Bottleneck mitigation through a variable speed limit system using connected vehicles

Ellen F. Grumert & Andreas Tapani

To cite this article: Ellen F. Grumert & Andreas Tapani (2018): Bottleneck mitigation through a variable speed limit system using connected vehicles, Transportmetrica A: Transport Science, DOI: 10.1080/23249935.2018.1547332

To link to this article: https://doi.org/10.1080/23249935.2018.1547332

© 2018 Hong Kong Society for Transportation Studies Limited. Published by Informa UK Limited, trading as Taylor & Francis Group

Accepted author version posted online: 09 Nov 2018.
Published online: 24 Nov 2018.

Submit your article to this journal 

Article views: 61

View Crossmark data
Bottleneck mitigation through a variable speed limit system using connected vehicles

Ellen F. Grumert a,b and Andreas Tapani a

aSwedish National Road and Transport Research Institute (VTI), Linköping, Sweden; bDepartment of Science and Technology, Linköping University, Norrköping, Sweden

ABSTRACT
Variable speed limit (VSL) systems are used to improve the traffic conditions by adjusting the speed limits based on the current traffic situation. Advances in vehicle technology have made it possible to use connected vehicles in VSL systems. Connected vehicles can continuously transmit information about their speed and location, which can be used to estimate the current traffic conditions at arbitrary locations. In this study, we propose a VSL system based on connected vehicles. The aim is to also allow application of VSLs for non-recurrent bottleneck mitigation at arbitrary locations, unlike today’s systems which require densely placed detectors or are limited to beforehand known bottleneck locations. The proposed system is evaluated by microscopic traffic simulation. The results indicate that the VSL system manage to improve traffic efficiency in a simulated incident scenario.

ARTICLE HISTORY
Received 23 January 2018
Accepted 8 November 2018

KEYWORDS
Connected vehicles; variable speed limit; traffic management; microscopic traffic simulation; traffic efficiency

1. Introduction

Variable speed limit (VSL) systems make use of an estimate of the current traffic state, described by traffic density, speed and flow, in order to adjust the speed limit. The estimate of the traffic state, used as input to the VSL system, is usually based on data from stationary detectors and the VSL is displayed on gantries over the road.

Most of the VSL systems in practical use are the so-called incident detection systems, implemented with the goal to improve traffic safety (van den Smulders and Smulders 1994; Lee, Hellinga, and Saccomanno. 2006; Allaby, Hellinga, and Bullock. 2007; Soriguera, Torné, and Rosas. 2013; Li, Ranjitkar, and Ceder 2014). In these systems, an incident is detected as a substantial and abrupt change in speed and flow. A large reduction in the speed limit is commonly applied by the system to limit the effects of the incident and thereby increase safety. However, this procedure is often leading to reduced traffic efficiency. Therefore, homogenization systems with the goal to improve traffic efficiency have been proposed in the literature, see e.g. Carlson, Papamichail, and Papageorgiou (2011) and Jin and Jin (2015). In homogenization systems, the speed limit is adjusted to reduce deviations in individual
vehicle speeds in order to maximize the throughput of vehicles at a bottleneck location. One common approach is to make use of control theory to decide on a suitable speed limit. The speed limit is adjusted to reflect a predefined target value. The target value can, for example, be the critical density, defined as the density associated with the maximum capacity. Since most of the homogenization systems are not implemented in real traffic, they are evaluated through simulation-based studies (Carlson, Papamichail, and Papageorgiou 2011; Jin and Jin 2015; Müller et al. 2015). An exception is the field trial presented by Hegyi and Hoogendoorn (2010). Both the field trial and the simulation-based studies show desirable effects on efficiency, but also positive effects on safety have been observed in the simulation-based study by Li et al. (2014). However, the systems commonly require predefined estimates of the specific traffic characteristics of the bottleneck, such as capacity, jam density, free flow speed, etc. These systems are therefore limited to beforehand known bottlenecks. Hence, with homogenization systems, it is often not possible to mitigate non-recurrent bottlenecks over a large area. Incident detection systems can cover a large area, but are limited in improving traffic efficiency due to the often large reduction in speed limit. Further, to be able to get information about the traffic state and the traffic characteristics at arbitrary locations, densely placed detectors are required.

The introduction of connected vehicles has enabled estimation of the traffic state and the traffic characteristics at arbitrary locations. The connected vehicles are able to continuously transmit and receive information about their current speed, position and speed limit. This allow identification of both recurrent and non-recurrent bottlenecks, as well as control of the connected vehicles towards a VSL, with limited use of expensive infrastructure equipment such as densely placed detectors and variable message signs installed on gantries. Further, as concluded by Wang et al. (2009), the traffic characteristics may change over time due to incidents and adverse weather conditions. Wang and Papageorgiou (2005), Wang et al. (2009) and Tampère and Immers (2007) propose filtering methods and real-time calibration of the fundamental diagrams to find time-dependent estimates of the traffic characteristics. Such time-dependent estimates based on connected vehicles can be used as input to a VSL algorithm to allow application of a suitable speed limit at arbitrary locations both for recurrent and non-recurrent bottlenecks and during normal conditions.

In this paper, we propose a VSL system that makes use of connected vehicles to allow application of VSLs for bottleneck mitigation at arbitrary locations. In the proposed VSL system, connected vehicles are used as sensors and the current speed limit is communicated directly to the connected vehicles instead of using variable message signs. The speed of the connected vehicles can thereby be adjusted towards the variable speed limit, similar to an adaptive cruise control. The proposed VSL system is an improvement to existing VSL systems by (1) allowing application of a VSL control strategy for non-recurrent, as well as recurrent, bottlenecks at arbitrary locations, (2) the application of a VSL control strategy that is based on a time- and space-dependent estimate of the critical density to account for changes in the traffic characteristics as a result of adverse weather, incidents, vehicle composition, etc., (3) adjustment of the speed of the connected vehicles through communication of the current speed limit and (4) limited use of stationary detectors.

We use microscopic traffic simulation for evaluation of the effects on traffic efficiency of the proposed VSL system. An incident scenario is simulated as an example of a change in the traffic conditions that will activate the VSL system. The traffic performance with the proposed VSL system is compared to the performance without the system. The results show
that the proposed VSL system manages to improve the individual travel time of the vehicles in all simulation runs. On average the individual travel time of the vehicles is reduced by 5% compared to the base case without the VSL system.

The reminder of the paper is organized as follows. In Section 2, a background to the methods used in the different parts of the VSL system is given. The proposed VSL system is presented in Section 3. In Section 4, the evaluation method and the simulated incident scenario are described. Simulation results are presented in Section 5. Finally, conclusions are given in Section 6.

2. Background

VSL systems can be categorized into two main types, incident detection systems and homogenization systems. The main difference between the systems is that homogenization systems are often applied before congested traffic states are reached, while incident detection systems are triggered when a breakdown, i.e. situations with very low speeds, is detected. For this reason, the speed limit is often reduced gradually in the homogenization systems, while more abrupt speed limit changes often are applied in incident detection systems.

Up until now, most implementations of VSL systems in practice are of incident detection systems, although combinations of incident detection and homogenization systems have also been implemented (Maunsell and Parkman 2007). Empirical studies of the effects of incident detection systems (van den Smulders and Smulders 1994; Smulders and Hellemann 1998; Maunsell and Parkman 2007; Nissan and Koutsopoulos 2011) showed a reduction in the number of incidents and a decreased variance in mean speed between lanes. No increase, or even a decrease, was seen in the throughput, i.e. traffic efficiency was not increased by the systems. Homogenization algorithms, such as the ones proposed by Hegyi, De Schutter and Hellendoorn (2005), Müller et al. (2015) and Jin and Jin (2015), have in simulation-based studies been shown to improve traffic efficiency. Moreover, a field trial by Hegyi and Hoogendoorn (2010) showed improvements in throughput as a result of the decreased time needed to resolve moving shockwaves. It should be noted that the base case without the homogenization algorithm in the field trial was estimated based on prior knowledge about the typical behaviour of the propagation of shockwaves, which resulted in estimated improvements rather than measurable improvements. Additionally to increased efficiency, Li et al. (2014) concluded that for the investigated homogenization system also safety was improved.

The growing amount of technology related to connected vehicles have resulted in recent studies where connected vehicles has been used as a part of VSL systems. Kattan et al. (2015) extended a VSL algorithm by Hegyi, De Schutter and Hellendoorn (2005) which is based on model predictive control. The aim was to minimize travel time by including measurements of speed from connected vehicles. Khondaker and Kattan (2015) considered estimates of each connected vehicle’s total travel time, time to collision and emission levels and optimized the corresponding aggregated values based on a model predictive control strategy. The goal of the optimization was to find the VSL to be displayed on variable message signs. Hegyi et al. (2013) used connected vehicle data, given as GPS position and in-vehicle speed, together with video-based monitoring, to reduce the time needed to detect shockwaves and to increase the precision of the application area of lower speed limits using the VSL
algorithm SPECIALIST proposed by Hegyi et al. (2008). Wang et al. (2016) introduced a car-following control algorithm taking into account the surrounding environment of the connected vehicles. Here, the desired speed of the individual vehicle control was based on the VSL algorithm SPECIALIST by Hegyi et al. (2008). Müller, Carlson, and Kraus (2016) used connected vehicles to apply speed limits, and thereby increase the compliance level, based on the control algorithm proposed by Müller et al. (2015) and Carlson, Papamichail, and Papageorgiou (2011). Han, Chen, and Ahn (2017) used connected vehicles to control the inflow at a recurrent one-lane bottleneck in such a way that the maximum throughput was guaranteed and the capacity drop was avoided. Han and Ahn (2018) proposed a method for calculating the probability of a breakdown in real time at a merge bottleneck based on the observed headways. Lower speed limits were applied to connected vehicles and on variable message signs to reduce the probability of a breakdown. Hence, connected vehicles have been shown to be useful for data collection and to estimate the traffic state, which is used as input to the VSL system (Hegyi et al. 2013; Kattan et al. 2015; Khondaker and Kattan 2015; Han and Ahn 2018), and for control of the speed of individual vehicles as part of the control strategy of the VSL system (Müller, Carlson, and Kraus 2016; Wang et al. 2016; Han, Chen, and Ahn 2017; Han and Ahn 2018).

The choice of a suitable speed limit is decided by the VSL algorithm. As input to the VSL algorithm, an estimate of the current traffic state and possibly also the traffic characteristics, such as critical density, free flow speed, jam density, wave propagation speed, etc. is required. The most common way to estimate the traffic state is by the use of data from stationary detectors, such as for example loop and radar detectors (Kurkjian et al. 1980; Coifman 2003; Singh and Li 2012). This is limiting the traffic state estimation to specific points in space, and the conditions in between detectors remains unknown. Data assimilation and fusion techniques based on filters and traffic modelling are common methods to get a more complete picture of the traffic state in between the detectors, see for example Muñoz et al. (2003), Mihaylova, Boel, and Hegyi. (2007), Wang and Papageorgiou (2005) and Duret, Leclercq, and El Faouzi (2016). As more sources of traffic data are becoming available, e.g. data from connected vehicles and mobile phones, these have been used as input to the filters to update the modelled traffic state. Examples are presented by Herrera and Bayen (2007), Work et al. (2010) and Seo, Kusakabe, and Asakura (2015b). Other methods have been making use of data from connected vehicles for traffic state estimation without the use of an underlying traffic model. Examples are presented by Herrera et al. (2010), van Lint and Hoogendoorn (2010), Ma et al. (2011), Seo, Kusakabe, and Asakura. (2015a), Montero et al. (2016) and Grumert and Tapani (2018).

Control algorithms with the goal to increase traffic efficiency, i.e. homogenization algorithms like the algorithm suggested by Carlson, Papamichail, and Papageorgiou (2011) and Müller et al. (2015), do often require an estimate of the traffic characteristics, such as the critical density, to be used as a target value. The critical density is defined as the density which correspond to the highest throughput of vehicles at the bottleneck, i.e. the bottleneck capacity. To find the critical density, the relation between the variables in the traffic state can be used. Hence, by the use of pairwise measurements of density and speed, or density and flow, the critical density can be identified. The relation between speed, density and flow can mathematically be expressed as something referred to as the fundamental diagram (Zhang, Kuhne, and Michalopoulos 2005; TRB 2011). The mathematical form of the fundamental diagram often consists of parameters describing the traffic characteristics.
of the traffic conditions, such as free flow speed, critical density and jam density. Hence, the critical density can be estimated through calibration of the fundamental diagram. A number of studies exist in the literature where various methods for calibration of the fundamental diagram were proposed, see e.g. Van Aerde and Rakha (1995), Dervisoglu et al. (2010), Qu, Wang, and Zhang (2015), Zhong et al. (2016), Qu, Zhang, and Wang (2017) and Knoop and Daamen (2017). In the proposed methods, optimization and regression techniques were commonly applied based on empirical observations from stationary detectors and a predefined functional form of the fundamental diagram. With the introduction of connected vehicles, methods applying calibration of the fundamental diagram based on connected vehicles were proposed (Seo, Kusakabe, and Asakura 2017; Clairais, Duret, and El Faouzi 2016). All of the previously mentioned methods utilized off-line calibration, meaning that the measurements are collected from the system in real time and the calibration process is done afterwards. However, the traffic characteristics are known to vary due to, for example, adverse weather conditions, incidents, etc. (Tampère and Immers 2007; Wang et al. 2009). A change in the traffic characteristics might imply a change in the critical density and thereby a change in the parameters related to the fundamental diagram. Further, the fundamental diagram might differ depending on the surrounding traffic environment, which is resulting in different parameters of the fundamental diagram depending on the location on the road. Hence, it is desirable to have parameter estimates of the fundamental diagram that are time and space dependent. This can be achieved by performing online, or real time, calibration of the parameters in the fundamental diagram for individual road segments. A few studies have suggested methods for performing real-time calibration of the parameters in the fundamental diagram. Examples are Wang and Papageorgiou (2005), Wang et al. (2009) and Tampère and Immers (2007).

As a conclusion, if the goal is to improve traffic efficiency a homogenization algorithm is suitable. Further, connected vehicles have been used in VSL systems for estimation of the traffic state or to control the speed of the connected vehicles towards the current VSL. However, to use connected vehicles for application of VSLs for mitigation of non-recurrent bottlenecks at arbitrary locations, in addition to recurrent bottlenecks, and to find a suitable speed limit based on time- and space-dependent estimates of the critical density has, to our knowledge, not been done before.

3. A VSL system using connected vehicles

The proposed VSL system makes use of connected vehicles to estimate the traffic state and the critical density and to allow application of variable speed limits for bottleneck mitigation at arbitrary locations. The traffic state is estimated by the use of connected vehicles and sparsely placed detectors following the approach developed in earlier research (Grumert and Tapani 2018). The identification of a critical capacity at arbitrary locations is based on a filtering approach and real-time estimation of the parameters in a triangular fundamental diagram, similar to what has been proposed by Wang and Papageorgiou (2005), Tampère and Immers (2007) and Wang et al. (2009). Finally, the VSLs are calculated according to the method by Müller et al. (2015). By using connected vehicles, it is possible to utilize a time- and space-dependent estimate of the critical density as a target value in the VSL control strategy. Hence, the VSL system is not limited to beforehand known bottlenecks but can also be applied for non-recurrent bottlenecks at arbitrary locations. This section includes
the methodology used for the different parts included in the proposed VSL system. The parts are: (1) traffic state estimation based on connected vehicles, (2) identification of a time and space dependent critical capacity, and (3) calculation of a suitable control strategy used for adjustment of the speed of the connected vehicles.

3.1. Traffic state estimation based on connected vehicles

The connected vehicles are used together with sparsely placed detectors, as suggested in earlier research (Grumert and Tapani 2018), to estimate traffic density and speed.

First, consider a road stretch divided into $K$ segments, with an arbitrary length $X_k$, where $k = 1, \ldots, K$. These segments are further divided into equally long sub-segments $i = 1, 2, \ldots, l_k$, where $l_k$ is the number of sub-segments in segment $k$. The length of sub-segment $i$ is calculated as $x_{k,i} = X_k/l_k$. See Figure 1 for an illustration of the division of a segment into sub-segments. Since connected vehicles are able to receive and transmit information continuously they can be used to estimate the traffic state for each sub-segment.

The detector data and the connected vehicle data are collected and processed over an aggregation time period $n$, resulting in estimates at discrete time-steps $nT$, where $T$ is the duration of the aggregation time period. The data is collected at time steps of $\Delta t_{CCV}$, resulting in $m = 1, \ldots, M$ collection instants within the aggregation time period. By making use of information on each connected vehicle’s location, the current segment and sub-segment for each connected vehicle can be identified and the current number of connected vehicles located on each sub-segment is given. Moreover, the speed of each connected vehicle is associated with its current segment and sub-segment. Hence, during the aggregation time period $[nT, (n + 1)T - \Delta t_{CCV}]$, the total number of connected vehicles, $C_k(t)$, and the speed of each connected vehicle $c$, $v_{k,i}(t)$, for sub-segment $i$ on segment $k$ is collected at $m = 1, \ldots, M$ collection instants and processed in the end of the aggregation time period. Further, the number of connected vehicles, $CV_k(n)$ and the total number of vehicles, $V_k(n)$ passing the detector in the beginning of road segment $k$ between the time $nT$ and the time $(n + 1)T - \Delta t_{CCV}$ is gathered.

In the end of the aggregation time period $(n + 1)T - \Delta t_{CCV}$, the data from the detectors and the connected vehicles is used to estimate the density and the speed for each sub-segment $i$ on segment $k$. The count connected vehicle (CCV) method presented by Grumert and Tapani estimates the traffic density based on the connected vehicle penetration rate for segment $k$, together with the average number of connected vehicles located within each sub-segment $i$ on segment $k$. Thereby, the density on each sub-segment can be estimated

![Figure 1](image). Illustration of one segment $k$ and how it is further divided into $l_k$ sub-segments indexed by $i$. A detector, shown as a thicker vertical line in the figure, is located in the beginning of each segment.
even though the penetration rate is only based on the detector data in the beginning of the segment. The penetration rate is calculated as \(CV_k(n)/V_k(n)\) and the average number of connected vehicles for sub-segment \(i\) on segment \(k\), including the \(M\) collection instants, becomes

\[
\bar{p}_{k,j}(n) = \frac{1}{M} \sum_{m=1}^{M} C_{k,j}(nT + m\Delta t_{CCV}).
\]  

(1)

Thereafter, the average number of connected vehicles is used together with the penetration rate to estimate the density for sub-segment \(i\) on segment \(k\) as

\[
\hat{\rho}_{k,j}(n) = \frac{\bar{p}_{k,j}(n)}{x_{k,j}} \frac{1}{CV_k(n)/V_k(n)}.
\]  

(2)

By assuming that the distribution of speed is the same for connected and non-connected vehicles, the speed estimate, \(\hat{v}_{k,j}(n)\), for a sub-segment is given as a time average, considering the \(M\) collection instants within an aggregation time period, over the speed of all connected vehicles located on the sub-segment,

\[
\hat{v}_{k,j}(n) = \frac{1}{M} \sum_{t=1}^{M} \frac{1}{c_{k,j}(nT + m\Delta t_{CCV})} \sum_{c=1}^{C_{k,j}(nT + m\Delta t_{CCV})} v_{k,j}^c(nT + m\Delta t_{CCV}).
\]  

(3)

Observe that the connected vehicles, \(CV_k(n)\), and all vehicles, \(V_k(n)\), used for estimating the penetration rate in Equation (2), use data from the detector in the beginning of segment \(k\) for the time period \(n\). Hence, the penetration rate is the same for all sub-segments \(i\). Whereas, the connected vehicles, \(C_{k,j}(t)\), is the connected vehicles, specific for sub-segment \(i\) at time \(t\), that contribute with a speed measurement to the average speed calculated in Equation (3).

### 3.2. Estimation of a time- and space-dependent critical density

In the proposed method, the density and speed estimates from the CCV method are used as input to an ensemble Kalman filter (Evensen 2003; van Lint and Hoogendoorn 2010). In an ensemble Kalman filter, an ensemble, represented by a set of realizations of the traffic state, is propagated in time by the use of a traffic model. The distribution of the ensemble, represented as a mean and a covariance matrix, is calculated in each time step. The mean is used to estimate the current traffic state, while the complete distribution is used for propagation of the ensemble to a future traffic state. In our case, the traffic evolution is described by the first-order CTM (Daganzo 1994; Lebacque 1996) and a triangular fundamental diagram, which is deemed to be sufficient to identify a time and space dependent critical capacity. This is supported by Tampère and Immers (2007), who conclude that a first-order model is enough to capture the capacity levels by real-time estimation of the model parameters. Hence, in the filtering approach, the traffic model is corrected by adjustments of the parameters of the triangular fundamental diagram and model outputs based on measurements corresponding to the estimates from the CCV. The model outputs are speed, \(\tilde{v}_{k,j}\),
and density, $\rho_{k,i}$. The triangular fundamental diagram is given as

$$Q_e(\rho_{k,i}) = \begin{cases} \tilde{V}_{f,k,i}, & \text{if } \rho_{k,i} \leq \rho_{k,i}^{\text{cap}} = \frac{1}{V_{f,k,i} \tau_{k,i} + 1 / \rho_{k,i}^{\text{max}}} \\ \frac{1}{\tau_{k,i}} \left(1 - \frac{\rho_{k,i}}{\rho_{k,i}^{\text{max}}}\right), & \text{if } \rho_{k,i}^{\text{cap}} < \rho_{k,i} \leq \rho_{k,i}^{\text{max}} \end{cases}$$

(4)

where $\rho_{k,i}^{\text{cap}}$ is the critical density, $\tilde{V}_{f,k,i}$ is the free flow speed, $\rho_{k,i}^{\text{max}}$ is the jam density and $\tau_{k,i}$ is the time gap. Here, the time gap is the desired distance, measured in time, between two vehicles. The model parameters are specific for each sub-segment $i$ on segment $k$. $Q_e(\rho_{k,i})$ is the flow in local equilibrium, which is assumed in the CTM.

The parameters in the triangular fundamental diagram are included in the traffic state, instead of being fixed and estimated prior to applying the ensemble Kalman filter. They are thereby allowed to vary over time by adding random variations to each of the ensemble realization of the parameters. Hence, the ensemble realizations with highest probability of representing the true state, in terms of speed and density, are the ones that are most probable to correspond to the best values of the model parameters as well.

### 3.3. Calculation and application of a VSL

The VSL system will continuously adjust the speed of the connected vehicles towards the current speed limit, where the current speed limit is calculated using the VSL algorithm presented by Müller et al. (2015). Hence, the need for variable message signs to display the current speed limit is reduced. Further, instead of applying the VSL algorithm to beforehand known bottlenecks and using the target value of critical density based on prior knowledge, as suggested by Müller et al. (2015), we extend the algorithm by the following: (1) application of the VSLs for bottlenecks at arbitrary locations, in addition to application of the VSLs for beforehand known bottlenecks, and (2) application of a target value, represented by the critical density, that is estimated in real time in order to identify and adjust for changes due to for example incidents, speed limit changes, adverse weather, etc. This is possible since we make use of connected vehicle measurements at arbitrary locations and sparsely located stationary detectors. Hence, we reduce the dependence of densely located detectors and prior knowledge of the traffic system. In the VSL algorithm, the difference between the estimated critical density, $\rho_{k,i}^{\text{cap}}(n)$, and the current estimated density, $\rho_{k,i}(n)$ on the road is used to adjust the speed limit. Let $e_{k,i}(n)$ be the difference between the estimated critical density and the estimated density at time period $n$ and for sub-segment $i$ on segment $k$. The final output of the VSL algorithm is the fraction of the basic speed limit on the road to be applied for time period $n$ and for sub-segment $i$ on segment $k$,

$$b_{k,i}(n) = b_{k,i}(n-1) + K_i e_{k,i}(n),$$

(5)

where $K_i$ is an integral gain. The critical density is in our case given from the real-time calibration of the model parameters in the triangular fundamental diagram of the CTM and the current density is given from the traffic state estimation of the CCV method. The final
variable speed limit becomes

\[ VSL_{k,i-2}(n) = b_{k,i}(n) VSL_{\text{max}}. \] (6)

Here, \( VSL_{\text{max}} \) is the basic speed limit on the road. The VSL is applied to connected vehicles at the sub-segment 500 m upstream of the considered sub-segment, i.e. at sub-segment \( i-2 \), based on the approach in Müller et al. (2015). Here, the length of a sub-segment \( i \) is assumed to be 250 m. By applying the speed limit 500 m upstream of the bottleneck, the connected vehicles are able to increase the speed to the original speed before reaching the bottleneck. At the edge of the segments, the VSL is applied to sub-segments on the upstream segment, \( k-1 \).

Outline of the algorithm

Figure 2 gives an illustration of the process for one aggregation time period \( n \), starting at time \( nT \) and ending at time \( (n+1)T - \min(\Delta t_{\text{CCV}}, \Delta t_{\text{CTM}}) \). During the aggregation time period \( n \), starting at time \( nT \), the following is performed. Measurements from the connected vehicles and the detectors are collected with a frequency of \( \Delta t_{\text{CCV}} \) (Box 1 and 2) and the CTM is propagated in time with an update frequency of \( \Delta t_{\text{CTM}} \) (Box 3). At time \( (n+1)T - \min(\Delta t_{\text{CCV}}, \Delta t_{\text{CTM}}) \), the connected vehicle and the detector measurements are being processed with the CCV method, resulting in an estimate of the traffic state from the connected vehicles (Box 4). The estimate of the traffic state using the CCV method and the current estimate of the traffic state from the CTM is used in the ensemble Kalman filter to estimate the parameters of the triangular fundamental diagram (Box 5). The fraction of current speed limit is calculated based on the method by Müller et al. (2015) (Box 6). The final output is the current speed limit used to adjust the speed of the connected vehicles.

The duration of the aggregation time period, \( T \), is chosen to be long enough to give stable estimates and short enough not to smooth out important changes in the traffic conditions (Grumert and Tapani 2018).

Note that the required communication technologies and infrastructure and vehicle equipment are not defined in this study. The reason for this is to show that the method is working as long as the requirements on the data and communication is present. Hence, the equipment inside the vehicle can be a unit able of collecting the current speed and position from the vehicle. Estimations of speed, current density, critical density and VSLs are assumed to be performed from a central unit. The central unit can for example be a cloud-based service.

4. Evaluation method

The proposed VSL system requires communication of information from, and adjustment of the speed of, individual connected vehicles. Therefore, a microscopic traffic simulator that is able to describe individual vehicles in the simulated traffic stream is used to evaluate the VSL system. We use the open-source microscopic traffic simulation tool SUMO version 0.31.0 (Krajzewicz et al. 2012; DLR and contributors 2017). The connected vehicle and the detector data is accessed during the simulation through SUMO’s Traffic Control Interface (TraCI). Python scripts are used to implement the VSL algorithm and for assigning VSLs to the connected vehicles during the simulation.
Figure 2. Illustration of the traffic state estimation, the estimation of the critical density and the application of a variable speed limit for one aggregation time period \( n \). The two boxes in the bottom illustrate the collection of connected vehicle measurements and detector measurements, and the propagation of the CTM at discrete time steps, \( \Delta t_{\text{CCV}} \) and \( \Delta t_{\text{CTM}} \), respectively. This is performed during the time period \([nT, \ldots, (n+1)T - \min(\Delta t_{\text{CCV}}, \Delta t_{\text{CTM}})]\). The box in the top illustrates the filtering process and the application of a VSL in the end of the aggregation time period, at time \((n+1)T - \min(\Delta t_{\text{CCV}}, \Delta t_{\text{CTM}})\). \( T \) is representing the duration of an aggregation time period.

The simulation scenario consists of a one-directional two-lane motorway, divided into two segments with a detector in the beginning of each segment. Each segment is further divided into 10 sub-segments, resulting in twenty 250 m sub-segments. Further, two segments for loading of vehicles and two end segments are included to avoid boundary effects, resulting in a 6 km long simulated road and 2.5 km between the detectors. Figure 3 gives an illustration of the simulated road stretch. The basic speed limit on the road is 100 km/h. The simulation is performed for a 60 min period, excluding 5 min for vehicle loading. The inflow is held constant at 4000 veh/h. This is corresponding to the maximum flow level observed during peak-hours of a two-lane urban motorway in Stockholm (Grumert and Tapani 2018). To trigger the VSL algorithm, an incident is modelled by letting one vehicle slow down to
20 km/h on sub-segment 12 after 10 min. The vehicle will increase the speed again when entering sub-segment 18. This will create a backwards propagating shockwave, which in turn will result in a substantial change in the traffic conditions on the affected segments.

The measurements from connected vehicles and the information about the current variable speed limit, communicated to each connected vehicle, is updated with a frequency of 1 Hz. (\( \Delta t_{CCV} = 1 \text{ s} \)). The duration of the aggregation time period, \( T \), i.e. the time interval over which the density and speed are estimated and given to the Kalman filter, is assumed to be 60 sec based on earlier research (Grumert and Tapani 2018). The time and space discretization of the CTM are \( \Delta t_{CTM} = 5 \text{ s} \) and \( \Delta x_{CTM} = 250 \text{ m} \), respectively. This guarantee that the Courant–Friedrich–Lévy (CFL) condition (Courant, Friedrichs, and Lewy 1967) is fulfilled, i.e. \( \Delta t_{CTM} < \min_{k,i} \left( x_{k,i} / \bar{V}_f^{k,i} \right) \). The number of ensemble realizations is set to 200 in order to have reasonable simulation run times and still assure a good representation of the ensembles. See Gillijns (2006) for further discussions on the required number of ensembles. The filtering approach is similar to what has been proposed by Seo, Kusakabe, and Asakura (2015b), Tampère and Immers (2007), Wang et al. (2009) and Wang and Papageorgiou (2005). The integral gain, \( K_I \), in the VSL algorithm is set to 0.003 based on the range suggested in Müller et al. (2015) and some fine-tuning to find the best value for the considered scenario. The speed limit is increased and decreased in increments of 10 km/h. Further, the variable speed limit is assumed to be at minimum the current estimated free flow speed, \( \bar{V}_f^{k,i} \), and at maximum the allowed speed limit on the road, which is 100 km/h.

The calibrated vehicle parameters used in the SUMO model are based on measurements consisting of 15 min averages of speed and flow for a two-lane urban motorway in Stockholm (Grumert and Tapani 2018). This is resulting in a scenario which is representative for the peak-hours of an urban motorway. The simulation parameters are given in Table 1. Further, Table 2 gives the composition of vehicles and the speed distribution for each vehicle class for the simulated scenario. These values are given from the same study and are therefore assumed to be representative for an urban motorway.

### Table 1. Vehicle parameters based on the calibrated scenario (Grumert and Tapani 2018).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Default</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accel. ability (m/s²)</td>
<td>2.6</td>
<td>0.8</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Decel. ability (m/s²)</td>
<td>4.5</td>
<td>8.0</td>
<td>8.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Imperfection (range 0-1)</td>
<td>0.5</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Reaction time (s)</td>
<td>1.0</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>lcCooperative (range 0-1)</td>
<td>1.0</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>lcSpeedGain (range 0-inf)</td>
<td>1.0</td>
<td>45.0</td>
<td>45.0</td>
<td>45.0</td>
<td>45.0</td>
</tr>
<tr>
<td>lcKeepRight (range 0-inf)</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>10.0</td>
<td>10.0</td>
</tr>
</tbody>
</table>
### Table 2. Free flow speed distribution for the different vehicle classes (Grumert and Tapani 2018).

<table>
<thead>
<tr>
<th>Class</th>
<th>Allowed speed limit (km/h)</th>
<th>Vehicle length (m)</th>
<th>Mean speed (km/h)</th>
<th>Std of speed (km/h)</th>
<th>Composition (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>110</td>
<td>0-8</td>
<td>109</td>
<td>13.13</td>
<td>91</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>8-12.5</td>
<td>92</td>
<td>10.33</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>12.5-24</td>
<td>90</td>
<td>8.35</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>24-36.5</td>
<td>88</td>
<td>5.55</td>
<td>1</td>
</tr>
</tbody>
</table>

The vehicles in the simulation are generated with exponentially distributed headways. To be able to evaluate how the VSL system perform under ideal conditions, it is assumed that all vehicles are able to send and retrieve information, as well as adapt to new speed limits, i.e. a 100% connected vehicle penetration rate is assumed. The speed of the connected vehicles are continuously adjusted towards the speed limit, corresponding to the current estimate of the critical density. This means that the adjustment of the speed is applied also for the maximum speed limit. However, the speed of the connected vehicles is limited by the maximum allowed speed, which differs depending on the vehicle type. The speed of the connected vehicles is adjusted by changing the maximum allowed speed in the car-following model in SUMO (DLR and contributors 2017). This means that each connected vehicle is not strictly following the current speed limit and adjustments towards new speed limits, but it is also taking a set of rules into account in order to perform a safe drive. Both the acceptable acceleration/deceleration levels of the vehicle and the behaviour of the surrounding traffic is considered in the car-following model.

### 5. Results

The aim of the proposed system is to increase traffic efficiency. Hence, traffic operations with the VSL system activated is compared to a situation without the system, referred to as the base case. By using the same random seed in the simulation, a comparison can be done to investigate if there are any improvements with respect to traffic efficiency when applying the proposed VSL compared to the base case. The average individual travel time and the total travel time of all vehicles per simulation run are used as a measure of effectiveness. Further, the average empirical cumulative distribution function of individual vehicle travel times is utilized to investigate how the individual travel times of the vehicles are distributed in the base case and when using the proposed VSL. The presented confidence intervals are for a 95% confidence level, assuming independent and identically normal distributed results from individual simulation runs. The average and the confidence intervals are based on 50 simulation runs since this is large enough to show significant differences between the base case and the VSL case.

Overall, considering all simulation runs, the average travel time per vehicle is significantly decreased by 14.7 ± 1.5 s when applying the VSL system. This is corresponding to a decrease of 16.5 ± 1.9 h in average total travel time for all vehicles during the simulated hour. The standard deviation is large, meaning that the improvement varies a lot between the simulation runs. However, the average travel time per vehicle is reduced in all of the simulation runs, and in 78% of the simulation runs the decrease is larger than 10 s. Figure 4 shows how the
improvements in average individual travel times between the base case and the VSL case are distributed for all simulation runs.

By comparing the empirical cumulative distribution functions of individual vehicle travel times in Figure 5 we conclude that it is a difference between the two functions. The VSL case is shifted to the left showing that the individual travel times are on average shorter when applying the VSL system compared to the base case. Significant differences are concluded for all the quantiles except for the 10% and the 100% quantile. The shift gives an indication on how the travel times are affected by applying the proposed VSL system and especially it is a way to compare the distribution of longer travel times for the two cases. The difference between the base case and the VSL case is largest at travel times around 4.5 min, which is 25% above the free flow travel time. Thus, the VSL system manages to decrease the percentage of long travel times (above 4.5 min). This indicates that the VSL system is able to decrease the difference in speeds between individual vehicles and, as a result of that, a more narrow distribution of individual travel times is observed. This result confirms that the applied VSL algorithm works as intended since the main goal is to get a more homogeneous distribution of speed to increase the throughput of vehicles.

Figure 6 shows an overview of how the congestion is propagating in the network for the two different simulation runs without and with the VSL system. Both the best and the worst case in terms of travel time are given. For the base case, this means that when one vehicle is starting to move slow on segment 12 (distance 2.5 km in the figure) after 10 min, the congestion is starting to propagate backwards in space represented by darker rectangles, i.e. lower mean speeds. The propagation rate is dependent on the arrival rate of the
Figure 5. Empirical cumulative distribution functions of individual vehicle travel time for the base case and with the proposed VSL system. Significant differences are observed for all quantiles except for 10% and 100%.

vehicles approaching the congestion. Hence, since the arrival rate at the time of the congestion might differ due to the randomness in the simulations different simulation runs result in different lengths of the queue. Further, the slow moving vehicle is continuing to move with a slow speed for 6 sub-segments and an increase in speed is seen again on sub-segment 18 (distance 4.0 km in the figure). After around 20–22 min when the slow moving vehicle is increasing its speed again the congestion is starting to dissolve, leading to higher speeds propagating backwards in time. At around 25–30 min the congestion is dissolved. The time when the congestion is dissolved is depending on the queue length of the congestion. The pattern is somewhat different for the VSL case since the change in speed limit affect the congestion as well. This leads in most case to reduced congestion for the VSL case, but as can be seen in Figure 6(d) the effects of the VSL is limited for the worst case. Figure 6(e,f) show the difference between the base case and the VSL case for the two simulation runs and for all aggregation time periods and sub-segments.

As shown in Figure 6(e) the mean speed is higher for the VSL case most of the time in the simulation run with the largest improvements. The largest increase is seen in the tail of the congestion, meaning that the propagation area and the large speed decrease in the tail is reduced. This is due to that the VSL system manage to control the traffic flow at the bottleneck towards a higher capacity than in the base case, which is corresponding to the goal of the VSL algorithm by Müller et al. (2015). The red area, just after the queue is dissolved, is the result of the time it takes for the VSL algorithm to adjust the speed limit to the basic speed limit for the uncongested situation. Figure 6(f) show a simulation scenario where the VSL system does not manage to improve the conditions on the road as much. Instead a larger decrease in mean speed is seen in the tail of the congestion and for a long time after the congestion is dissolved. In situations where the VSL system is not as effective in improving the traffic efficiency, the tail of the congestion is usually the part where the algorithm fail to improve the traffic conditions. The reasons for this might be twofold, either the traffic state estimation is too simple and fail to accurately estimate the critical density; or since
The simulation run with the largest decrease in travel time is represented in the left column (a, c and e) and the simulation run with the smallest decrease is represented in the right column (b, d and f). The best and the worst run are given for the base case (top row) and for the VSL case (middle row). The random seed is the same for the base case and the VSL case for each of run. The mean speed for each sub-segment $i$ over each minute shown in the top and the middle row is represented by rectangles in the figures and is ranging from 0 to 110 km/h. The bottom row gives the difference between the base case and the VSL case. The difference is ranging from -40 to 40 km/h, where 0 means no change, a positive value an increase in mean speed with the VSL system and a negative value a decrease in mean speed with the VSL system.

Further, it is concluded that the propagation of the congestion is very different for the two simulation runs as a result of the randomness in the simulation. In the simulation run with smallest impact of the VSL system, the tail does not have the same propagation rate as in the simulation run resulting in an improvement, which is a result of the
Figure 7. The VSL applied at different points in time based on the proposed VSL algorithm for the best (a) and the worst (b) simulation run. The speed limit for each sub-segment $i$ over each minute is represented by rectangles in the figures and is ranging from 0 to 110 km/h. The lowest applied speed limit is 50 km/h.

Traffic composition in the simulation at this point in time. This indicates that in cases when the congestion can be resolved efficiently without applying VSL, the application of a VSL algorithm will result in an overreaction with too low speed levels as a result. On the other hand, the VSL system seems to be effective in cases when the backwards propagation rate of the tail is high. This corresponds well with the other simulation runs, where similar behaviour of the propagation of the tail can be observed. Also, the adjustment of the speed of the vehicles towards the current speed limit is active also during uncongested situations. This results in reduced differences in speed between vehicles which in turn reduces the faster backwards propagating shockwaves due to differences in speed between the vehicles.

Finally, to evaluate the fulfilment of the goal of the algorithm, that is to increase throughput at the bottleneck, Figure 8 shows the average difference in flow between the VSL case and the base case. The blue area in the head of the congestion, from 12 to 25 min and at a distance of 3.0 – 5.5 km in the figure, indicate a higher throughput for the VSL case compared to the base case. The higher throughput is observable over a 2.5 km long road stretch as a result of the moving bottleneck caused by the slow moving vehicle. Further, a blue area is apparent when the congestion starts to propagate backwards at 3 km after 12 min. When the congestion starts to dissolve at 25 min, the blue areas has reached 1.75 km. The width of the blue area is on average 0.5 km due to the less severe congestion and thereby a higher flow for the VSL case in this area, which is a direct result of the higher throughput at the moving bottleneck. The distinct red area shown between 25 and 33 min (at a distance of 1.75 to 5.5 km) when the congestion is resolved indicate a higher flow for the base case, which is the effect of a longer recovery period with a more severe congestion. Since the results is an average over the 50 simulation runs it is concluded that the VSL case does on average improve the throughput. The differences are significant at a 95% confidence level during the incident and for the affected segments.
Figure 8. Difference in flow between the VSL case and the base case. The difference in flow for each sub-segment $i$ over each minute is represented by rectangles in the figure and is ranging from $-1000$ to $1000$ veh/h, where 0 means no change, a positive value a larger flow with the VSL system and a negative value a smaller flow with the VSL system. The results are an average over the 50 simulation runs.

6. Conclusions

We propose a VSL system including connected vehicles for bottleneck mitigation. The contribution of the system is both related to the design of the algorithm and the way of applying the variable speed limits. By the use of connected vehicles, the dependence on stationary detectors and variable message signs can be greatly limited without losing the ability to estimate the current traffic state at arbitrary locations. An estimate of the traffic state based on the method by Grumert and Tapani (2018) and an ensemble Kalman filter are used to find the current critical density, which is utilized as a target value in the VSL control strategy based on the approach by Müller et al. (2015). The difference, compared to the approach by Müller et al. (2015), is that the target value is allowed to change over time and in space. Hence, the control strategy can be applied at arbitrary locations and thereby also for non-recurrent bottlenecks. The VSL control strategy is applied by continuous control of the speeds of the connected vehicles towards the speed limit. The improvements make it possible to apply a VSL system at locations where it today is impossible to apply VSL systems without the use of densely placed detectors and variable message signs or a beforehand known bottleneck location.

The results indicate that, by using connected vehicles to find the current critical density at an arbitrary location and then adjusting the speeds of the connected vehicles towards a suitable speed limit, the traffic performance can be increased. The individual travel times of the connected vehicles, as well as the total travel time for all vehicles, are significantly reduced in all simulation runs. The VSL system manage to reduce the negative consequences of an incident and also to improve congested situations by reducing the difference between vehicle speeds due to the adjustment of the speeds of the connected vehicles. A
direct result of this is the higher throughput at the head of congestion, showing that the control algorithm works as intended. As a conclusion, it is shown that even for a simple road design, traffic efficiency can be increased by using connected vehicles in a VSL system.

In this study, we use a simple road design to investigate the properties and potential of the proposed system. The next step is to test the method for a more complex road design. This will show if the proposed VSL system has potential to improve traffic efficiency also in more complex scenarios. The applied VSL control algorithm can also be subject to further studies, both with respect to the sensitivity of the results to the parameter values in the VSL control algorithm, and by considering other VSL algorithms. It should be noted that the effectiveness of the VSL control algorithm is highly dependent on the estimate of the traffic state and the critical density. Hence, a more complex model of the traffic state can lead to larger benefits. In the study, a connected vehicle penetration rate of 100% is assumed in order to study the potential of the VSL system under ideal conditions. The traffic performance for other connected vehicle penetration rates has to be investigated to evaluate how large the benefits of the VSL system can be for a mixed traffic flow, including both connected and non-connected vehicles. The density and speed estimation is done using the method proposed by Grumert and Tapani (2018). However, if a lower penetration rate of connected vehicles is assumed, the speed distribution of the connected vehicles may differ from the speed distribution of the regular vehicles due to the adjustment of the speeds of the connected vehicles. Hence, assumptions in the traffic model used in the ensemble Kalman filter have to be changed to take this into account. The critical density can be estimated at arbitrary locations and is allowed to change over time. This makes the method suitable, not only for VSL systems, but for other traffic management applications where identification of non-recurrent bottlenecks is required, i.e. incident detection. Finally, more research is needed on the technology and equipment required. Both with respect to the precision in the required data, but also with respect to the required type of equipment and the limitations coming with that.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors would like to thank the Swedish Transportation Administration (Trafikverket) through the Center for Traffic Research for the financial support.

ORCID

Ellen F. Grumert https://orcid.org/0000-0001-5531-0274
Andreas Tapani https://orcid.org/0000-0001-9635-5233

References


