Mainstream traffic flow control at sags

Bernat Goñi Ros
Delft University of Technology
Department of Transport and Planning
Stevinweg 1, 2628 CN Delft, The Netherlands
+31 15 278 4912
b.goniros@tudelft.nl
(Corresponding author)

Victor L. Knoop
Delft University of Technology
Department of Transport and Planning
Stevinweg 1, 2628 CN Delft, The Netherlands
v.l.knoop@tudelft.nl

Bart van Arem
Delft University of Technology
Department of Transport and Planning
Stevinweg 1, 2628 CN Delft, The Netherlands
b.vanarem@tudelft.nl

Serge P. Hoogendoorn
Delft University of Technology
Department of Transport and Planning
Stevinweg 1, 2628 CN Delft, The Netherlands
s.p.hoogendoorn@tudelft.nl

November 12, 2013

Word count:

Number of words in abstract	243
Number of words in text (including abstract)	6250
Number of figures and tables	5 * 250 = 1250
Total	7500

Submitted to the 93rd Annual Meeting of the Transportation Research Board

ABSTRACT

Sags are freeway sections along which the gradient changes significantly from downwards to upwards. The capacity of sags is significantly lower than the capacity of normal sections. As a result, sags often become bottlenecks in freeway networks, causing the formation of congestion in conditions of high traffic demand. Congestion results in a further decrease in capacity. Recently, several control measures have been proposed to improve traffic flow efficiency at sags. Those measures generally aim to increase the capacity of the bottleneck and/or to prevent the formation of traffic flow perturbations in nearly-saturated conditions. In this contribution, we present an alternative type of measure based on the concept of mainstream traffic flow control. The proposed control strategy regulates the traffic density at the bottleneck area in order to keep it slightly below the critical density, hence preventing traffic from breaking down while maximizing outflow. Density is regulated by means of a variable speed limit section that regulates the inflow to the bottleneck. Speed limits are set based on a proportional feedback control law. We evaluate the effectiveness of the proposed control strategy by means of a simple case study using microscopic traffic simulation. The results show a significant increase in bottleneck outflow, particularly during periods of very high demand, which leads to a considerable decrease in total delay. This finding suggests that mainstream traffic flow control strategies using variable speed limits have the potential to substantially improve the performance of freeway networks containing sags.

1 INTRODUCTION

Sags are freeway sections along which the gradient changes significantly from downwards to upwards in the direction of traffic *(1)*. The capacity of sags is generally lower than that of flat sections *(2, 3)*. In general, the bottleneck is located at the beginning of the uphill section, i.e., on the first 0.5-1.0 km downstream of the bottom of the sag *(4)*. As a consequence of the reduced capacity, traffic often breaks down at sags in conditions of high demand. The formation of congestion results in a further decrease in bottleneck capacity *(2)*. Recently, various control measures have been proposed to improve traffic flow efficiency at sags. In general, those measures aim to increase the capacity of the bottleneck and/or to prevent the formation of traffic flow perturbations in nearly-saturated conditions.

The objective of this paper is to present an alternative type of control strategy and to evaluate its potential effectiveness, performing a proof of principle. The proposed control strategy is based on the concept of mainstream traffic flow control *(5)*. The traffic density at the bottleneck area is regulated in order to keep it below the critical density, hence preventing traffic from breaking down. The capacity drop due to congestion does not occur, so the outflow from the bottleneck can be higher. The density at the bottleneck area is regulated by means of a variable speed limit section that regulates the inflow to the bottleneck. Speed limits are set based on a proportional feedback control law.

The effectiveness of the control strategy is evaluated by means of a simple case study using microscopic traffic simulation. Traffic flow is simulated in a single-lane freeway stretch containing a sag, with and without implementing the control strategy. The results show a significant increase in bottleneck outflow (particularly in periods of very high demand), which leads to a considerable decrease in total delay. This finding suggests that mainstream traffic flow control strategies using variable speed limits can considerably improve traffic flow efficiency in freeway networks containing sags.

The rest of this paper is structured as follows. Section 2 contains a literature review on the characteristics of traffic flow at sags and on types of control measures to mitigate congestion at sag bottlenecks. Section 3 describes the proposed control strategy. Section 4 describes the method used to evaluate the effectiveness of the control strategy. Section 5 presents the results of the evaluation. Section 6 contains a sensitivity analysis of the results of the evaluation. Section 7 presents the conclusions of this study as well as some suggestions for future research.

2 BACKGROUND

2.1 Sags as freeway bottlenecks

Bottlenecks are freeway sections that have a lower capacity than the immediate upstream section. Generally, the causes of that lower capacity are: i) spatial inhomogeneities (such as lane drops, ramps, curves, tunnels, and changes in gradient); ii) traffic conditions (e.g., slow vehicles or accidents); and/or iii) environmental conditions (e.g., adverse weather conditions) *(6, 7)*. It is important to remark that the capacity of a bottleneck depends on the traffic state: the capacity in congested traffic conditions (queue discharge capacity) is significantly lower than the capacity in uncongested traffic conditions (free flow capacity). This difference, which is called *capacity drop*, ranges from 3% to 20% according to different studies *(8, 9, 10)*. When traffic demand exceeds the free flow capacity of a bottleneck, congestion forms upstream of the bottleneck. As a result of the capacity drop, the formation of congestion causes the capacity of the bottleneck to decrease further, to the queue discharge capacity.

Several empirical studies show that the capacity of sags can be significantly lower than the capacity of flat sections having the same number of lanes *(2, 3)*. In general, the lower part of the uphill section (i.e., first 0.5-1.0 km downstream of the bottom of the sag) is the main bottleneck *(4)*. Xing *et al. (3)* present empirical measurements of the free flow capacities and the queue discharge capacities of various sag sections of Japanese freeways. Most of the measurements were taken on holidays, when traffic demand consists mainly of passenger cars and the percentage of heavy vehicles is relatively low. According to the data presented in that study, the average free flow capacity is 3150 veh/h at two-lane sags and 5340 veh/h at three-lane sags. The average queue discharge capacity is 2780 veh/h at two-lane sags and 4600 veh/h at three-lane sags, which means that the capacity drop is -12% and -14%, respectively. Similar capacity estimates have been reported by other authors *(2, 11)*.

If we compare the capacities of sags with those of flat sections, we can observe that the free flow capacity and the queue discharge capacity of sags are considerably lower. At flat sections, free flow capacities are generally around 4000 pcu/h (two lanes) and 6000 pcu/h (three lanes) *(2)*. Assuming a 10% capacity drop, we obtain queue discharge capacities for flat sections of 3600 pcu/h (two lanes) and 5400 pcu/h (three lanes). Therefore, the free flow capacity and the queue discharge capacity of two-lane freeways are around 20% lower at sags than at normal sections (10-15% lower in three-lane freeways). As a result, sags frequently become bottlenecks in freeway networks.

The main cause of capacity reduction at sags seems to be related to the impact that the change in freeway gradient has on the longitudinal driving behavior of drivers. Several empirical studies show that two important changes in longitudinal driving behavior occur when vehicles reach the uphill section. First, drivers tend to reduce speed *(1, 4)*. Second, drivers tend to keep longer distance headways than expected given their speed *(12, 13)*. These local changes in longitudinal driving behavior seem to be caused by the fact that drivers are unable to accelerate sufficiently and compensate for the increase in resistance force resulting from the increase in slope *(14)*.

2.2 Control measures to mitigate congestion at sags

In the last two decades, several measures have been proposed to prevent or delay the formation of congestion at sags, and to reduce its severity. In general, those measures can be sorted into three categories: a) measures that aim to increase the free flow capacity of sag bottlenecks; b) measures that aim to prevent the formation of traffic flow perturbations at sag bottlenecks in nearly-saturated conditions; and c) measures that aim to increase the queue discharge capacity of active sag bottlenecks. An example of a measure from the first category is equipping vehicles with adaptive cruise control systems, which perform the acceleration task more efficiently than human drivers at sags *(15)*. Another example is distributing the traffic flow more evenly across lanes in order to use the bottleneck capacity more efficiently *(3, 16)*. The second category comprises measures such as preventing the formation of long vehicle platoons *(16)* and discouraging drivers from performing lane changes to the busiest lanes *(11, 16)*. The third category comprises measures such as giving information to drivers about the location of the head of the queue, encouraging them to recover speed after leaving congestion *(17, 18)*. Also, control measures belonging to the above-mentioned categories have been proposed for other types of bottlenecks besides sags, such as on-ramp bottlenecks *(19)* and weaving sections *(20)*. The potential effectiveness of most of those measures has been demonstrated by means of empirical data analysis or simulation.

However, there is an additional category of measures that could significantly improve traffic flow efficiency at sags but has received little attention in the recent literature, namely mainstream traffic flow control measures. In mainstream traffic flow control, the inflow to a given bottleneck is regulated by creating a controlled section upstream. The traffic density at the bottleneck area is kept below the critical density. As a result, when demand gets very high, traffic does not break down at the bottleneck and the capacity drop does not occur, hence the outflow from the bottleneck can be higher than its queue discharge capacity. Mainstream traffic flow control is a concept that was first applied in the 1950s and 1960s *(21)*. More recently, it has been presented as an effective measure to mitigate congestion at on-ramp bottlenecks *(5)*. We argue that mainstream traffic flow control can also be used to improve traffic flow efficiency at sags, either by itself or in combination with other types of measures. It is important to note that this control concept can only result in relevant improvements in traffic flow efficiency if the queue discharge capacity of the bottleneck is significantly lower than the queue discharge capacity of the controlled section. This is usually the case with sag bottlenecks, according to the capacity values presented in Section 2.1. In the next section, we present a control strategy based on the concept of mainstream traffic flow control.

3 CONTROL STRATEGY

This section describes the characteristics of our mainstream traffic flow control strategy to mitigate congestion at sags. The control goal is to minimize the total time spent by vehicles in the network over a certain time period. Note that if we assume that the flow entering the network cannot be influenced by any control measure, then minimizing the total time spent over a certain period is equivalent to maximizing the time-weighted sum of exit flows over the same period *(22)*. For the sake of simplicity, we consider a simple network consisting of a freeway stretch with a sag (bottleneck), without any on-ramps or off-ramps. Hence, the network that we aim to control has a single entry point and a single exit point. However, the control strategy described in this section could be generalized to more complex networks, possibly in combination with other control measures.

3.1 Control concept: mainstream traffic flow control

The outflow from a sag bottleneck (q_{b}) is lower or equal to its capacity ($q_{\text{b,max}}$) regardless of the traffic demand. Therefore, if there is no other bottleneck within the network or downstream of it, then the network exit flow (s) is mainly constrained by the capacity of the sag bottleneck.

$$s \simeq q_{\text{b}} \leq q_{\text{b,max}} \tag{1}$$

As mentioned in Section 2, the capacity of a bottleneck depends on the traffic state: the queue discharge capacity of the bottleneck ($q^{\text{c}}_{\text{b,max}}$) is lower than the free flow capacity ($q^{\text{f}}_{\text{b,max}}$).

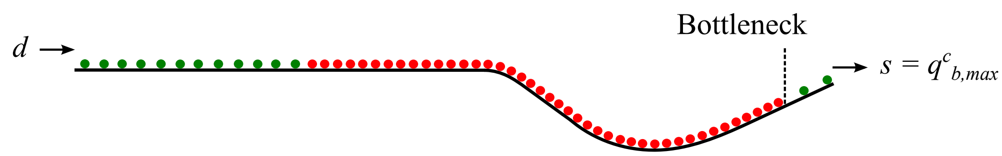
$$q_{\text{b,max}} = \begin{cases} q^{\text{f}}_{\text{b,max}} & \text{in uncongested traffic conditions} \\ q^{\text{c}}_{\text{b,max}} & \text{in congested traffic conditions} \end{cases} \tag{2}$$

where:

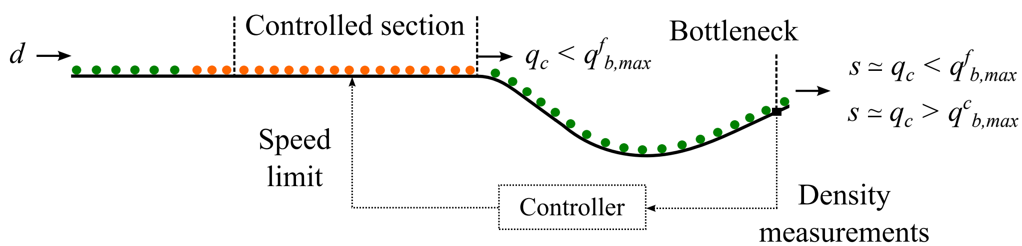
$$q^{\text{c}}_{\text{b,max}} < q^{\text{f}}_{\text{b,max}} \tag{3}$$

Since network exit flows (s) can be higher if traffic flow at the bottleneck is uncongested than if it is congested, a way to maximize the time-weighted sum of exit flows in our network (control goal) is to prevent traffic from breaking down at the sag bottleneck area. To that end, we propose a control strategy based on the concept of *mainstream traffic flow control* (see also Section 2.2).

The control strategy aims to regulate the traffic inflow to the sag bottleneck ($q_{\text{b,in}}$) in order to achieve a desired traffic state at the bottleneck that maximizes outflow. The inflow to the sag bottleneck is regulated by means of a controlled section upstream of the bottleneck (see Figure 1). On that controlled section, the speed limit is variable. Speed limits are set by the controller based on measurements of the traffic conditions (density) at the bottleneck (as explained in Section 3.2). As a result of the fundamental relation between traffic speed and flow, the outflow from the controlled section (q_c) depends on the speed limit (assuming that drivers comply with it). The inflow to the bottleneck is approximately equal to the outflow from the controlled section ($q_{\text{b,in}} \simeq q_c$). By applying an appropriate speed limit on the controlled section, the inflow to the bottleneck can be kept slightly below its free flow capacity ($q_c \simeq q_{\text{b,in}} < q^{\text{f}}_{\text{b,max}}$). Therefore, even in conditions of high demand, the density at the bottleneck does not go above the critical density and traffic does not break down at the bottleneck area (see Figure 1). Note that the formation of congestion is not completely avoided: congestion forms on the controlled section and upstream of it. However, if an appropriate speed limit is applied, the outflow from the controlled section can be higher than the queue discharge capacity of the bottleneck ($q_c > q^{\mathrm{c}}_{\mathrm{b,max}}$). As a result, we can obtain higher exit flows (s) than if traffic flow becomes congested at the bottleneck area (see Figure 1). This should result in a higher time-weighted sum of exit flows, and a lower total time spent by vehicles in the network.



(a) Scenario without controlled section. Traffic flow is congested at the sag bottleneck area, so the exit flow is equal to the queue discarge capacity of the bottleneck.



(b) Scenario with controlled section. Traffic flow is uncongested at the sag bottleneck area, so the exit flow can be higher than the queue discarge capacity of the bottleneck.

FIGURE 1 Flows in the network (d is demand flow, s is exit flow, q_c is outflow from the controlled section, $q^{\mathbf{f}}_{\mathbf{b,max}}$ is free flow capacity, and $q^{\mathbf{c}}_{\mathbf{b,max}}$ is queue discharge capacity). Green circles represent vehicles in uncongested traffic conditions; red and orange circles represent vehicles in severe and less severe congested traffic conditions, respectively.

3.2 Control law: proportional feedback

The controller determines the speed limits to be applied on the controlled section by means of a proportional feedback control law that is similar in nature to the one used by the ramp metering control algorithm ALINEA *(23)*. The control law requires: a) the specification of a target traffic density at the sag bottleneck; and b) the availability of real-time measurements of the traffic density at the bottleneck area. As explained in Section 3.1, the target density should be slightly below the critical density of the bottleneck. The density at the bottleneck is measured in real time by means of loop detectors.

The control law determines the speed limit to be applied on the controlled section ($v^{\mathrm{VSL}}_{\mathrm{lim}}$) based on the difference between the target density ($\widehat{\rho_b}$) and the measured density (ρ_b). The speed limit is re-evaluated each time that the controller receives a new density measurement; hence the control time step period (T_c) is equal to the sampling time period of the detector (T_s). However, there is a delay ($r \cdot T_c$) between the time when the detector time sampling period finishes and the time when the new speed limit is actually applied on the control section.

$$v^{\mathrm{VSL}}_{\mathrm{lim}}(k) = v^{\mathrm{VSL}}_{\mathrm{lim,0}} + K_p \cdot [\widehat{\rho_b} - \rho_b(k-r)] \tag{4}$$

where: k is the control time step index; K_p is the proportional gain; r is the control time step delay; and $v^{\mathrm{VSL}}_{\mathrm{lim,0}}$ is the target speed limit when $\rho_b(k-r) = \widehat{\rho_b}$.

Additionally, we imposed three constraints on the variable speed limits displayed on the message signs in order to make it easier for drivers to comply with them. First, the value of $v^{\mathrm{VSL}}_{\mathrm{lim}}(k)$ is always rounded to a value multiple of 10. Second, the displayed speed limit cannot be lower than a minimum threshold ($v_{\text{lim}}^{\text{VSL}}(k) \geq v_{\text{lim,min}}^{\text{VSL}}$). Third, the speed limit change between two consecutive control steps cannot be higher than a maximum change rate ($|v_{\text{lim}}^{\text{VSL}}(k) - v_{\text{lim}}^{\text{VSL}}(k+1)| \leq \Delta v_{\text{lim}}^{\text{VSL}}$).

By means of the feedback control law described above, the controller should be able to dynamically regulate the speed limit on the controlled section so that the outflow from the bottleneck is maximized. In stationary high demand conditions, the controller maintains the density (ρ_b) near the target value ($\widehat{\rho_b}$) in order to avoid traffic from breaking down at the bottleneck. Furthermore, the controller should be able to react immediately to density deviations. If the measured density is significantly lower than the target density (e.g., because the demand is low), the controller will choose to apply a high speed limit (or even the regular speed limit) in order to maximize the outflow from the bottleneck. If the measured density is higher than the target density (e.g., because traffic has broken down at the bottleneck), the controller will choose to apply a lower speed limit in order to reduce the density at the bottleneck to the target value. The latter is very important, because traffic flow in nearly-saturated conditions can easily destabilize and become congested, and the controller must be able to react to that. Finally, note that the controller reacts to density deviations with a certain delay. This delay is due to the control delay ($r \cdot T_c$), but also to the time needed by drivers to cover the distance between the controlled section and the bottleneck.

4 PERFORMANCE EVALUATION METHOD

A case study was carried out to evaluate the performance of the control strategy presented in Section 3. A microscopic modeling approach specifically developed to model traffic flow at sags *(24)* was used to simulate traffic flow in two scenarios: a) no-control scenario (in which no control measures are implemented); and b) control scenario (in which the proposed control strategy is operative). The performance of the controller was assessed by comparing the total delay experienced by drivers in the two scenarios.

4.1 Traffic flow model

The traffic flow model consists of two sub-models: a) network model; and b) longitudinal driving behavior model. Note that the traffic flow model is face-valid and in line with empirical findings *(24)*, but has not been calibrated yet.

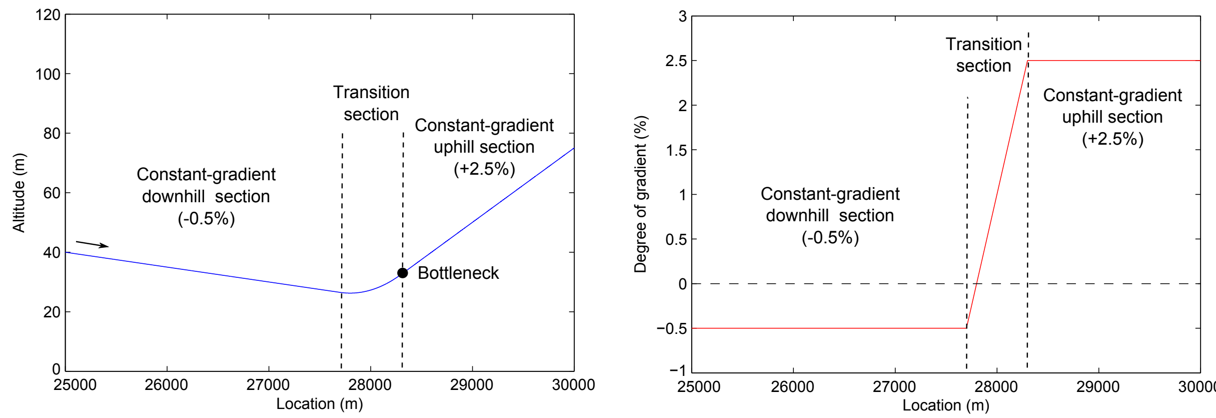
4.1.1 Network model

The network model describes the characteristics of the freeway network. With regard to gradient, a sag is modeled as a combination of three sections (see example in Figure 2): a) constant-gradient downhill section (i.e., section with constant negative slope); b) transition section (i.e., section along which the slope increases linearly from the negative value of the constant-gradient downhill section to the positive value of the constant-gradient uphill section); and c) constant-gradient uphill section (i.e., section with constant positive slope). The network model does not take the freeway horizontal curvature into account.

4.1.2 Longitudinal driving behavior model

The longitudinal driving behavior model describes the acceleration of vehicles ($\dot{v}$) based on a function with two terms. The first term (f_r) describes regular car-following behavior. The second term (f_g) accounts for the influence of freeway gradient. The acceleration of each vehicle is computed each simulation time step and is assumed to stay constant over the period $[t, t + \Delta t]$, where Δt is the simulation step period.

$$\dot{v}(t) = f_r(t) + f_g(t) \tag{5}$$



(a) Altitude vs. distance.

(b) Degree of gradient vs. distance.

FIGURE 2 Vertical alignment profile of the network (from $x = 25.0$ km to $x = 30.0$ km).

The first term (f_r) describes regular car-following behavior, accounting for the influence of vehicle speed (v), relative speed to the leading vehicle (Δv) and distance headway (s) on vehicle acceleration. The formulation of the first term is based on the IDM+ model *(25)*.

$$f_r(t) = a \cdot \min \left[1 - \left(\frac{v(t)}{v_{\text{des}}(t)} \right)^4, \; 1 - \left(\frac{s^*(t)}{s(t) - l} \right)^2 \right] \tag{6}$$

where the dynamic desired net distance headway (s^*) is:

$$s^*(t) = s_0 + v(t) \cdot \tau(t) + \frac{v(t) \cdot \Delta v(t)}{2 \cdot \sqrt{ab}} \tag{7}$$

The parameters in Equations 6 and 7 are: desired speed (v_{des}), vehicle length (l), maximum acceleration (a), desired deceleration (b), net distance headway at standstill (s_0), and safe time headway (τ). In order to model the capacity drop, the value of parameter τ depends on the traffic state. In congested traffic conditions (i.e., below the critical speed v_{crit}), the value of τ is higher than in uncongested conditions.

$$\tau(t) = \begin{cases} \tau_0 & \text{if } v(t) \geq v_{\text{crit}} \\ \gamma \cdot \tau_0 & \text{if } v(t) < v_{\text{crit}} \end{cases} \tag{8}$$

where:

$$\gamma > 1 \tag{9}$$

The second term (f_g) is added in order to account for the influence of changes in gradient on vehicle acceleration. At a given time t, this influence is the gravity acceleration ($g = 9.81$ m/s^2) multiplied by the difference between the gradient at the location where the vehicle is at that time ($G(t)$) and the gradient compensated by the driver until that time ($G_c(t)$).

$$f_g(t) = -g \cdot [G(t) - G_c(t)] \tag{10}$$

The compensated gradient (G_c) is a variable that accounts for the fact that drivers have a limited ability to accelerate on freeway sections where the slope increases (e.g., the transition section of sags). We assume that drivers compensate for positive changes in slope linearly over time (with a maximum gradient compensation rate defined by parameter c). Furthermore, we assume that drivers can fully compensate for negative changes in gradient.

$$G_c(t) = \begin{cases} G(t) & \text{if } G(t) \leq G(t_c) + c \cdot (t - t_c) \\ G(t_c) + c \cdot (t - t_c) & \text{if } G(t) > G(t_c) + c \cdot (t - t_c) \end{cases} \tag{11}$$

where:

$$t_c = \max(t \mid G_c(t) = G(t)) \tag{12}$$

If the rate at which the freeway slope increases over time is lower than the driver's maximum gradient compensation rate (c), then $G_c = G$ for any t. Therefore, $f_g = 0$ for any t, which means that vehicle acceleration is not affected by the increase in gradient. However, if the rate at which the freeway slope increases over time is higher than the driver's maximum gradient compensation rate (c) (as is usually the case on the transition section of sags), then $G_c < G$. As a result, $f_g < 0$, which limits vehicle acceleration ($\dot{v}$). This limitation in vehicle acceleration seems to be the main cause of the local changes in longitudinal driving behavior that reduce the capacity of sags *(14)*. Note that the longitudinal driving model generates the main bottleneck of sags at the end of the transition section (see Figure 2), because the maximum difference between G_c and G occurs at that location. This is in line with empirical observations *(2, 4)*.

4.2 Simulation settings

4.2.1 Network characteristics

The simulated network is a 30 km long freeway stretch that contains a sag. The constant-gradient downhill section goes from $x = 0$ to $x = 27.7$ km; the transition section goes from $x = 27.7$ to $x = 28.3$ km; and the constant-gradient uphill section goes from $x = 28.3$ to $x = 30.0$ km (see Figure 2). The long length of the freeway stretch ensures that traffic flow at the network entry point is not influenced by traffic conditions at the sag bottleneck area. The regular speed limit on the whole network is 120 km/h. The network has only one lane (with no overtaking possibilities). There are no on-ramps or off-ramps. There are three detectors in the network, which are used to monitor traffic conditions at key locations: i) the network entry area ($x = 0.3$ km); ii) the area where the controlled section is located in the control scenario ($x = 27.0$ km); and iii) the network exit area ($x = 29.9$ km).

4.2.2 Longitudinal driving behavior

The parameters of the longitudinal driving behavior model are shown in Table 1. For the sake of simplicity, we do not take into account driver and vehicle heterogeneity.

4.2.3 Traffic demand

The simulation period is 10000 s. At t = 0, there are no vehicles in the network. Network loading starts in the first simulation time step. The demand profile (i.e., flow at $x = 0$ over time) contains three periods that are relevant to test the proposed control strategy. First, from $t = 0$ to $t = 2000$ s, demand increases and goes above the capacity of the sag bottleneck. Second, from $t = 2000$ s to $t = 3000$ s, demand decreases significantly. Third, from $t = 3000$ s to $t = 7000$ s, demand increases again, goes above the capacity of the sag bottleneck and stays at that level. The controller should be able to control traffic adequately in periods of high and low demand, and it should be able to react adequately to demand fluctuations. Note that from $t = 9000$ s to $t = 10000$ s, demand is zero. This end period of zero demand is necessary to ensure that all vehicles are able to exit the network before the end of the simulation period, which allows us to compare network performance in different scenarios. The demand profile can be seen in Figure 3, which shows the flows measured by the detector located at $x = 0.3$ km during the whole simulation period.

TABLE 1 Parameter values.

Long. driv. beh. model		Control strategy	
Parameter	Value	Parameter	Value
v_{des} (km/h)	120	T_s (s)	30
l (m)	4	T_c (s)	30
a (m/s^2)	1.45	$v_{\text{lim},0}^{\text{VSL}}$ (km/h)	60
b (m/s^2)	2.10	K_p (h/veh)	4.8
τ_0 (s)	1.20	$\widehat{\rho}_b$ (veh/km)	18.0
s_0 (m)	3	r (-)	2
v_{crit} (km/h)	65	$v_{\text{lim,min}}^{\text{VSL}}$ (km/h)	20
γ (-)	1.15	$\Delta v_{\text{lim}}^{\text{VSL}}$ (km/h)	20
c (s^{-1})	0.0001	v_{lim} (km/h)	120
Δt (s)	0.5		

4.2.4 Control

Two scenarios were defined: a) no-control scenario; and b) control scenario. In the *no-control scenario*, it is assumed that no control measures are implemented. In the *control scenario*, it is assumed that the control strategy described in Section 3 is in operation; for that reason, the traffic flow model (see Section 4.1) is extended in the following ways:

- A *controlled section* is added to the network. On that section, the speed limit is variable. Speed limits are displayed on message signs. The controlled section is 1.0 km long. That length gives sufficient time to drivers to adapt to the speed limit before leaving the controlled section. The controlled section is located between $x = 26.3$ km and $x = 27.3$ km. The downstream end of the controlled section is 1.0 km upstream of the bottleneck in order to make sure that drivers have sufficient time to accelerate before reaching the bottleneck, so vehicle speeds on the bottleneck are not influenced by the speed limit on the controlled section. There are three message signs located in different points of the controlled section: i) the upstream end ($x = 26.3$ km); ii) the center point ($x = 26.8$ km); and iii) the downstream end ($x = 27.3$ km). Only the first two message signs display the variable speed limits ($v_{\text{lim}}^{\text{VSL}}$). The sign at the downstream end of the controlled section always displays the regular speed limit of the freeway (v_{lim}).

- A *detector* is added to the network. The detector is located at the bottleneck, which is the end of the transition section (see Figure 2). Density measurements from that detector are used as input by the controller.

- The *longitudinal driving behavior model* is extended based on two assumptions. First, we assumed that drivers notice the message signs displaying the variable speed limits when the distance between driver and sign is 300 m or shorter. Second, we assumed that longitudinal driving behavior after noticing a message sign can be adequately reproduced by changing the value of the desired speed parameter (v_{des}) to the displayed speed limit (we assumed that all drivers fully comply with speed limits), keeping the remaining parameter values unchanged. Note that a change in the desired speed parameter does not result in an instantaneous change in the vehicle speed.

- The *controller* that sets the variable speed limits is added to the traffic flow model. The controller uses the proportional feedback control law described in Section 3.2 to select the speed limit to be applied on the controlled section. The values of the control law parameters are shown in Table 1. The parameter values were selected after analyzing the controller performance for different sets of values. No optimization method was used to tune the controller.

4.3 Performance indicator: total delay

The performance of the proposed control strategy is evaluated by comparing the total delay experienced by drivers in the no-control scenario and in the control scenario. The *total delay* (TD) in a given scenario is defined as:

$$\text{TD} = \text{TTS} - \text{TTS}_{\text{ref}} \tag{13}$$

where: TTS is the total time spent by vehicles in the network in that scenario; and TTS_{ref} is the total time spent by vehicles in the network in the reference scenario. Total time spent is calculated based on the demand and exit flows *(22)*. The demand flows are the flows measured by the detector located at $x = 0.3$ km, and the exit flows are the flows measured by the detector located at $x = 29.9$ km.

The *reference scenario* is a hypothetical scenario in which the freeway vertical alignment is assumed to have no influence on the acceleration behavior of drivers ($f_g = 0$). Drivers are able to fully compensate for any increase in slope. This hypothetical behavior is modeled by setting the value of the maximum gradient compensation rate parameter to a very high value: $c = 999$ s^{-1}. As a result, in the reference scenario, the sag is not a bottleneck.

5 RESULTS

5.1 Reference scenario

In the reference scenario, traffic flow remains uncongested everywhere in the network during the whole simulation period. Thus, the exit flow profile over time is similar to the demand flow profile, with an offset of around 900 s (see Figure 3a). The total time spent by vehicles in the network is 1035 veh·h.

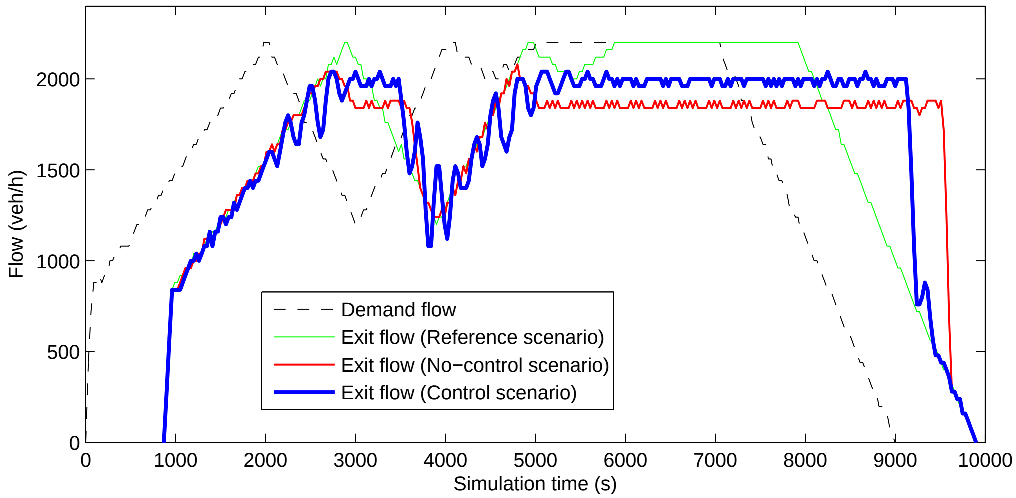
5.2 No-control scenario

In the no-control scenario, traffic breaks down at the sag bottleneck when the inflow goes above 2050 veh/h (which can be considered as the free flow capacity of the bottleneck). When traffic breaks down, the outflow from the bottleneck decreases to around 1855 veh/h (which can be considered as the queue discharge capacity), reducing the network exit flow to 1855 veh/h as well (see Figure 3a). During the simulation period, traffic breaks down two times. After the first breakdown, the demand flow decreases considerably, allowing the first queue to dissolve. Afterwards, the demand flow increases again above the free flow capacity of the bottleneck, causing a second breakdown (Figure 3a). In both cases, since the demand flow is higher than the exit flow, the number of vehicles within the network increases. This accumulation of vehicles results in a higher total time spent than in the reference scenario. The total time spent by vehicles in the network in the no-control scenario is 1237 veh·h, so the total delay is 202 veh·h.

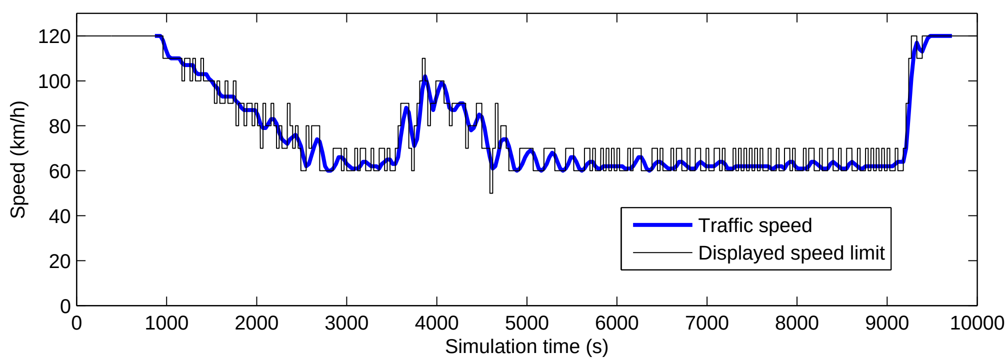
5.3 Control scenario

In the control scenario, the outflow from the controlled section is regulated so that it does not go above the free flow capacity of the bottleneck. Because of that, traffic does not break down at the bottleneck during the whole simulation period. In conditions of high demand, congestion forms on the controlled section. However, the outflow from the controlled section is higher (around 1985 veh/h) than the queue discharge capacity of the bottleneck (which is around 1855 veh/h) (see Figure 3a). As a result, in the periods of high demand, network exit flows are around 1985 veh/h, i.e., 7% higher than in the no-control scenario (see

Figure 3a). Therefore, less vehicles accumulate in the network, which results in a considerably lower total delay. In the control scenario, the total time spent by vehicles in the network is 1177 veh·h (5% lower than in the no-control scenario), so the total delay is 142 veh·h (30% lower than in the no-control scenario).



(a) Demand and exit flows over time in all scenarios. The demand flows are the flows measured by the detector located at $x = 0.3$ km. The exit flows are the flows measured by the detector located at $x = 29.9$ km. Flows are smoothed by using a simple moving average method: the flow for a given sampling period is the unweighted mean of the measured flow on that sampling period and the measured flows on the previous and next sampling period.



(b) Speed limit and average traffic speed over time at location $x = 27.0$ km (i.e., 300 m before the end of the controlled section) in the control scenario. Traffic speeds are smoothed by using a simple moving average method: the speed for a given sampling period is the unweighted mean of the measured average speed on that sampling period and the measured average speeds on the previous and next sampling period.

FIGURE 3 Simulation results: demand and exit flows over time in all scenarios; speed limit and traffic speed on the controlled section over time in the control scenario.

The controller is able to react adequately to fluctuations in demand (see Figure 3). Demand flows reach high levels before $t = 2000$ s (see Figure 3a). When density at the bottleneck gets close to the target density, the controller sets a speed limit of 60-70 km/h on the controlled section (around $t = 2700$ s in Figure 3b). Between $t = 2000$ s and $t = 3000$ s, demand significantly decreases (see Figure 3a), which results in low densities at the bottleneck. When such low densities are measured, the controller increases the speed limit on the controlled section (see Figure 3b). The reason is that the demand is too low to cause traffic to break down at the bottleneck, so there is less need to restrict the inflow. Afterwards, between $t = 3000$ s and

$t = 4000$ s, the demand increases again (Figure 3a). The controller responds by decreasing the speed limit on the controlled section to 60-70 km/h again (see Figure 3b), in order to prevent traffic from breaking down at the bottleneck. Note that due to the proportional structure of the controller, demand fluctuations result in variable speed limit oscillations (see Figure 3b). However, in our case study, oscillations seem to dampen out with time, so the system does not become unstable.

6 SENSITIVITY ANALYSIS

We selected the values of the controller parameters (see Table 1) to ensure high controller performance under the assumption that traffic behaves according to the traffic flow model described in Section 4. However, we also analyzed the performance of the controller assuming that traffic does not behave exactly as described by our traffic flow model. More specifically, we investigated the sensitivity of the controller performance to two key parameters of the longitudinal driving behavior model that have a significant influence on the capacity of the sag bottleneck. Those parameters are the maximum gradient compensation rate (c) and the congestion factor on safe time headway (γ). First, we evaluated the performance of the controller assuming a lower and a higher value for parameter c (i.e., 0.00005 s^{-1} and 0.00015 s^{-1}, respectively), whereas the other parameters remained unchanged. Second, we evaluated the performance of the controller assuming a lower and a higher value for parameter γ (i.e., 1.12 and 1.18, respectively), whereas the other parameters remained unchanged.

The results indicate that the reduction in total delay resulting from the implementation of the proposed control strategy significantly depends on the value of parameter c. If $c = 0.00010$ s^{-1} (default value), the total delay in the control scenario is 30% lower than in the no-control scenario. If $c = 0.00005$ s^{-1}, that percentage is 36%, whereas if $c = 0.00015$ s^{-1}, that percentage is 23% (see Table 2). The main reason for those differences is that a higher (lower) value of c results in a higher (lower) queue discharge capacity of the sag bottleneck, hence it also results in higher (lower) exit flows in the no-control scenario. Instead, in the control scenario, exit flows are almost the same regardless of the value of c. Therefore, the controller reduces total delay to a larger extent if the value of c is lower. The reduction in total delay resulting from the implementation of the controller does not significantly depend on the value of parameter γ. If $\gamma = 1.15$ (default value), the total delay in the control scenario is 30% lower than in the no-control scenario. If $\gamma = 1.12$, that percentage is 31%, whereas if $\gamma = 1.18$, that percentage is 29% (see Table 2). The main reason why the percentages are similar is that a higher (lower) value of γ results in a higher (lower) queue discharge capacity of both the sag bottleneck and the controlled section. Therefore, a higher (lower) value of γ results in higher (lower) exit flows in both the no-control scenario and the control scenario.

To conclude, the sensitivity analysis shows that the results of the evaluation of the controller performance depend on the specification of the traffic flow model. However, the sensitivity analysis also shows that the controller is able to significantly reduce total delay even after changing the values of key model parameters.

TABLE 2 Controller performance (sensitivity analysis).

	Model parameter values				
Parameter c (s^{-1})	0.00010	0.00005	0.00015	0.00010	0.00010
Parameter γ (-)	1.15	1.15	1.15	1.12	1.18
Total delay in the no-control scenario (veh·h)	202	227	177	157	244
Total delay in the control scenario (veh·h)	142	145	137	108	173
Absolute difference (veh·h)	−60	−82	−40	−49	−71
Relative difference (%)	−29.7	−36.1	−22.6	−31.2	−29.0

7 CONCLUSIONS AND OUTLOOK

The capacity of sags is considerably lower than the capacity of normal freeway sections. As a result, sags often become bottlenecks in freeway networks, causing the formation of congestion in high traffic demand conditions. Congestion results in a further decrease in bottleneck capacity. This paper presented a new control strategy to mitigate congestion at sags, based on the concept of mainstream traffic flow control. By limiting the traffic speed (and hence the flow) on a controlled section upstream of the bottleneck, the proposed strategy regulates the density at the bottleneck area in order to keep it slightly below the critical density, hence preventing traffic from breaking down. The capacity drop due to congestion does not occur, so the outflow from the bottleneck can be higher. The speed limit on the controlled section is set using a proportional feedback control law. The performance of the proposed control strategy was evaluated by means of a simple case study using microscopic traffic simulation. The results show a considerable improvement in traffic flow efficiency. In periods of high demand, the flow exiting the network is around 7% higher in the control scenario than in the no-control scenario, which reduces the total delay by around 30%. A sensitivity analysis shows that the controller is able to considerably reduce total delay even if we assume different values for key parameters of the traffic flow model. In spite of the simplicity of the case study, our findings show for the first time that mainstream traffic flow control strategies using variable speed limits have the potential to considerably improve traffic flow efficiency in freeway networks containing sags.

Further research is necessary to make a more thorough evaluation of the performance of the proposed control strategy. Such evaluation requires extending the case study to include a multi-lane network and heterogeneous traffic. In addition, the driving behavior model used for traffic simulation should take into account the level of compliance of drivers to variable speed limits, which may have a strong influence on the performance of the control strategy. Also, the model should be calibrated and validated. Moreover, further research should be carried out to refine the controller design and improve its performance. For example, the oscillatory behavior of the controller could be mitigated by using an alternative type of control law (e.g., proportional-integral feedback). Also, other means to regulate the speed on the controlled section could be tested. An alternative to displaying variable speed limits on message signs could be to regulate the speed of vehicles equipped with cooperative adaptive cruise control systems (via infrastructure-to-vehicle communication). Finally, the controller design could be extended in order to make it operational in more complex networks (e.g., networks with ramps and/or other types of bottlenecks). This may require combining the control strategy presented in this paper with other types of control measures.

ACKNOWLEDGMENT

This research was sponsored by Toyota Motor Europe.

REFERENCES

[1] Furuichi, T., S. Yamamoto, M. Kotani, and M. Iwasaki. Characteristics of spatial speed change at motorway sag sections and capacity bottlenecks. Presented at 82nd Annual Meeting of the Transportation Research Board, Washington, D.C., 2003.

[2] Koshi, M., M. Kuwahara, and H. Akahane. Capacity of sags and tunnels on Japanese motorways. *ITE Journal*, Vol. 62, No. 5, 1992, pp. 17–22.

[3] Xing, J., K. Sagae, and E. Muramatsu. Balance lane use of traffic to mitigate motorway traffic congestion with roadside variable message signs. Presented at ITS World Congress, Busan, South Korea, 2010.

[4] Brilon, W., and A. Bressler. Traffic Flow on Freeway Upgrades. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1883, Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 112–121.

[5] Carlson, R., I. Papamichail, and M. Papageorgiou. Local Feedback-Based Mainstream Traffic Flow Control on Motorways Using Variable Speed Limits. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, No. 4, 2011, pp. 1261–1276.

[6] Helbing, D. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, Vol. 73, 2001, pp. 1067–1141.

[7] Schönhof, M., and D. Helbing. Empirical Features of Congested Traffic States and Their Implications for Traffic Modeling. *Transportation Science*, Vol. 41, No. 2, 2007, pp. 135–166.

[8] Hall, F. L., and K. Agyemang-Duah. Freeway Capacity Drop and the Definition of Capacity. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1320, Transportation Research Board of the National Academies, Washington, D.C., 1991, pp. 91–98.

[9] Cassidy, M., and R. Bertini. Some traffic features at freeway bottlenecks. *Transportation Research Part B: Methodological*, Vol. 33, No. 1, 1999, pp. 25–42.

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

[11] Patire, A. D., and M. J. Cassidy. Lane changing patterns of bane and benefit: Observations of an uphill expressway. *Transportation Research Part B: Methodological*, Vol. 45, No. 4, 2011, pp. 656–666.

[12] Koshi, M. An Interpretation of a Traffic Engineer on Vehicular Traffic Flow. In: M. Fukui, Y. Sugiyama, M. Schreckenberg, and D. E. Wolf (Eds.). *Traffic and Granular Flow'01*. Springer, Berlin, 2001, pp. 199–210.

[13] Goñi Ros, B., V. L. Knoop, B. van Arem, and S. P. Hoogendoorn. 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.

[14] Yoshizawa, R., Y. Shiomi, N. Uno, K. Iida, and M. Yamaguchi. Analysis of Car-following Behavior on Sag and Curve Sections at Intercity Expressways with Driving Simulator. *International Journal of Intelligent Transportation Systems Research*, Vol. 10, No. 2, 2011, pp. 56–65.

[15] Ozaki, H. Modeling of Vehicular Behavior from Road Traffic Engineering Perspectives. In: M. Fukui, Y. Sugiyama, M. Schreckenberg, and D. E. Wolf (Eds.). *Traffic and Granular Flow'01*. Springer, Berlin, 2001, pp. 281–292.

[16] Hatakenaka, H., T. Hirasawa, K. Yamada, H. Yamada, Y. Katayama, and M. Maeda. Development of AHS for traffic congestion in sag sections. Presented at ITS World Congress, London, U.K., 2006.

[17] Xing, J., H. Takahashi, and T. Takeuchi. Increasing Bottleneck Capacity through Provision of Bottleneck Location Information. Presented at 11th World Conference on Transport Research, Berkeley, California, 2007.

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

[19] Laval, J. A., M. J. Cassidy, and C. F. Daganzo. Impacts of Lane Changes at Merge Bottlenecks: A Theory and Strategies to Maximize Capacity. In: T. Schadschneider, A. Pöschel, R. Kühne, M. Schreckenberg, and D. E. Wolf (Eds.), *Traffic and Granular Flow'05*. Springer, Berlin, 2005, pp. 577–586.

[20] Daganzo, C. F., J. A. Laval, and J. C. Muñoz. Some ideas for freeway congestion mitigation with advanced technologies. *Traffic Engineering and Control*, Vol. 10, No. 43, 2002, pp. 397–403.

[21] Greensberg, H., and A. Daou. The control of traffic flow to increase flow. *Operations Research*, Vol. 8, No. 4, 1960, pp. 524–532.

[22] Papageorgiou, M., C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang. Review of road traffic control strategies. *Proceedings of the IEEE*, Vol. 91, 2003, pp. 2043–2067.

[23] Papageorgiou, M., H. Hadj-Salem, and J. M. Blosseville. ALINEA: A local feedback control law for on-ramp metering. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1320, Transportation Research Board of the National Academies, Washington, D.C., 1991, pp. 58–64.

[24] Goñi Ros, B., V. L. Knoop, B. van Arem, and S. P. Hoogendoorn. Modeling driving behavior and traffic flow at sags. Presented at ITS World Congress, Tokyo, Japan, 2013.

[25] Schakel, W., V. L. Knoop, and B. van Arem. Integrated Lane Change Model with Relaxation and Synchronization. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2316, Transportation Research Board of the National Academies, Washington, D.C., 2012, pp. 47–57.