in Belgium is studied and a comparison is made between a simulation of a morning rush hour situation without control and a simulation of a morning rush hour situation with ramp metering implemented. Two types of controllers are compared: a traditional ALINEA based controller and a model predictive control based ramp metering controller. In order to evaluate the controllers in a realistic framework, the simulations presented in this paper are based on real-life traffic measurements, and constraints on the maximal allowed queue lengths at the on-ramps are imposed. The presented simulations are indicative for the reduction in the total time spent (on the studied motorway and on the on-ramps) that can be achieved by ramp metering during a typical morning rush hour.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Model predictive control; Traffic control; Ramp metering

1. Introduction

Many countries around the world invest large amounts of resources in attempts to reduce the occurrence of congestion and as such its negative impact on e.g., traffic safety, the environment (air pollution, wasted fuel) and the quality of life (health problems, noise, stress). Since the construction of new roads is not always a viable option due to economical and environmental issues, other solutions are needed. One solution that can be implemented in the short term is dynamic traffic control. Dynamic traffic control is a traffic responsive control method that takes the variations of the traffic situation over time into account. These variations in the traffic state can result from a changing

traffic demand during rush hours, accidents, manifestations, ... The research on and implementation of dynamic traffic control systems is aimed at increasing the traffic operation efficiency without building new roads.

This paper focuses on increasing the efficiency of traffic operations on motorways using dynamic control. One way to control traffic on a motorway is ramp metering or admission control. The control signals or metering rates can be obtained by using traditional PID-like control (ALINEA) or they can be optimized in a receding horizon framework as will be presented later.

This paper is organized as follows: First, the concept of ramp metering for control of motorway traffic is presented followed by two methods to calculate dynamic or traffic responsive metering rates. The method presented first consists of a traffic regulator ALINEA while the second method relies on a receding horizon framework. In the third section a motorway in Belgium is considered as a real-life case study. Traffic simulations

*Corresponding author. Tel.: +32 16 32 17 09; fax: +32 16 32 19 70.

E-mail addresses: Tom.Bellemans@esat.kuleuven.ac.be (T. Bellemans), b.deschutter@dsc.tudelft.nl (B. De Schutter), Bart.DeMoor@esat.kuleuven.ac.be (B. De Moor).

of the case study illustrate the positive impact of ramp metering on the traffic situation and allow for a comparison of the performance of an ALINEA based controller with a controller based on model predictive control.

2. Motorway control using ramp metering

This section deals with ramp metering as a means to control traffic operations on a motorway. Before discussing ramp metering, the fundamental diagrams from traffic theory are presented in order to get a better understanding of the concept behind ramp metering. As far as determining the appropriate control signals is concerned, the ALINEA controller is presented followed by a presentation of the model predictive control approach.

2.1. The fundamental diagrams

Observations and measurements of traffic on motorways show that traffic behaves approximately according to what are known as the fundamental diagrams in traffic flow theory (May, 1990). The fundamental diagrams plot the relations between the traffic density, the average speed, and the traffic flow. A typical flow-density fundamental diagram is presented in Fig. 1. In low traffic conditions, the traffic flow increases in a nearly proportional way with increasing traffic density. If the traffic density keeps increasing, the traffic flow starts saturating until a maximal flow is reached at the critical density ρ_{cr} . The maximal flow associated with the critical density ρ_{cr} is called the capacity q_{cap} of the motorway. A typical value of ρ_{cr} is 34 vehicles per kilometer and per lane. The capacity q_{cap} of a three-lane

motorway is typically around 6000 vehicles per hour. Once the critical density is reached, traffic breakdown occurs and the traffic flow starts decreasing with further increasing traffic density. As soon as breakdown of the traffic flow at ρ_{cr} occurs, congestion sets in and traffic starts operating in a congested regime. This congested regime is unstable in the sense that a perturbation that momentarily increases the density on the motorway section will cause the traffic flow to decrease, thus giving rise to an even larger traffic density. The traffic density in congested regime where the average traffic speed is zero or, in other words, where the traffic comes to a stand-still is called the jam density ρ_{jam} (see Fig. 1). A typical value of ρ_{jam} is 180 vehicles per hour and per lane. The values of ρ_{cr} , ρ_{jam} and q_{cap} depend on the motorway characteristics such as e.g., the curvature, the speed limits, the slope, ... Stable, free flowing traffic operation can only occur at densities below the critical density.

2.2. Ramp metering

A ramp metering set-up consists of a traffic light that is placed at the on-ramp of a motorway as schematically represented in Fig. 2. The traffic light alternates between the red and the green phase. During the green phase only one vehicle is allowed to enter the motorway using the on-ramp. By varying the timing of the red and the green phases, the number of vehicles that enters the motorway through the on-ramp is controlled. When the traffic density on the motorway tends to exceed the critical density, the ramp metering set-up limits the inflow of vehicles onto the motorway in order to keep the traffic density below the critical density, thus avoiding traffic breakdown and congestion. Whenever the traffic demand is larger than the number of cars that is allowed to enter the motorway, a waiting queue of vehicles is formed at the on-ramp.

By keeping the traffic state on the motorway in the region of stable operation, ramp metering tries to

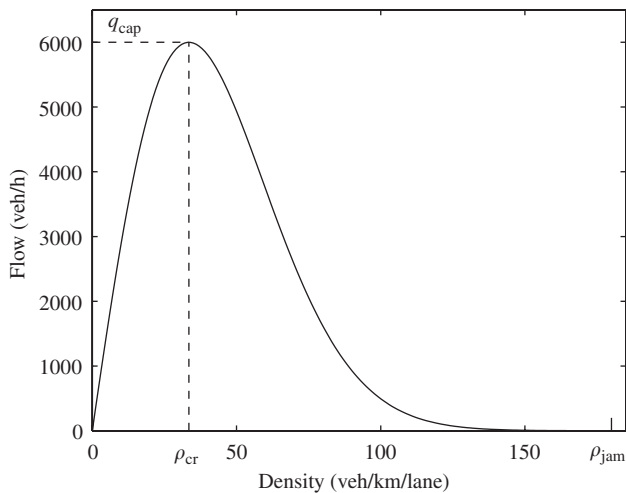


Fig. 1. The flow-density fundamental diagram showing the relation between the traffic flow and the traffic density on a motorway.

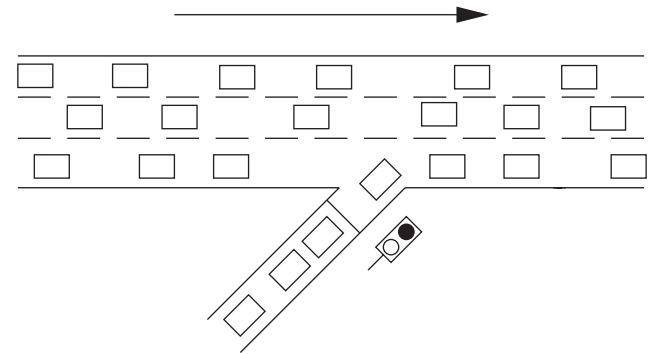


Fig. 2. Schematic representation of ramp metering. The arrow denotes the direction of the traffic flow.

prevent the occurrence of traffic breakdown and congestion. This way, the throughput on the motorway is as high as possible. The preservation of the higher throughput results in a smaller travel time for the vehicles on the motorway. The travel time of the vehicles on the on-ramp increases due to the presence of the on-ramp queue. If the metering rates are determined in a proper way, ramp metering optimizes the traffic flow in the system, resulting in a net decrease of the travel time. Sections 2.3 and 2.4 present two methods to determine the appropriate metering rates.

The control of a single, locally controlled on-ramp can be extended to the coordinated dynamic control of several on-ramps in a motorway network. The coordination of the metering rates of the different on-ramps assures that the control actions taken at different locations in the network reinforce rather than cancel each other. This way, coordinated ramp metering often leads to better results than the combination of multiple independently locally controlled ramp metering set-ups. Sometimes, ramp metering set-ups are integrated into a larger traffic control framework together with control measures such as e.g., traffic density dependent speed limits and route guidance. As the focus of this paper is on-ramp metering control, integrated control of motorway traffic (Hegyi, De Schutter, Hellendoorn, & van den Boom, 2002; Kotsialos, Papageorgiou, & Messmer, 1999; Kotsialos, Papageorgiou, Mangeas, & Haj-Salem, 2002) is beyond the scope of this paper.

The remainder of this section discusses two ways to determine the appropriate on-ramp metering rates: ALINEA and model predictive control.

2.3. ALINEA

ALINEA is the acronym for ‘Asservissement linéaire d’entrée autoroutière’,¹ a feedback control methodology for ramp metering presented by Papageorgiou (Papageorgiou, Hadj-Salem, & Blosseville, 1991). A schematic representation of ALINEA is given in Fig. 3. The goal of ALINEA is to maintain the traffic density on the motorway equal to a preset value $\hat{\rho}$. The value of $\hat{\rho}$ can be chosen to be equal to the critical density ρ_{cr} where the traffic flow on the motorway is maximal and equal to the capacity q_{cap} . This results in a controller that optimizes the traffic flow on the motorway. In discrete time, ALINEA’s control law reads:

$$r(k) = r(k-1) + K_R(\hat{\rho} - \rho(k)), \quad (1)$$

where $r(k)$ is the metering rate at sample step k , K_R is a positive constant, $\rho(k)$ is the traffic density measured downstream of the ramp metering set-up, and $\hat{\rho}$ is the set-point. The metering rate $r(k)$ is confined to the interval $[0, 1]$. ALINEA’s control results are insensitive

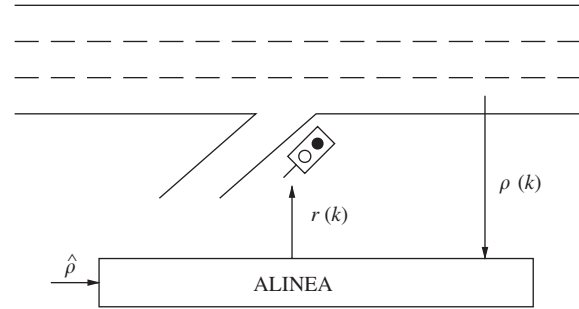


Fig. 3. Control scheme for ALINEA.

to the value of the heuristically determined parameter K_R for a wide range of values (Hasan, Jha, & Ben-Akiva, 2002).

Eq. (1) represents an integrating feedback controller, i.e., a special type of PID control. If the traffic density on the motorway becomes too high (larger than $\hat{\rho}$), the metering rate is reduced and vice versa.

2.4. Model predictive control

Model predictive control (MPC) (Camacho & Bordons, 1995; Maciejowski, 2002) is a flexible approach towards the ramp metering control problem that optimizes an objective or cost function using a motorway traffic model in a receding horizon framework. The choice of the objective function allows for a customization of the controller to a desired policy. By merely changing the cost function, the implemented policy can be altered. Other advantages of MPC based ramp metering are the ability to take constraints into account and the ability to deal with slow changes in the behavior of the traffic system.

2.4.1. General description of model predictive control

The main ingredients of MPC are that it is an on-line control approach in which a model is used to predict the future behavior of the system for a given input sequence and in which a cost criterion is optimized subject to constraints on the inputs and outputs. In addition, MPC uses a receding horizon strategy.

In a receding horizon framework, a prediction horizon N_p is defined and at each sample step k the metering rates for the time period $[k\Delta T_{ctrl}, (k + N_p)\Delta T_{ctrl})$ are determined by minimizing an objective function over this period. The controller time step ΔT_{ctrl} is the rate at which the control signals are updated. A typical value of the controller time step is 1 min. During the optimization, the objective function, which will be discussed in detail in Section 2.4.3, is evaluated based on a prediction of the future traffic behavior of the studied traffic system. The future traffic behavior is simulated using a traffic model. An example of such a model is discussed in Section 2.4.2.

¹French for ‘Linear ramp metering control’.

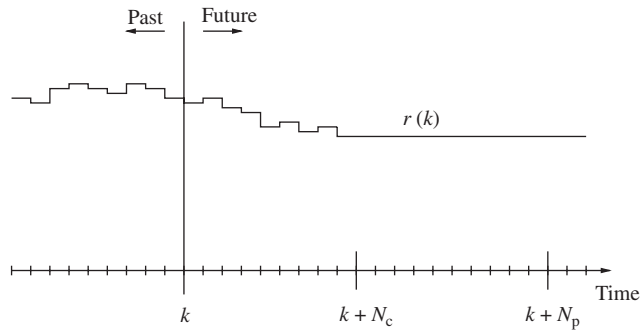


Fig. 4. Schematic representation of the model predictive control concept.

In order to reduce the computational complexity of the optimization, the control horizon N_c ($N_c \leq N_p$) is defined (Fig. 4). The metering rate is only allowed to change during the period $[k\Delta T_{\text{ctrl}}, (k + N_c)\Delta T_{\text{ctrl}})$, after which it is considered to remain constant. In a receding horizon framework, only the first calculated metering rate is applied to the ramp metering set-up. The other metering rates are discarded and recalculated during the next iteration. Once the metering rate is implemented, the state of the traffic in the studied area is updated using measurements, and the whole process starts all over again with the prediction and the control horizon shifted one sample step forward.

The parameters N_p and N_c need to be carefully chosen in order to make a trade-off between the computational complexity and the accuracy of the controller. The larger the prediction horizon, the further the controller can look ahead. This allows the controller to foresee certain events that can be predicted such as e.g., congestion due to increased traffic demand. A larger prediction horizon also implies a larger computational complexity. For the choice of N_c a similar trade-off is made. The length of the control horizon is directly related to the number of metering rates that needs to be optimized. Since the computational complexity of the optimization increases strongly with the number of parameters to be optimized, a trade-off between the computational complexity and the accuracy of the controller can be made. In Section 3.3.3 the values of N_p and N_c will be determined for the E17 motorway case study.

2.4.2. Motorway traffic flow model

During optimization of the metering rates over the prediction horizon, the controller uses a prediction of the traffic behavior generated by a traffic model. This section presents the traffic flow model that will be used in the simulations discussed in Section 3. For the sake of brevity, only those parts of the model that are relevant for interpreting the simulation results of the benchmark motorway are described. It is important to note that the

choice of the specific traffic model used is not imposed by the model predictive framework but by considerations concerning accuracy and computational requirements. Other traffic models than the one presented here, such as e.g., the models discussed in (Treiber & Helbing, 2001) and in (Barceló, 2002), can also be used in the model predictive control framework.

Payne (1971) described a second-order traffic flow model that was extended later on by Papageorgiou et al. to the METANET model (TUC, 2000; Messmer & Papageorgiou, 1990; Papageorgiou, Blosseville, & Hadj-Salem, 1990). The METANET model is a second-order model that is discrete in both space and time. The modeled motorway is discretized in consecutive sections. Typical values for the discretization in time and space are $\Delta T_{\text{sim}} = 10$ s and 500 m respectively (Papageorgiou, Blosseville, & Hadj-Salem, 1989).

The model equations can be written down for every section of the motorway. The first equation expresses the conservation of the number of vehicles in a section:

$$\rho_j(l+1) = \rho_j(l) + \frac{\Delta T_{\text{sim}}}{n_j l_j} [q_{\text{in},j}(l) - q_{\text{out},j}(l)], \quad (2)$$

where $\rho_j(l+1)$ is the traffic density in section j at sample step $l+1$. The traffic density $\rho_j(l+1)$ depends on the traffic density $\rho_j(l)$ at sample step l and on the net inflow into the section during the time interval $[l\Delta T_{\text{sim}}, (l+1)\Delta T_{\text{sim}})$. The net inflow into section j equals the inflow $q_{\text{in},j}(l)$ minus the outflow $q_{\text{out},j}(l)$. The number of lanes in section j is denoted by n_j , and the length of the sections by l_j .

It is important to note that the simulation time step ΔT_{sim} of the traffic simulation model will in general be different from the control time step ΔT_{ctrl} . In order to emphasize the difference between the simulation time step and the control time step, the simulation step counter is denoted in this paper by l , and the control step counter by k .

The average speed in section j at time $l+1$ is given by

$$\begin{aligned} v_j(l+1) = & v_j(l) \\ & + \frac{\Delta T_{\text{sim}}}{\tau} [V[\rho_j(l)] - v_j(l)] \quad \text{Relaxation} \\ & + \frac{\Delta T_{\text{sim}}}{l_j} v_j(l) [v_{j-1}(l) - v_j(l)] \quad \text{Convection} \\ & - \frac{v \Delta T_{\text{sim}} [\rho_{j+1}(l) - \rho_j(l)]}{\tau l_j [\rho_j(l) + \kappa]}, \quad \text{Anticipation} \end{aligned} \quad (3)$$

where τ , v and κ are parameters that can be fitted to traffic data using conventional identification techniques (Ljung, 1999). Three phenomena contribute to the change of the average speed in a section: relaxation, convection and anticipation. The relaxation term states that the average speed in every section tends to evolve towards a density dependent equilibrium value $V[\rho_j(l)]$. An empirical expression for the relation between $V[\rho_j(l)]$

and the traffic density is given by (May, 1990):

$$V[\rho_j(l)] = v_f \exp\left(-\frac{1}{a_m} \left(\frac{\rho_j(l)}{\rho_{cr,j}}\right)^{a_m}\right). \quad (4)$$

The parameter v_f is the free flow speed or the speed that a vehicle obtains in a section if there are no interactions between vehicles. The parameter a_m is a model parameter.

The traffic flow $q_j(l)$ in section j can be expressed in terms of the traffic density and the average speed in the section

$$q_j(l) = \rho_j(l)v_j(l)n_j. \quad (5)$$

When the traffic demand $D_m(l)$ at on-ramp m exceeds the service rate of the on-ramp $q_{on,m}(l)$, a queue is formed. The evolution in time of the queue length $w_m(l)$ is given by

$$w_m(l+1) = w_m(l) + \Delta T_{sim}(D_m(l) - q_{on,m}(l)). \quad (6)$$

The service rate of the on-ramp is the minimum of the number of cars that want to enter and the number of cars that can enter the motorway. This leads to the following expression:

$$q_{on,m}(l) = \min\left[D_m(l) + \frac{w_m(l)}{\Delta T_{sim}}, Q_m \min\left(r_m(l), \frac{\rho_{max,j} - \rho_j(l)}{\rho_{max,j} - \rho_{cr,j}}\right)\right], \quad (7)$$

where Q_m is the capacity of the on-ramp (veh/h) and $\rho_{max,j}$ is the maximal possible density in the section the on-ramp feeds into (here section j). Through the metering rate $r_m(l)$, the service rate of the on-ramp can be limited. The metering rate $r_m(l)$ theoretically lies in the interval $[0, 1]$, but often a lower bound is imposed on the metering rate such that $r_m(l) \in [r_{min}, 1]$.

2.4.3. Objective function

The objective function assigns a cost to every possible traffic state on the studied motorway. This cost can be composed of several components such as e.g., economic, social, environmental, ... terms that can be weighted according to the studied area or according to local policies.

The objective function used in this paper consists of the total time spent by all vehicles in the studied area and of a term that penalizes fluctuations in the control

signal. In the receding horizon framework this leads to the following expression:

$$J(k_0) = \sum_{k=k_0}^{k_0+N_p-1} \left[\sum_{j \in \mathcal{J}_s} \rho_j(k) l_j n_j + \alpha \sum_{m \in \mathcal{J}_o} w_m(k) + \alpha_{ramp} (r(k) - r(k-1))^2 \right] \Delta T_{ctrl}, \quad (8)$$

where \mathcal{J}_s denotes the set of all motorway sections and \mathcal{J}_o is the set of all on-ramps. The total time spent by all the vehicles in the studied area consists of the total time spent by all the vehicles on the motorway sections (the first term in (8)) plus the total time spent by the vehicles in the queues at the on-ramps (the second term in (8)). The parameter α allows for putting more or less emphasis on the time spent by the vehicles in the queues. In order to smooth the control signal, a penalty on variations of the control signal is added to the objective function (the third term in (8)). The value of the parameter α_{ramp} determines the relative importance of this smoothing term.

3. Case study

This section presents simulation results of ramp metering in a real-life situation in Belgium. The performance of MPC based ramp metering is compared with the performance of ALINEA based ramp metering and with the no-control case.

3.1. Set-up

In order to assess the performance of ramp metering, a stretch of 9 km of the E17 motorway Ghent–Antwerp is considered as a case study in this paper. Only traffic in the direction of Antwerp is considered. The studied motorway stretch, which is schematically presented in Fig. 5, has some interesting features such as two on-ramps on the left-hand side of the motorway, a tunnel underneath the river Scheldt, and recurrent congestion during the morning rush hour. The case study contains four off-ramps and five on-ramps. The traffic measurements for the E17 motorway stretch are available on a minute by minute basis and consist of the traffic flow, the average speed, and the occupancy for every of the three lanes.

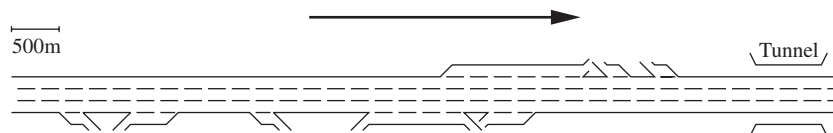


Fig. 5. Schematic representation of the E17 motorway Ghent–Antwerp in Belgium. The arrow denotes the direction of the traffic flow.

Traffic drives on the right-hand side in Belgium. Hence, on-ramps are normally on the right-hand side of the motorway. In the case study presented here, the last two on-ramps are located on the left-hand side of the motorway. The last kilometer of the motorway in Fig. 5 is a tunnel underneath the river Scheldt. Both the on-ramps on the left-hand side of the motorway and the tunnel influence the traffic behavior in the corresponding motorway sections. This behavior is captured by the model parameters, which are identified based on the real-life measurement data.

The motorway in Fig. 5 is divided in 18 sections of 500 m length. In order to model this motorway, Eqs. (2)–(5) are written down for every section. For the sections with an on-ramp equations (6) and (7) are added. In case no ramp metering is implemented at the on-ramp the metering rate in Eq. (7) is set equal to 1.

3.2. Experiment description

Based on the available traffic flow measurements, an estimation of the traffic flows on the on-ramps and the off-ramps can be made. The off-ramp traffic flows are expressed as turning fractions or as the fractions of the vehicle flows on the motorway that leave the motorway through the off-ramps. The on-ramp flows result from the on-ramp traffic demands. The traffic demands and the turning fractions used in the simulations in this paper are based on traffic measurements during a test

period ranging from Tuesday, February 22, 2000 up to and including Friday, February 25, 2000. As an example, Fig. 6 (left) shows the evolution over time of the traffic flow on Thursday, February 24, 2000. For simplicity, the evolution of the traffic demands and the turning fractions over time are approximated by a piecewise affine function as illustrated in Fig. 6 (right).

Given the traffic demands and the turning fractions, the parameters of the traffic flow model can be fitted using standard nonlinear identification techniques (Ljung, 1999).

3.3. Simulation results

Recurrent traffic congestion occurs on the studied motorway near the entrance of the tunnel underneath the river Scheldt i.e., near the fifth and the sixth on-ramp. Therefore, ramp metering will be investigated at these two large volume (two-lane) on-ramps.

3.3.1. No control case

In order to have a point of reference, a simulation of the traffic model for the morning rush hour is ran. The simulated rush hour ranges from 5 a.m. till 10 a.m. The traffic flow model provides the traffic density, the average speed and the traffic flow in every section and for every simulation step. These traffic states are plotted in Fig. 7.

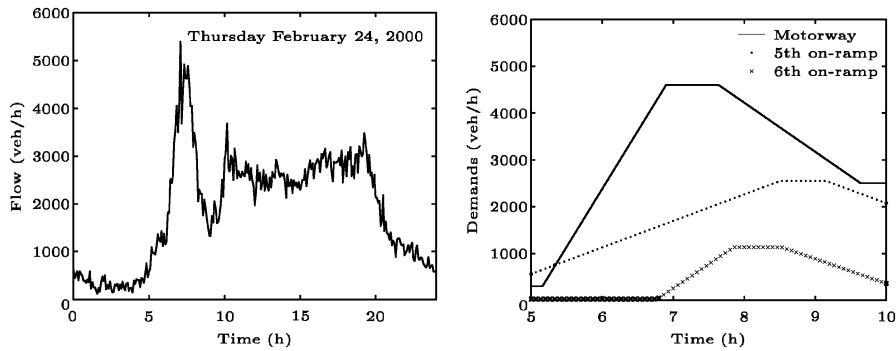


Fig. 6. Evolution of the measured traffic flow (left) at the entrance of the E17 motorway stretch presented in Fig. 5. Piecewise affine approximation of the traffic demands on the E17 motorway Ghent–Antwerp during the morning rush hour on a typical working day (right).

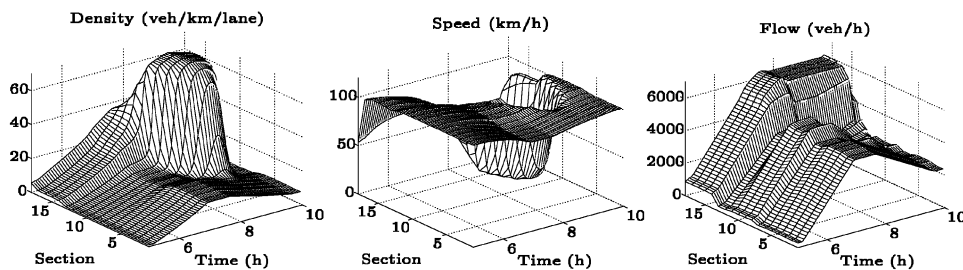


Fig. 7. Plots of the traffic density (left), the average speed (center) and the traffic flow (right) for every section and time step.

In the left plot in Fig. 7 the evolution of the traffic density is presented for every of the 18 sections. The traffic density in the first section increases as the traffic demand increases and decreases again after the peak traffic demand is over. This plot also illustrates the occurrence of congestion in the last sections of the motorway. For example, in section 16 the traffic density grows larger than the critical density ($\rho_{cr} = 34 \text{ veh/km/lane}$) as the rush hour progresses. Congestion sets in. This can be observed in the center plot in Fig. 7 where the evolution of the average speed in the sections over time is plotted. The average speed in section 16 decreases drastically when the traffic density in the section is larger than ρ_{cr} . In the right plot of Fig. 7, the evolution of the traffic flow in every section over time is shown. Note that although the average speed decreases drastically in the congested motorway sections, the traffic flow decrease is limited due to the increased traffic density in the congested motorway sections.

In order to be able to compare the results of different simulations and different controllers, the total time spent (TTS) by all vehicles on the motorway and in the queues at the on-ramps during the simulated morning rush hour is considered. The lower the TTS during the simulated 5 h period, the higher the performance of the motorway system. The TTS by the vehicles during the simulated five hour period is given by:

$$TTS = \sum_{l=1}^{N_{sim}} \left[\sum_{j \in \mathcal{J}_s} \rho_j(l) l_j n_j + \sum_{m \in \mathcal{J}_o} w_m(l) \right] \Delta T_{sim}, \quad (9)$$

where N_{sim} is the number of simulation steps in the simulated period. The TTS in the network without traffic control is equal to 2960 vehicle hours (veh.h).

3.3.2. ALINEA

Using Eq. (1) and the traffic flow model of the motorway, a simulation can be run to assess the performance of ALINEA as a traffic control measure on the E17 motorway. Since during ramp metering a queue can form and since this queue cannot be allowed to grow larger than the available storage capacity of vehicles at the on-ramp, the queue length must be limited. ALINEA does not take the queue length into account in its control law. The constraint on the queue length can be imposed by overriding the metering rate provided by ALINEA once a threshold queue length is

exceeded. In this case study this threshold value is set to 100 vehicles. The metering rate is set equal to 1 for as long as the queue length is above the threshold. Once the queue length is smaller than the threshold, the metering rates provided by ALINEA are implemented again.

ALINEA’s integrating control law of Eq. (1) contains two parameters $\hat{\rho}$ and K_R . Some simulations of the morning rush hour period were ran to determine appropriate values of these parameters. The performance of ALINEA was found not to be very sensitive to the choice of the parameter $\hat{\rho}$. Hence, $\hat{\rho} = \rho_{cr}$ was chosen since the motorway operates at full capacity at the critical density.

The experiments with varying value of K_R are summarized in Table 1. Fig. 8 (left) shows the simulation results of an ALINEA controller with $K_R = 0.0005$ and Fig. 8 (right) the results of an ALINEA controller with $K_R = 0.005$. From Table 1 it can be seen that the controller with the larger gain ($K_R = 0.005$) outperforms the controller with the smaller gain ($K_R = 0.0005$) considering TTS.

In the upper plots in Fig. 8 the traffic demand at the fifth on-ramp is presented as a solid line and the real on-ramp flow (according to Eq. (7)) is presented as a dotted line. The controller with the smaller gain (left) results in a much smoother on-ramp flow than the controller with the larger gain. The controller with the larger gain oscillates at a higher frequency than the one with the smaller gain. This is a known phenomenon in control theory (Dorf & Bishop, 1995). The larger the gain, the faster the switching between ALINEA control and the overriding of the control due to the constraint on the queue length. These oscillations in the metering rate have an influence on the traffic conditions on the motorway. In the lower two plots in Fig. 8, the traffic density and the average speed in the section fed by the fifth on-ramp are presented. The larger the gain of the ALINEA controller, the larger the oscillations in the average speed and the traffic density. These oscillations in the average speed on the motorway need to be suppressed as much as possible since they can become dangerous if their amplitude becomes too large.

A trade-off needs to be made between the performance (TTS) of the controller and the suppression of oscillations in the metering rate. An ALINEA controller with a controller gain of $K_R = 0.001$ seems to provide a

Table 1
Overview of the TTS by the vehicles on the E17 motorway and its on-ramps realized by an ALINEA ramp metering controller on the fifth on-ramp

K_R	0.0001	0.0005	0.001	0.005	0.01	0.05	0.1	0.5	1
TTS (veh.h)	2960	2943	2931	2902	2890	2856	2854	2864	2858

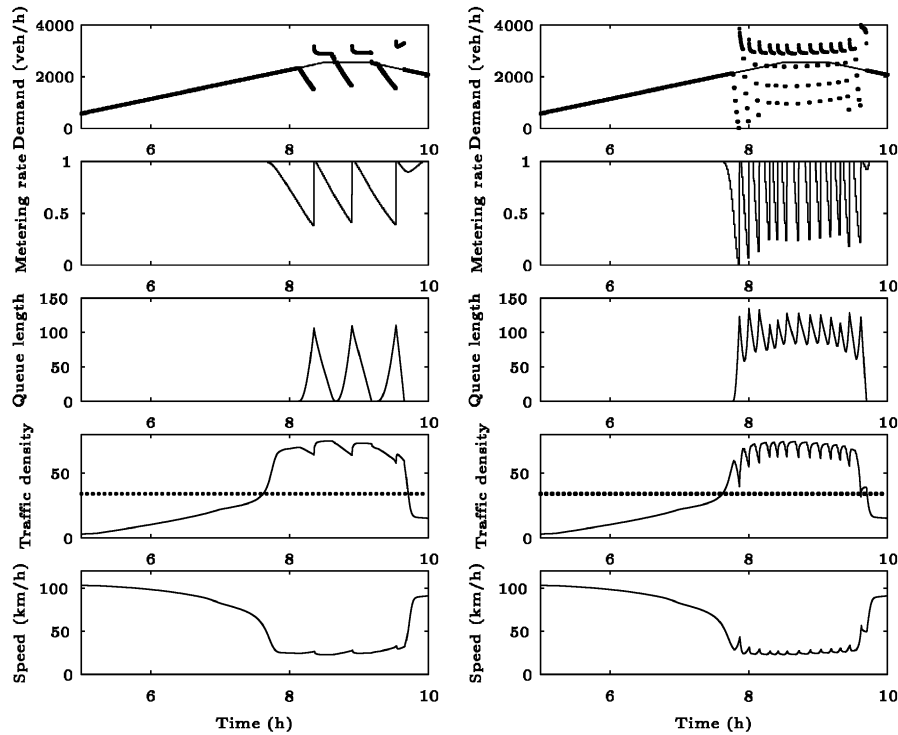


Fig. 8. ALINEA based ramp metering control for $K_R = 0.0005$ (left) and $K_R = 0.005$ (right) on the fifth on-ramp. The dotted line in the upper plots represents the real on-ramp traffic flow. The dotted line in the density graphs represents the critical density ρ_{cr} .

good trade-off for the fifth and the sixth on-ramp of the E17 motorway.

3.3.3. Model predictive control

In a model predictive control framework as described in Section 2.4.1, the metering rates for the controlled on-ramps are computed using (numerical) optimization. For every metered on-ramp, N_c metering rates are computed such that an objective function is minimized for the prediction horizon N_p . Only the first metering rate found for every on-ramp is applied to the system. Next, the prediction and the control horizon are shifted one sample forward and the traffic states are updated after which the whole process is started all over again during the next iteration step.

To allow for a traffic simulation, the studied E17 motorway stretch was implemented in Matlab using the METANET traffic flow model presented in Section 2.4.2. The simulation time step ΔT_{sim} of the model was 10 s. The metering rates provided by the controller are only updated every minute, resulting in a controller time step ΔT_{ctrl} of 1 min. The optimization of the objective function over the prediction horizon was conducted using the function *fmincon* from Matlab's optimization toolbox (The Mathworks, 2002). This optimization routine uses sequential quadratic programming and allows for an easy incorporation of (nonlinear) constraints in the optimization. The fact that a queue at an

on-ramp cannot grow larger than the physically available space is added as a hard constraint to the optimization by limiting the queue lengths at the on-ramps to 100 vehicles or less.

In a receding horizon framework, an optimization of the metering rates is conducted during every iteration step. During the optimization of the metering rates, the optimization algorithm can get stuck in a local minimum and return locally optimal values of the metering rates. This problem can be overcome by restarting the optimization process several times with different starting values of the optimization variables. The number of restarts, and also the computational complexity, can be reduced by carefully choosing the starting values of the optimization based on the optimization results of the previous iteration step. For the simulations presented here, it was found that three restarts suffice. This was verified by looking at the values of the objective function for the different restarts. The following three sets of starting values for the optimization of the metering rates were chosen:

- (1) A set of initial values based on the optimization results from the previous controller step. The first sample of the metering rates found during the previous controller step was applied to the system in the previous iteration. In the current controller step, the discarded metering rates from the previous

controller step can be used as initial values in the current step.² The last value is repeated in order to obtain the required N_c initial values per metered on-ramp.

- (2) The second set of initial metering rates consists of the minimal metering rates for every metered on-ramp. This way, the optimization starts from the state where the on-ramp traffic is maximally restricted at all metered on-ramps. This will generally cause queues to form at the on-ramps and force the optimization to relax the metering rates to a higher value.
- (3) Finally, also a random sequence of initial values confined within the interval $[r_{\min}, 1]$ is used.

After completion of all the optimizations, the metering rate values resulting in the lowest value of the objective function are chosen.

In the MPC framework of Section 2.4, four tuning parameters were defined: N_p , N_c , α and α_{ramp} . The parameter α_{ramp} in (8) assigns a weight to the total time spent by the vehicles in the on-ramp queues relative to the total time spent by the vehicles on the motorway. The choice of the value of α_{ramp} is a policy decision, and therefore α_{ramp} is chosen equal to 1 in this paper. In the next two paragraphs, a description is given of how appropriate values of the remaining parameters can be determined.

The values of the parameters N_p and N_c are a trade-off between the computational complexity and the performance of the controller. In order to determine appropriate lengths of the prediction horizon N_p and the control horizon N_c , several simulations were ran and the performance of the controllers was observed. MPC based ramp metering was simulated on the fifth, on the sixth and on both the fifth and the sixth on-ramps. In order not to bias the performance of the controller, the weighting parameter α_{ramp} in the objective function (8) was chosen equal to 0 while determining appropriate values for N_p and N_c . The performance of the controllers was found not to be very sensitive to the length of the prediction horizon. A prediction horizon of 10 min, which is roughly the travel time through the studied motorway at congested traffic operation, was chosen. In order to limit the computational complexity of the optimization, the control horizon was chosen to be 5 min, reducing the number of parameters to be optimized per metered on-ramp to 5, while preserving the ability of the controller to ‘see’ 10 min ahead in time.

The objective function used is the total time spent with an additional penalty term on variations in the metering rate as presented in (8). The penalty term is weighted by a factor α_{ramp} which allows to put more or

less emphasis on a smooth control signal. In order to determine an appropriate value for α_{ramp} , MPC based ramp metering was simulated on the fifth on-ramp, the sixth on-ramp and both the fifth and the sixth on-ramps with varying values of α_{ramp} . It was found that the value of α_{ramp} had very little impact on the simulation results for the traffic demands as presented in Fig. 6. Only if an excessively large value of α_{ramp} was chosen, e.g., $\alpha_{\text{ramp}} = 1000$, the performance of the controller decreased. A value of $\alpha_{\text{ramp}} = 40$ was found to smooth the metering rates with little or no loss of controller performance.

Fig. 9 shows the results of a simulation of the E17 motorway with two independent MPC controllers, one on the fifth on-ramp and one on the sixth on-ramp. The values of the parameters $N_p = 10$, $N_c = 5$, $\alpha = 1$ and $\alpha_{\text{ramp}} = 40$ were chosen based on the reasoning presented above.

The metering rates in Fig. 9 remain 1 until the traffic density on the motorway reaches the critical density ρ_{cr} and the metering rates drop thus limiting the number of vehicles allowed to enter the motorway. This is illustrated in the upper plots of Fig. 9, where the solid line represents the on-ramp traffic demand and the dotted line represents the real on-ramp flow. A queue starts to form at the on-ramps until the maximal queue length is reached. It can be seen in Fig. 9 that the maximal queue length of 100 vehicles is never exceeded as opposed to the ALINEA case from the previous section. This means that once the queue length is 100 vehicles, the metering rate must be adjusted such that the service rate of the queue equals the traffic demand of the on-ramp. Looking at the traffic density in Fig. 9, it is observed that the ramp metering controller postpones the occurrence of too high a traffic density with corresponding lower traffic flow. Once the queue reaches its maximal length, the density cannot be further controlled, the metering rate increases and the traffic density on the motorway increases accordingly. Comparing the two lower left plots in Fig. 9, which show the traffic density and the average speed in the section fed by the fifth on-ramp, with the plots of the average speed and the traffic density in the same section of the E17 motorway using ALINEA control (Fig. 8), it is clear that the metering rates and the evolution of the average speed and the traffic density are much smoother in the MPC case than in the ALINEA case.

By adding the queue length constraints as hard constraints to the optimization problem, the queue length constraints are strictly respected using MPC as opposed to the ALINEA case presented in the previous section. In case the constraints on the queue lengths are not added as hard constraints to the optimization problem but as an additional penalty term in the objective function as was suggested by Kotsialos et al. (2002), the compliance with the constraints cannot be guaranteed. Adding the

²If $\tilde{r}(k-1|k-1), \dots, \tilde{r}(k+N_c-2|k-1)$ is the optimal metering rate sequence for control step $k-1$ then $\tilde{r}(k|k-1), \dots, \tilde{r}(k+N_c-2|k-1), \tilde{r}(k+N_c-2|k-1)$ can be used as the starting point for the optimization at control step k .

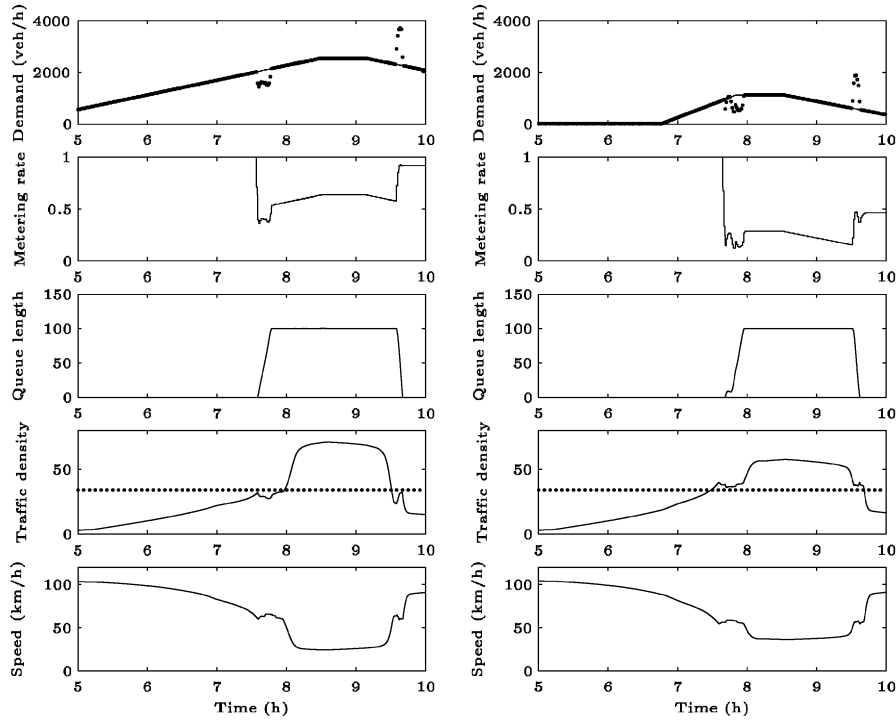


Fig. 9. Simulation results of non-coordinated model predictive control based ramp metering on the fifth and the sixth on-ramp of the E17 Ghent–Antwerp. The dotted line in the upper plots represents the real on-ramp traffic flow while the dotted line in the traffic density plots represents the critical density ρ_{cr} .

Table 2
Overview of the TTS on the E17 motorway for different ramp metering controllers

Controller	No control	ALINEA			MPC		
		5	6	5 & 6	5	6	5 & 6
On-ramp	–						
TTS (veh.h)	2960	2931	2953	2922	2843	2856	2757
TTS red. (%)	0.0	0.9	0.2	1.3	4.0	3.5	6.9

The last row presents the realized reduction in TTS relative to the no control case.

constraints to the optimization problem as was done in this paper guarantees compliance of the MPC based controller with the queue length constraints.

In general, coordination of control can contribute to an additional increase in performance of the system. The MPC based controllers for the fifth and the sixth on-ramp presented in Fig. 9 were not coordinated. Coordination of the controllers of these on-ramps was investigated and was found not to enhance the performance of the system for the traffic demands presented in Fig. 6, although the computational complexity did increase. This can be understood when looking at the controllers in Fig. 9. Both controllers become active at about the same time and due to the intensity of the traffic demands both controllers reach the queue length constraint rather quickly. Hence, there is no way one controller could help the other by being more restrictive (since they both are already maximally restrictive in the non-coordinated case).

3.3.4. Summary of the simulation results

To conclude this section the simulation results are summarized in Table 2.

In Table 2 an overview of the total time spent associated with the different simulations discussed above is presented. The implementation of ALINEA based ramp metering on the fifth on-ramp results in a gain of 29 veh.h during the simulated rush hour compared to the no-control case. The performance gain resulting from ALINEA based ramp metering implemented on the sixth on-ramp is smaller (7 veh.h). When implementing ALINEA based ramp metering on both the fifth and the sixth on-ramps, the gain compared to the no-control case is 38 veh.h.

The gain in TTS by implementing MPC based ramp metering on the fifth on-ramp is 117 veh.h, or more than three times the gain achieved by ALINEA. For the sixth on-ramp, the gain achieved by the MPC based controller compared to the no control case is 104 veh.h,

versus 7 veh.h for the ALINEA controller. MPC based control on both the fifth and the sixth on-ramp results in a reduction of 203 veh.h compared to the no control case or a performance improvement of nearly 7 percent. Besides the improved performance of the MPC based ramp metering compared to the ALINEA based controllers, the control signal of the MPC based controller is smooth (see Fig. 9) and the constraints on the queue lengths are strictly respected.

4. Conclusions

The motorway E17 Ghent–Antwerp in Belgium was modeled and ALINEA based ramp metering control was compared to MPC based ramp metering control using simulations.

First, the controller parameters of the ALINEA and the MPC based controllers were determined after which the performance of a number of controllers was compared.

It was observed that the ALINEA based controllers resulted in a gain in TTS during the simulated rush hour. The controllers were able to limit the queues approximately to the maximal queue length. The metering rates oscillated due to the constraints on the maximal queue lengths overriding the controller. This resulted in oscillations in the traffic density and the average speed in the section fed by the on-ramp. These oscillations need to be suppressed as much as possible. A very strong point of ALINEA is its limited computational complexity.

The MPC based ramp metering controllers were observed to realize a higher performance (smaller TTS) than the ALINEA based controllers. Moreover, the control signals of the MPC based controllers were very smooth. The traffic density and the average speed in the section fed by the on-ramps were behaving very smoothly, even during the rush hour. No oscillations occurred. The constraints on the queue lengths at the on-ramps were strictly respected by the MPC based controllers. The computational complexity of the MPC based controllers is larger than the computational complexity of ALINEA due to the optimization problem that needs to be solved at every iteration step. By appropriately choosing the prediction and the control horizons, a trade-off was made between the performance and the computational complexity of MPC based ramp metering on the E17 motorway in Belgium.

Acknowledgements

Our research is supported by grants from several funding agencies and sources: Research Council KUL: GOA-Mefisto 666, IDO; Flemish Government:

FWO G.0256.97, G.0115.01, G.0240.99, G.0197.02, G.0407.02, ICCoS, ANMMM, IWT, STWW Genprom, GBOU McKnow, Eureka; Belgian Federal Government: OSTC (IUAP IV-02 and IUAP V-10-29), PODO-II (CP-TR-18); Direct contract research.

This research is also partially funded by the NWO-CONNEKT project AMICI (014-34-523), the BSIK project TRANSUMO, the STW-VIDI project “Multi-agent control of large-scale hybrid systems” (DWV.6188), and the Transport Research Centre Delft.

References

- Barceló, J. (2002). Dynamic network simulation with Aimsun. In *Proceedings of the international symposium on transport simulation*, Yokohama, Japan. <http://www.aimsun.com/documents.html>.
- Camacho, E., & Bordons, C. (1995). *Model predictive control in the process industry*. Berlin, Germany: Springer.
- Dorf, R. C., & Bishop, R. H. (1995). *Modern control systems*. Addison-Wesley.
- Hasan, M., Jha, M., & Ben-Akiva, M. (2002). Evaluation of ramp control algorithms using microscopic traffic simulation. *Transportation Research C*, 10, 229–256.
- Hegyi, A., De Schutter, B., Hellendoorn, H. & van den Boom, T. (2002). Optimal coordination of ramp metering and variable speed control—An MPC Approach. In *Proceedings of the 2002 American Control Conference*. (pp. 3600–3605). Anchorage, Alaska.
- Kotsialos, A., Papageorgiou, M., Mangeas, M., & Haj-Salem, H. (2002). Coordinated and integrated control of motorway networks via non-linear optimal control. *Transportation Research C*, 10(1), 65–84.
- Kotsialos, A., Papageorgiou, M., Messmer, A. (1999). Optimal coordinated and integrated control motorway traffic control. In A. Ceder, (Ed.), *Proceedings of the 14th international symposium on transportation and traffic theory*. (pp. 621–644). Jerusalem.
- Ljung, L. (1999). *System identification: theory for the user* (2nd ed.). Upper Saddle River, New Jersey: Prentice-Hall.
- Maciejowski, J. (2002). *Predictive control with constraints*. Harlow, England: Prentice Hall.
- May, A. D. (1990). *Traffic flow fundamentals*. Englewood Cliffs, NJ: Prentice-Hall.
- Messmer, A., & Papageorgiou, M. (1990). METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering and Control*, 31(9), 466–470.
- Papageorgiou, M., Blosseville, J. M., & Hadj-Salem, H. (1989). Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris. *Transportation Research B*, 23B(1), 29–47.
- Papageorgiou, M., Blosseville, J.-M., & Hadj-Salem, H. (1990). Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris: Part I: Modelling. *Transportation Research A*, 24A(5), 345–359.
- Papageorgiou, M., Hadj-Salem, H., & Blosseville, J.-M. (1991). ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record*, 1320, 58–64.
- Payne, H. J. (1971). Models of freeway traffic and control. In G. A. Bekey (Ed.), *Mathematical models of public systems, simulation council proceedings series* (Vol. 1) (pp. 51–61), La Jolia, CA.
- The Mathworks (2002). *Optimization Toolbox User's Guide v. 2.2*.
- Treiber, M., & Helbing, D. (2001). Microsimulations of freeway traffic including control measures. *Automatisierungstechnik*, 49, 478–484.
- TUC (2000). *METANET—A simulation program for motorway networks*.