Optimization of traffic flow at freeway sags by controlling the acceleration of vehicles equipped with in-car systems

Bernat **Goñi-Ros**[a,x], Victor L. **Knoop**[a], Toshimichi **Takahashi**[b], Ichiro **Sakata**[b], Bart **van Arem**[a] and Serge P. **Hoogendoorn**[a]

[a] Delft University of Technology, Faculty of Civil Engineering, Department of Transport and Planning
Stevinweg 1, 2628 CN, Delft, The Netherlands. Tel: +31 15 278 4912
E-mail: {b.goniros,v.l.knoop,b.vanarem,s.p.hoogendoorn}@tudelft.nl
[b] Toyota Motor Europe, Technical Center
Hoge Wei 33, 1930 Zaventem, Belgium. Tel.: +32 2 712 3416
E-mail: taka@vdlab.tytlabs.co.jp,ichiro.sakata@toyota-europe.com
[x] Corresponding author

May 28, 2016

Abstract

Sags are bottlenecks in freeway networks. According to previous research, the main cause is that most drivers do not accelerate enough at sags. Consequently, they keep longer headways than expected given their speed, which leads to congestion in high demand conditions. Nowadays, there is growing interest in the development of traffic control measures for sags based on the use of in-car systems. This paper aims to determine the optimal acceleration behavior of vehicles equipped with in-car systems at sags and the related effects on traffic flow, thereby laying the theoretical foundation for developing effective traffic management applications. We formulate an optimal control problem in which a centralized controller regulates the acceleration of some vehicles of a traffic stream moving along a single-lane freeway stretch with a sag. The control objective is to minimize total travel time. The problem is solved for scenarios with different numbers of controlled vehicles and positions in the stream, assuming low penetration rates. The results indicate that the optimal behavior involves performing a deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. This maneuver induces the first vehicles located behind the controlled vehicle to accelerate fast along the vertical curve. As a result, traffic speed and flow at the end of the sag (bottleneck) increase for a time. The maneuver also triggers a stop-and-go wave that temporarily limits the inflow into the sag, slowing down the formation of congestion at the bottleneck. Moreover, in some cases controlled vehicles perform one or more deceleration-acceleration maneuvers upstream of the sag. This additional strategy is used to manage congestion so that inflow is regulated more effectively. Although we cannot guarantee global optimality, our findings reveal a potentially highly effective and innovative way to reduce congestion at sags, which could possibly be implemented using cooperative adaptive cruise control systems.

Keywords

Traffic management, In-car system, Sag vertical curve, Stop-and-go wave, Car-following model, Optimal control

1 Introduction

Sags (or sag vertical curves) are freeway sections along which the gradient increases gradually in the direction of traffic (see Figure 1). The capacity of sags is generally lower than that of freeway sections with other vertical profiles [1]. Therefore, sags constitute bottlenecks in freeway networks. Indeed, traffic often becomes congested at sags in high demand conditions [2, 3, 4]. In some countries, such as Japan, sags are one of the most common types of freeway bottleneck [5, 6]. The scientific literature suggests that the main cause of traffic congestion at sags is that most drivers do not accelerate enough as they move along the vertical curve. Generally, drivers do not compensate instantaneously for the increase in resistance force resulting from the increase in gradient, which limits the acceleration of their vehicles [7]. As a consequence, most drivers keep significantly longer distance headways than expected given their speed [8, 9]. This leads to periodic formation of stop-and-go waves when traffic demand is sufficiently high [2, 4, 10]. The bottleneck is generally the end of the vertical curve [3, 4].

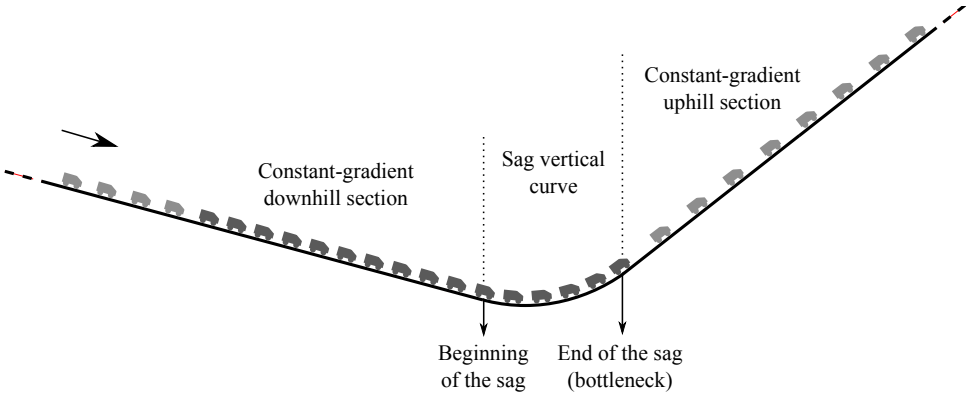


Figure 1: Freeway stretch with a sag: vertical profile and typical traffic conditions with high demand (light-gray vehicles are driving in free flow; dark-gray vehicles are driving in congested flow).

During the last two decades, various traffic management measures have been proposed for mitigating congestion at sags. In general, the goals of those measures are: *a*) to increase the free-flow capacity of sags; *b*) to prevent the formation of congestion at sags in nearly-saturated conditions; and/or *c*) to increase the queue discharge capacity of sags. The proposed measures use very diverse strategies to achieve those goals, such as improving the ability of drivers to compensate for the increased grade resistance force at sags [11], limiting the inflow into the vertical curve [12], and encouraging drivers to accelerate fast after leaving congestion at sags [13]. Most proposed measures, particularly those which are in more advanced development stages, use variable message signs (VMS) as actuators [5, 12, 13]. However, in recent years there has been growing interest in the development of traffic management measures that use in-car systems as actuators. The systems used for that purpose are mainly advisory systems [6] or basic/advanced adaptive cruise control (ACC) systems [11, 14, 15]. Although traffic management measures based on the use of in-vehicle systems have great potential, they are mostly in early phases of development. We argue that, at this stage, it is important to determine how vehicles equipped with this type of systems should behave at sags under various circumstances in order to generate the greatest possible reduction in congestion. This would clarify what are the most effective strategies to

mitigate congestion at sags via in-car systems and what are the mechanisms by which these strategies reduce congestion, thereby laying the theoretical foundation for developing effective traffic management applications. The present study constitutes a first step in this direction.

The main goal of the study is to identify the way in which vehicles equipped with in-car systems need to move at sags in the longitudinal direction in order to reduce congestion as much as possible, assuming penetration rates that are realistic for the coming years (i.e., limited numbers of equipped vehicles). To this end, we formulate an optimal control problem in which a centralized controller regulates the acceleration of some vehicles belonging to a traffic stream that moves along a single-lane freeway stretch with a sag. The objective of the controller is to minimize the total travel time of all vehicles. The problem is solved for various scenarios defined by the number of controlled vehicles and their positions in the stream. By analyzing the results, we identify the main strategies that vehicles equipped with in-car systems should use at sags to reduce congestion to the greatest possible degree. Therefore, the main contribution of this paper is twofold. First, we present a new optimization-based method for identifying the best way to manage traffic by means of in-vehicle systems. Second, we determine the most effective strategies to manage traffic via in-car systems at sags, providing a detailed description of the principles of these strategies. In this respect, we are mostly interested in identifying the optimal trajectories of individual vehicles (which are assumed to be equipped with some kind of in-car systems), and their effects on traffic flow. The development of specific traffic management measures to make individual vehicles move in the identified optimal way is out of the scope of this paper and is left for future work. For this reason, the control problem formulation does not specify the type of in-car system that vehicles are equipped with, nor does it specify any technical aspects of this system: the formulation is based on the assumption that the acceleration of some vehicles can be regulated or influenced via some type of in-car system.

The rest of this paper is structured as follows. Section 2 presents the optimal control framework for traffic flow optimization. Section 3 presents the setup of the experiments that were carried out to identify the optimal acceleration behavior of vehicles equipped with in-car systems at sags. Section 4 reports the results of the experiments, focusing on the main strategies used by the controlled vehicles and their effects on traffic flow. Finally, Section 5 presents the conclusions of this study.

2 Optimal control framework for traffic flow optimization

This section presents an optimal control framework aimed to determine the optimal trajectories of a set of vehicles as they move along a freeway stretch with a sag. Section 2.1 specifies the elements of the system to be controlled. Section 2.2 states the assumptions of the optimal control framework. Section 2.3 presents the formulation of the control problem.

2.1 System elements

The system to be controlled consists of a stream of vehicles moving along a freeway stretch. In total, the stream contains n vehicles. Each of them is assigned a number i that corresponds to its position in

the stream ($i = 1, 2, \ldots, n$). The set that contains all vehicle numbers i is denoted by N. There are two types of vehicles in the stream: *a*) non-controlled vehicles; and *b*) controlled vehicles. The total number of controlled vehicles is denoted by m, and the subset of N that contains the numbers i of all controlled vehicles is denoted by M. Each controlled vehicle is assigned a number j that corresponds to its position in relation to the other controlled vehicles ($j = 1, 2, \ldots, m$). For instance, in the vehicle stream shown in Figure 2, $n = 8$, $N = \{1, 2, 3, 4, 5, 6, 7, 8\}$, $m = 2$ and $M = \{3, 7\}$.

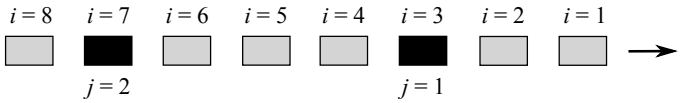


Figure 2: Example of vehicle stream. Gray rectangles represent non-controlled vehicles and black rectangles represent controlled vehicles. The arrow shows the direction of traffic.

The freeway stretch contains a sag vertical curve. The relation between freeway gradient and location along the freeway is known. For the sake of simplicity, we assume that the freeway stretch has one lane (i.e., overtaking is not possible) and there are no on-ramps, off-ramps, merges or diverges.

2.2 Assumptions

The optimal control framework is based on four assumptions:

1. A stream of vehicles moving along a freeway can be considered a discrete-time dynamical system with constant-acceleration time steps, provided that the time step length is sufficiently small.

2. The acceleration behavior of the drivers can be correctly described by the car-following model presented in [16]. That model assumes that drivers accelerate with the objective of reaching/keeping their desired speed (in unconstrained driving conditions) or desired distance headway (in constrained conditions). The model also assumes that drivers operate their vehicle in such a way that they compensate gradually (not instantaneously) for the increased grade resistance force at sags. As shown in [16], the model is able to reproduce the main characteristics of traffic flow at sags.

3. Within a certain freeway section, the acceleration of some vehicles of the stream can be regulated by a centralized controller that can perfectly predict the state of the traffic stream at any time step.

4. The controller does not impose at any time step a higher acceleration than the one that would result from human driving given the current traffic conditions and position along the road.

Note that, given the objectives and scope of this research (see Section 1), it is not necessary to specify how the controller acts on vehicle acceleration or how the controller determines the acceleration that would be realized by a human driver under certain circumstances. In practical applications, the acceleration of the controlled vehicles could be regulated either by automatic systems (such as cooperative ACC systems) or by human drivers (assisted by in-car advisory systems) (see Section 1). In general, vehicle acceleration would be more limited in the latter case, particularly at sags. The fourth assumption is

necessary to prevent very short headways and collisions, and to not exclude the possibility of controlled vehicles representing vehicles whose acceleration depends on human driving behavior.

2.3 Control problem formulation

This section formulates the optimal control problem. Sections 2.3.1 and 2.3.2 define the state and control variables. Section 2.3.3 presents the equations that determine the dynamics of the vehicle stream. Sections 2.3.4 and 2.3.5 define the initial state, the admissible control region and the cost function. Finally, Section 2.3.6 formulates the control problem as a mathematical program.

2.3.1 State variables

The state variables of a dynamical system can be defined as the minimum set of variables that contain enough information about the system to determine its future behavior. In our case, the system is a stream of vehicles moving along a freeway stretch (see Section 2.1) and the state variables are the variables needed to determine the future trajectory of every vehicle in the space-time plane.

As explained in Section 2.2, we assume that: *a*) the trajectories of human-driven vehicles are determined by the car-following model presented in [16]; *b*) the trajectories of controlled vehicles are regulated by the controller, which uses the car-following model to impose upper bounds on the acceleration of these vehicles. The car-following model presented in [16] calculates the acceleration of a vehicle at any time step based on the following variables: position and speed of the target vehicle, position and speed of the preceding vehicle, freeway gradient at the position of the target vehicle, and amount of freeway gradient compensated by the driver of the target vehicle. The latter is a variable that represents the amount of freeway gradient that does not cause a limitation in vehicle acceleration because its corresponding amount of grade resistance force is compensated for by the traction force exerted on the vehicle (see more details in [16]). Note that freeway gradient is a variable that depends directly on vehicle position.

Therefore, the state variables of our system are: *a*) longitudinal position of (the rear-bumper of) all vehicles along the freeway (r_i, $\forall i$); *b*) speed of all vehicles (v_i, $\forall i$); and *c*) amount of freeway gradient compensated by the drivers of all vehicles ($G_{\text{com},i}$, $\forall i$). More specifically, the state at simulation time step τ is defined by the following matrix:

$$\mathbf{x}(\tau) = \begin{bmatrix} r_1(\tau) & r_2(\tau) & \dots & r_n(\tau) \\ v_1(\tau) & v_2(\tau) & \dots & v_n(\tau) \\ G_{\text{com},1}(\tau) & G_{\text{com},2}(\tau) & \dots & G_{\text{com},n}(\tau) \end{bmatrix} \tag{1}$$

2.3.2 Control variables

The control variables can be defined as the variables that can be manipulated by the controller in order to influence the dynamics of the system. In accordance with the third and fourth assumptions stated in Section 2.2, in our case we chose the control variables to be the maximum accelerations of all controlled

vehicles (u_j, $\forall j$). Therefore, the control input at control time step κ is defined by the following vector:

$$\mathbf{u}(\kappa) = \begin{bmatrix} u_1(\kappa) & u_2(\kappa) & \dots & u_m(\kappa) \end{bmatrix} \tag{2}$$

Different counters are used for simulation time step and control time step (τ and κ) because the control time step length (T_c) can be assigned a different value than the simulation time step length (T_s), as long as T_c is a multiple of T_s and the multiplier is a natural number ($T_\mathrm{c} = \eta \cdot T_\mathrm{s}$, where $\eta \in \mathbb{N}^+$). If $T_\mathrm{c} > T_\mathrm{s}$, the control input stays constant over multiple simulation time steps.

2.3.3 State dynamics

Figure 3 shows an overview of the inputs that are necessary to update the values of the state variables (i.e., r_i, v_i and $G_{\mathrm{com},i}$, $\forall i$) at every simulation time step. The position and speed of the vehicles change over time according to the following equations of motion:

$$r_i(\tau+1) = r_i(\tau) + v_i(\tau) \cdot T_\mathrm{s} + \frac{a_i(\tau)}{2} \cdot T_\mathrm{s}^2 \tag{3}$$

$$v_i(\tau+1) = v_i(\tau) + a_i(\tau) \cdot T_\mathrm{s} \tag{4}$$

where $a_i(\tau)$ denotes the acceleration of vehicle i at simulation time step τ.

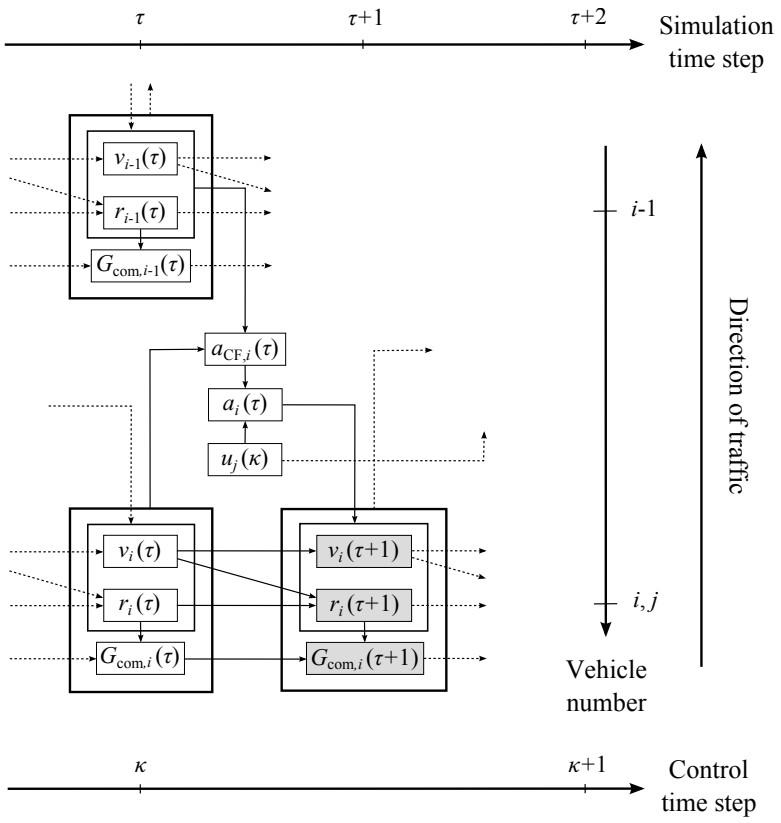


Figure 3: Overview of the inputs needed to calculate the state dynamics.

The compensated gradient changes over time as follows (see also [16]): a) drivers compensate for an increase in freeway gradient with a maximum compensation rate denoted by λ; b) once drivers have fully compensated for an increase in gradient, G_{com} equals the actual gradient (G); and c) drivers fully compensate for any decrease in gradient instantaneously.

$$G_{\text{com},i}(\tau+1) = \begin{cases} G(r_i(\tau+1)) & \text{if } G(r_i(\tau+1)) \leq G_{\text{com},i}(\tau) + \lambda_i \cdot T_{\text{s}} \\ G_{\text{com},i}(\tau) + \lambda_i \cdot T_{\text{s}} & \text{if } G(r_i(\tau+1)) > G_{\text{com},i}(\tau) + \lambda_i \cdot T_{\text{s}} \end{cases} \tag{5}$$

Vehicle accelerations (a_i) are calculated as follows. For non-controlled vehicles, the acceleration of vehicle i at time step τ ($a_i(\tau)$) is equal to the acceleration given by the car-following model at that time step ($a^{\text{h}}_{\text{CF},i}(\tau)$). For controlled vehicles, $a_i(\tau)$ is equal to the acceleration determined by the controller for time step τ ($a_{\text{c},i}(\tau)$), which in turn is defined as the minimum of the control input ($u_j(\kappa)$) and the acceleration given by the car-following model ($a^{\text{c}}_{\text{CF},i}(\tau)$). By defining $a_{\text{c},i}$ in this way, we ensure that the fourth assumption stated in Section 2.2 holds without having to explicitly specify a state constraint in the control problem. That simplifies the optimization procedure. Thus, accelerations are calculated as follows:

$$a_i(\tau) = \begin{cases} a^{\text{h}}_{\text{CF},i}(\tau) & \text{if } i \notin M \\ a_{\text{c},i}(\tau) & \text{if } i \in M \end{cases} \tag{6}$$

where:

$$a_{\text{c},i}(\tau) = \begin{cases} \min(u_j(\kappa), a^{\text{c}}_{\text{CF},i}(\tau)) & \text{if } r_i(\tau) \in [r^{\text{c}}_0, r^{\text{c}}_{\text{f}}] \\ a^{\text{c}}_{\text{CF},i}(\tau) & \text{if } r_i(\tau) \notin [r^{\text{c}}_0, r^{\text{c}}_{\text{f}}] \end{cases} \tag{7}$$

In Equation 7: r^{c}_0 and r^{c}_{f} denote the points that delimit the area where acceleration control is applicable (see Figure 4 for illustration); i and j correspond to the same vehicle; and κ is such that $\tau \cdot T_{\text{s}} \in [\kappa \cdot T_{\text{c}}, (\kappa+1) \cdot T_{\text{c}})$.

The car-following model is the same for non-controlled vehicles and controlled vehicles, thus $a^{\text{h}}_{\text{CF},i}$ and $a^{\text{c}}_{\text{CF},i}$ are calculated in the same way. However, there are two important differences between those two variables. First, $a^{\text{h}}_{\text{CF},i}(\tau)$ is always the actual acceleration of vehicle i at simulation time step τ, whereas $a^{\text{c}}_{\text{CF},i}(\tau)$ is only the realized acceleration if the control input is greater or not applicable (see Equations 6 and 7). Second, $a^{\text{h}}_{\text{CF},i}(\tau)$ is always calculated based on the state of the vehicle stream at simulation time step τ, but that is not true for $a^{\text{c}}_{\text{CF},i}(\tau)$. If $T_{\text{c}} > T_{\text{s}}$, then $a^{\text{c}}_{\text{CF},i}(\tau)$ is calculated based on the current state only at simulation time steps when $\tau \cdot T_{\text{s}} = \kappa \cdot T_{\text{c}}$ holds (i.e., at the beginning of every control time step), or when the realized accelerations (a_i) at all previous simulation time steps within the current control step were equal to $a^{\text{c}}_{\text{CF},i}(\tau)$. At all other simulation time steps, $a^{\text{c}}_{\text{CF},i}(\tau)$ is calculated based on the state that would have been obtained if the realized vehicle accelerations (a_i) at all previous simulation time steps within the same control step had been equal to $a^{\text{c}}_{\text{CF},i}(\tau)$ (for example, see Figure 7(a)). This formulation makes $a^{\text{c}}_{\text{CF},i}$ independent of u_j within every control step, which simplifies the optimization procedure.

The car-following model presented in [16] is used to calculate $a^{\text{h}}_{\text{CF},i}$ and $a^{\text{c}}_{\text{CF},i}$. Here we only present the main features of this model. We refer to [16] for a discussion of its properties and validity, also in comparison with other existing models. Let $a_{\text{CF},i}$ be a variable that corresponds to $a^{\text{h}}_{\text{CF},i}$ if vehicle i is non-controlled and to $a^{\text{c}}_{\text{CF},i}$ if vehicle i is controlled. Then, the car-following model calculates vehicle accelerations as follows:

$$a_{\text{CF},i}(\tau) = \max(a_{\text{des},i}(\tau) + \delta_i(\tau), a_{\text{min},i}, -\frac{v_i(\tau)}{T_{\text{s}}}) \tag{8}$$

In Equation 8, $a_{\text{min},i}$ and $-v_i(\tau)/T_s$ define lower bounds that prevent unrealistically strong decelerations and negative speeds, respectively. The influence of traffic conditions and freeway gradient on longitudinal driving behavior is mainly described by the expression $a_{\text{des},i}(\tau) + \delta_i(\tau)$. The desired acceleration is calculated as follows:

$$a_{\text{des},i}(\tau) = \alpha_i \cdot \min\left[1 - \left(\frac{v_i(\tau)}{v_{\text{des},i}}\right)^4, 1 - \left(\frac{s_{\text{des},i}(v_i(\tau), \Delta v_i(\tau))}{s_i(\tau)}\right)^2\right] \tag{9}$$

where s_i denotes net distance headway, and the relative speed (Δv_i) and desired net distance headway ($s_{\text{des},i}$) are defined as:

$$\Delta v_i(\tau) = v_i(\tau) - v_{i-1}(\tau) \tag{10}$$

$$s_{\text{des},i}(v_i(\tau), \Delta v_i(\tau)) = s_{\text{s},i} + v_i(\tau) \cdot H_i + \frac{v_i(\tau) \cdot \Delta v_i(\tau)}{2 \cdot \sqrt{\alpha_i \beta_i}} \tag{11}$$

The parameters in Equations 9-11 are: desired speed ($v_{\text{des},i}$), net distance headway at standstill ($s_{\text{s},i}$), safe time headway (H_i), maximum acceleration (α_i) and maximum comfortable deceleration (β_i).

The term δ_i in Equation 8 accounts for the influence of the freeway's vertical alignment on longitudinal driving behavior. At a given simulation time step, $\delta_i(\tau)$ is equal to the difference between the gradient in the freeway location where the vehicle is at that time step and the amount of gradient compensated by the driver until that time step, multiplied by a sensitivity parameter (θ_i) (see more details in [16]):

$$\delta_i(\tau) = -\theta_i \cdot (G(r_i(\tau)) - G_{\text{com},i}(\tau)) \tag{12}$$

2.3.4 Initial state and admissible control region

The initial state of the stream of vehicles ($\mathbf{x}(0)$) is known and equal to $\mathbf{x}_0$. The admissible control region $\mathcal{U}$ contains all maximum acceleration (u_j) values that can be chosen by the controller for every controlled vehicle j. In our case, the set of admissible maximum acceleration values is the same for all controlled vehicles and for all control time steps. That set contains all real numbers between a lower bound denoted by u_{min} and an upper bound denoted by u_{max}. Therefore, $\mathcal{U}$ is such that:

$$u_j(\kappa) \in [u_{\text{min}}, u_{\text{max}}]\,,\ \forall j \in [1, m],\ \forall \kappa \in [0, \tfrac{T}{T_{\text{c}}}] \tag{13}$$

2.3.5 Cost function

The cost function (J) is defined as the total travel time of all vehicles from their initial positions to the arrival point R:

$$J(\mathbf{x}(0), \mathbf{x}(1), \ldots, \mathbf{x}(\tfrac{T}{T_\mathrm{s}}), \mathbf{u}(0), \mathbf{u}(1), \ldots, \mathbf{u}(\tfrac{T}{T_\mathrm{c}})) = \sum_{i=1}^{n} (T_s \cdot \tau_{R,i} + \Delta t_i) \tag{14}$$

where $\tau_{R,i}$ denotes the last simulation time step at which vehicle i is located upstream of point R:

$$\tau_{R,i} = \max(\tau \mid r_i(\tau) \leq R) \tag{15}$$

and Δt_i denotes the additional time required by vehicle i to move from its position at simulation time step $\tau_{R,i}$ to position R, which is calculated by solving the following quadratic equation:

$$\frac{a_i(\tau_{R,i})}{2} \cdot (\Delta t_i)^2 + v_i(\tau_{R,i}) \cdot \Delta t_i + (r_i(\tau_{R,i}) - R) = 0 \tag{16}$$

2.3.6 Optimal control problem

The discrete-time optimal control problem, which is non-linear and non-convex, can be formulated as the following mathematical program:

Find $\mathbf{u}^*(0), \mathbf{u}^*(1), \ldots, \mathbf{u}^*(\frac{T}{T_\mathrm{c}})$

that minimize $J(\mathbf{x}(0), \mathbf{x}(1), \ldots, \mathbf{x}(\frac{T}{T_\mathrm{s}}), \mathbf{u}(0), \mathbf{u}(1), \ldots, \mathbf{u}(\frac{T}{T_\mathrm{c}}))$

subject to:

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{17}$$

$$\mathbf{u}(\kappa) \in \mathcal{U}, \text{ for } \kappa = 0, 1, 2, \ldots, \tfrac{T}{T_\mathrm{c}} \tag{18}$$

$$\mathbf{x}(\tau + 1) = \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\kappa)), \text{ for } \tau = 0, 1, 2, \ldots, \tfrac{T}{T_\mathrm{s}} \tag{19}$$

where κ is such that $\tau \cdot T_\mathrm{s} \in [\kappa \cdot T_\mathrm{c}, (\kappa + 1) \cdot T_\mathrm{c})$.

The state matrix ($\mathbf{x}(\tau)$) and the control input vector ($\mathbf{u}(\kappa)$) are defined in Equations 1 and 2, respectively. The objective function (J) is calculated by using Equations 14-16. The admissible control region ($\mathcal{U}$) is defined in Equation 13. The vehicle stream state dynamics are determined by Equations 3-12.

3 Experimental setup

We carried out a series of optimization experiments that entailed solving the discrete-time optimal control problem presented in the previous section for various scenarios. Section 3.1 specifies the objectives of the experiments, Section 3.2 defines the scenarios, Section 3.3 describes the method used to solve the problem, and Section 3.4 defines the indicator used to evaluate the effectiveness of the optimal controller.

3.1 Objectives

The objectives of the optimization experiments were as follows:

1. To determine the optimal acceleration behavior of controlled vehicles as they move along a one-lane freeway with a sag in nearly-saturated traffic conditions, assuming low penetration rates.

2. To determine the effects that the optimal acceleration behavior of controlled vehicles has on traffic flow dynamics, and the reasons why the total travel time decreases as a result of those effects.

3.2 Scenarios

We defined eight scenarios in which the traffic stream contains 300 vehicles ($n = 300$). The scenarios differ only in two inputs: number of controlled vehicles and positions of those vehicles in the stream (see Section 3.2.1). All other inputs are the same in all scenarios (Section 3.2.2).

3.2.1 Distinctive inputs

The scenarios differ in the number of controlled vehicles (m) and the positions of those vehicles in the stream (set M). To define the scenarios, we stipulated that the number of controlled vehicles should be 0, 1, 2 or 3, and the positions in the stream should be $\frac{n}{4}$, $\frac{2n}{4}$ and/or $\frac{3n}{4}$. Therefore, we assumed that the number of controlled vehicles is limited and their positions in the stream are relatively far apart. A scenario was defined for every possible configuration of set M, which makes a total of eight scenarios, including seven control scenarios and a no-control scenario (scenario with $m = 0$).

3.2.2 Common inputs

The simulation time step length (T_s) is 0.5 s and the control time step length (T_c) is 8 s. The total simulation period length (T) is 800 s. Note that T is a multiple of T_s and T_c, and it is large enough to ensure that all vehicles pass point R before the last simulation time step.

The freeway stretch has one lane and can be divided in three consecutive sections on the basis of its vertical profile (see Figure 4): a) constant-gradient downhill section; b) sag vertical curve; and c) constant-gradient uphill section. The vertical curve starts in point r_d and ends in point r_u. The gradient upstream of point r_d (which is denoted by G_d) is constant and negative, whereas the gradient downstream of point r_u (G_u) is constant and positive. Between points r_d and r_u, the freeway slope increases linearly over distance (Figure 4). Table 1 shows the values of parameters r_d, r_u, G_d and G_u, as well as the limits

of the freeway section within which the control input is applicable (r_0^c and r_f^c) and the location of the arrival point used to calculate travel times (R). Note that point R is far enough from the sag vertical curve so as to ensure that traffic is in stationary conditions at that location.

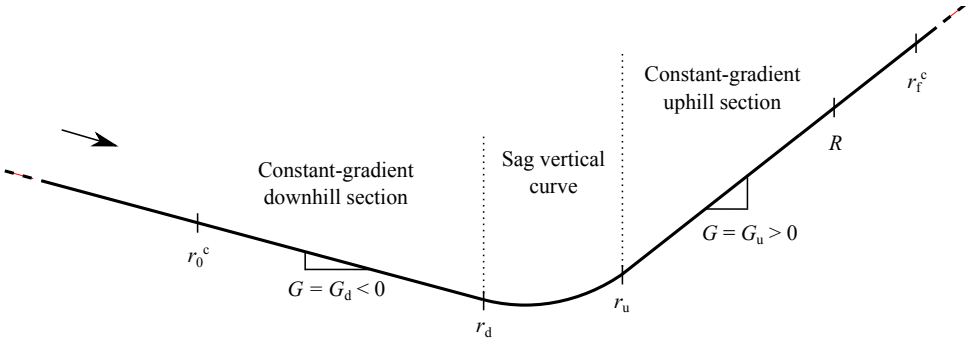


Figure 4: Vertical profile of the freeway stretch.

Table 1: Values of the parameters that describe the system.

Network parameters		CF model parameters	
Parameter	Value	Parameter	Value
G_d (%)	-0.5	$v_{des,i}$, $\forall i$ (km/h)	120
G_u (%)	$+2.5$	α_i, $\forall i$ (m/s^2)	1.4
r_d (m)	1000	β_i, $\forall i$ (m/s^2)	2.1
r_u (m)	1600	H_i, $\forall i$ (s)	1.20
R (m)	5000	$s_{s,i}$, $\forall i$ (m)	3
r_0^c (m)	-2000	θ_i, $\forall i$ (m/s^2)	22
r_f^c (m)	7000	λ_i, $\forall i$ (s^{-1})	0.0004
		$a_{min,i}$, $\forall i$ (m/s^2)	-8

All vehicle-driver units are 4 m long and are assigned the same value for every parameter of the car-following model (Table 1). Hence, the car-following behavior of all units is based on the same rules. This setup reduces the complexity of traffic dynamics and makes it easier to compare different scenarios.

At time zero, the stream of vehicles is in homogeneous and stationary conditions, the initial speed of all vehicle-driver units is equal to the desired speed ($v_{des,i}$), and the first vehicle of the stream is located at point r_0^c. It is assumed that the traffic density at time zero is the critical density of the constant-gradient downhill section. Therefore, since the values of parameters $s_{s,i}$, $v_{des,i}$ and H_i are the same for all vehicles i, the initial position of every vehicle i is defined as follows:

$$r_i(0) = r_0^c - (i-1) \cdot (s_{s,i} + v_{des,i} \cdot H_i) \tag{20}$$

Since $r_i(0) < r_d, \forall i$, the freeway gradient at the initial location of all vehicles is equal to the slope of the constant-gradient downhill section (G_d). Initially, the compensated gradient is equal to G_d for all vehicles, so the freeway gradient has no influence on vehicle acceleration ($\delta_i(0) = 0, \forall i$).

The lower bound (u_{min}) and the upper bound (u_{max}) that define the admissible control region are set to -0.5 m/s^2 and 1.4 m/s^2, respectively. The lower bound corresponds to a moderate deceleration

rate, which prevents too strong decelerations. The upper bound is equal to the value of the car-following model parameter α (see Table 1), which determines the maximum vehicle acceleration.

3.3 Solution method

We developed a MATLAB program that solves the optimal control problem for all scenarios. The core of the program is the optimization function `fmincon`, which uses sequential quadratic programming to iteratively solve nonlinear optimization problems. The `fmincon` function was set to use the `sqp` solution algorithm. Note that the algorithm has a limit of cost function evaluations and requires an initial solution guess as input. In our case, we define the initial solution guess as follows. The control input is equal to the maximum acceleration parameter (see Table 1) at all control time steps for all controlled vehicles:

$$\mathbf{u}^{\mathrm{g}}(\kappa) = \begin{bmatrix} \alpha_1(\kappa) & \alpha_2(\kappa) & \dots & \alpha_m(\kappa) \end{bmatrix}, \text{ for } \kappa = 1, 2, \dots, \tfrac{T}{T_{\mathrm{c}}} \tag{21}$$

With this initial solution guess, the control input in the first optimization iteration is not lower than the acceleration given by the car-following model at any control step; therefore, the trajectories of controlled vehicles are the same as in the no-control scenario (see Equations 6-12).

3.4 Performance Indicator

To evaluate the extent to which the optimal controller reduces congestion, the average vehicle delay (AVD) in every control scenario is compared to the AVD in the scenario with no controlled vehicles (no-control scenario). The AVD in a given scenario is calculated as follows:

$$\mathrm{AVD} = \frac{\mathrm{TTT} - \mathrm{TTT}_{\mathrm{ref}}}{n} \tag{22}$$

where TTT is the total travel time in that scenario, and $\mathrm{TTS}_{\mathrm{ref}}$ is the total travel time in the reference scenario. The reference scenario is a hypothetical one in which there are no controlled vehicles and the gradient is constant along the whole freeway stretch (i.e., $G_{\mathrm{u}} = G_{\mathrm{d}}$).

In addition, the AVD values in the seven control scenarios (see Section 3.2) are compared to the AVD values in seven analogous scenarios in which it is assumed that the acceleration of the controlled vehicles is not regulated by the optimal controller, but by a basic ACC system. In general, the acceleration behavior of ACC-equipped vehicles is not affected by the vertical profile of the freeway; thus, equipping vehicles with this type of systems can reduce congestion at sags [11]. Therefore, this additional comparison allows us to compare the effectiveness of the optimal controller with that of a simple traffic management strategy, which is used as a benchmark.

We assume that the ACC system operates in two modes: cruising mode and following mode. The system operates in cruising mode when there is no vehicle within the detection range of the ACC sensor (r^{ACC}). In this mode, the system tries to maintain a user-defined speed ($v_{\mathrm{des},i}^{\mathrm{ACC}}$). When there is a vehicle within range, the system operates in following mode. Then, it tries to maintain a user-defined time

headway (H_i^{ACC}). More specifically, the ACC system regulates vehicle acceleration on the basis of the following control law (see more details in [15]):

$$a_i^{\text{ACC}}(p) = \begin{cases} K_1 \cdot \left(v_{\text{input},i}(s_i(p)) - v_i(p)\right) - K_2 \cdot \dfrac{\Delta v_i(p)}{s_i(p)} & \text{if } s_i(p) \leq r^{\text{ACC}} \\ K_1 \cdot \left(v_{\text{des},i}^{\text{ACC}} - v_i(p)\right) & \text{if } s_i(p) > r^{\text{ACC}} \end{cases} \tag{23}$$

where p is the control time step counter, and K_1 and K_2 are feedback gains. Variable $v_{\text{input},i}$ is the target speed in following mode, which is gap-dependent:

$$v_{\text{input},i}(s_i(p)) = \min\left(\frac{s_i(p) - s_{\text{s},i}}{H_i^{\text{ACC}}}, v_{\text{des},i}^{\text{ACC}}\right) \tag{24}$$

This control law is subject to a non-collision constraint ($s_i(p) > 0, \forall p, \forall i \in M$) and a constraint that defines an admissible acceleration range ($a_{\text{min},i}^{\text{ACC}} \leq a_i^{\text{ACC}}(p) \leq a_{\text{max},i}^{\text{ACC}}, \forall p, \forall i \in M$). We chose the following values for the ACC parameters (which are the same for all ACC-equipped vehicles): $v_{\text{des},i}^{\text{ACC}} = 120$ km/h; $H_i^{\text{ACC}} = 1.20$ s; $a_{\text{min}}^{\text{ACC}} = -8$ m/s^2; $a_{\text{max}}^{\text{ACC}} = 1.4$ m/s^2; $r^{\text{ACC}} = 150$ m; $K_1 = 0.2$ s^{-1}; and $K_2 = 15$ m/s. The ACC control time step length is 0.05 s.

4 Results

This section presents the results of the optimization experiments. Section 4.1 provides an overview of the results in all scenarios. Section 4.2 describes the primary strategy applied by the controlled vehicles, including the characteristics of the maneuver and its effects on traffic flow dynamics and total travel time. Section 4.3 describes what we called the supporting strategy.

4.1 Overview of the optimization results in all scenarios

This section provides an overview of the results in all scenarios. Section 4.1.1 discusses the optimality of the solutions. Section 4.1.2 briefly describes the two main strategies that define the optimal acceleration behavior of controlled vehicles in all scenarios (i.e., primary strategy and supporting strategy). Section 4.1.3 discusses the effectiveness of the controller in reducing the AVD in every scenario. Section 4.1.4 compares the influence that the primary and supporting strategies have on the average vehicle delay (AVD). Finally, Section 4.1.5 compares the effectiveness of the optimal controller with that of a simple traffic management strategy based on equipping vehicles with ACC systems.

4.1.1 Optimality of the solutions

The solver finds an optimal solution before reaching the maximum number of iterations for four of the seven control scenarios, namely the scenarios with M equal to $\{75\}$, $\{150\}$, $\{225\}$ and $\{75, 225\}$. Note that, given the characteristics of the optimal control problem and the solution algorithm, we cannot ensure that the solutions found for those scenarios correspond to global optima (strictly speaking, they can only be considered local optima). For the remaining three control scenarios, the solver does not find an

optimal solution before reaching the maximum number of iterations. However, in these cases, the value of the cost function in the last iterations is very stable, which indicates that the solution given as output in the last iteration is probably close to the (local) optimal solution. Interestingly, controlled vehicles behave in a similar way in all scenarios, which suggests that the solutions may be (close to) global optima. Be it as it may, the solutions obtained reveal a potentially greatly effective (and innovative) way to mitigate congestion at sags via in-car systems.

4.1.2 Main strategies applied by the optimal controller

The results show that the optimal acceleration behavior of controlled vehicles is defined by two main strategies, which we call primary strategy and supporting strategy. The characteristics of those two strategies and their effects on traffic flow dynamics are described in detail in Sections 4.2 and 4.3, respectively. Here we briefly describe their main principles. The *primary* strategy consists in performing a deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. Remarkably, that strategy is used by all controlled vehicles in all scenarios. In all cases, the primary strategy limits the inflow to the sag and increases its outflow for a particular time interval, which produces considerable total travel time savings. The *supporting* strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the vertical curve. That strategy is only used by some controlled vehicles in some scenarios. The role of the supporting strategy is to manage congestion upstream of the sag in such a way that the inflow to the vertical curve is regulated more effectively, which increases the effectiveness of the primary strategy.

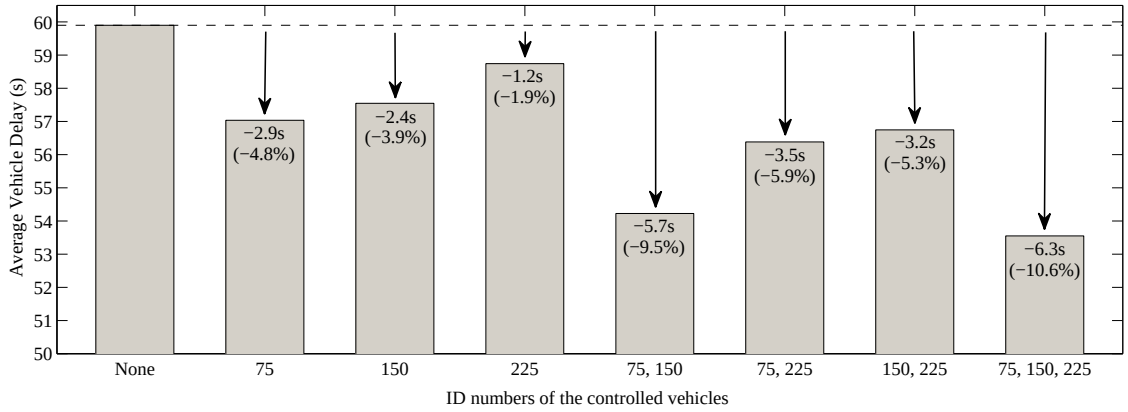
4.1.3 Reduction in average vehicle delay (AVD)

Figure 5(a) shows the average vehicle delay (AVD) in all scenarios. As seen in the figure, the optimal controller reduces the AVD by 2 to 11% (depending on the control scenario) in comparison with the no-control scenario. In our experiments, the greatest reduction in AVD is achieved in the scenarios with more controlled vehicles (Figure 5(a)). The reason is that every controlled vehicle causes a temporary period of increased sag outflow. Those periods do not overlap with each other. Therefore, the more controlled vehicles, the longer the total period of high outflow and, hence, the lower the total travel time. Note, however, that this rule might not apply if the controlled vehicles were located closer to each other and/or the number of controlled vehicles was higher. Under those circumstances, the periods of increased sag outflow caused by the controlled vehicles might overlap; hence having more vehicles behave as described in Section 4.1.2 might not lead to significant extra AVD savings.

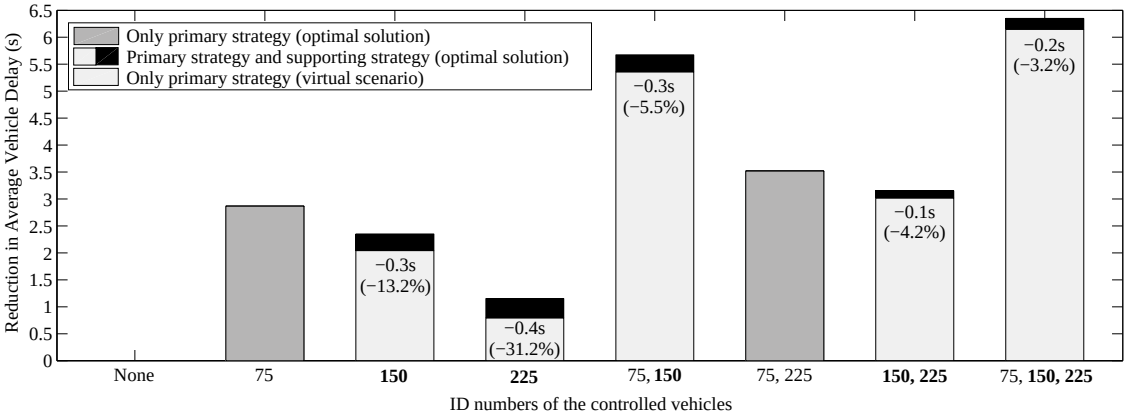
If we compare scenarios with the same number of vehicles, the greatest reduction in AVD is observed in the scenarios in which the controlled vehicles are closer to the first vehicle of the stream (see Figure 5(a)). The reason is as follows. In all control scenarios, the controlled vehicles' behavior has a similar impact on the time headways of the following vehicles at point R. Therefore, the contribution to total travel time of those following vehicles decreases to a greater extent in comparison with the no-control scenario if they are closer to the first vehicle of the stream, because then they influence the travel time of a greater number of vehicles (see Appendix).

4.1.4 Influence of each strategy on AVD reduction

Let us now compare the AVD in every scenario in which some or all controlled vehicles use the supporting strategy with the AVD in corresponding *virtual* scenarios in which the supporting strategy is omitted. The behavior of controlled vehicles in the virtual scenarios is defined on the basis of the solutions obtained for the control scenarios: in the virtual scenarios, controlled vehicles perform a DADA maneuver in the sag area that has exactly the same characteristics as in the control scenario, but they do not perform any additional maneuver upstream of the sag. As shown in Figure 5(b), in the scenarios in which some controlled vehicles use the supporting strategy, the decrease in AVD in comparison with the



(a) Average vehicle delay (AVD) in every scenario. The numbers inside the bars indicate the absolute and relative difference in comparison with the no-control scenario.



(b) Reduction in average vehicle delay (AVD) in every control scenario in comparison with the no-control scenario. In the horizontal axis, numbers in normal font indicate controlled vehicles that use only the primary strategy, whereas numbers in bold indicate controlled vehicles that use both the primary strategy and the supporting strategy. For the control scenarios in which controlled vehicles use the supporting control strategy, the figure also shows the AVD reduction in the corresponding virtual scenarios (in which the supporting strategy is not applied). For those scenarios, the numbers inside the bars indicate the absolute and relative difference in AVD reduction between the virtual scenario (without supporting strategy) and the control scenario with the optimal solution (which includes the supporting strategy).

Figure 5: Average vehicle delay in every scenario and differences in comparison with the no-control scenario.

no-control scenario would be 0.1 to 0.4 s lower if those vehicles did not use that strategy. Except for the control scenario with $M = \{225\}$, this represents a small (although significant) decrease in AVD reduction (3-13% depending on the scenario). We conclude that making some controlled vehicles apply the supporting strategy in combination with the primary strategy can reduce the AVD to a slightly greater extent than if those vehicles use only the primary strategy. However, in all scenarios, the primary strategy is the one that is more responsible for the reduction in average vehicle delay (see Figure 5(b)).

4.1.5 Comparison with scenarios with ACC-equipped vehicles

As shown in Figure 6, the optimal controller is considerably more effective than the ACC systems in reducing the AVD, at least at penetration rates lower or equal than 1%. In our experiments, the optimal controller reduces the AVD by 2 to 10% more (depending on the scenario) than the ACC systems in comparison with the no-control scenario. The reason for this is as follows. Since the acceleration behavior of ACC-equipped vehicles is not affected by the sag vertical curve, ACC-equipped vehicles reduce their time headways at point R in comparison with the no-control scenario. However, ACC-equipped vehicles do not have any significant influence on the headways of the following vehicles at this point, which implies that they do not cause any significant increase in outflow. For this reason, the headway reduction of the ACC-equipped vehicles has a marginal effect on the AVD. Instead, vehicles controlled by the optimal controller do have an influence on the headways of the following vehicles at point R, which causes a temporary increase in outflow. The ability of the optimally-controlled vehicles to influence the behavior of the following vehicles and thereby cause a significant increase in bottleneck outflow is what makes this controller reduce the AVD to a greater extent than a basic ACC controller.

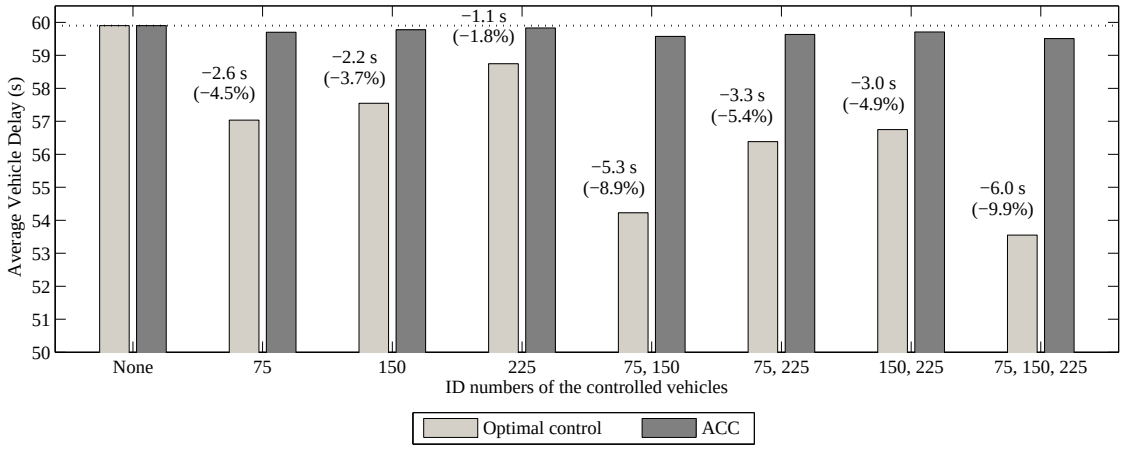


Figure 6: Average vehicle delay (AVD) in the no-control scenario, the optimal control scenarios and the ACC scenarios. The numbers on top of the light-gray bars indicate the difference in AVD and the difference in relative AVD reduction in comparison with the no-control scenario between the optimal control scenarios and the ACC scenarios with the same M.

4.2 Primary strategy: deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area

This section describes the primary strategy applied by controlled vehicles. Section 4.2.1 describes the main characteristics of the maneuver. Sections 4.2.2 and 4.2.3 describe the typical effects of the maneuver on traffic flow and total travel time, respectively. For illustrative purposes, throughout this section we show the results of the optimization experiment corresponding to the scenario with $M = \{75\}$, which constitute a good example of the typical acceleration behavior of controlled vehicles in the sag area and the resulting effects.

4.2.1 Characteristics of the DADA maneuver

In all scenarios, all controlled vehicles behave similarly when they reach the sag area. They perform a maneuver consisting of four phases (see, for example, Figure 7): first deceleration phase, first positive-acceleration phase, second deceleration phase and second positive-acceleration phase.

The *first deceleration phase* (D1) begins upstream of the sag or right after entering it (depending on the case). In this phase, controlled vehicles decelerate moderately (normally at around the minimum acceleration rate allowed by the controller, $u_{\min}$), even though they are not obliged to decelerate that fast in order to adjust to the behavior of the leader. During this phase, the distance headway of controlled vehicles increases considerably.

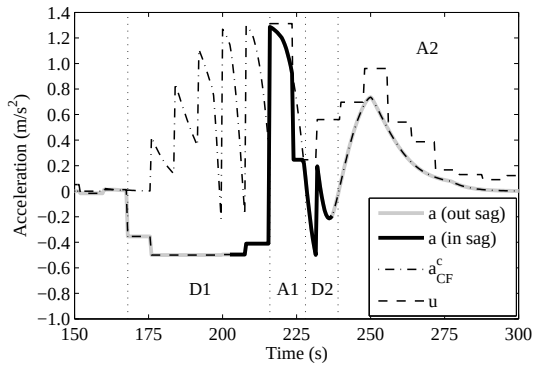
The *first positive-acceleration phase* (A1) begins somewhere halfway through the vertical curve. During this phase, controlled vehicles accelerate fast (with maximum acceleration rates up to 1 m/s^2 or higher). Fast acceleration is possible because the distance between controlled vehicles and their predecessor is quite long. The distance gap decreases quickly during this phase.

When controlled vehicles reach the last part of the vertical curve, they undergo a *second deceleration phase* (D2). At this point, their distance headway has become short enough to force them to decelerate a bit in order to adjust to the behavior of their leader. However, during this phase, the distance headway of controlled vehicles continues to decrease, because their predecessor moves at a lower speed.

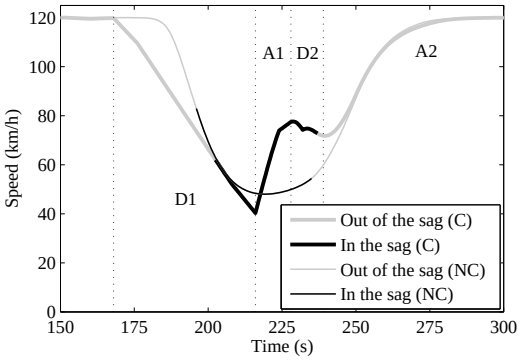
Eventually, controlled vehicles catch up with their leader. This generally occurs around the end of the vertical curve. At that point, the *second positive-acceleration phase* (A2) begins. In this phase, controlled vehicles simply accelerate to the desired speed. Acceleration is lower than in phase A1. Note that, after exiting the sag, the trajectories of controlled vehicles are almost the same as in the no-control scenario, although in some cases the distance headway is slightly longer.

4.2.2 Effects on traffic flow dynamics in the sag area

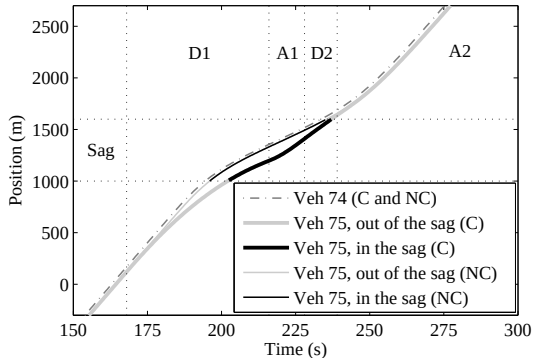
In all cases, the DADA maneuvers have three main effects on traffic flow dynamics. Firstly, the first positive-acceleration phase (A1) of the maneuver induces the first vehicles located behind the controlled vehicle (up to 85 vehicles in some cases) to accelerate fast along the sag (see, for example, Figure 8). As a result, traffic speed at the end of the sag (bottleneck) increases and stays moderately high (70-90 km/h) for a particular period (2-3 min), contrary to what happens in the no-control scenario (compare



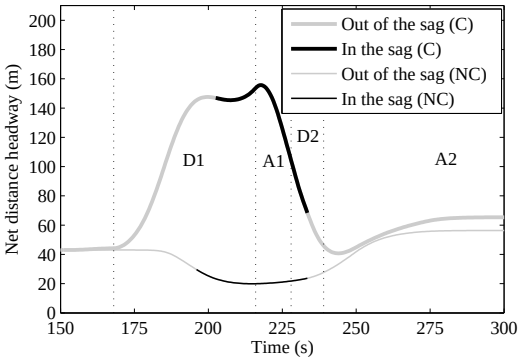
(a) Acceleration of vehicle 75 over time in the scenario with $M = \{75\}$ (a is the realized acceleration, u denotes the control input, and a^c_{CF} denotes the acceleration given by the car-following model). Note that a increases and decreases rapidly between instants 232 and 236 s; that is just a by-product of the interaction between u and a^c_{CF} at the previous control time step (the average acceleration rate in phase D2 is -0.13 m/s^2). This phenomenon is also observed in other scenarios.



(b) Speed of vehicle 75 over time in the scenario with $M = \{75\}$ (C) and the no-control scenario (NC).



(c) Positions of vehicle 75 and its leader over time in the scenario with $M = \{75\}$ (C) and the no-control scenario (NC).



(d) Net distance headway of vehicle 75 over time in the scenario with $M = \{75\}$ (C) and the no-control scenario (NC).

Figure 7: Deceleration-acceleration-deceleration-acceleration (DADA) maneuver of the controlled vehicle in the scenario with $M = \{75\}$.

Figures 9(a) and 9(b)). That has important implications for the sag outflow. The car-following model assumes that the desired time headway of drivers decreases if the speed increases. Hence, during the period of high traffic speeds, time headways at the end of the vertical curve (bottleneck) are considerably shorter than in the no-control scenario, which implies that the sag outflow is higher (generally, around 5% higher).

Secondly, the first deceleration phase (D1) of the maneuver triggers a stop-and-go wave on the first part of the sag (see Figure 9(b)). That stop-and-go wave temporarily limits the inflow to the vertical curve. The reason why that is beneficial is as follows. The positive effects that phase A1 has on traffic speed and flow at the bottleneck are only temporary. Shortly after a DADA maneuver, traffic speed at the end of the sag begins to decrease (due, mainly, to the limiting effect that the vertical curve has on vehicle acceleration). As a result, traffic eventually becomes congested at that location (see Figure 9(b)),

at which point the outflow from the sag becomes similar to that in the no-control scenario. Limiting the inflow to the vertical curve slows down the process of formation of congestion at the end of the sag [12], hence high levels of sag outflow can be maintained for a longer period of time.

Finally, the second deceleration phase (D2) of the maneuver causes a small speed disturbance on the last part of the sag, which fades away as it propagates downstream (see Figure 8). The reason why that is beneficial is as follows. The occurrence of the speed disturbance makes the traffic speed at the end of the vertical curve increase for a short period of time right after the controlled vehicle passes that location, thus delaying the formation of congestion at the bottleneck (see Figure 8). In addition, as explained in the next section, the speed disturbance influences the time headways of the first group of vehicles located behind the controlled vehicle, which has relevant effects on total travel time.

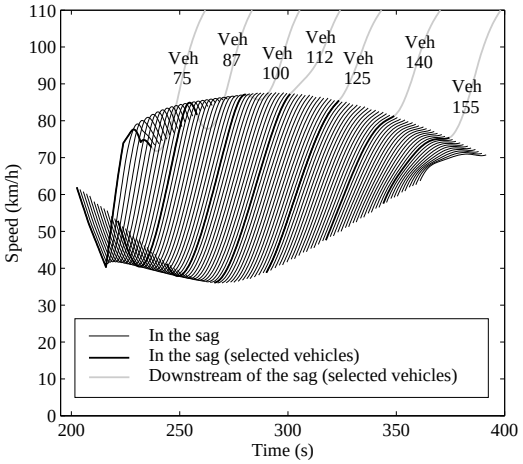
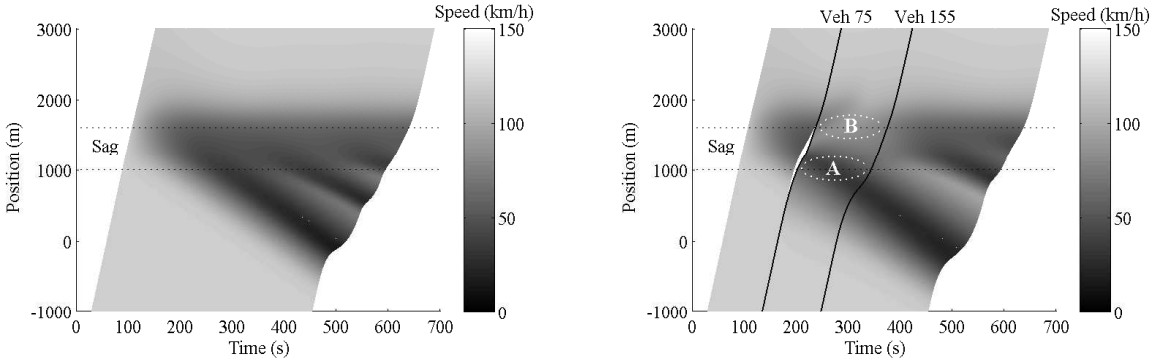


Figure 8: Effect of the DADA maneuver of the controlled vehicle on the behavior of the first 90 following vehicles in the control scenario with $M = \{75\}$.



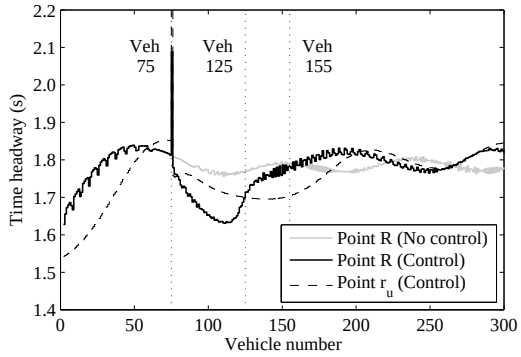
(a) No-control scenario.

(b) Scenario with $M = \{75\}$. *A* is a region of moderately low traffic speed and limited flow around the beginning of the sag; *B* is a region of moderately high traffic speed and high flow around the end of the sag.

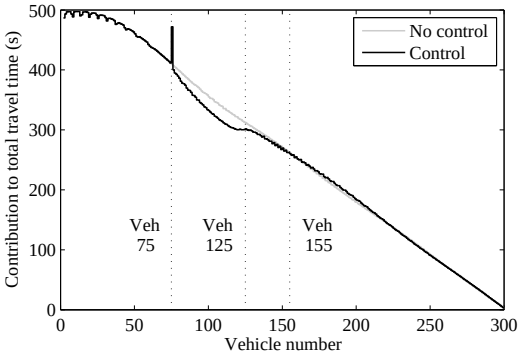
Figure 9: Speed contour plots of the no-control scenario and the control scenario with $M = \{75\}$.

4.2.3 Effects on total travel time

In all control scenarios, the vehicles that cause the decrease in total travel time in comparison with the no-control scenario are primarily the ones that exit the sag during the periods of high outflow induced by controlled vehicles. As explained in the Appendix, that is deduced from the fact that those vehicles keep shorter time headways at the arrival point (R) than in the no-control scenario, whereas most other vehicles do not (see, for example, Figure 10(a)).



(a) Time headway of every vehicle at points r_u and R. Time headways at point R oscillate due to differences in speed profile between consecutive vehicles as they accelerate to the desired speed after exiting the sag.



(b) Contribution to total travel time (ψ_i) of vehicles 2 to 300 (see Appendix). The area below each line corresponds to the total travel time in each scenario minus the contribution of vehicle 1 (which is the same in both scenarios).

Figure 10: Effect of the DADA maneuver of the controlled vehicle on the total travel time in the control scenario with $M = \{75\}$.

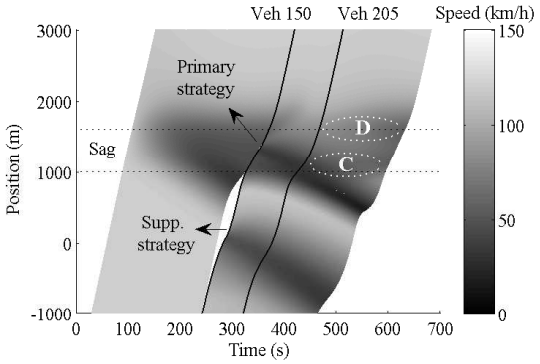
Furthermore, among the vehicles that exit the sag during the periods of high outflow, the ones that reduce the total travel time to a greater extent are generally the first 40-50 vehicles located behind each controlled vehicle (Figure 10(b)). Those vehicles show the greatest differences in time headway at point R in comparison with the no-control scenario (Figure 10(a)). The main reason is that their time headways decrease after exiting the sag (Figure 10(a)), whereas the time headways of the next group of vehicles increase. This occurs because the first 40-50 vehicles located behind a controlled vehicle accelerate to the desired speed a little faster than their immediate predecessors, since they are affected by the speed disturbance caused by the second deceleration phase of the controlled vehicle (Figure 8). Instead, the next vehicles accelerate to the desired speed a bit more slowly than their immediate predecessors, because they exit the sag at lower speeds (due to the vertical curvature) and they are not affected by the aforementioned speed disturbance (Figure 8).

We conclude that an important reason why DADA maneuvers include a second deceleration phase (D2) is that the resulting speed disturbance makes the first group of following vehicles reduce their time headways after exiting the vertical curve. The vehicles that are not affected by the disturbance increase their headways after exiting the sag. Therefore, the second deceleration phase increases the effectiveness of controlled vehicles in reducing total travel time. It is worth noting that, in all control scenarios, the time headway of some vehicles at point R (including the controlled vehicle, in some cases) is longer than in the no-control scenario (see, for example, Figure 10(a)). Overall, however, the increase in the

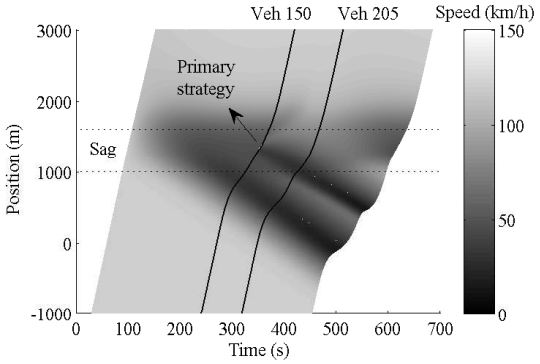
contributions to total travel time of those vehicles is compensated by the decrease in the contributions of the other vehicles (Figure 10(b)).

4.3 Supporting strategy: deceleration-acceleration maneuvers upstream of the sag area

In some scenarios, some controlled vehicles apply an additional strategy besides performing a DADA maneuver in the sag area, which we call *supporting* strategy. The supporting strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the sag, catching up with the preceding vehicle before entering the vertical curve (an example is shown in Figure 11(a)). The characteristics of these maneuvers are very situation-specific: in different scenarios, they start in different locations, they are quicker or slower, they are performed consecutively or separately, and they last for different periods of time. However, their overall effect on traffic flow dynamics is similar.



(a) Control scenario with $M = \{150\}$. *C* is a region located on the first part of the sag that has slightly lower speed and flow than in the virtual scenario; *D* is a region around the end of the sag that has slightly higher speed and flow than in the virtual scenario.



(b) Virtual scenario corresponding to the control scenario with $M = \{150\}$ (the supporting strategy is omitted).

Figure 11: Speed contour plots of the control scenario with $M = \{150\}$ and the corresponding virtual scenario in which the supporting strategy is omitted. In this control scenario, the controlled vehicle uses both the primary and the supporting strategy; note, however, that the DADA maneuver corresponding to the primary strategy is shorter than in the control scenario with $M = \{75\}$ (see Figure 9(b)).

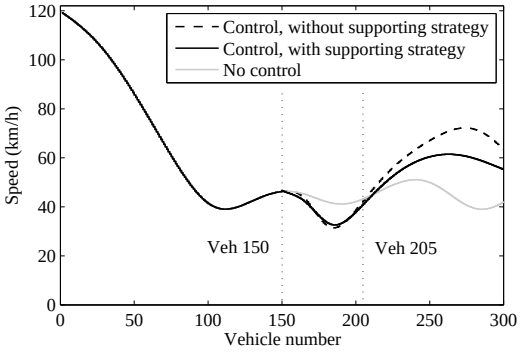
Recall from Section 4.2.2 that the primary strategy triggers a stop-and-go wave on the first part of the sag. This is beneficial because it limits the inflow into the vertical curve, which slows down the formation of congestion at the bottleneck. However, this limitation in inflow is only temporary. Once the stop-and-go wave moves away from the vertical curve, the inflow increases again, and this coincides with the formation of congestion at the bottleneck (see Figure 9(b)).

The controlled vehicles use the supporting strategy to change the location and severity of congestion upstream of the vertical curve (compare Figures 11(a) and 11(b)) in such a way that, after the stop-and-go wave caused by the primary strategy moves away from the sag, the inflow into the bottleneck stays somewhat lower than if the supporting strategy was not applied. This can be seen in Figures 12(a) and 12(c), which show that, in the control scenario corresponding to Figure 11(a), vehicles 205 to 300 drive a bit more slowly and keep slightly longer headways on the first part of the sag than in the

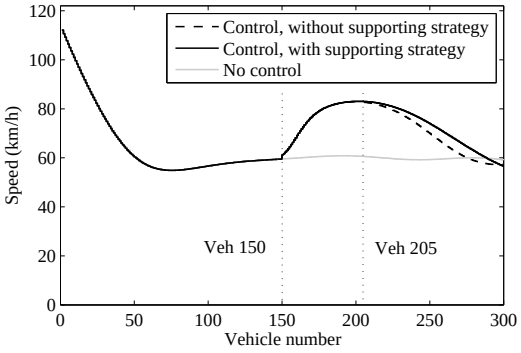
corresponding virtual scenario without supporting strategy (Figure 11(b)).

This reduction in inflow slows down the formation of congestion at the bottleneck compared to if the supporting strategy was not applied. As a result, the primary strategy is able to produce higher traffic speeds and flows at the end of the sag after the above-mentioned stop-and-go wave moves away from the vertical curve (see Figures 12(b) and 12(d), which show that, in the control scenario corresponding to Figure 11(a), vehicles 205 to 300 drive a bit faster and keep slightly shorter headways at the end of the sag than in the corresponding virtual scenario without supporting strategy). This implies a small increase in bottleneck outflow, which brings about some extra total travel time savings (see Figure 5(b)).

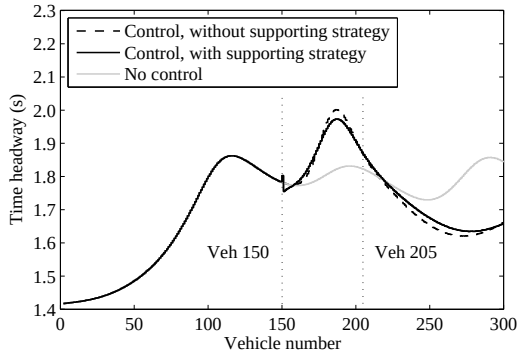
It is important to highlight that the supporting strategy is not applied alone in any scenario: it is always applied in combination with the primary strategy. In fact, as explained in Section 4.1.4, if we compare the average vehicle delay (AVD) in every scenario in which some or all controlled vehicles use the supporting strategy with the AVD in the corresponding virtual scenarios without supporting strategy, we can observe that the primary strategy is the one that contributes the most to reduce the total travel time (see Figure 5(b)). The function of the supporting strategy seems to be merely to increase the effectiveness of the primary strategy.



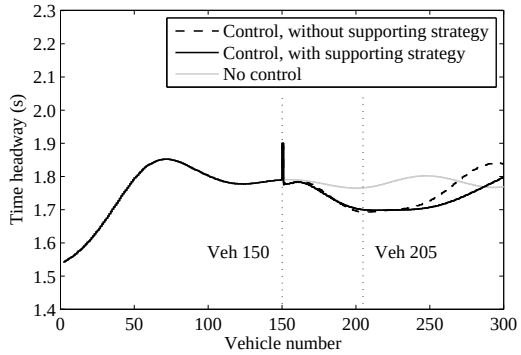
(a) Speed of every vehicle at point r_{d}+150 m.



(b) Speed of every vehicle at the end of the sag (r_{u}).



(c) Time headway of every vehicle at point r_{d}+150 m.



(d) Time headway of every vehicle at the end of the sag (r_{u}).

Figure 12: Speed and time headway of every vehicle on the first part of the sag and at the end of the sag in the no-control scenario, in the control scenario with $M = \{150\}$ (optimal solution) and in the virtual scenario without supporting strategy corresponding to that control scenario.

5 Conclusions

The main goal of this study was to identify the way in which vehicles equipped with in-car systems need to move at sags in the longitudinal direction in order to reduce congestion as much as possible, assuming low penetration rates. To this end, we formulated an optimal control problem in which a centralized controller regulates the acceleration of some vehicles belonging to a traffic stream that moves along a single-lane freeway stretch with a sag. The control objective is to minimize the total travel time. The problem was solved for various scenarios defined by the number of controlled vehicles and their positions in the stream. Although we cannot guarantee global optimality, the solutions obtained reveal a potentially greatly effective and innovative way to manage traffic at sags using in-car systems.

The analysis of the solutions indicates that the optimal acceleration behavior of controlled vehicles is defined by two strategies, which we called primary strategy and supporting strategy. The *primary* strategy is applied in all cases and is the one that contributes the most to reduce the total travel time. It involves performing a specific type of deceleration-acceleration-deceleration-acceleration (DADA) maneuver in the sag area. This maneuver induces the first group of following vehicles to accelerate fast along the sag. As a result, traffic speeds at the end of the sag (bottleneck) increase for a particular time interval. During this period, the sag outflow increases up to 5% in comparison with the no-control scenario, which produces substantial travel time savings for all vehicles located behind the equipped vehicle. Interestingly, every DADA maneuver triggers a stop-and-go wave on the first part of the sag that temporarily limits the inflow to the vertical curve. Limiting the inflow is beneficial because it slows down the process of formation of congestion at the end of the sag (bottleneck), hence high levels of sag outflow can be maintained for a longer period of time. The *supporting* strategy consists in performing one or more deceleration-acceleration maneuvers upstream of the sag. Controlled vehicles use the supporting strategy to manage congestion upstream of the sag in such a way that the inflow to the vertical curve is regulated more effectively, thus enhancing the effectiveness of the primary strategy.

Our findings provide valuable insight into how congestion can be reduced at sags by means of traffic management measures using in-car systems. More specifically, they highlight the relevance of limiting the inflow into sags and of motivating drivers to accelerate fast along the vertical curve; in addition, they indicate ways to do that by regulating/influencing the acceleration of vehicles equipped with in-car systems. Note that a feedback control measure that regulates the inflow into sags (in this case, through variable speed limits) was already presented and evaluated in [12]. In this respect, our experiments indicate that a combined strategy aimed at both limiting inflow and inducing fast acceleration along the vertical curve can reduce congestion at sags to a greater extent than a strategy aimed solely at limiting the inflow. Another type of traffic management measure that has been proposed for sags is to equip vehicles with ACC systems [11]. In general, the acceleration behavior of these vehicles is not affected by the vertical profile of the freeway, hence their trajectories are smoother than those of vehicles driven by human drivers. Interestingly, our findings show that making individual vehicles perform a DADA maneuver at the sag area reduces congestion to a greater degree than equipping those vehicles with basic ACC systems (at least if low penetration rates are assumed). The main reason is that ACC-equipped vehicles do not have any significant influence on the behavior of their followers, unlike vehicles that

perform a DADA maneuver. Furthermore, our findings prove the usefulness of the proposed optimization method as a tool for control measure development. We conclude that this method can be used to identify effective traffic management strategies for other types of bottlenecks besides sags.

Further research is necessary to determine whether the traffic management strategies identified in this paper are really globally optimal and, if this is the case, whether they are also optimal in other relevant scenarios (such as scenarios with multi-lane freeways, different positions of the equipped vehicles in the stream, higher penetration rates and/or lower traffic demand). In order to do this, additional optimization experiments need to be carried out. Note that performing optimization experiments for multi-lane scenarios will require modifying the formulation of the optimal control problem in order to include the possibility that vehicles change lanes and to specify the lateral driving behavior of drivers. Our expectation is that the traffic management strategies identified in this paper (i.e., motivating drivers to accelerate fast along the sag, and limiting inflow) will also be effective and optimal in those other relevant scenarios, since these strategies seem to maximize the bottleneck outflow.

Also, further research is necessary to translate the identified strategies into effective traffic control measures. Mainly, these measures should induce vehicles equipped with specific in-car systems to perform a DADA maneuver at sags in high demand conditions. The characteristics of the DADA maneuvers may depend on the freeway layout and the traffic conditions. Also, supplementary strategies may be necessary to facilitate the DADA maneuvers and to enhance their effectiveness in reducing congestion. In that respect, the use of special ACC systems capable of communicating with the infrastructure constitutes a promising direction. In order to make vehicles equipped with such systems perform a DADA maneuver at sags, a local road-side controller could increase the target time headway of the ACC systems right before equipped vehicles enter the vertical curve and decrease it again when they are halfway through the sag. This type of measure, which is currently under investigation, may need to be combined with inflow-regulation measures and/or lane-change management measures in order to improve its effectiveness.

Appendix

In Section 2.3.5, total travel time (TTT) was defined as the sum of the travel times of all vehicles i from their initial positions to point R (which are denoted by TT_i):

$$\text{TTT} = \sum_{i=1}^{n} \text{TT}_i \tag{25}$$

The time headway of vehicle i at point R can be defined as the travel time of vehicle i minus the travel time of its predecessor:

$$h_{i,R} = \text{TT}_i - \text{TT}_{i-1}, \quad \text{for } i = 2, \ldots, n \tag{26}$$

From Equation 26, it can be derived that the travel time of vehicle $i = 2, \ldots, n$ is equal to the travel

time of vehicle 1 plus the headways of vehicles 2 to i:

$$\mathrm{TT}_i = \begin{cases} \mathrm{TT}_1 & \text{if } i = 1 \\ \mathrm{TT}_1 + \sum_{\nu=2}^{i} h_{\nu,R} & \text{if } i = 2, \ldots, n \end{cases} \tag{27}$$

By substituting Equation 27 into Equation 25 and rearranging the terms, the total travel time can be expressed as a function of the travel time of vehicle 1 and the headways of vehicles 2 to n at point R:

$$\mathrm{TTT} = n \cdot \mathrm{TT}_1 + \sum_{i=2}^{n} \left((n - i + 1) \cdot h_{i,R} \right) \tag{28}$$

From Equation 28, it follows that the contribution of each vehicle to total travel time (ψ_i) is:

$$\psi_i = \begin{cases} n \cdot \mathrm{TT}_1 & \text{if } i = 1 \\ (n - i + 1) \cdot h_{i,R} & \text{if } i = 2, \ldots, n \end{cases} \tag{29}$$

The sum of the contributions to total travel time of all vehicles equals the total travel time:

$$\mathrm{TTT} = \sum_{i=1}^{n} \psi_i \tag{30}$$

We conclude that the vehicles that reduce the total travel time the most in a given control scenario are the ones whose ψ_i decreases to a greater extent in comparison with the no-control scenario. From Equation 29, it follows that ψ_i decreases more for vehicles whose time headways at point R decrease to a greater extent and are located closer to the first vehicle of the stream.

Acknowledgments

This work was supported by Toyota Motor Europe. The authors would like to thank Andreas Hegyi for many helpful comments and suggestions.

References

[1] Xing, J., Sagae, K., Muramatsu, E. Balance Lane Use of Traffic to Mitigate Motorway Traffic Congestion with Roadside Variable Message Signs. Presented at 17th ITS World Congress, Busan, South Korea, 2010.

[2] Koshi, M., Kuwahara, M., Akahane, H., 1992. Capacity of Sags and Tunnels on Japanese Motorways. ITE Journal, 62(5), 17–22.

[3] Brilon, W., Bressler, A., 2004. Traffic Flow on Freeway Upgrades. Transportation Research Record: Journal of the Transportation Research Board, 1883, 112–121.

[4] Patire, A.D., Cassidy, M.J., 2011. Lane Changing Patterns of Bane and Benefit: Observations of an Uphill Expressway. Transportation Research Part B: Methodological, 45(4), 656–666.

[5] Xing, J., Muramatsu, E., Harayama, T., 2014. Balance Lane Use with VMS to Mitigate Motorway Traffic Congestion. International Journal of Intelligent Transportation Systems Research, 12(1), 26–35.

[6] Hatakenaka, H., Hirasawa, T., Yamada, K., Yamada, H., Katayama, Y., Maeda, M. Development of AHS for Traffic Congestion in Sag Sections. Presented at 13th ITS World Congress, London, UK, 2006.

[7] Yoshizawa, R., Shiomi, Y., Uno, N., Iida, K., Yamaguchi, M., 2012. Analysis of Car-following Behavior on Sag and Curve Sections at Intercity Expressways with Driving Simulator. International Journal of Intelligent Transportation Systems Research, 10(2), 56–65.

[8] Koshi, M., 2003. An Interpretation of a Traffic Engineer on Vehicular Traffic Flow. In: Fukui, M., Sugiyama, Y., Schreckenberg, M., Wolf, D.E. (Eds.), Traffic and Granular Flow'01, Springer, Berlin, 199–210.

[9] Goñi-Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P. Car-following Behavior at Sags and its Impacts on Traffic Flow. Presented at 92nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2013.

[10] Goñi-Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P., 2014. Empirical Analysis of the Causes of Stop-and-go Waves at Sags. IET Intelligent Transport Systems, 8(5), 499–506.

[11] Ozaki, H., 2003. Modeling of Vehicular Behavior from Road Traffic Engineering Perspectives. In: Fukui, M., Sugiyama, Y., Schreckenberg, M., Wolf, D.E. (Eds.), Traffic and Granular Flow'01, Springer, Berlin, 281–292.

[12] Goñi-Ros, B., Knoop, V.L., van Arem, B., Hoogendoorn, S.P., 2014. Mainstream Traffic Flow Control at Sags. Transportation Research Record: Journal of the Transportation Research Board, 2470, 57–64.

[13] Sato, H., Xing, J., Tanaka, S., Watauchi, T. An Automatic Traffic Congestion Mitigation System by Providing Real Time Information on Head of Queue. Presented at 16th ITS World Congress, Stockholm, Sweden, 2009.

[14] Kesting, A., Treiber, M., Lämmer, S., Schönhof, M., Helbing, D. Decentralized Approaches to Adaptive Traffic Control and an Extended Level of Service Concept. Presented at 17th International Conference on the Application of Computer Science and Mathematics in Architecture and Civil Engineering, Weimar, Germany, 2006.

[15] Papacharalampous, A.E., Wang, M., Knoop, V.L., Goñi-Ros, B., Takahashi, T., Sakata, I., van Arem, B., Hoogendoorn, S.P. Mitigating Congestion at Sags with Adaptive Cruise Control Systems. Presented at 18th IEEE International Conference on Intelligent Transportation Systems, Las Palmas de Gran Canaria, Spain, 2015.

[16] Goñi-Ros, B., Knoop, V.L., Shiomi, Y., Takahashi, T., van Arem, B., Hoogendoorn, S.P., 2016. Modeling Traffic at Sags. International Journal of Intelligent Transportation Systems Research, 14(1), 64–74.